Article Semantics in Second Language Acquisition

by

Tania Ruth Ionin

B.A. Linguistics and Psychology
University of Michigan, 1998

Submitted to the Department of Brain and Cognitive Sciences in Partial Fulfillment
of the Requirements for the Degree of

Doctor of Philosophy

at the

Massachusetts Institute of Technology

September 2003

© Massachusetts Institute of Technology 2003

Signature of Author:_____
Department of Brain and Cognitive Sciences
July 22, 2003

Certified by:_____
Kenneth Wexler
Professor of Psychology and Linguistics
Thesis Supervisor

Accepted by:_____
Earl Miller
Professor of Neuroscience
Chairman, Department Graduate Committee
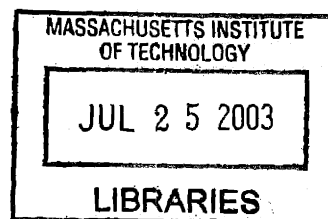
Article Semantics in Second Language Acquisition

by

Tania Ruth Ionin

Submitted to the Department of Brain and Cognitive Sciences
on July 22, 2003 in Partial Fulfillment of the
Requirements for the degree of Doctor of Philosophy in Cognitive Science

ABSTRACT

This thesis examines article choice and parameter-setting in second language (L2) acquisition. It argues, on the basis of L2-English elicitation and production data, that L2-learners have access to UG-based semantic distinctions governing article choice, but do not know which distinction is appropriate for English. A *Fluctuation Hypothesis (FH)* is proposed, according to which L2-learners fluctuate between different parameter settings until the input leads them to set the parameter to the target value.

The thesis proposes that articles cross-linguistically may encode definiteness or specificity. The definition of *specificity* that is adopted is based on Fodor and Sag's (1982) view of specificity as *speaker intent to refer*. The behavior of referential *this*, a specificity marker in colloquial English, is examined, and it is proposed that the definition of specificity incorporates the concept of *noteworthy property*. An *Article Choice Parameter* is next proposed, which governs whether articles in a given language are distinguished on the basis of definiteness or on the basis of specificity. While English has the Definiteness setting of this parameter, it is suggested, on the basis of data from Mosel and Hovdhaugen (1992), that Samoan has the Specificity setting.

It is hypothesized, in accordance with the FH, that L2-learners fluctuate between the two settings of the Article Choice Parameter. This hypothesis leads to the prediction that L2-English errors of article use should come in two types: overuse of *the* with specific indefinites and overuse of *a* with non-specific definites. These predictions are examined in a series of studies with adult speakers of Russian and Korean, two languages with no articles. The empirical data confirm the predictions, and show that L2-English article choice is not random but reflects access to the two settings of the Article Choice Parameter. The same patterns of results are found for L1-Russian and L1-Korean speakers, and it is shown that the results are not attributable to L1-transfer.

On the basis of these findings, it is concluded that L2-learners have direct UG-access to semantic distinctions underlying article choice. The data also provide evidence for the existence of a specificity distinction which cross-cuts the definiteness distinction.

Thesis Supervisor: Kenneth Wexler
Title: Professor of Psychology and Linguistics

Acknowledgements

Table of Contents

Chapter 4: Experiment 1

Chapter 5: Experiment 2

Chapter 6: Experiment 3

Chapter 1: Introduction

1. Goals of this dissertation

The main goal of this dissertation is to examine the relationship between linguistic theory and second language (L2) acquisition in the domain of article semantics. The specific goals are two-fold. The first goal is to examine the issues of UG access and parameter-setting in L2-acquisition in the domain of article choice. It is proposed that L2-learners have access to UG parameters, and, in the absence of sufficient input, undergo fluctuation between different parameter settings, including those not instantiated in either the L1 or the L2. It will be shown that data from L2-English article choice provide support for this proposal.

The second goal is to use L2-English data as a means of investigating article semantics, and to argue that *definiteness* and *specificity* are two independent features that enter into the lexical specifications of articles. This dissertation proposes that *specificity as speaker intent to refer* (cf. Fodor and Sag 1982) is morphologically encoded cross-linguistically and that it cross-cuts the definiteness distinction. L2-English data provide support for this proposal: it is shown that L2-learners use articles to distinguish between specific and non-specific indefinites, as well as between specific and non-specific definites.

The empirical data for this thesis come from three studies with adult L2-English learners: one study with L1-Russian learners of English, and two studies with both L1-Russian and L1-Korean learners of English. The studies involve elicitation tasks as well as a collection of written production data.

2. Some background: Universal Grammar in L2-acquisition

Young children almost always succeed in learning their first language (L1): in the absence of physiological abnormalities, children nearly always attain full mastery of their native language. Yet children frequently do not receive explicit instruction in their L1, and the L1 input is deficient, underdetermining the target grammar. This "poverty of the stimulus" problem led Chomsky (1965) to propose that L1 acquisition is constrained by Universal Grammar (UG), which restricts the range of possible grammars. In the *Principles and Parameters* framework (Chomsky 1981), UG is hypothesized to consist of universal *principles*, which are operative in all languages, and *parameters,* aspects on which languages vary. Children are born with the knowledge of principles and parameters "built in"; they then use the input to establish which principles are parameter settings are instantiated in their L1.

A question asked by much L2-literature is whether L2-acquisition is similarly constrained by UG. While children nearly always succeed in fully acquiring their L1, adult L2-learners often do not acquire their L2 fully. For instance, Johnson and Newport (1989, 1991) found that the older L2-learners were at the start of acquisition, the less accurate they were on a variety of tests of English. L2-English learners who had arrived in the US (and hence begun intensive exposure to English) as children performed better than those L2-learners who had arrived as adolescents, and these in turn performed better than those learners who had arrived as adults. Johnson and Newport argued for a *critical period* for L2-acquisition. On this view, the ability to acquire a language declines with

age. However, other researchers (e.g., Birdsong 1992, White and Genesee 1996, Birdsong and Molis 1998) have found near-native performance among adult L2-learners, and/or absence of age effects on acquisition, casting doubt on the Critical Period Hypothesis for L2-acquisition. (For more work on the subject of age effects in L2-acquisition, see the papers in Singleton and Lengyel 1995).

The question of whether adult L2-acquisition is constrained by UG is far from uncontroversial. I will now give a brief overview of the major positions on UG-access in L2-acquisition, and discuss such issues as parameter setting, transfer, and optionality. I will argue that there is compelling evidence in favor of UG-access in L2-acquisition.

2.1. Models of L2-acquisition: UG-access

Proposals concerning UG access in L2-acquisition are often grouped (e.g., by Flynn 1996; White 1996, 2000) into three major types: the No Access position, the Partial Access position, and the Full Access position.

2.1.1. The No Access position

Researchers who have taken the No Access position (e.g., Bley-Vroman 1989, Clahsen 1988, Clahsen and Muysken 1986) argue that L1 and L2 acquisition are fundamentally different processes which cannot be accounted for in the same framework. As Flynn (1996:122) points out, these researchers interpret any differences between L1 and L2 acquisition as evidence that the two processes are fundamentally different.

In early work in favor of the No Access position, Clahsen and Muysken (1986) argued that L2-learners do not have a parameter-based grammar. Studying the acquisition of L2-German, Clahsen and Muysken argued that L2-learners have an "unnatural" grammar which allows non-finite verbs to move to the end of the sentence, while the natural L1-grammar has finite verbs moving leftwards. Clahsen and Muysken argued that L2-learners lack parameters associated with word order. For arguments against this position, see du Plessis, Solin, Travis and White (1987) and Schwartz and Tomaselli (1990), who argue that word order in L2-German is in fact UG-constrained. Later work arguing for UG-constraints on L2-German word order includes Vainikka and Young-Scholten (1994, 1996) and Schwartz and Sprouse (1994, 1996).

Another proposal for No Access in L2-acquisition is Bley-Vroman (1989), who proposed the *Fundamental Difference Hypothesis*. According this hypothesis, L2-learners have no UG access but rely on general learning strategies, including distributional analysis and analogy. For arguments against this hypothesis, see Epstein, Flynn and Martohardjono (1996), among others.

The No Access position has difficulty accounting for grammatical knowledge of the part of L2-learners, such as knowledge of Structure-Dependence. In a small study with L1-Japanese learners of English, Otsu and Naoi (1986) found that these learners exhibited knowledge of structure-dependence when forming English questions. Since Japanese does not have movement in questions, the results cannot be due to L1-transfer.

When forming questions from sentences such as (1a), the learners in Otsu and Naoi's study correctly formed questions of the form of the form in (1b) rather than (1c). If the learners were merely analogizing from simple question formation (e.g., *The boy is laughing* becomes *Is the boy laughing?*), they would have no reason to prefer (1b) over (1c). In fact, on a strategy-based view, the learners might well use the strategy "move the

first auxiliary to the front of the sentence in order to form a question" – in which case they would use (1c) rather than (1b). Otsu and Naoi's findings provide evidence that L2-learners pay attention to the underlying structure rather than the surface string – a finding that cannot easily be explained under the No Access view.

1. a) The boy who is in the room is laughing.
    b) Is the boy who is in the room ___ laughing?
    c) *Is the boy who ___ in the room is laughing?

It should be noted that the participants in this study were adolescent L2-learners, so the results may not have anything to say about adult L2-acquisition. However, there is much evidence for access to UG principles and parameters on the part of adult L2-learners as well. I discuss some of this evidence in the next section.

2.1.2. The Partial Access position

According to the Partial Access position (e.g., Schachter 1989, 1990), L2-learners have access only to those principles and parameter settings of UG that are instantiated in their L1. Evidence for this position comes from Schachter's work on Subjacency in L2-acquisition. For instance, Schachter (1990) looked at whether adult speakers of Dutch, Chinese, Indonesian and Korean could detect Subjacency violations in English sentences such as (2).

2. a) *What did Susan visit the store that had in stock?
    b) *Who did the Senator ask the President where he would send?

The native languages of the L2-learners' in this study differ in how they instantiate Subjacency. According to Schachter (1990, 1996), Subjacency is instantiated the same way in Dutch as it is in English. In Chinese and Indonesian, Subjacency effects are more limited, and Korean shows no Subjacency effects at all. Schachter found that the degree to which Subjacency is instantiated in the learners' L1 had a direct relationship to the learners' ability to detect Subjacency violations in English. Dutch speakers performed as well as L1-English controls; Chinese and Indonesian speakers did not reject all of the ungrammatical sentences such as (2); and Korean speakers performed completely randomly on the ungrammatical sentences. Schachter (1990, 1996) argues that these results show lack of direct UG access in L2-acquisition: a principle, such as Subjacency, constrains a learner's L2-grammar only when it is already incorporated in the L1-grammar. Since Subjacency is not incorporated into the L1-grammar of Korean speakers, these speakers' L2-grammar is not constrained by Subjacency.

However, later studies provided evidence that L2-learners do show knowledge of Subjacency even when their L1's are not constrained by this principle. An example is the study of Martohardjono (1992, 1993), who tested Chinese and Indonesian speakers on their knowledge of Subjacency violations in questions (3). Chinese and Indonesian do not have overt wh-movement, and the equivalents of (3) in these languages are grammatical: no overt wh-movement is involved, and Subjacency is not violated.

3. a) *Which man did Tom fix the door that ___ had broken?
    b) *Which mayor did Mary read the book that praised ___?

Martohardjono found that these L2-learners were very accurate at identifying ungrammatical wh-questions like (3): for the L1-Indonesian speakers, in particular, the

accuracy rate was around 90%. Since the learners' knowledge of Subjacency violations in sentences like (3) cannot be coming from the L1, these results are problematic for the Partial Access hypothesis. Similarly, White and Juffs (1998) found that proficient adult L2-English learners whose L1 was Chinese did not differ significantly from native English speakers in detecting Subjacency violations. These findings provide evidence that adult L2-learners have access to aspects of UG that are not instantiated in their L1.

2.1.3. The Full Access position

Researchers who take the Full Access view of L2-acquisition (e.g., Epstein et al. 1996, 1998; Flynn and Martohardjono, 1994; Flynn, 1996; Schwartz and Sprouse 1994, 1996) argue that L2-acquisition, like L1-acquisition, is fully constrained by UG. Support for this view comes from a number of studies which show L2-learners successfully acquiring principles and parameter settings which are not instantiated in their L1. I have already discussed evidence for knowledge of UG principles (Structure-Dependency, Subjacency) in the previous sections. In this section, I will present some data showing that L2-learners are able to acquire new parameter settings as well.

Early work by Flynn (1983, 1984, 1987; summarized in Flynn 1996, Epstein et al. 1996) showed that Japanese-speaking L2-English learners are able to assign a new value to the head-direction parameter. Flynn used an elicited imitation and comprehension study to test earners on preposed and postposed subordinate clauses, as in (4). Preposed subordinate clauses correspond to head-final structures, and postposed subordinate clauses – to head-initial structures.

4. a) When the actor finished the book, the woman called the professor.
    b) The worker called the owner when the engineer finished the plans.

English is head-initial while Japanese is head-final. Therefore, according to Epstein et al. (1996:687), "if L2-learners had access to their L1 parameter values alone, then the Japanese speakers tested in this study would have access only to a head-final parameter value. If this were so, we would expect to find no evidence that these learners were able to identify and assign a new value to the head-direction parameter for the L2... We might even expect that those structures that follow from the L1 parameter setting would be more accessible to the Japanese learner than those that follow from the L2 parameter setting; that is, they might show a preference for preposed sentence structures rather than postposed sentence structures." However, this was not the case. Flynn found that Japanese adult L2-English learners did not find preposed structure significantly easier than postposed structures to either imitate or comprehend. Moreover, L2-learners with the highest level of proficiency tested showed a significant preference for postposed sentence structures, suggesting "that these L2-learners had assigned a value to the head-direction parameter in conformity with the English value. These results suggest that UG remains available to the L2-learner" (Epstein et al. 1996:687). (See Section 2.2.2 for more discussion of this parameter in L2-acquisition).

Evidence for the acquisition of new parameter values exists in other domains as well. On example comes from the domain of reflexive binding. A number of studies, including Finer and Broselow (1986) and Thomas (1991), found that speakers of Korean and Chinese, which allow long-distance reflexive-binding, correctly disallowed long-distance reflexive binding in their L2, English. This evidence suggests that L2-learners were able to reset the Governing Category Parameter related to reflexive binding (see Section 2.3

for more discussion, as well as evidence of L2-learners accessing multiple parameter options related to reflexive-binding).

2.2. The role of transfer in L2-acquisition

Under the Full Access position, L2-learners are like L1-learners in that their grammar is UG-constrained. However, L2-acquisition differs from L1-acquisition in one very obvious way: L2-learners already know one language. Thus, discussion of UG-access in L2-acquisition typically involves a discussion of L1-transfer as well.

As seen in the previous section, under the Partial Access view, L1-transfer determines L2-acquisition: L2-learners have access only to those aspects of UG that are instantiated in the L1. The Full Access view, on the other hand, predicts that L2-learners can access principles and parameter settings that are not part of the L1. However, the Full Access view does not preclude the possibility of L1-transfer: it is in principle possible that L2-learners transfer all of the principles and parameter settings from their L1, and subsequently reset the parameters to the L2-values (this is the position of Schwartz and Sprouse 1994, 1996, among others). A different possibility, argued for by Epstein et al. (1996), and Flynn (1996), among others, is that while transfer is seen in some aspects of L2-acquisition, it is not operative across the board: for instance, languages-specific aspects such as lexical learning are not subject to transfer (see below for more discussion).

The evidence concerning L1-transfer in L2-acquisition is mixed. I now discuss evidence for transfer as well as evidence for lack of transfer in L2-acquisition.

2.2.1. Evidence for transfer in L2-acquisition

Schwartz and Sprouse (1994, 1996) argue that the initial state of L2-acquisition is fully determined by L1-transfer; parameters are subsequently reset from the L1 to the target L2 value. Schwartz (1998) argues that support for this position comes from the findings of Hulk (1991). Hulk found that beginner L1-Dutch learners of French transferred Dutch word order onto French and treated French as an SOV/V2 language. The tendency to treat French as an SOV/V2 language decreased with proficiency, suggesting that the learners were able to reset the parameter governing verb placement to the French value.

Robertson and Sorace (1999) similarly found transfer of the V2 constraint among L1-German learners of English. They found that advanced L2-learners sometimes produced sentences like (5a) instead of (5b), which suggests a residual V2-constraint in the learners' L2-English. A grammaticality judgment test with L1-German learners of English showed that a minority of the learners preferred ungrammatical V2 order to grammatical English order; the proportion of individuals showing this preference declined with the number of years of English instruction (Robertson and Sorace 1999:333).

5. a) For many kids is living with their parents a nightmare.
    b) For many kids living with their parents is a nightmare.

Other areas in which both transfer and UG access have been argued to take place are placement of adjectives (see Parodi, Schwarz and Clahsen 1997, Schwartz 1998) and placement of negation (see Sprouse and Schwartz 1998).

Vainikka and Young-Scholten (1994, 1996) take a different view, arguing for only partial transfer from the L1. Under their *Minimal Trees* model, L2-learners transfer lexical categories but not functional categories from their L1. For instance, learners transfer the VP from their L1, but have to gradually build the higher functional projections. Vainikka and Young-Scholten (1994, 1996) provide evidence that L2-learners of German, a head-final language, transfer VP-headedness from their L1 – thus, L1-Italian and L1-Spanish speakers initially treat German as a head-initial language, while L1-Korean and L1-Turkish speakers treat it as head-final from the start. On the other hand, Vainikka and Young-Scholten argue, L2-learners do not exhibit evidence of transfer of functional projections.

2.2.2. *Evidence for lack of transfer in L2-acquisition*

On the other hand, some other domains of L2-acquisition do not show evidence of L1-transfer at all. An example is the work of Flynn on the CP-headedness parameter, discussed in Section 2.1.3 above. L1-Japanese speakers did not appear to transfer the setting of this parameter onto English. Additional evidence that Japanese speakers do not transfer the setting of this parameter onto English is shown by the work of Flynn (1987) and Flynn, Foley and Lust (2000) on relative clause formation.

Spanish is a head-initial language like English; as in English, relative clauses in Spanish follow the head noun. In contrast, Japanese is head-final, and the relative clause precedes the head noun. Flynn (1987) and Flynn et al. (2000) found that transfer was operative among L1-Spanish speakers, who produced lexically headed relative clauses in the appropriate position. Japanese speakers, on the other hand, did not immediately acquire lexically headed relative clauses in English, but also did not show transfer from Japanese: instead, they went through a developmental stage during which they used free relatives in place of lexically headed relative clauses. The same developmental pattern has been attested for L1-English learners (Flynn and Lust 1980). See Section 3.2.2 for more discussion.

Another domain in which there is evidence for lack of L1-transfer is L2-acquisition of control verbs like *promise, tell,* and *remind.* Flynn, Foley and Lardiere (1991) tested L1-Spanish speakers on their acquisition of English control verbs. They tested L2-learners in an elicited imitation task with sentences such as (6): structures in which the control verb takes a finite-clause complement (6a) and structures where it takes an infinitival complement (6b).

6. a) John promises/tells/reminds Henry that he will go to the store.
   b) John promises/tells/reminds Henry to go to the store.

Spanish, like English, allows both finite and infinitival clause complements for *promise*, but only finite clause complements for *tell* and *remind*. If L1-transfer takes place in this domain, then Spanish speakers should find finite clause complements easier to acquire than infinitival clause complements for the verbs *tell* and *remind* in English. However, Flynn et al. found that the learners showed significant preference for infinitival complements over finite clause complements with all three verbs. Similar results were obtained with L1-Japanese and L1-Chinese speakers.

Earlier studies of control verbs in L2-acquisition (e.g., Cooper, Olshtain, Tucker and Waterbury 1979, among others) found that L2-learners in early stages of acquisition interpreted subject control verbs (e.g., *promise*) as if they were object control verbs (e.g.,

tell). Interestingly, both preference for object control verbs and preference for infinitival complements have been attested for L1-acquisition (see Sherman and Lust 1993 for an overview). Thus, in the domain of control verbs, L2-learners appear to follow the same developmental pattern as L1-learners, rather than relying on the properties of their L1.

Other domains in which L1-transfer does not appear to play a role are verb-raising and reflexive binding. In these domains, there is an additional interesting property of L2-learners apparently accessing parameter settings that are coming from neither the L1 nor the L2. I discuss this in the next section.

2.3. Access to non-L1/non-L2 parameter settings in L2-acquisition

There is evidence from some domains of L2-acquisition that L2-learners access UG parameter options that are coming from neither L1-transfer nor L2-input. The most robust evidence for this comes from the domains of verb-raising and reflexive binding, which I will now discuss in some detail.

2.3.1. Evidence from verb-raising

Much work on L2-acquisition of verb-raising is based in the framework of Pollock (1989). Pollock investigated the difference between English and French illustrated in (7) and (8): in English, verbs follow negation and adverbs, while in French, verbs precede negation and adverbs.

7. a) Mary does not like Jack / *Mary likes not Jack
    b) Mary often eats soup / *Mary eats often soup

8. a) Marie n'aime   pas Jacques / * Marie ne pas aime Jacques
        M.   neg-likes   not J.
    b) Marie mange souvent   la soup / * Marie souvent mange la soup
        M.   eats   often   the soup

Pollock proposed that there is parametric variation in verb raising: in French, verbs move to INFL, while in English, lexical verbs stay within the VP. This is illustrated in (9) for adverbs. The same holds for negation, which is higher in the tree than adverbs.
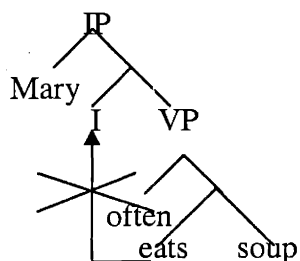
9. a) English   b) French
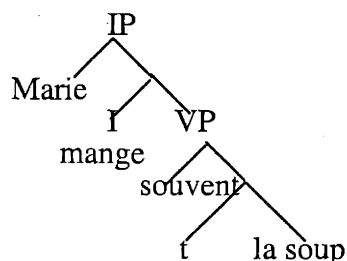
We can thus think of a *verb raising parameter*: in languages with the positive setting of this parameter, such as French, verbs raise out of the VP, and in languages with the negative setting, such as English, lexical verbs stay *in situ*. In the Minimalist framework (Chomsky 1993, 1995), this means that a functional head is strong if verbs move to it and weak if they do not.

There is additional evidence that multiple functional heads underlie verb-movement. There is evidence that non-finite verbs (infinitives and participles) in French and Italian move out of the VP to a head below INFL (see Pollock 1989, Belletti 1990, Cinque 1999; see also Pesetsky (1989) for a proposal of short verb movement in English). Thus, there may be more than one verb raising parameter: if multiple functional heads exist cross-linguistically (per Cinque 1999), then it is possible that the feature strength of each functional head is subject to parametric variation.

L2-researchers starting with White (1990/91, 1992) have been interested in how L2-learners acquire the verb raising properties of their target language. In her studies with adolescent L1-French speakers acquiring English as an L2, White found evidence for verb raising, specifically, verb-adverb order, in L2-English. Since French is a [+verb raising] language, this finding was naturally attributable to L1-transfer.

However, later researchers found that optional verb raising in L2-English cannot be attributed to transfer. Eubank, Bischof, Huffstutler, Leek and West (1997) found that speakers of Chinese, a language with no verb raising, who were acquiring English, another language with no verb raising, nevertheless allowed verbs to be placed before adverbs in English.[1] In Ionin and Wexler (2002), we found the same to be true for L1-Russian learners of English, even though neither Russian nor English has verb-raising; we found that the learners optionally raised both finite verbs and participles past adverbs, but not past negation.

The above evidence suggests that L2-learners optionally allow verb movement out of the VP even when this is not allowed in either the L1 or the L2. The finding that L2-learners do not raise verbs past negation suggests that verb-raising in L2-English is to an intermediate functional head, which is above (VP-attached) adverbs but below negation (see Section 3.2.4 for more discussion).

2.3.2. Evidence from reflexive binding

Another domain in which L2-learners appear to access options which are coming from neither the L1 nor the L2 is reflexive binding.

Much work in L2-acquisition of reflexives concerns binding domains. Wexler and Manzini (1987) proposed the Governing Category Parameter (GCP), which captures cross-linguistic differences in binding domains. The GCP is given in (10) (from Wexler and Manzini 1987:419). It provides five possibilities for governing categories cross-linguistically.

10. γ is a governing category for α if γ is the minimal category that contains α and has
   a) a subject; or
   b) an Infl; or
   c) a Tense; or
   d) a 'referential' Tense; or
   e) a 'root' Tense

---

[1] On the other hand, Yuan (2000) found that L1-English speakers learning Chinese as an L2 (i.e., the reverse scenario of Eubank et al. 1997) did *not* allow verb-raising in Chinese. This fact has not been explained.

English has the most restrictive setting of this parameter (10a): an anaphor must be found within the smallest domain containing a subject. Thus, English allows *Mary* to bind *herself* in sentences like (11a), but not in sentences like (11b) or (11c). In (11b-c), there is an intervening subject (*Tom*) between *Mary* and the anaphor, so *Mary* cannot bind *herself* and the sentences are ungrammatical.

Other languages have less restrictive settings of the CGP. For instance, Finer and Broselow (1986) proposed that Russian has the setting in (10c). Russian allows the equivalents of both (11a) and (11b), but not of (11c): an anaphor must be bound within the smallest tensed clause, so the subject of the matrix clause can bind an anaphor inside an infinitival embedded clause (11b) but not an anaphor inside a finite embedded clause (11c). Finally, the Japanese anaphor *zibun* is subject to the least restrictive setting (10e). The equivalents of all of the sentences in (11) are grammatical in Japanese, since the anaphor is bound inside the root clause in all of them. The same holds for anaphors in Korean and Chinese.

11. a) Mary painted herself.
    b) Mary asked Tom to paint herself.
    c) Mary said that Tom painted herself.

Investigations into L2-acquisition of reflexive binding have found that L2-learners access options of the GCP instantiated in neither their L1 nor their L2. Finer and Broselow (1986) as well as Finer (1991) found that Japanese and Korean speakers acquiring English as an L2 usually disallowed non-local binding in sentences with the structure in (12a)[2]: only *Mary* would be allowed to bind *herself* in (12a). Thus, the learners were not transferring the L1-value of the GCP (10e) onto English. However, the learners were often allowed non-local binding in sentences like (12b), allowing *Sarah* rather than *Mary* to bind the anaphor. Thus, they were also not correctly choosing the L2-value of the GCP (10a). Finer and Broselow (1986) and Finer (1991) proposed that the learners were choosing the intermediate Russian value (10c).

12. a) Sarah said that Mary painted herself.
    b) Sarah told Mary to paint herself.

On the other hand, Hirakawa (1990), who tested Japanese speakers acquiring English, found that these L2-learners allowed non-local binding in both (12a) and (12b), although they allowed it to a greater extent with infinitival embedded clauses (12b) than with finite embedded clauses (12a). Hirakawa argued that learners transfer the L1 parameter setting onto their L2.

Finally, Thomas (1989b, 1991) examined reflexive binding among L1-Spanish as well as L1-Japanese and L1-Chinese speakers. Since Spanish allows only local binding, like English, L1-transfer should lead Spanish speakers to correctly allow only local binding in sentences such as (12) (i.e., only binding by *Mary* and not by *Sarah* in (12)).

Nevertheless, the L1-Spanish speakers allowed some non-local binding in sentences like (12a). Their acceptance rates of non-local binding in this structure type did not differ significantly from the acceptance rates of L1-Japanese and L1-Chinese speakers.

---

[2] The sentences in (12) are not actual examples from any of the studies cited here. They are simply illustrations of the structure types used in these studies.

To sum up, there is evidence that L2-learners access settings of the GCP that are not coming from either their L1 or their L2: L1-Korean and L1-Japanese learners of English allow the "Russian" setting, and L1-Spanish learners of English allow the "Japanese" setting. Thus, learners appear to access three different GCP settings: (10a), (10c), and (10e), regardless of which settings are instantiated in their L1. (The other settings of the GCP, (10b) and (10d), have not been examined for L2-acquisition).

2.3.3. Evidence from other domains

While verb-raising and reflexive binding present the clearest cases of non-L1/non-L2 parameter settings in L2-acquisition, there is suggestive evidence from other domains as well. White (2000) suggests, interpreting the findings of Clahsen and Hong (1995), that another domain in which L2-learners may access non-L1/non-L2 parameter settings is the null subject parameter. Clahsen and Hong found that Korean-speaking learners of German do not reset the null subject parameter from the [+null subject] to the [-null subject] setting. Specifically, there is no relationship between the acquisition of agreement and the loss of null subjects in the L2-data, while this relationship is observed in L1-German acquisition data. White suggests that the L2-learners who failed to acquire the German setting of the null subject parameter have chosen the Italian setting instead – the setting which allows both rich agreement and null subjects (this possibility is not discussed in Clahsen and Hong, but is White's interpretation of their data).

Finally, Broselow and Finer (1991) investigated the phonological Minimal Sonority Distance parameter, and argued that Korean and Japanese speakers learning English adopted a parameter value that is more marked than the setting in their L1's but less marked than the one required by their L2. (See White 1996 for more discussion).

2.4. Summary: parameter-setting in L2-acquisition

The main points of the brief review in this section can be summarized as follows:

I. There is evidence that L2-learners have access to UG principles and parameter settings that are not instantiated in their L1.

Evidence for Point 1 comes from a variety of domains including wh-movement, word order, reflexive binding, and acquisition of control verbs. Given this evidence, I will from now on assume the Full Access position for L2-acquisition rather than the Partial or No Access positions. The Partial Access position cannot explain how L2-learners come to acquire aspects of the L2 not instantiated in their L1. The No Access view accounts for L2-knowledge in terms of general cognitive strategies and does not make testable predictions for L2-acquisition (see Chapter 8 for specific arguments against strategy-based explanations in the domain of article choice). The Full Access view, on the other hand, makes testable predictions, since it predicts L2-acquisition to always be UG-constrained.

II. Transfer of parameter setting exists for some (but not all) domains of L2-acquisition, but parameter resetting is possible.

The evidence concerning transfer in L2-acquisition is fairly mixed, as discussed above. However, some general patterns appear to emerge. First, transfer does not appear to take place in domains concerned with the specifications of lexical items. As shown by the study of Flynn et al. (1991), L2-learners do not transfer the lexical specifications of control verbs like *promise* and *tell* from their L1 onto their L2, even in those cases where the L1 and the L2 match. Similarly, the specifications of inflectional morphemes do not appear to be subject to transfer: speakers of inflectionally rich languages at least initially omit verbal inflection in their L2 (see, e.g., the data in Prévost and White (2000) for omission of finiteness among a Spanish speaker and a Portuguese speaker acquiring L2-German).

In domains connected to directionality and syntactic movement, the data are more mixed. There is evidence that transfer is operative for VP-headedness (e.g., Vainikka and Young-Scholten 1994, 1996) but that it is not operative for CP-headedness except in those cases where the L1 and L2 values match (e.g., Flynn 1987, Flynn et al. 2000). There is evidence for transfer in the domain of noun-adjective order (Parodi et al. 1997), but not in the domain of verb-adverb order (e.g., Eubank et al. 1997). Given the conflicting evidence, I will not be taking a position concerning transfer in this work. I will look at the acquisition of articles by learners of article-less languages: a domain where transfer cannot be at work, since learners do not have articles in their L1 to help guide their acquisition of articles in their L2 (see Chapter 3 for more evidence that transfer does not indeed play a role).

III. In some domains of L2-acquisition, there is evidence that L2-learners have access to parameter settings instantiated in neither their L1 nor their L2.

This is a particularly relevant point for my proposal. Not only do L2-learners often fail to transfer the L1 parameter setting to the L2 (see above), but they sometimes access parameter settings that are instantiated in neither the L1 nor the L2. As discussed in the previous section, this is particularly visible in the domains of verb-raising (e.g., Eubank et al. 1997, Ionin and Wexler 2002) and reflexive-binding (e.g., Finer and Broselow 1986, Finer 1991, Thomas 1989b, 1991, Hirakawa 1990). In this thesis, I will show that article choice is another area in which access to non-L1/non-L2 settings takes place. In formulating my proposal, I will be particularly concerned with capturing the fact that L2-learners have access to non-L1/non-L2 settings.

2.5. Optionality in parameter-setting

A final issue that I will consider before formulating my proposal is *optionality* in L2-acquisition. This issue has attracted attention in recent L2-literature (e.g., Eubank 1993/94, Prévost and White 2000, Sorace 2000). The type of optionality that I will be particularly concerned with here is *optional adherence to parameter settings*. By this I mean cases when L2-learners' verbal behavior seems to reflect more than one setting of some Parameter X at the same time: i.e., when during the course of the same study, learners show evidence of adherence to Setting 1 of Parameter X some of the time, to Setting 2 of Parameter X some of the time, to Setting 3 some of the time, and so on.

Some of these cases involve parameter resetting from the L1 value to the L2 value: L2-learners' behavior in such cases is neither 100% consistent with the L1 parameter

setting, nor 100% consistent with the L2 parameter setting. For instance, speakers of SVO languages who are acquiring SOV languages go through a stage during which they use both SVO and SOV constructions (see Vainikka and Young-Scholten 1996:15). Similarly, speakers of V2 languages who have acquired a non-V2 language produce sentences which exhibit V2 word orders (e.g., (5a)) alongside target-like sentences which do not exhibit V2 word order (Robertson and Sorace 1999). Robertson and Sorace (1999:333) also show that in a grammaticality judgment test, "the majority of learners do not make their judgments consistently in conformity with the dictates of one grammar or the other" (i.e., of either a V2 or a non-V2 grammar).

Optionality in L2-acquisition has been much discussed in the domain of verb-raising. As discussed in the previous section, L2-learners sometimes raise verbs out of the VP, past adverbs (though not past negation), and sometimes leave them *in situ*. This optional verb-raising past adverbs occurs when speakers of a [+verb-raising] L1 acquire a [-verb-raising] L2 (White 1990/91, 1992); when speakers of a [-verb-raising] L1 acquire a [+verb-raising] L2 (Beck 1998); and when speakers of a [-verb-raising] L1 acquire [-verb-raising] L2 (Eubank et al. 1997, Ionin and Wexler 2002). In the last case, learners optionally adhere to a parameter setting which allows short verb-raising, even though it is instantiated in neither their L1 nor their L2[3].

This kind of optional adherence does not occur across all domains of L2-acquisition, and cannot explain all cases of optionality in L2-acquisition (see Section 3.2.2). However, the fact that it does occur suggests that L2-learners may access multiple parameter settings at the same time. While most evidence for this phenomenon comes from syntactic domains, I will show that access to multiple parameter settings also occurs in the semantic domain of article choice.

---

[3] Reflexive binding is another domain where L2-learners exhibit optionality, as we have seen. However, it is not at present possible to tell whether this optionality is due to optional adherence to parameter settings, since the settings of the GCP form a subset relation. For instance, consider Thomas's (1989b) finding that L2-English learners allow both local binding and non-local binding in sentences such as (12a): sometimes they choose *Mary* and sometimes they choose *Sarah* as the antecedent for the anaphor. A possible explanation for this is that the learners have adopted the parameter setting in (10e): both local and non-local binding are licensed under this parameter setting. Alternatively, it might be that learners are accessing both the setting in (10a) (which allows local binding) and the setting in (10e) (which allows local as well as non-local binding). It is impossible to tell which alternative is correct.

However, there is some indication that L2-learners access both settings. If learners uniformly adopted the setting in (10e), they should allow both local and non-local binding in sentences such as (12a) freely; however, the degree to which non-local binding is allowed by L2-learners is in fact very small. This suggests that learners usually access setting (10a), which allows only local binding, and that they rarely access setting (10e), which allows both local and non-local binding. In order to know whether this optional adherence to parameter settings occurs at the level of individual learners, it is necessary to look at individual results. Thomas (1991) reports individual results, showing that the majority of L2-learners allow local binding only (the setting in (10a)) and a sizeable minority allow both local and non-local binding (the setting in (10e)). However, Thomas has only a single stimulus sentence of the relevant type (12a). Thus, we do not know whether each individual learner obligatorily adheres to either the setting in (10a) or the setting in (10e): for instance, if presented with ten sentences of the type in (12a), would an individual learner consistently allow only local binding, consistently allow both local and non-local binding, or vary in her judgments? Variation in judgments would suggest that the learner is accessing both the setting in (10a) and the setting in (10e), going back and forth between them. This remains an open question.

3. The proposal: fluctuation in L2-acquisition

Two important points emerge from the above discussion. First, as Broselow and Finer (1986) originally showed, L2-learners show evidence of accessing parameter settings that are instantiated in neither their L1 nor their L2. Second, in some domains, L2-learners show optional adherence to parameter settings: their behavior suggests that they sometimes adopt one setting of the parameter, and sometimes another. Both findings are fully consistent with the Full Access view of L2-acquisition. If L2-learners have access to UG, they have access to multiple settings of each UG parameter, not only those instantiated in the L1 and/or the L2. And in the absence of sufficient input, learners may go back and forth between multiple settings. This brings me to my proposal. I call this proposal the *Fluctuation Hypothesis,* and give it in (13).

13. **The Fluctuation Hypothesis (FH):**
    1) L2-learners have full access to UG principles and parameter settings.
    2) L2-learners fluctuate between different parameter settings until the input leads them to set the parameter to the appropriate value.

3.1. The specifics of the FH

3.1.1. UG-constrained grammar

Under the FH, L2-grammar is UG-constrained. L2-learners' errors are predicted to be non-random, but to reflect possible UG parameter settings. On this view, L2-learners' errors represent possible linguistic options which exist in natural language and/or are predicted by linguistic theory. This position is supported by the evidence discussed in Section 2 concerning L2-learners' linguistic knowledge.

3.1.2. Access to multiple parameter settings

The FH states that errors in L2-data stem from the learners fluctuating between two or more parameter settings, some of which are not appropriate for the target language. The fluctuation between multiple parameter settings can be seen in several domains, most markedly in the domain of verb-raising.

Moreover, L2-grammar is not constrained by the L1: L2-learners may adopt parameter settings that are instantiated in *neither* the L1 *nor* the L2 – but that may be instantiated in some third language. We have seen evidence that this is the case for the parameters governing verb-raising and reflexive-binding.

The logic behind this proposal is as follows. Suppose that L2-learners have to set a particular parameter, Parameter X. Let's leave aside for the moment the case where the L1 and L2 settings of the parameter match, and transfer is known to take place (see Section 3.2.1 for more discussion of transfer). Suppose that L2-learners have to acquire a new setting of Parameter X: either Parameter X has different settings in the L1 and the L2, or Parameter X is not instantiated in the L1 at all. Given sufficient evidence in the input, the learners may succeed in setting Parameter X to the target value. But what happens until they can do so? I suggest that until then, learners access all of the possible parameter settings. They might initially give preference to a particular setting if it is instantiated in their L1 (see Section 3.2.1) or because certain syntactic constraints make this setting unmarked or default (see footnote 5). In the absence of either transfer or markedness effects, learners will not give initial preference to a particular setting.

Crucially, in all cases, L2-learners will go through a period during which they access all of the possible parameter values, until the input leads to converge on the target value.

3.1.3. The FH and optionality

The main contribution of the FH is to provide a principled way of looking at optionality in L2-acquisition. When we see L2-learners showing optional adherence to the target grammar in some domain, it is possible that the optionality stems from fluctuation between parameter settings. While not all cases of optional adherence are tied to parameter setting (see 3.2.2 for other possible sources), in those cases where parameter setting *is* involved, the FH predicts the range of possible grammars. The FH predicts that L2-learners' errors should be constrained by the possible parameter settings, and should therefore be non-random.

3.2. Sources of errors in L2-acquisition and the FH

The FH is not intended to cover all cases of errors in all domains of L2-acquisition. In this section, I will discuss when the FH is and is not operative.

3.2.1. The FH and transfer

The FH as formulated above does not say anything specific about transfer. For domains where transfer takes place (e.g., VP-headedness), the FH needs to be modified so that it assigns a special role to the L1 parameter setting[4]. When transfer occurs, learners initially assign the L1 value to the parameter. As they subsequently reset the parameter to its L2 value, they undergo fluctuation between the L1 and L2 settings. However, for those domains where transfer is operative, if the L1 and L2 values are the same, no fluctuation occurs. For instance, Vainikka and Young-Scholten (1996) found that speakers of SOV languages acquiring German (another SOV language) produced head-final VPs 98% of the time; i.e., there was no fluctuation. In contrast, speakers of SVO languages who were acquiring German fluctuated between VP-final and VP-initial orders, suggesting that they were in the process of resetting the VP-headedness parameter.

I will not address the "transfer version" of the FH in this work. I will be concerned with the predictions of the FH in areas which are unaffected by transfer: where multiple parameter settings are equally available to the learners.

3.2.2. The FH and errors in L2-acquisition

The FH is not designed to cover all cases of errors that occur in L2-acquisition. It is concerned only with parameter setting. However, parameter-setting is not the only task faced by L2-learners. Learners also have the tasks of acquiring complex syntactic structures and morphological mapping.

---

[4] This is not to say that transfer and subsequent fluctuation must take place whenever a particular parameter is instantiated in both the L1 and the L2. As discussed earlier, many parameter settings (e.g., those related to CP-headedness, reflexive binding, and acquisition of control verbs) are apparently not subject to transfer. There is to my knowledge currently no account in L2-literature for why some domains are more subject to transfer than others. The only claim that I am making in this section is that in those domains where there *is* evidence for transfer, learners should go through a stage during which they fluctuate between the L1 and L2 settings.

For instance, there is evidence from several domains in L2-acquisition that L2-learners follow a developmental pattern similar to L1-learners. This has been discussed above for acquisition of control verbs (Flynn et al. 1991) and acquisition of relative clauses (Flynn 1987). Another example concerns embedding. Flynn and Martohardjono (1994) report on studies of Flynn (1983, 1987, 1991) which found that L1-Spanish learners of English initially convert structures containing embedding into coordination structures. The same developmental pattern has been attested for L1-acquisition (e.g., Lust 1981, among others; see Lust 1999 for an overview).

These developmental patterns are arguably not a result of parameter setting alone. For instance, in the case of control verbs, Flynn et al. (1991) proposed that L2-learners prefer object control with infinitival complements because in such structures the object antecedent minimally c-commands PRO (see also Epstein et al. 1996, Flynn 1996). A similar locality principle has been proposed for L1-acquisition (see Sherman and Lust 1993). On this view, acquiring the properties of control verbs is not simply a matter of throwing setting the parameter which determines whether a given verb is object-control or subject-control, and whether it takes infinitival complements or finite-clause complements (or both). Rather, locality considerations determine that a particular option (object control, infinitival complements) is, in a sense, default[5].

Similarly, the acquisition of relative clauses is more than a matter of setting the CP-direction parameter to head-final or head-initial. The studies of Flynn (1983, 1984, 1987) show that this parameter is set correctly early on in the course of L2-acquisition, but that the setting of this parameter to the target value does not automatically lead to appropriate use of relative clauses. The actual syntactic structure must be acquired. The data from Lust (1980) and Flynn (1987) provide evidence that both L1-learners and L2-learners go through a developmental stage during which free relatives are used in place of lexically headed relatives. (See Lust 1994, 1999 for more discussion of L1-acquisition, and Flynn, Vinnitskaya and Foley 2001 for a discussion of the developmental process underlying the acquisition of relative clauses).

Another domain which arguably falls outside of the scope of the FH concerns acquisition of inflectional morphology. There is evidence from studies by Lardiere (1998, 2000) and Prévost and White (2000), among others, that L2-learners often omit inflectional morphology and use non-finite verbs in place of finite verbs. However, the non-finite verbs in these learners' data behave as if they were syntactically finite. Prévost

---

[5] We might still think of parameter-setting in this domain, treating the subcategorization properties of each control verb in terms of parameters. The locality principle discussed here determines that a particular setting of this parameter is default or unmarked: the object-control, infinitival complement setting. The learner therefore initially adopts this setting. If a given verb indeed has such a setting in the target grammar, the learner will not need to reset the parameter. If a given verb is a subject-control verb and/or has finite clause complements, parameter resetting will take place. In this case, we should see fluctuation. For instance, we should see both L1-learners and L2-learners go through a stage during which they optionally treat *promise* as an object-control verb, and optionally – as a subject-control verb. This pattern was indeed attested in Sherman and Lust's (1993) study of control verbs in L1-acquisition: children allowed both object and subject antecedents for PRO in the infinitival complements of *promise*; the preference for subject antecedents grew stronger with age.

This suggests that the FH may be operative in domains where a particular parameter setting is default or unmarked, with the additional constraint that fluctuation will take place only when learners undergo parameter resetting from the default to the non-default value. This is an interesting subject for further investigation.

and White (2000), building upon the proposal of Lardiere (1998, 2000), argue that L2-learners have unimpaired functional categories and features (a position consistent with the Full Access view) but have difficulty with morphological mapping from the underlying features to their surface representations. As a result, they use default uninflected forms. The problem with morphological mapping appears quite distinct from the problem of parameter-setting: the learners have acquired lexical specifications on inflectional morphemes but nevertheless often omit these morphemes. Thus, optional use of inflectional morphology is a case of optionality that is not related to the FH.

To sum up, the FH is not intended to cover all acquisition patterns and error types attested for L2-acquisition. Rather, it is intended as a principled way of capturing parameter setting in L2-acquisition. Whenever we can isolate a parameter relevant for L2-acquisition, and specify the settings of this parameter, we should expect to see learners fluctuating between the settings until the input leads them to choose the target value of the parameter. The FH predicts that there should be no instantaneous parameter resetting: L2-learners should not go from 100% adherence to parameter value A to 100% adherence to parameter value B. Rather, they should go through a period in which both settings are employed. Data from such domains as VP-headedness and verb raising support this prediction. Finally, there may be some parameters which are set correctly from the very start of acquisition, so that we do not see fluctuation. See the next two sections for more discussion.

3.2.3. The FH and ultimate attainment

The next point to address is the persistence of fluctuation. For how long should L2-learners go back and forth between different parameter settings?

Some researchers argue that optionality is a permanent property of some domains of L2-acquisition. This is the case for the proposals of Eubank et al. (1997) and Beck (1998). These researchers argued that optional verb-raising in L2-acquisition arises from the learners' inability to set the feature strength on a functional head (Tense and/or Agr) to either [weak] or [strong]. Eubank et al. propose that in L2-acquisition, Tense is permanently underspecified: since it is neither [weak] nor [strong], verbs can optionally raise or stay *in situ*. Beck (1998) similarly proposes that feature strength on Tense is permanently impaired.

These proposals could be phrased in terms of the FH. Instead of saying that Tense is unspecified for strength, we could say that it is optionally specified as [weak] vs. [strong][6]. The proposals of Eubank et al. (1997) and Beck (1998) would then say that L2-learners are in a permanent state of fluctuation between the two possibilities.

However, I will not argue for the view that fluctuation is a permanent property of L2-acquisition. In our study of verb-raising with L1-Russian learners of English (Ionin and Wexler 2002), we found that overall L2-proficiency (as measured by the Michigan test) correlated positively with accuracy: more advanced L2-learners were more likely to appropriately place verbs after adverbs in English. This suggests that it is after all possible to set parameters in L2-acquisition.

---

[6] Given the findings that L2-learners allow optional verb raising only past adverbs, not past negation, the functional head in question actually needs to be a head lower than Tense – see Ionin and Wexler 2002 for more discussion.

Moreover, in some domains, L2-learners do not show fluctuation at all but appear to choose the appropriate value from the start. For instance, Flynn's studies of the CP-direction parameter in L2-acquisition provide evidence that L2-learners correctly set this parameter from the start of acquisition. Similar accuracy exists in the placement of verbs with respect to negation: optional verb-raising in L2-acquisition is confined to short raising past adverbs, not raising all the way past negation (see White 1992; Ionin and Wexler 2002). This suggests that L2-learners successfully set the functional strength on Tense from the start of acquisition, but do not set the strength of a lower functional head.

Thus, parameter setting to the target value is possible, at least for some parameters. On the other hand, Sorace (2000:98), in discussing optionality in L2-acquisition, states that "[w]hat can be observed for L2 optionality... is that, as in L1 acquisition, the pattern of preferences for one option over the other changes over time, until a potentially permanent stage is reached at which the target option is strongly, but not categorically, preferred, and the dispreferred non-target option is never completely expunged, but still surfaces in some circumstances." Under this view, L2-learners never quite set the parameter: as the learners become more advanced, they become more likely to choose the target value of the parameter, but still once in a while fluctuate towards a different value.

Since my studies do not look at L2-learners who are unambiguously at the end-state of their acquisition process, I remain agnostic about the issue of whether ultimate attainment is possible for all learners. The results of my studies will show, however, that parameter setting (in the domain of article choice) is possible for at least some learners.

3.2.4. The FH and triggers

In order to set a given parameter, L2-learners need to attend to the input data and generalize from the input trigger related to the parameter. For instance, take a parameter like VP-headedness, and a head-final language like German. Learners need to attend to the input data which tells them that objects come before verbs (in embedded clauses); they subsequently need to generalize from the multiple instances of hearing the object-verb order, and set the parameter to the "VP-final" value[7].

The differential success at setting parameters in L2-acquisition suggests that not all triggers necessary for parameter-setting are equally available to L2-learners: i.e., fluctuation is especially pronounced in those domains where L2-learners have difficulty generalizing from the input triggers necessarily to set the parameter. For instance, in Ionin and Wexler (2002), we argued that L2-learners have immediate access to triggers needed to set the strength of Tense, but not to triggers needed to set the strength on lower functional heads – hence the availability of short verb-raising past adverbs, but not long verb-raising to Tense, past negation. We suggested that the trigger for setting the strength of Tense in English could be the placement of negation or the presence of do-support (see also Sprouse and Schwartz 1998). However, the triggers related to the strength on lower functional heads are less obvious. Thus, L2-learners have more difficulty setting the strength on lower functional heads, compared to setting the strength on Tense.

There are many possible reasons for why certain triggers are more available to L2-learners than others. Ambiguity may be a factor, for instance. In order to set the feature

---

[7] All discussion of triggers in this section is fairly simplified. The triggers that I suggest for the various parameters may not be the relevant, or the only, triggers for those parameters. This discussion is intended for illustration purposes only.

strength of Tense, learners need to attend to the placement of negation: since negation necessarily precedes the lexical verb in English, this is a fairly unambiguous trigger. In order to set the feature strength on lower functional heads, on the other hand, learners presumably need to attend to adverb-verb order. This means that they first need to learn which lexical items are adverbs. They then need to attend only to those adverbs which are not placed sentence-finally: for instance, adverb placement in *I ate the soup slowly* will not inform learners about verb-raising in English. Moreover, a sentence such as *I ate quickly* may cause the learners to mis-analyze English as a verb-raising language. The triggering data related to adverb-verb placement in English are thus sufficiently ambiguous to prevent quick setting of the parameter.

One might make similar arguments in other domains, showing why some triggers are more readily available than others. For instance, the triggers related to VP-headedness in German may be ambiguous because the underlying SOV order is seen in embedded clauses, but SVO order often surfaces in main clauses. In contrast, the triggers related to CP-directionality may be easily available from a variety of different constructions, such as embedding and relative clauses. Finally, in the semantic domain of article choice, the triggers are quite different than in any of the syntactic domains – they depend not on a particular syntactic construction, but rather on a careful evaluation of the discourse. I will suggest in Chapter 3 that such discourse-based triggers may be less available than syntactic triggers.

Ideally, it would be necessary to construct a model which takes all the factors related to parameter-setting into consideration and shows why certain triggers are more available than others. The construction of such a model is far beyond the scope of this thesis.

On a final note, I would like to suggest that difficulty in setting parameters is one possible source of problems for L2-learners. On the one hand, there is much evidence that L2-learners have access to UG principles and parameter settings. On the other hand, we know that L2-learners are not as successful as L1-learners at acquiring the target grammar: not all learners reach ultimate attainment, and optional adherence to parameter settings is seen even in fairly advanced L2-learners (see the discussion in Sorace 2000). A possible (though by no means exhaustive) explanation of this would be to say that while L2-learners have full access to UG principles and parameter settings, their ability to generalize from the input triggers is decreased compared to L1-learners. Thus, L2-learners may remain in a state of fluctuation indefinitely, especially in those domains where the triggers are particularly subtle or ambiguous. The ability to generalize from the input is not gone, since many L2-learners are able to set parameters; it is, however, somewhat diminished. This is a speculative account and cannot be tested without an extensive model of triggers and parameter-setting.

3.2.5. The FH and its relation to L1-acquisition

Since parameter-setting takes place in both L1 and L2 acquisition, should the FH be equally applicable to both?

One possibility is that the FH is indeed applicable to both: whenever we see fluctuation in adult L2-acquisition, we should see similar fluctuation in child L1-acquisition. However, even if the FH is operative in both L1 and L2 acquisition, there is an important differences between the two types of processes: ultimate attainment. L1-learners nearly always successfully acquire their L1; on the other hand, many L2-learners

never fully acquire their L2. A possible source of this difference is that L1-learners are more successful at parameter-setting. Even if both L1 and L2 learners undergo fluctuation between parameter settings, L1-learners should eventually set the parameter to the target value, while L2-learners may remain in a state of fluctuation indefinitely. The ability to generalize from the input triggers necessary to set a parameter may decline with age, as suggested in the previous section. This would result in more fluctuation (and hence more persistent errors) in adult L2-learners than in child L1-learners.

An alternative answer is that the FH is applicable only to L2-acquisition: L1-learners do not go through a process of fluctuation but set parameters to the target value from the start. Since young children are exposed to L1-input for months before they start talking, it is possible that they successfully set parameters before, or shortly after, their first utterances. This position of early parameter-setting has been espoused, to a greater or lesser degree, by Wexler (1998), Lust (in prep) and Snyder (2002), among others. On this view, errors in child language are due to sources other than lack of parameter setting – e.g., to gradual UG maturation (Wexler 1998), or to development in grammatical mapping (Lust 1999). Some errors of child language may not be linguistic in nature at all, but may reflect children's egocentricity or lack of knowledge of pragmatic principles; proposals along these lines have been put forth in the domain of article choice (e.g., Maratsos 1976) and binding (e.g., Thornton and Wexler 1999). It is beyond the scope of this thesis to discuss all of the proposals that have been put forth in the literature regarding errors and parameter-setting in L1-acquisition, and I will not be taking a particular view regarding parameter-setting in L1-acquisition.

However, it important to consider the FH with regards to both L1 and L2 acquisition and examine whether it holds for both. If the FH holds for the acquisition of a particular Parameter X by L2-learners, it is necessary to examine whether it also holds for Parameter X with L1-learners, and to distinguish between the following three possibilities. (1) Only L2-learners undergo fluctuation between the settings of Parameter X. L1-learners correctly set Parameter X from the start and make no errors. (2) Only L2-learners undergo fluctuation between the settings of Parameter X. L1-learners also make errors in the domain related to Parameter X, but their errors are attributable to some other source (e.g., lack of pragmatic knowledge) and there is evidence that, despite the errors, L1-learners have in fact set Parameter X to the target value from the start. (3) Both L2-learners and L1-learners undergo fluctuation between the settings of Parameter X.

In some domains, there is clear evidence that either (1) or (2) is correct – children set parameters early and the FH does not apply. There is evidence that this is the case for the parameters related to verb movement (see Wexler 1998) and preposition stranding / pied-piping (see Snyder 2002), to name just a couple. However, in other domains, such as article choice, the data are not as clear.

Since article choice is the primary subject of this thesis, I will examine articles in L1-acquisition and consider which of the three positions above explain the data best. Possibility (1) above will be easily ruled out: L1-learners of English, like L2-learners of English, misuse English articles (see Maratsos 1976, among others). In Chapter 3, I will examine previous studies article acquisition, and examine whether article errors in L1 vs. L2 acquisition stem from different sources (possibility 2 above) or whether both can be explained under the FH (possibility 3 above). I will suggest that discourse-based parameters (such as the parameter I will propose for article choice) may be set somewhat

later than other parameters by child learners, and therefore provide a good testing ground for the FH in L1-acquisition.

To sum up, from the standpoint of the FH, L1 and L2 acquisition may differ from each other in at least two different ways: (1) fluctuation between parameter settings is an inherent property of L2-acquisition; L1-learners set parameters very early on and do not undergo fluctuation; or (2) both L1 and L2 learners undergo fluctuation between parameter settings, but L1-learners are more successful at setting parameters than L2-learners.

3.3. The FH and article choice

In this chapter, I have given a brief overview of some issues in L2-acquisition, and have focused on L2-learners' optional adherence to parameter settings in certain domains. The data concerning parameter setting in L2-acquisition led me to propose a Fluctuation Hypothesis for L2-acquisition.

The goal of this thesis is to show that the Fluctuation Hypothesis predicts the distribution of articles in L2-English. In order to show that this is the case, I will do two things: first, I will establish the relevant parameter[8] which governs article choice; and second, I will show that L2-English learners exhibit fluctuation between the settings of this parameter.

The term "parameter" in this context refers to constraints on lexical specifications. I will provide evidence in Chapter 2 that articles cross-linguistically can receive different lexical specifications related to *definiteness* and *specificity*. I will propose that these differences can be captured via parametric variation, with languages varying on whether they use articles to encode the [+definite] feature, the [+specific] feature, or both.

The Article Choice Parameter that I will propose is different from parameters that are usually examined in L2-acquisition studies[9]. It does not deal with any syntactic properties such as directionality, movement, or locality. Its focus is entirely on lexical specifications of articles, and how these specifications interact with the properties of the discourse. The Article Choice Parameter is discourse-related: it dictates whether articles encode the state of hearer knowledge or the state of speaker knowledge.

My focus throughout will be on article *use* in L2-English. While Chapter 7 briefly considers article omission in L2-English production data, the accuracy of article use in different kinds of discourse contexts will be the primary issue. The main body of the empirical data comes from elicitation tasks, in which article omission with singular DPs is very low.

The focus of the discussion will be on group results. However, there is always the possibility that group results obscure individual performance. It is important to show that individual learners do not behave randomly, and that the predicted patterns of article use are seen at the individual level. Therefore, I will always report individual as well as group

---

[8] The parameter that I will establish is not, however, the *only* parameter governing article choice. As discussed in Chapter 2, other parameters may be related to article choice as well, such as Chierchia's (1998) parameter related to article use in generics. I am concerned primarily with the parameter governing discourse-based distinctions – see Chapter 2 for more discussion.

[9] But see Matthewson and Schaeffer's 2000 work on articles in L1-English in the context of a differently formulated Article Choice Parameter. Their study is summarized in Chapter 3.

results. The individual results will show that very few learners show random patterns, with the vast majority confirming to the predicted patterns.

4. Overview of this thesis

This thesis is organized as follows. Chapter 2 lays out the theoretical background for my proposal. The first portion of this chapter gives a brief overview of the semantic concepts of *definiteness, scope,* and *referentiality.* In the second half of the chapter, I formulate a lexical entry for *specificity,* which is based on Fodor and Sag's (1982) proposal concerning referentiality, but which also incorporates the concept of *noteworthy property.* I discuss the relationship between definiteness and specificity, and lay out the predictions for possible article classifications. I then provide some cross-linguistic data that support these predictions. Chapter 2 ends with a formulation of the Article Choice Parameter.

In Chapter 3, I advance a proposal for article choice in L2-English. This proposal falls under the Fluctuation Hypothesis: after specifying the relevant parameter related to article choice, I argue that L2-English learners fluctuate between the two settings of this parameter. I examine my hypothesis in light of previous studies of both L1 and L2 acquisition. I also discuss the potential for transfer in the acquisition of English articles by L1-Russian and L1-Korean speakers, and show that transfer is not predicted to take place for these learners.

Chapters 4 through 7 provide empirical data in support of my proposal. Chapter 4 reports on an elicitation study of article choice in the L2-English of L1-Russian speakers, and shows that overuse of *the* with indefinites in L2-English is tied to specificity, as predicted. Chapter 5 reports on a second elicitation study, which expands on the first study; this chapter shows that the patterns of article use and misuse in L2-English hold for speakers of very different L1's, Russian and Korean. Chapter 6 extends the discussion to definites, and reports on a third elicitation study of L2-English article choice. This chapter shows that L1-Russian and L1-Korean speakers distinguish both definites and indefinites on the basis of specificity. Chapter 7 complements the three elicitation studies by reporting on articles in L2-English written production data.

In the concluding chapter, Chapter 8, some alternative explanations for my findings are considered, and arguments are advanced against them. Finally, this chapter discusses some directions for future research that are related to article choice and/or fluctuation in acquisition.

Chapter 2: Article semantics

1. Introduction

In Chapter 1, I hypothesized that L2-English learners undergo fluctuation between parameter settings in the domain of article choice. The main goal of the present chapter is to propose the relevant parameter governing article choice, and to specify the possible settings of this parameter. Subsequent chapters will provide evidence that L2-learners do indeed undergo fluctuation between the settings of this parameter.

In this chapter, I will propose the Article Choice Parameter, which is given in (1). In order to motivate the existence of this parameter, I will provide evidence that articles cross-linguistically can encode *definiteness* or *specificity*. While there are many definitions of specificity in the literature, the view of specificity that I will assume is based on Fodor and Sag's (1982) view of *referentiality*.

1. The Article Choice Parameter
    A language which has two articles distinguishes them as follows:
    Setting I. Articles are distinguished on the basis of specificity.
    Setting II. Articles are distinguished on the basis of definiteness.

In order to determine the lexical entry for specific DPs, I will examine the behavior of the English referential *this*, and show that use of *this* carries two requirements: first, that the speaker intend to refer to a particular individual (per Fodor and Sag 1982); and second, that the speaker have in mind a noteworthy property that describes this individual. I will propose that these requirements constitute *felicity* conditions on specificity. I will then contrast the felicity conditions carried by specific DPs with the presuppositions carried by definite DPs, and argue that the latter are more informative from the standpoint of the discourse. The distinction between presuppositions and felicity conditions allows me to make predictions for article distribution cross-linguistically. These predictions are supported by data from English, Hebrew, and Samoan. I will capture cross-linguistic variation in definiteness / specificity marking by proposing the Article Choice Parameter in (1), which will be central to the discussion of L2-acquisition of articles in the rest of this thesis.

This chapter is organized as follows. First, I provide a brief overview of the background for my proposal. I look at three ways in which DPs can potentially be distinguished: *definiteness, scope,* and *referentiality*. These are discussed in Section 2, Section 3, and Section 4, respectively.

In Section 5, I advance my proposal concerning the nature of *specificity*. This proposal is built upon Fodor and Sag's view of referentiality, but with modifications that specify the felicity condition on specificity markers. I also show that the specificity distinction is predicted to cut across the definiteness distinction. In Section 6, I examine cross-linguistic data that support my proposal. In Section 7, I bring the discussion of the preceding sections together and propose the Article Choice Parameter.

2. Definiteness: an overview

One of the ways in which DPs can be distinguished from one another is *definiteness*. This distinction is morphologically encoded in English: definites take *the,* while

indefinites take *a* in the singular and no article in the plural. In this section, I will provide a very brief overview of the semantics of definiteness, based on the discussion in Heim (1991).

2.1. The Russellian analysis of definites

One of the classical analyses of definites comes from Russell; it is given in (1).

2. <u>Russellian analysis:</u>
    A sentence of the form [*the* ζ] ξ expresses that proposition which is true if there exists exactly one ζ and it is ξ, and which is false otherwise.
    <div align="right">(from Heim 1991:2)</div>

The Russellian analysis says that a sentence such as (3), which contains a definite DP, will be true if and only if there exists a unique cat, and that cat is drinking milk. If no cat exists, or if multiple cats exist, or if a unique cat exists but is not drinking milk, the sentence will be false.

3. <u>The cat</u> is drinking milk.

However, a problem for this analysis has been pointed out in the literature. Take a sentence like (4a). We know that there is no king of France. Under Russell's analysis, (4a) is therefore predicted to have the reading in (4b) available to it. Under this reading, (4a) is true: it is indeed not the case that I am meeting the king of France tonight, for the simple reason that the king of France does not exist. Under the Russellian analysis, speakers should have no problem accessing this reading and considering (4a) to be true. However, the preferred reading for (4a) is actually the bizarre one in (4c), which states that the king of France exists. (See Heim 1991 for more examples and discussion).

4. a) I am not meeting <u>the king of France</u> tonight.
    b) It is not the case that there exists a unique king of France and I am meeting him tonight.
    c) There exists a unique king of France, and it is not the case that I am meeting him tonight.

This behavior of definites has led to the *presuppositional* view of definites, described in the next section.

2.2. The Fregean analysis of definites

Informally, a presupposition is a statement that must be true in order for another statement to have a truth-value at all. This is stated more formally in (5).

5. Let p and q be (possibly partial) propositions. Then q is a <u>semantic presupposition</u> of p iff q is true at every world-time pair where p is true or false. (Heim 1991:8)

As an illustration, consider the statement in (6a). The proposition expressed by this statement will be true if Miss Bock used to drink champagne in the mornings and has stopped doing so. The proposition will be false if Ms. Bock is continuing to drink champagne in the mornings. But what if Ms. Bock has never drunk champagne in the mornings in the first place? The proposition in (6a) will then not have a truth-value – it will be neither true nor false. This means that the proposition in (6b) is a presupposition of (6a): in order to for (6a) to have a truth value it is necessary that (6b) be true. Thus,

when one utters (6a), one *presupposes* that Ms. Bock used to drink champagne in the mornings, and *asserts* that she has stopped doing so.

6. a) Miss Bock has stopped drinking champagne in the mornings.
    b) At some time prior to the present, Miss Bock drank champagne in the mornings.

The presuppositional analysis has been applied to definites. In what is called the Fregean analysis of definites, the definite article carries presuppositions of *existence* and *uniqueness,* as stated in (7).

7. Fregean analysis: (from Heim 1991:9)
    [*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 on 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.

- Take again a sentence like (4b): this sentence now *presupposes* that a unique king of France exists, and *asserts* that I am not meeting him tonight. Since there is no king of France, (4a) results in *presupposition failure* – it is neither true nor false.

Consider next a simple sentence like (3). This sentence will be true whenever a unique cat exists, and it is drinking milk; it is false if a unique cat exists, but is not drinking milk; and it has no truth-value if there is no cat, or there are multiple cats.

Now, we live in a world which has multiple cats. There is always more then one cat in existence, so (3) should have no truth-value. However, (3) can clearly be a felicitous utterance as long as there is a unique cat in some *contextually given domain*. The standard analysis in the literature is to assume a contextually-given restriction on the set denoted by the NP (see Evans 1980, Kadmon 1990, Roberts 2000, among others). Uniqueness is then computed with respect to a contextually relevant domain, as stated in (8) (from the lecture notes of Heim and Wexler (2000)).

8. *Domain Selection*: "Uniqueness is computed with respect to a contextually given domain, which typically is a proper subset of the set of all individuals."

Crucially, the contextually given domain must be part of the shared knowledge of speaker and listener: it is infelicitous for me talk about *the cat* unless my listener is aware of a unique, contextually salient cat, or can at least accommodate knowledge of such a cat[1].

Previous discourse is not always necessary for establishing uniqueness. In some cases, the uniqueness presupposition is satisfied as a result of mutual world knowledge. For instance, in order for (9) to be felicitous, it is not necessary that the speaker and hearer be talking about some salient winner. Given our world knowledge that a tournament typically has only one winner, the uniqueness presupposition is satisfied.

9. The winner of this tournament will receive a prize.

---

[1] An example of accommodation is if I say "I should go feed *the cat*" while standing in my house. Even if my listener has never seen my cat, and does not know that I have one, he can accommodate the knowledge that there is a unique, contextually salient cat under discussion – the one that I own. (See Chapter 8 for more discussion).

Finally, both the Fregean analysis and the Russellian analyses of definites can be generalized to plurals (as first shown by Sharvy 1980). If the Fregean analysis is extended to plurals, a definite plural like *the cats* is presupposed to contain a "greatest" element, where "a greatest element of a set M is an element of M which has all other elements of M as parts" (Heim 1991:22).

In the case of singular definites, the greatest element of the set denoted by the NP is unique. On the generalized Fregean analysis, (10a) presupposes that the set denoted by *cat* has one and only one member, while (10b) presupposes that the set denoted by *cats* has at least two members and has a greatest element: this is the *maximality presupposition*. The predicate must hold of this greatest element – i.e., of all members of the set. In the case of singulars, there is only one member in the set, so the predicate must hold of this one member.

10. a) The cat is asleep.
     b) The cats are asleep.

2.3. The quantificational analysis of indefinites

The traditional view of indefinites, under both Frege and Russell, treats them as existential quantifiers, as shown in (11). Indefinites neither entail nor presuppose uniqueness. A sentence containing an indefinite DP, such as (12), is always either true or false: it is true if there exists at least one cat which is drinking milk, and it is false otherwise (e.g., if there is no cat, or if there are cats, but none of them are drinking milk)[2].

11. A sentence of the form [a ζ] ξ expresses that proposition which is true if there is at least one individual which is both ζ and ξ, and false otherwise.

     (Heim 1991:26)

12. A cat is drinking milk.

2.3.1. Use of *a* and non-uniqueness

One of the problems faced by the quantificational analysis of indefinites is that, as noted by Hawkins 1978, it does not explain a "non-uniqueness" condition that seems to go along with *a*. The indefinite article cannot be used when there is clearly a unique referent in the discourse, as in (13). The definition in (11), however, would predict the sentences in (13) to be well-formed: for instance, (13b) should express the proposition which is true if there is at least one father of the victim, and I interviewed him.

13. a) #A weight of our tent is under 4 lbs.
     b) #I interviewed a father of the victim.   (Heim 1991:27)

---

[2] There are sentences in which indefinites in subject position appear to carry an existence presupposition. For instance, (i) (due to David Pesetsky, p.c.), on the non-generic reading, appears to presuppose the existence of Roman consuls. The presupposition seems absent if the indefinite is in object position, as in (ii), which is false rather than truth-valueless; this means that *presuppositionality* is not an inherent property of indefinites. See Diesing (1992) for a proposal linking the subject ([Spec, IP]) position and presuppositionality.

(i) A Roman consul is bald.
(ii) I saw a Roman consul yesterday.

A possible explanation for cases such as (13) would be to say that *a* in fact carries a "non-uniqueness" condition. However, Heim 1991 shows that this would make incorrect predictions. She considers two possibilities. The first possibility is that the truth-conditions of *a* require that the set denoted by the restrictor NP have at least two members. This would predict that the sentences in (13) are false, since the sets *weight of our tent* and *father of the victim* each contain only one member. Heim notes that intuitively, the sentences in (13) appear to be infelicitous rather than false, which suggests that non-uniqueness is probably not part of the truth-conditions on *a*.

A second possibility that Heim discusses, based on Hawkins 1978, is that *a* carries a presupposition of non-uniqueness: a sentence containing an indefinite headed by *a* has a truth-value only if there are at least two members in the set denoted by the restrictor NP of the indefinite. The reason (13b) is infelicitous is that the set *father of the victim* has only one member. While this proposal explains the facts in (13), Heim shows that it makes the wrong predictions for sentences such as (14). Under the presuppositional analysis of *a*, these sentences presuppose that there exist at least two 20 ft. long catfish -and at least two pathologically nosy neighbors of mine, respectively. However, these sentences do not, intuitively, carry such presuppositions.

14. a) Robert caught a 20 ft. long catfish.
     b) A pathologically nosy neighbor of mine broke into the attic.

2.3.2. *"Maximize Presupposition"*

Given the problems of associated with positing a non-uniqueness condition on *a*, Heim (1991) proposes that *a* always has the standard quantificational analysis in (11). She proposes that the reason the sentences in (13) are infelicitous is that in these sentences, the presuppositions for *the* have been met: there is necessarily a unique father to any victim, and a unique weight to any tent. She formulates the rule in (15).

15. In utterance situations where the presupposition for [*the* ζ] ξ is already known to be satisfied, it is not permitted to utter [*a* ζ] ξ.   (Heim 1991:27)

Heim points out that (15) cannot be straightforwardly derived from Gricean principles. It needs to stem from an additional maxim such as (16), termed the "Maximize Presupposition" Principle by Heim and Wexler (2000).

16. The Maximize Presupposition Principle:
     "Make your contribution presuppose as much as possible!" (Heim 1991:28)

Since *the* carries more presuppositions than *a*, it is preferable to use *the* whenever its presuppositions are met.

For the rest of this thesis, I will adopt the Fregean (presuppositional) analysis of definites and the quantificational analysis of indefinites as the standard analyses.

3. Scope: an overview

In addition to definiteness, DPs can be classified on the basis of *scope*.

3.1. Scope in indefinites and definites

An indefinite DP may take wide or narrow scope with respect to an operator such as an intensional verb or a modal. In (17a), for instance, the DP *a book about fruit flies* can

scope over the intensional verb *wants,* getting the reading in (17b), which may be paraphrased as follows: there exists a particular book about fruit flies which Sarah wants to read. Alternatively, the DP can scope under *wants,* getting the reading in (17c), which will be true as long as in all the worlds in which Sarah's wishes are satisfied, Sarah reads some book or other about fruit flies.

17. a) Sarah wants to read a book about fruit flies.
     b) $\exists$x [x is a book about fruit flies and Sarah wants to read x]
     c) Sarah wants $\exists$x [x is a book about fruit flies and Sarah reads x]

Similar effects obtain when an indefinite enters into a scope relation with a quantifier such as *every,* as in (18). When the indefinite in (18a) takes wide scope over the universal, the reading in (18b) obtains, which will be true if every entomology student read the same book about fruit flies. When the indefinite scopes under the universal quantifier, the reading in (18c) obtains, which will be true as long as each student read some book or other about fruit flies, with the books potentially not being the same.

18. a) Every entomology student read a book about fruit flies.
     b) $\exists$x[x is a book about fruit flies and $\forall$y[y is an entomology student→y read x]]
     c) $\forall$y[y is an entomology student→$\exists$x[x is a book about fruit flies and y read x]]

Whenever an indefinite occurs in a sentence with two or more other scopal elements, such as intensional operators or quantifiers, it may have intermediate scope readings available to it in addition to the widest and narrowest scope readings. This is discussed in detail in Section 4.1.2. In this section, I restrict my attention to examples which contain only one scopal element besides the indefinite.

When a sentence contains an intensional verb or modal as well as an indefinite, there is a further terminology of *de re / de dicto* readings: an indefinite is *de re* if it scopes over an intensional or modal operator (17b), and it is *de dicto* when it scopes under an operator, as in (17c).

When no intensional/modal operators or other quantifiers are present, indefinites take wide scope by default. (19a) has the reading in (19b), which will be true if Sarah read a book about fruit flies.

19. a) Sarah read a book about fruit flies.
     b) $\exists$x [x is about fruit flies and Sarah read x]

While different scope readings are most evident for indefinites, they are also found with definites. Consider (20). On the wide-scope reading of the definite in (20a), there is a presupposition that a unique individual exists in the actual world who won yesterday's race. On the narrow-scope reading of the definite in (20b), there is a presupposition that in each possible world in which my wishes are satisfied, there is a unique individual who won yesterday's race. See Section 4.2 for more discussion.

20. a) I want to talk to the winner of yesterday's race.
     b) the unique x: x is the winner of yesterday's race in w [$\forall$w' [my wishes in w are satisfied in w' → I talk to x in w']]
     c) $\forall$w' [my wishes in w are satisfied in w' → I talk in w' to the unique x: x is the winner of yesterday's race in w']

3.2. Scope and parametric variation

Not all languages encode definiteness in their article system. There is evidence that some encode a distinction related to scope instead. Matthewson (1998) argues that a distinction related to scope is encoded in some Salish languages, specifically in St'át'imcets (Lillooet Salish). (But see Matthewson (1999) for a different analysis of Salish in terms of choice functions).

3.2.1. Articles and assertion of existence in St'át'imcets

Matthewson (1998) shows that St'át'imcets distinguishes between two types of articles. The first determiner type (*ti...a, i...a,* and a variety of others) are determiners which, according to Matthewson, *assert existence*; these determiners obligatorily take wide scope over an intensional verb or modal.

The other determiner type consists of the article *ku,* which does not assert existence, and which must be licensed by a higher scope-bearing element. The sentences in (21) illustrate the difference between the two determiner types. In a declarative sentence with no intensional verbs, modals, or negation, such as (21a), only an *assertion of existence* determiner, such as *ti...a,* may be used. When negation is present, a DP headed by *ti...a* must take wide scope over negation, as in (21b); a DP headed by the *ku* must take narrow scope under negation, as in (21c).

21. a) *az'-en-as*   *[ti*   *sts'úqwaz'-a]* *kw-s*   *Sophie.*
         buy-Tr-3erg   [Det   fish-Det]   Det-Nmlzr Sophie
         "Sofie bought [a fish]."
         ∃x, x is a fish, Sofie bought x.

     b) *cw7aoz*   *kw-s*   *az'-en-as*   *[ti*   *sts'úqwaz'-a]* *kw-s*
         Neg   Det-Nmlzr   buy-Tr-3erg   [Det   fish-Det]   Det-Nmlzr
         Sophie.
         Sophie
         "Sofie didn't buy [a fish]."
         ∃x, x is a fish, ~Sofie bought x.

     c) *cw7aoz*   *kw-s*   *az'-en-as*   *[ku*   *sts'úqwaz']*   *kw-s*
         Neg   Det-Nmlzr   buy-Tr-3erg   [Det   fish]   Det-Nmlzr
         Sophie.
         Sophie
         "Sofie didn't buy [a fish]."
         ~∃x, x is a fish, Sofie bought x.   (Matthewson 1998:55-56, ex. 62)

While the focus of Matthewson's (1998) work is on scope distinctions in indefinites, she notes that definites are also compatible with both determiner types. Typically, DPs that would take the definite determiner in English are used with the *assertion of existence* determiners in St'át'imcets. However, *ku* must be used with narrow-scope definite DPs in

examples such as (22), "in accordance with the non-existence of the individual in present time" (Matthewson 1998:57)[3].

22. *cúz'-lhkan   melyi-s   [ku   cuz'   kúkwpi7 láku7Fountain]*
    going.to-1Sg.Sub marry-Caus [Non.Exis.Det going.to chief   Deic Fountain]
    "I will marry the next chief of Fountain." (whoever it is)

    (Matthewson 1998:57, ex. 67)

3.2.2. English vs. Salish as a case of parametric variation

Matthewson (1998) proposes that the difference between English, in which articles are distinguished on the basis of definiteness, and St'át'imcets, in which articles are distinguished on the basis of *assertion of existence,* can be captured under the *Common Ground Parameter (CGP)* in (23).

23. *Common Ground Parameter*
    Determiners may access the common ground.
    Yes: {English,...}
    No: {Salish,...}   (Matthewson 1998:114, ex. 61)

By *common ground,* Matthewson means the knowledge and beliefs that are accessible to both the speaker and the hearer of a given utterance. Presuppositions, and hence definite DPs, necessarily access the common ground: both the speaker and the hearer must presuppose that the definite description has a unique referent in order for use of the definite to be felicitous. Indefinites, on the other hand, do not access the common ground.

Matthewson (1998) does not expressly specify what the negative setting of the CGP is: in principle, there can be many different features that determiners could encode that would not be related to the common ground. Matthewson proposes that determiners in Salish, which has the negative setting of the CGP, access only speaker knowledge, encoding *assertion of existence* on the part of the speaker.

This idea is adopted by Matthewson and Schaeffer (2000), who rename the CGP *The Article Choice Parameter* and give it two settings, as shown in (24). While the term *speaker beliefs* can in principle be interpreted in a variety of ways, Matthewson and Schaeffer use it to mean that the speaker has *grounds for an existential assertion.*

24. *Article Choice Parameter*
    If the language semantically distinguishes more than one article, the distinction is based on:
    I. Speaker Beliefs
    II. Common Ground   (Matthewson and Schaeffer 2000:23)

For Matthewson and Schaeffer, a language such as St'át'imcets has Setting I of the Article Choice Parameter and distinguishes between definites and wide-scope indefinites on the one hand, and narrow-scope indefinites on the other. They do not consider narrow-scope definites such as (22).

---

[3] It is of course quite possible (indeed, most likely) that there does exist, at the time of utterance, the individual who at some future time will become the next chief of Fountain. Presumably, the reason that *ku* is licensed in (22) is that there exists no individual who is *the next chief of Fountain* at the present time – i.e., *ku* is licensed by being in the scope of a temporal operator. Matthewson does not discuss examples such as (22) in detail.

3.2.3. More on scope in St'át'imcets

Matthewson (1999) provides a different analysis of determiners in St'át'imcets. She looks at interactions between determiners and distributive operators. As shown in (25a), *ku* must scope under a distributive operator. On the other hand, determiners like *ti...a* cannot be in the scope of a distributive operator, as shown by (25b).

25. a) *[i zí7zeg'-a smelhmúlhats]ú7stek [ku sts'úqwaz']*
  [DET.PL each-DET woman(PL)] catch.fish [DET fish]
  "Each woman caught a (different) fish." (Matthewson 1999:102, ex. (43a))

 b) *[zí7zeg' smelhmúlhats]met'-en-ítas [ta máw-a]*
  [each woman(PL)] pet-TR-3PL.ERG [DET cat-DET]
  "Each woman petted a cat."
  i. Accepted in context: Each woman petted the same cat.
  ii. Rejected in context: Each woman petted a different cat.
             (Matthewson 1999:103, ex. (46))

These data present a problem for the Article Choice Parameter of Matthewson and Schaeffer: in (25a), the existence of at least one fish is asserted, yet the determiner *ku,* which is supposed to be a *non-assertion of existence* determiner, is used. Moreover, Matthewson (1999) showed that when a DP with a normally wide-scope determiner such as *i...a* contains a bound pronoun, it *can* take scope under the distributive operator, as in (26).

26. [tákem i smelhmúlhats-a]wa7 xwey-s-twítas [i kalhélhs-a
 [all DetPl woman(Pl)-Det] PROG love-CAUS-3PL.ERG [DetPl three(ANIM)-DET
 máw-i]
 cat-3PL.POSS]
 "All (the) ladies like their three kittens."
 Accepted in context: Each woman loves three different kittens.
 Consultant's comment: "Could be three different ones for each lady. It's
 THEIRS."   (Matthewson 1999:116, ex. (83b))

These data show that St'át'imcets does not after all distinguish articles on the basis of *assertion of existence* or on the basis of scope: determiners which are normally wide-scope (such as *ti...a*) can take narrow scope (with respect to a higher quantifier) under special circumstances. Matthewson (1999) proposes an analysis in terms of choice functions – see Section 4.1.2 for more discussion.

4. Referentiality: an overview

Another distinction that has been proposed for DPs is *referentiality*. This distinction has been proposed both for indefinites (Fodor and Sag 1982) and for definites (Donnellan 1966). In this section, I examine these proposals and the evidence both for and against them. In Section 5, I will formulate a proposal for *specificity* which builds on the discussion of referentiality in this section.

4.1. The ambiguity of the English indefinite

Traditionally, indefinites are considered to be quantifiers, as discussed in Section 2.3. The lexical entry for an indefinite DP is repeated in (27). Given this standard quantificational analysis of indefinites, we should expect them to behave just like other quantifiers (*every, each, many*, etc.)

27. A sentence of the form [a ζ] ξ expresses that proposition which is true if there is at least one individual which is both ζ and ξ, and false otherwise.

(Heim 1991:26)

However, it has long been noted that indefinites do not in fact behave like other quantifiers. As Fodor and Sag (1982) first showed, indefinites are able to scope out of *if*-clauses, *that*-clauses, and other scope islands. As an illustration, consider Fodor and Sag's example in (28). In (28a), the indefinite *a friend of mine from Texas* can take scope over *if*, allowing for the following paraphrase of the sentence: there is a particular friend of mine from Texas such that, if that friend had died in the fire, I would have inherited a fortune. This is in fact the salient reading for this sentence (on the narrow-scope reading of the indefinite, the sentence would state that I would have inherited a fortune if *any* friend of mine from Texas had died in the fire).

Now consider (28b). If *each friend of mine from Texas* took wide scope over *if*, the sentence would mean that for each of my Texan friends x, if x had died in the fire, I would have inherited the fortune. That is, it would not be necessary for *all* of my Texan friends to die: I would inherit a fortune if any one of them died. However, this reading is unavailable: (28b) can only mean that I would have inherited a fortune if all of my friends from Texas had died.

28. a) If a friend of mine from Texas had died in the fire, I would have inherited a fortune. (Fodor and Sag 1982, ex. 60)
    b) If each friend of mine from Texas had died in the fire, I would have inherited a fortune. (Fodor and Sag 1982, ex. 62)

A similar contrast obtains in example (29), which is often used in the literature. (29a) has a reading which can be paraphrased as follows: Mary read every book recommended by one particular professor (e.g., Professor Smith); on this reading, the indefinite scopes over the relative clause. In contrast, (29b) does not have the reading on which *every teacher* scopes over the relative clause. This reading, if it existed, would be paraphrased as follows: for every teacher x, x recommended a book and Mary read that book. This reading is clearly unavailable for (29b). Indefinites which scope out of islands, as in (28a) and (29a), are termed *long-distance indefinites* in the literature.

29. a) Mary read every book that *a / some teacher* had recommended.
    b) Mary read a book that *every teacher* had recommended.

The exceptional scope-taking properties of long-distance indefinites would not be a problem if we simply said that indefinites are not subject to locality constraints; the wide-scope readings in (28a) and (29a) are due to long-distance scope-shifting, as illustrated in (30) for (29a). I follow Schwarz (2001) in calling this the *scope shifting analysis*.

30. [a/some/a certain teacher] λ_1[Mary read every book that t_1 had recommended]

The problem with the scope-shifting analysis is that it requires a stipulation: it says that for some reason, indefinites differ from all other quantifiers (*every, each, at least one,* etc.) in not being subject to locality constraints. This stipulation has no independent motivation. The rest of this section is concerned with various alternative analyses that have been proposed to account for long-distance indefinites. These proposals give long-distance indefinites readings that are not quantificational.

4.1.1. Referential analysis of long-distance indefinites

One of the first proposals concerning long-distance indefinites is due to Fodor and Sag (1982), who argued that English indefinites are ambiguous between a referential and a quantificational reading. On the *referential analysis*, the indefinites in (28a) and (29a) are not scoping out of islands at all: rather, they are referring expressions, and as such, give the appearance of widest scope (similarly to demonstratives like *this book* or proper names like *Mary*). Heim's (1991) formulation of the semantics of Fodor and Sag's referential indefinites is given in (31). Crucially, the extension of a referential indefinite is defined if and only if the speaker intends to refer to the unique individual denoted by the indefinite. The referential indefinite is an indexical: it is not influenced by the index of evaluation, so it has the same reading regardless of the presence of an intensional operator. Under this analysis, (29a) has the reading in (32).

31. A sentence of the form $[a_r \, \zeta] \, \xi$ expresses a proposition only in those utterance contexts c where the speaker intends to refer to exactly one individual **a** and **a** is ζ in c. When this condition is fulfilled, $[a_r \, \zeta] \, \xi$ expresses that proposition which is true at an index i if **a** is ξ at i and false otherwise.   (Heim 1991:30)

32. Mary read every book that [the unique teacher that the speaker intends to refer to in the world and time of utterance] had recommended.

Under Fodor and Sag's analysis, indefinites are ambiguous between the referential reading in (31) and the standard quantificational reading in (27). On the latter, they behave like other quantifiers, taking wide or narrow scope and being subject to locality constraints. Thus, *a / some teacher* in (29a) can also scope inside the *that*-clause.

Fodor and Sag extend the referential/quantificational ambiguity to sentences with no intensional operators, such as the pair of sentences in (33). In (33a), the speaker most likely has a particular referent in mind when she utters *a man*: this is suggested by the fact that the speaker is clearly aware of the identity of the man who proposed to her, and in fact wishes to hide his identity from the hearer. The man's identity is relevant from the speaker's viewpoint, and the condition in (31) is likely to be satisfied.

In contrast, the speaker has no particular referent in mind in (33b), where she is concerned simply with the presence of some man or other in the women's bathroom. For Fodor and Sag, (33a) contains a referential indefinite while (33b) contains a quantificational indefinite.

33. a) A man just proposed to me in the orangery (though I'm much too embarrassed to tell you who it was).
    b) A man is in the women's bathroom (but I haven't dared to go in there to see who it is).   (Fodor and Sag 1982, ex. 7 and 8)

How can we know whether an indefinite is referential or quantificational? First of all, given the lexical entry in (31), a referential indefinite always gives the appearance of taking widest scope: if the speaker intends to refer to a particular individual in the world of utterance, that individual must exist in the world of utterance. On the other hand, not all wide-scope indefinites are referential. This is illustrated in (34).

In (34a), the wide-scope indefinite may be referential: it is quite plausible that the speaker intends to refer to a particular colleague, as evidenced by the fact that the speaker subsequently identifies a particular colleague, Jane Brown. In contrast, (34b) is a case of a non-referential indefinite (since the speaker has no particular colleague in mind) which nevertheless takes wide scope. Finally, (34c) contains a narrow-scope indefinite which is obligatorily non-referential: the speaker is not talking about any particular colleague in the actual world.

34. a) Sarah wants to talk to a colleague of mine – Jane Brown, who is very famous.
    b) Sarah wants to talk to a colleague of mine, but I don't know which one.
    c) Sarah wants to talk to a colleague of mine – any colleague will do.

Thus, when the speaker makes it clear that she is talking about a particular individual, and has knowledge of this individual, it is likely that she intends to refer to this individual, and that the indefinite is referential.

Fodor and Sag provide some other potential diagnostics for referentiality besides explicitly stated speaker knowledge. They point out that a restrictive relative clause modifying an indefinite tends to favor the referential reading (35a), and that the effect is even stronger for non-restrictive relatives (35b). Since the speaker is in possession of some knowledge of a particular individual (i.e., that this individual is a student in the syntax class who furthermore has a Ph.D. in astrophysics), and is conveying this information to the hearer, it is likely that the speaker in fact intends to refer to a particular individual[4].

35. a) A student in the syntax class who has a Ph.D. in astrophysics cheated on the exam.
    b) A student in the syntax class, who has a Ph.D. in astrophysics, cheated on the exam.   (Fodor and Sag 1982, ex. 22-23)

Non-restrictive modification strongly biases speakers into treating an indefinite as referential, and hence compatible with speaker knowledge. For instance, I have found that some speakers easily allow a denial of speaker knowledge after an indefinite with restrictive modification (36a) but find such denial infelicitous after an indefinite with non-restrictive modification (36b).

36. a) The university awarded the first prize in a writing competition to a psychology student who lives in my dorm. I have no idea which student it was.

    b) The university awarded the first prize in a writing competition to a psychology student, who lives in my dorm. #I have no idea which student it was.

---

[4] For more discussion of why a relative clause should facilitate the referential reading, see Section 5.1.2.

However, it is possible to construct contexts in which most speakers find non-restrictive modification completely compatible with a denial of speaker knowledge, as in (37). While use a non-restrictive relative typically implies that the speaker possesses some information about the referent of the indefinite, this implicature can be negated, which means that non-restrictive modification is not a diagnostic for referentiality.

37. The university awarded the first prize in a writing competition to a student from my dorm, who must be feeling very lucky! I haven't heard which student it actually was.

A clearer diagnostic for referentiality that Fodor and Sag provide is the use of modifiers like *certain, specific,* and *particular.* For instance, the use of *certain* in (38) suggests that the speaker intends to refer to a particular student. That *certain* is a marker of referentiality is further shown in (39): indefinites headed by *a certain* can take wide scope (39a) but are incompatible with denial of speaker knowledge (39b) and cannot take narrow scope under an intensional operator (39c)[5].

38. I accused a certain student of cheating.   (Fodor and Sag 1982, ex. 27)

39. a) Sarah wants to talk to a certain colleague of mine – Jane Brown.
    b) #Sarah wants to talk to a certain colleague of mine, but I don't know which one.
    c) #Sarah wants to talk to a certain colleague of mine – any colleague will do.

Finally, as Fodor and Sag note, indefinites headed by referential *this* (hence: *this*$_\text{ref}$) are obligatorily referential. This indefinite use of *this* is crucially different from its demonstrative, definite use. I will now show that *this*-indefinites are referential in the sense of Fodor and Sag.

First of all, *this*-indefinites are impossible in the scope of an intensional operator, as shown in (40).

40. a) Sarah wants to talk to this colleague of mine – Jane Brown, who is very famous.
    b) #Sarah wants to talk to this colleague of mine – any colleague will do.

Secondly, as Maclaran (1982:90) notes, use of *this*$_\text{ref}$ "draws attention to the fact that the speaker has a particular referent in mind, about which further information may be given". This is shown in (41): the speaker intends to draw attention to a particular telephone in (41b), but not in (41a).

41. a) John has {a, #this} telephone, so you can reach me there.
    b) John has {a, this} weird purple telephone. (Maclaran 1982 :88, ex. (85))

Statistical support for the view that *this*$_\text{ref}$ draws attention to a particular referent comes from Prince (1981), who found that 209 out of 243 instances of *this*$_\text{ref}$ (86%) introduced a referent which was referred to again within a few clauses, either explicitly or

---

[5] To the extent that (39b) is acceptable, it seems to pick up on the attitude of Sarah rather than on the speaker: *a certain colleague* then refers to the individual that Sarah has in mind. See Abusch and Rooth (1997), and footnote 9 in this chapter. Also, *a certain* indefinites can take scope under a higher quantifier – see the discussion in Section 4.1.2, which shows that *a certain* indefinites cannot in fact be treated as Fodor and Sag's referential indefinites.

implicitly. (Unfortunately, Prince does not report a corresponding statistic for *a*-indefinites). I will discuss the properties of *this*_{ref} in more detail in Section 5.1.

The different properties of referential vs. non-referential indefinites are summarized in (42).

42. Summary: referential vs. non-referential indefinites

diagnostic	**referential indefinites**	**non-referential indefinites**
speaker knowledge	yes	not necessarily
scope in simple clauses	wide	wide or narrow
scope out of islands?	yes	no
compatible with *this*_{ref}?	yes	no
compatible with *certain*?	yes	no

Fodor and Sag's analysis has been challenged a number of times. Ludlow and Neale (1991) argue that there is no need to posit an ambiguity for English indefinites: they propose that there is only one semantic analysis of indefinite descriptions, and that the different uses of indefinites (e.g., referential vs. non-referential) can be derived pragmatically from Gricean principles.

Other researchers, however, such as Reinhart 1997 and Kratzer 1998, have followed Fodor and Sag in positing a semantic ambiguity, but have argued against Fodor and Sag's implementation of the ambiguity. The challenge to Fodor and Sag has to do with the availability of intermediate scope readings. Fodor and Sag's analysis makes a clear prediction: a long-distance indefinite can never give the appearance of being in the scope of another quantifier. Either an indefinite is quantificational, and therefore subject to island constraints (i.e., it's not a long-distance indefinite); or, it's referential and therefore gives the appearance of widest possible scope. Fodor and Sag argued that long-distance indefinites in fact cannot take intermediate scope.

4.1.2. Choice-function analyses

However, more recent literature (cf. Ruys 1992, Abusch 1994, Reinhart 1997, Winter 1997, Kratzer 1998, i.a.) has shown that intermediate scope readings do exist. Thus, (43a) has the reading which is true if every student read every book that had been recommended by a particular teacher: e.g., Mary read every book recommended by Dr. Smith, Alice read every book recommended by Dr. Brown, etc. Under the scope-shifting analysis, (43a) would have the reading in (43b) – a reading in which *some teacher* is clearly taking intermediate scope, below *every student* but above *every book* and the *that*-clause.

43. a) Every student read every book that some teacher had recommended.
    b) [every student] λ_1 [[some teacher] λ_2 [t$_1$ read every book that t$_2$ had recommended]]

The referential analysis cannot account for such readings. Recent analyses that have accounted for these readings are *choice function* analyses. On these analyses, the indefinite article is translated as the variable *f*, which ranges over choice-functions, functions which map any non-empty set in their domain to an element of this set.

One variant of the choice-function analysis is that of Reinhart (1997) / Winter (1997). On their view, the choice function variables are bound by Existential Closure (EC), which is not syntactically restricted and is not subject to locality constraints. On this view, (29a), repeated here as (44a), receives the LF in (44b). This reading can be paraphrased as follows: there exists a way of choosing an element from a set, which, when applied to the set of teachers, chooses the teacher x such that Mary read every book which had been recommended by x.

44. a) Mary read every book that *a / some teacher* had recommended.
     b) $\exists f$ [Mary read every book f(teacher) had recommended]
     c) [every student] λ_1 [$\exists f$ [t_1 read every book f(teacher) had recommended]]

A different variant of the choice function analysis is that of Kratzer (1998), for whom choice function variables are free variables, not subject to Existential Closure. Thus, (44a) receives the LF in (45). The choice function variable is contextually determined: it is necessary that the speaker have a particular function in mind, which she does not reveal to the audience. For instance, in the case of (45), the speaker may have in mind a function which, when applied to the set of teachers, picks out Dr. Smith. I will from now on use f^s to refer to free choice function variables, with the superscript s standing for the speaker. Kratzer's analysis is close to Fodor and Sag's original analysis in that it ties the long-distance reading of the indefinite to the speaker's state of mind.

45. Mary read every book f^s(teacher) had recommended

A minimally different version of Kratzer's proposal was formulated by Matthewson (1999): choice function variables on her analysis are not left free, but are bound by EC; however, unlike the Reinhart / Winter free EC analysis, Matthewson's analysis requires EC to occur in the topmost position. The LF in (45) then becomes the LF in (46).

46. $\exists f$ [Mary read every book f(teacher) had recommended]

Kratzer's analysis accounts for intermediate scope readings by assuming that the function denoted by f takes an additional argument, which appears as an index on the function variable at LF. Chierchia (2001) explicitly formulated this proposal for the topmost EC analysis of Matthewson[6], and called the index on the function variable a *Skolem* index. Chierchia's analysis of the intermediate scope reading of the indefinite in (43a) is given in (47a), with topmost EC[7]; the minimally different analysis with a free choice function variable is given in (47b).

47. a) $\exists f$ [[every student] λ_1 [t_1 read every book [f_1 teacher] had recommended]
                                                              (from Schwarz 2001, ex. 7a)
     b) [every student] λ_1 [t_1 read every book [f^s_1 teacher] had recommended]

In both LFs in (47), the variable f is intended to range over *Skolemized* choice functions of type $\langle e, \langle et, e \rangle \rangle$: functions from individuals to choice functions. The Skolem

---

[6] Matthewson (1999) herself, using data from Salish (see Section 3.2.3), argued that intermediate scope readings are only possible when there is an *explicit* bound variable inside the indefinite.

[7] See Chierchia (2001) and Schwarz (2001) for problems with the predictions of the topmost EC analysis when the higher quantifier is downward entailing.

index is bound by the higher quantifier *every student,* so that the choice of teachers varies with the students. Under the Kratzer analysis, this means that there is a contextually determined relationship between students and teachers – e.g., each student is mapped to a choice function which, when applied to the set of teachers, picks out the student's favorite teacher.

Like Fodor and Sag's referential analysis, the choice-function analyses have to account for why indefinites headed by *a certain* do not take narrow scope under intensional verbs, as shown in (48). While *a book* in (48a) can scope under *wants,* this is not a possibility for *a certain book* in (48b): (48b) cannot mean that Mary wants to read *any* book.

48. a) Mary wants to read *a book.*
     b) Mary wants to read *a certain book.*

The topmost EC analysis and Kratzer's free variable analysis have no difficulty explaining this. These analyses simply have to say that indefinites headed by *a certain* obligatorily receive a choice-function interpretation. If the choice-function variable is free, as in (49a), then the speaker has a contextually salient function in mind which picks out a particular book from the set of all books, and Mary wants to read the book so chosen. If the choice-function variable is bound by EC at the topmost level, as in (49b), then there exists a way of choosing an element from a set such that Mary wants to read the book so chosen. In both cases, *a certain book* gives the appearance of scoping over *wants.*[8]

49. a) Mary wants to read f^s(book)
     b) $\exists f$ [Mary wants to read f(book)]

On the other hand, indefinites headed by *a* or *some* are ambiguous between a choice-function reading (which is obligatory when the indefinite is long-distance) and a quantificational reading; on the latter reading, they can scope under *wants.*

While not being able to scope under an intensional operator, indefinites headed by *a certain* do scope under higher quantifiers, as shown in (50). This is easily explained under Skolemization: *a certain*-indefinites can denote Skolemized choice functions. In (50), a Skolemized choice function maps each boy to a choice function which, when applied to the set of dates, picks out the date that stands in a particular relationship to the boy (e.g., the date is his mother's birthday). As Schwarz (2001) argues, *a certain* indefinites have all the characteristics of Kratzer-style (Skolemized) choice functions.

50. Every boy forgot a certain date (his mother's birthday).

On the other hand, the Reinhart/Winter free EC analysis cannot easily account for why *a certain* indefinites cannot scope under intensional verbs. Since EC is not syntactically restricted on this view, nothing prevents (48b) (as well as (48a)) from having the reading in (51). This reading can be paraphrased roughly as follows: in each world w' compatible with Mary's desires in the actual world w, there exists a way of

---

[8] Irene Heim (p.c.) points out that the narrow-scope reading of *a certain book* under an intensional operator could be derived on the topmost EC or free variable analysis if Skolemization with a world variable is allowed. The topmost EC and free variable analyses would have to rule out this possibility in order to explain the absence of the narrow scope reading of *a certain book* in (48b).

choosing an element from a set such that Mary reads (in w') the book so chosen from the set of all books.

51. Mary wants [∃f PRO to read f(book)]

In fact, in a version of the free EC analysis, Winter (1997) argues that all indefinites are interpreted through choice-functions – that there are no quantificational indefinites (Reinhart 1997, on the other hand, allows for the possibility of ambiguity). Then, it is not clear why (51) is a possible LF for (48a), but not for (48b).

Finally, indefinites headed by *this*$_{ref}$ do not allow intermediate scope: (52a) does not have the reading in which the professors vary with the students. In fact, when a bound variable inside the indefinite rules out the widest-scope reading of the indefinite, forcing an intermediate or narrow scope reading, use of *this*$_{ref}$ becomes infelicitous, as in (52b).

52. a) Every student read every book that this professor had recommended.
     b)#Every student$_1$ read every book that this professor of his$_1$ had recommended.

Thus, if indefinites headed by *this*$_{ref}$ are choice functions, they obligatorily have topmost EC (or no EC at all) and do not allow Skolemization and cannot be bound by a higher quantifier (unlike indefinites headed by *a certain*). The obligatorily wide scope readings of *this*-indefinites suggest that *this*-indefinites are better analyzed on a referential analysis like Fodor and Sag's than on a choice-function analysis.

4.1.3. Long-distance indefinites and speaker knowledge

To summarize, there is evidence that English indefinites are ambiguous between a quantificational reading and a different reading: referential or choice-function. (One exception is Winter 1997, for whom there are no quantificational readings of indefinites at all). The latter reading gives an indefinite the appearance of taking long-distance scope. Some analyses (Fodor and Sag 1982, Kratzer 1998) tie the readings of long-distance indefinites to some form of speaker knowledge, while others (Winter 1997, Reinhart 1997) do not consider speaker knowledge at all.

Both Fodor and Sag's and Kratzer's analyses must face the challenge of *anti-referentiality* (von Fintel and Fox 2002 Lecture Notes). Long-distance indefinites headed by *some* seem quite compatible with the denial of speaker knowledge, as demonstrated by (53). Since long-distance indefinites cannot be quantificational, this means that a referential or choice-function reading is compatible with the denial of speaker knowledge. It is not clear whether Fodor and Sag's definition of referentiality can be reconciled with denial of speaker knowledge.

53. Mary read every book that some teacher had recommended. I have no idea which teacher it was.

The same is true for *a*-indefinites. An unmodified *a*-indefinite does not easily allow for a long-distance reading, as shown in (54a) – compare to the *some*-indefinite in (53). A modified *a*-indefinite can have a long-distance reading and seems compatible both with explicit statement of speaker knowledge (54b) and denial of speaker knowledge (54c).

54. a) Mary read every book recommended by a professor.
     b) Mary read every book recommended by a professor that she knows. This professor's name is Helen Brown.

c) Mary read every book recommended by a professor that she knows. I don't know which professor it was.

On the other hand, indefinites headed by $this_{ref}$ or *a certain* seem to carry an implication of speaker knowledge, as illustrated again in (55). This fact cannot be easily accounted for under Reinhart's or Winter's approaches, but is compatible with Fodor and Sag's and Kratzer's approaches.

55. a)# Mary got a certain book from the library, but I don't know which one. [9]
    b)#Mary got this book from the library, but I don't know which one. [10]

It is possible that long-distance *some*-indefinites and *a*-indefinites arise from a different mechanism than do *this*-indefinites and indefinites headed by *a certain*. A proposal along these lines is made in Schwarz (2001), who argues, based on the truth-conditions of intermediate scope readings, that indefinites headed by *a certain* receive a Kratzer-style free variable analysis while other long-distance indefinites must be interpreted through some other mechanism, such as scope-shifting or the Reinhart / Winter choice function analysis.

To conclude, indefinites have readings that cannot be accounted for under the standard quantificational analysis of indefinites. At the same time, no one analysis can account for all of the readings of long-distance indefinites: Reinhart and Winter can't easily explain why indefinites headed by *a certain* and $this_{ref}$ cannot take narrow scope under intensional verbs, and why they imply speaker knowledge; Fodor and Sag as well as Kratzer can't easily explain why long-distance indefinites headed by *a* or *some* are compatible with a denial of speaker knowledge.

In this thesis, I will adopt Schwarz's view that there are indeed more than one kind of long-distance indefinites. I will disregard long-distance indefinites which are not directly related to speaker knowledge, assuming that they are derived through a choice function mechanism such as Reinhart's or Winter's[11]. I will focus on indefinites headed by $this_{ref}$ and *a certain*. The exceptional scope-taking properties of these indefinites, along with their apparent requirement of "speaker knowledge", suggest that they require an analysis along the lines of Fodor and Sag's or Kratzer's. In Section 5, I will define "speaker knowledge" more stringently and give a modified lexical entry for specific (referential)

---

[9] In the case of *a certain*, speaker knowledge becomes unnecessary if the indefinite is embedded under an attitude verb and picks up on the attitude of the speaker of the embedded clause, as in (i). See the discussion in Abusch and Rooth 1997 who argue that *certain* and some other adjectives, including *undisclosed*, have the property of picking up on the attitude of someone in the context (which by default is that of the matrix speaker). This is irrelevant for my present purposes.
    (i) a) There was a story in Spy about Solange. According to the story, she has moved to a certain remote island in the Pacific. I don't know which one, it was some exotic-sound place.
        b) Claude evidently believes that Solange is involved with a certain ballet dancer. I have no way of telling who this dancer is supposed to be.   (Abusch and Rooth 1997, ex. 74-75)
[10] See Section 5.1 for evidence that under the right conditions (a statement of a noteworthy property), a *this*-indefinite is actually compatible with denial of speaker knowledge. For the purposes of the present section, it is important that there is a requirement of some speaker knowledge on *this*-indefinites but not on *a*-indefinites.
[11] Schwarz (2001) shows that indefinites headed by *a* or *some* are compatible with the Reinhart / Winter choice function analysis, but, crucially, not with the Kratzer-style choice function analysis, which is possible only for indefinites headed by *a certain*.

indefinites, of which *this*-indefinites are a prime example. Right now, I turn to definites and examine the case for referential definites.

4.2. The ambiguity of the English definite

The traditional lexical entries for definite description, such as the Fregean (given again in (56) and the Russellian, are termed *attributive* readings of the definite.

56.   Fregean analysis: (from Heim 1991:9)
       [*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 on 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.

4.2.1. Referential vs. attributive definites

Donnellan (1966) argued that the attributive reading is insufficient to account for all readings that definites have, and that definites are in fact ambiguous between a referential and an attributive reading. Donnellan's famous example illustrating the two readings is given in (57).

57.   Smith's murderer is insane.

The DP in (57) is clearly definite (it can be rephrased as *the murderer of Smith*). The sentence can be uttered when the speaker believes that whoever murdered Smith must have been insane; it is not known who actually murdered Smith, but Smith was such a sweet, harmless person that his murderer (whoever that is) is clearly insane. In contrast, on the referential reading, the speaker identifies a particular individual, e.g., Brown, as the murderer of Smith. The speaker has grounds to believe that Brown is insane and therefore utters (57). It may be that Brown is not in fact Smith's murderer, and that the actual murderer is quite sane. In this scenario, Donnellan argues, (57) would still be true on the referential reading of the definite, since the referent whom the speaker identified as *Smith's murderer* (i.e., Brown) is in fact insane.

The difference between referential and attributive definites is more visible in intensional contexts, as illustrated in (58). In (58a), the speaker wants to talk to Samantha Jones, whom the speaker identifies as *the winner of this race*. The speaker may be mistaken – perhaps Samantha is not the winner; however, the speaker still wants to talk to Samantha. In (58a), the definite has a referential reading. In contrast, in (58b), the speaker wants to talk to the winner whoever that is: what matters is not the winner's identity but the description of her as the winner. This is the attributive reading.

58.   a) I want to talk to the winner of this race; her name is Samantha Jones, and I've
           been wanting to talk to Samantha for a long time.
       b) I want to talk to the winner of this race; I'm writing a story about this race for
           tomorrow's newspaper.

The semantics for Donnellan's referential definites, as formulated in Kripke 1977, is given in (59). Slightly different lexical entries for referential definites were proposed in Stalnaker (1970) and Kaplan (1978); these are given in (60a) and (60b), respectively. Crucially, in all these lexical entries, the referential definite is an indexical. All of them state that the speaker is talking about a particular individual in the world of utterance,

who furthermore has (or is presupposed / believed to have) the property denoted by the restrictor NP.

59. Donnellan (1966), as formulated in Kripke (1977):

$[[\text{the}_{ref} \ \zeta]]^{c, i} = \tau x.$ s_c in w_c at t_c takes x to have property $[[\zeta]]^c_{\cancel{\phi}}$ and s_c in w_c at t_c intends to refer to x. (from von Fintel and Fox 2002 Lecture Notes)

60. a) Kaplan (1978): *the* ζ, uttered in a world w at a time t, denotes something only if exactly one ζ exists in w at t. If this condition is fulfilled, it denotes that ζ.
(Heim 1991:3)

b) Stalnaker (1970): *the* ζ, uttered in context c, denotes something only if the speaker of c presupposes of exactly one individual that it is ζ. If so, the utterance of *the* ζ denotes that individual. (Heim 1991:12)

The ambiguity view of definites has been challenged in the literature. For instance, Kripke (1977) proposed that both attributive and referential readings of definites can be derived from the quantificational (Russellian) reading of definites via Gricean principles[12].

4.2.2. *Arguments for and against the referential reading of definites*

Heim (1991) goes through some arguments in favor of the ambiguity view of English definites, and shows that there is no real motivation to posit such an ambiguity. The main argument comes from the exceptional scope-taking properties of definites. Consider (61). Both sentences in (61) have two readings available to them. One is the narrow-scope reading of the definite, which can be paraphrased roughly as follows: it is always the case that the person who wins is the one who was the player on the left (maybe there is something lucky about the left side). The other is the wide-scope reading of the definite: the person who happens to be the player on the left (e.g., Mary) has the property of always winning. On the classical analysis of definites, (61a) is predicted to have both of these readings, but (61b) is predicted to have only the narrow-scope reading: the wide-scope reading would require the definite to scope out of a *that*-clause, which is an island. As shown in (62), other scope-bearing elements cannot scope out of a *that*-clause: only (62a), but not (62b), has the wide-scope reading of *one or two people* available to it.

61. a) The player on the left always wins.
    b) It is always the case that the player on the left wins. (Heim 1991:14)

62. a) One or two people are always late.
    b) It always happens that one or two people are late. (Heim 1991:14)

The ambiguity view of definites can easily explain the exceptional scope-taking properties of definites: in additional to the classical attributive reading, definites have a referential reading which gives the appearance of widest scope. This is the same argument that Fodor and Sag (1982) give for the existence of referential definites. It also

---

[12] Ludlow and Segal (2002) go even further and argue that definite and indefinite descriptions have the same semantics, and that the uniqueness implicature on definiteness is pragmatically derived. I do not discuss their proposal here.

runs into the same problem as does Fodor and Sag's analysis: the availability of intermediate scope readings. Consider Heim's example in (63).

63. Each time, it could have happened just as easily that the player on the left would have been on the right. (Heim 1991:16)

One of the readings available to this sentence can be paraphrased as follows: "for every time t, the player who in fact is on the left at t could just as easily have wound up on the right" (Heim 1991:16). On this reading, the definite is not referential, since the players vary with the times t. However, the definite still scopes out of a *that*-island: it must scope above the operator *it could have happened just as easily*: as Heim (1991:14) explains, "we are not, after all, discussing possible worlds in which the player who is on the left there is simultaneously on the right." If definites on the *non-referential* reading can scope out of islands, as this example shows, then there is no need to posit an additional referential reading.

Heim shows that it is possible to explain the exceptional scope-taking properties of definites by giving nouns and verbs argument positions for worlds and times, and by allowing the world and time arguments of a noun to be bound at different levels, or to be interpreted deictically. This analysis allows (61b) to have a reading in which the definite apparently takes wide scope without any violation of scope barriers.

The argument slots of *player-on-the-left* in (64) could be filled in a variety of ways. If the slots are filled with the variables (w, t'), then the original narrow-scope reading of the definite obtains, and the sentence can be paraphrased as follows: it is the case for all times t' that the person who plays on the left at time t' wins at time t' – i.e., whoever is on the left always wins.

On the other hand, if the slots are filled with the variables (w, t), the seemingly referential reading results. The meaning of the sentence would be paraphrased as follows: it is the case for all times t' that the unique person who plays on the left at time t wins at time t'. Heim shows that the same logic can be applied to derive the intermediate reading of the definite in (63).

64. λw, t [it is-always-the-case(w, t)$_{t'}$ that [[the player-on-the-left(_,_)] wins(w, t']] (Heim 1991:17)

To sum up, there does not seem to be any good reason to believe that definites in English are ambiguous. It is best to say that *the* is uniformly marked as [+definite], where the feature [+definite] corresponds to the standard Fregean definite entry (for arguments for why the Fregean entry is preferable to the Russellian, see Section 2.2).

5. The proposal: specificity as noteworthiness

In contrast, referentiality does play a role with English indefinites. Referentiality in indefinites receives morphological expression in *this*$_{ref}$. Contrary to Fodor and Sag's proposal, I maintain that indefinites headed by *a* are not ambiguous between a referential and a non-referential reading. As discussed in Section 4, long-distance *a*-indefinites (i.e., Fodor and Sag's original referential indefinites) allow for intermediate scope readings. Furthermore, they carry no implication of speaker knowledge – unlike *this*-indefinites, which do carry such an implication. I will assume that long-distance readings of *a*-

indefinites can be derived on a choice function analysis such as Reinhart's or Winter's, or through scope shifting.

From now on, I will consider only *this*-indefinites to be referential, and will build my proposal of what referentiality is on the basis of the behavior of *this*-indefinites. I will use the term *specific* instead of *referential* throughout the discussion, since *speaker intent to refer* is not as central to my proposal as it is to Fodor and Sag's. On the basis of the behavior of *this*-indefinites, I will propose a lexical entry for specific DPs in English. I will then examine the relationship between specificity and definiteness marking in English.

The discussion in this section establishes the original motivation for the Article Choice Parameter by showing that definiteness and specificity are two independent semantic features. This section lays the groundwork for my proposal concerning articles in L2-acquisition, since I will argue that L2-learners' use of *the* with indefinites parallels L1-English speakers' use of *this*$_{ref}$.

5.1. Conditions on the use of *this*-indefinites

In Section 4.1.1, I followed Prince (1981) and Fodor and Sag (1982) in arguing that *this*-indefinites are obligatorily specific. I will now explore what this means: what conditions must be satisfied in order for a *this*-indefinite to be felicitous.

5.1.1. This-*indefinites and noteworthiness*

As shown in Section 4.1.1, *this*-indefinites cannot be in the scope of an intensional operator, and carry an implication of speaker knowledge of the referent. I will now attempt to define more precisely what this "speaker knowledge" is.

At first glance, it looks like *this*-indefinites are incompatible with an explicit denial of speaker knowledge of the referent – this is shown in (65a). However, the same speakers who considered (65a) infelicitous considered (65b) and (65c) perfect or nearly perfect. Yet, in neither (65b) nor (65c) can the speaker actually name the movie under discussion, or even say anything about its content – which is also the case for the infelicitous (65a).

65.   a)# Mary wants to see this new movie; I don't know which movie it is.
       b) Mary wants to see this new movie; I don't know which movie it is, but she's been all excited about seeing it for weeks now.
       c) I want to see this new movie – I can't remember its name and I have no idea what it's about, but someone mentioned to me that it's really interesting.

Crucially, in (65b) and (65c), the speaker is able to say *something* about the movie, even if it is not something about its name or content. In (65b), the speaker can state that the relevant movie has *the property of being a movie that Mary has been talking about for weeks,* and in (65c), the speaker can state that the relevant movie has *the property of having been described to me as interesting.* In contrast, in (65a), the speaker does not exhibit knowledge of anything connected to the movie – hence the infelicity.

Another example of a use of *this*$_{ref}$ along with a denial of speaker knowledge was suggested by Martha McGinnis (p.c.) in (66a). Here, the speaker has no familiarity with the colleague in question – it could be anyone. But the speaker does know that the relevant colleague has *the property of being an old friend of John's.* (66a) is more

felicitous than (66b), in which the speaker does not exhibit knowledge of any properties of the individual in question.

66. a) John is having dinner with this colleague of mine tonight. I don't know which one it is, but John said that she is an old friend of his. If it turns out to be Sally then I'm in big trouble!

b)#John is having dinner with this colleague of mine tonight. I don't know which one it is.

The felicitous sentences in (65) and (66) all contain statements of some *noteworthy property* that holds of the referent of the *this*-indefinite[13]. Crucially, no such requirement of *noteworthiness* exists for *a*-indefinites, as shown by (67).

67. Mary wants to see a new movie. I don't know which movie it is.

5.1.2. *Where can noteworthiness come from?*

In (65) and (66), *this*-indefinites are licensed only when the speaker follows the indefinite with a statement of some *noteworthy property* that holds of its referent.

However, there are also sentences in which a *this*-indefinite is felicitous even though it is not followed by any additional statement. As an illustration, consider (68). Speakers note a contrast between (68a), which is rather infelicitous, and (68b), which is fine. The sentence in (68b), in which the *this*-indefinite bears RC-modification, is as felicitous as (68c), in which the *this*-indefinite is followed by a separate statement conveying speaker knowledge. Crucially, in both (68b) and (68c), what makes a particular new movie important for the purposes of the discourse is that *it is a movie that my friends have been recommending to me for ages*. In (68b), the noteworthy property is conveyed by the restrictor NP itself. The presence of RC-modification makes this possible: the property *x is a new movie that my friends have been recommending to me for ages* is more likely to be construed as noteworthy than the property *x is a new movie*.

68. 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.

However, RC-modification is not necessary for noteworthiness. Use of a *this*-indefinite in (69a) (suggested by David Pesetsky, p.c.) is much more felicitous than in (69b), despite having no RC-modification or explicit statement of speaker knowledge. What makes the referent of the indefinite in (69a) noteworthy is that it is blue. Apples are not normally blue, and the "unexpectedness" conveyed by the DP *this blue apple* is sufficient to make the referent noteworthy. (The same point is made by Maclaran's *telephone* example in (41b)).

69. a) I found this blue apple on my plate!
b)#I found this apple on my plate!

---

[13] Thanks to Danny Fox (p.c.) for initially suggesting this line of thought to me, and to David Pesetsky (p.c.) for coining the term *noteworthiness* to describe the phenomena discussed here.

Finally, consider (70) (suggested by Irene Heim, p.c.), in which use of a *this*-indefinite is quite felicitous. Being a cat is not particularly noteworthy, and the speaker makes no follow-up statement of a noteworthy property. The noteworthy property of the cat is that *it jumped onto the roof of my car.* Thus in this case, the noteworthy property is conveyed by the predicate.

70. I was driving down the road, and suddenly, this cat ran out of the bushes and jumped onto the roof of my car!

5.1.3. *Previous statement of noteworthiness*

The previous sections make clear that use of a *this*-indefinite is felicitous only when the speaker acquaints the listener with some noteworthy property which holds of the DP's referent. This noteworthy property can be part of the restrictor NP (as in (69a)); it can be part of the predicate (as in (70)); or it can be stated in a separate statement (as in (65b,c)). What's crucial is that the hearer be acquainted with the noteworthy property at some point.

This predicts that if the hearer has been acquainted with the noteworthy property *prior* to hearing a *this*-indefinite, the speaker can felicitously utter a statement like (68a), with no continuation. This is indeed the case. An illustration is (71), due to Martha McGinnis (p.c.), who points out that here, the use of a *this*-indefinite with no follow-up is more felicitous than in (68a). In (71), the noteworthy property of the movie is *x is such that I am leaving because of x / because I want to see x.* It is clear to A that the noteworthiness of the movie comes from the fact that the movie is causing B to leave. Thus, A does not necessarily expect B to say anything else noteworthy about the movie.

71. A: Are you leaving?
     B: Yeah, I want to see this new movie.

This example shows that use of a specificity marker on the part of the speaker is only felicitous if the hearer understands why the speaker is using it. If the speaker has used a specific indefinite, there must be something important about the individual that the speaker is talking about. As discussed above, this "something important" does not have to be directly related to the identity of the individual[14]. As an additional illustration, consider the example in (72), due to Seth Cable (p.c.). Here, use of *this* is felicitous because Speaker A understands that what's noteworthy about a particular movie x, from Speaker B's perspective, is the fact that *the speaker does not have time to see x.*

72. Speaker A: It's so frustrating being so busy with final papers. I can't go out at all!
     Speaker B: I know what you mean. I want to see this new movie, but I just don't have the time.

---

[14] Another possible way to license a *this*-indefinite may be through intonation. Consider the following scenario, suggested by David Pesetsky (p.c.). Mary is missing an important meeting at work because she has decided to go call a caterer about an event not related to her work. A colleague of hers, Laura, in explaining Mary's absence to the others, says "Mary went to call *this caterer*", with stress on the indefinite and an intonation pattern which conveys indignation and/or contempt. The noteworthy property of the caterer x is that *Mary should not be calling x during work hours.* This noteworthy property can be conveyed without any explicit statements, but rather through the intonation pattern.

5.2. The specifics of specificity

The previous section showed that *noteworthiness* plays an important role in the licensing of a specific indefinite. I therefore propose that Fodor and Sag's entry for specific (referential) indefinites be amended to include the concept of *noteworthy property*.

5.2.1. Lexical entry

The lexical entry for specific indefinites is given in (73)[15]. I am using *sp* to mean *specificity marker*; in English, the morphological manifestation of *sp* is *this*$_{ref}$.

73. 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.

As an illustration of how the definition in (73) works for indefinites, consider (74a), with the relevant lexical entries given in (74b). Here, the morphological expression of *sp* is *this*.

74. a) This strange letter just came in the mail.

   b) $[[α]]^{c, i} = λw. λx.$ x is a strange letter in w
   $[[ζ]]^{c, i} = λx.$ x just came in the mail in w
   [[this strange letter just came in the mail]]$^{c, i}$ 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, and there exists a property φ which the speaker considers noteworthy in c, and x_c is a strange letter in c and x is φ in c. If this condition is fulfilled, [[this strange letter just came in the mail]]$^{c, i} = 1$ iff x_c just came in the mail.

The term *noteworthy* as I use it here means roughly *important / relevant for the purposes of the discourse*. A property that is considered noteworthy in one discourse setting may be completely irrelevant in another. For instance, in the scenario described by (69), the property of *being a blue apple* noteworthy, since we do not expect apples to be blue. On the other hand, suppose that I am participating in a game in which the goal is to collect plastic apples of various colors. If I want to express the fact that I have collected yet another apple, which happens to be blue, it would be infelicitous for me to say *I found this blue apple*. The property of *being a blue apple* would not be considered noteworthy.

The lexical entry in (73) can 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

---

[15] The more technical variant of (73) is given in (i) below:
   (i) $λi.$ [[sp α]]$^{c, i}$ is defined for a given context c if the following felicity condition is fulfilled: s_c in w_c at t_c intends to refer to exactly one individual x_c, and $∃φ_{⟨s, et⟩}$ which s_c in w_c at t_c considers noteworthy, and $α(w_c)(x_c) = φ(w_c)(x_c) = 1$. If this condition is fulfilled, $λi.$ [[sp α]]$^{c, i} = λi.\ x_c$.

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.

5.2.2. Noteworthiness as a felicity condition

The lexical entry in (73) contains a *felicity condition* on the context. A felicity condition is crucially different from a presupposition. A presupposition is a statement presupposed to be true by *both speaker and listener*. For instance, a definite carries the presuppositions of existence and uniqueness: this means that the speaker can use a definite DP felicitously only if she can reasonably assume that her listener shares the knowledge that the set denoted by the restrictor NP contains a unique member (or that the listener can accommodate this knowledge).

In contrast, a felicity condition focuses on the knowledge state of the speaker. In deciding to use a specific DP, the speaker considers only her own view of what's noteworthy, and not the state of her listener's knowledge. Crucially, the speaker needs to acquaint her listener with what makes the DP's referent noteworthy: as the examples in Section 5.1 show, use of the specificity marker *this*$_{ref}$ is infelicitous unless the hearer is given some indication as to what the noteworthy property in question is. However, the speaker does not need to assume any prior knowledge on the part of her listener. The speaker can use a specific indefinite any time she wishes to convey something noteworthy about the referent. The speaker cannot, however, use a definite any time she wishes – she must evaluate her listener's state of knowledge first.

From the standpoint of the discourse, a felicity condition is weaker than a presupposition: the former takes only speaker knowledge into account, and the latter takes both speaker and hearer knowledge into account. This distinction between felicity conditions and presuppositions will become crucial in the account of English determiners.

It is necessary that the entry in (73) include a felicity condition: the conditions on the individual x_c and the noteworthy property φ have to be part of a felicity condition rather than part of the truth-conditions. As an illustration, take a sentence like (75a). Suppose that the speaker considers the property *x is about dolphins* noteworthy, and is using the *this*-indefinite in (75a) to draw attention to Mary having read a really interesting book about dolphins. Suppose the speaker is lying, or mistaken: Mary in fact did not read the really interesting book about dolphins – she only leafed through it. The sentence is clearly false.

Suppose, on the other hand, that the speaker utters (75b). Here, the speaker is denying all knowledge of the book's identity – i.e., there is no property φ which the speaker considers noteworthy and which holds of the book that Mary read. The sentence isn't false under these conditions, but it is infelicitous. This means that the existence of φ is part of the felicity conditions rather than the truth-conditions.

75. a) Mary read this really interesting book. (It was about dolphins.)
    b)#Mary read this really interesting book, but I don't know which one.

5.2.3. Noteworthiness: the details

The lexical entry in (73) includes both the concepts of *speaker intent to refer* and *noteworthiness*. Noteworthiness is necessary because *speaker intent to refer* cannot, on its own, account for the felicity conditions on *this*$_{ref}$. We have seen that a *this*-indefinite is infelicitous if the speaker denies all knowledge of the referent (76a), or if the speaker

states nothing noteworthy about the referent (76b). However, in both sentences in (76), we could say that the speaker *intends to refer to a particular individual*. One might argue that the speaker cannot intend to refer to a particular movie in (76a), on the grounds that one cannot refer to something about which one lacks all knowledge. However, the speaker of (76b) clearly has knowledge of a particular movie, so there is no *a priori* reason to think that the speaker cannot refer to that movie.

76. a)# Mary saw this interesting movie, but I don't know which one.
    b)#I just saw this interesting movie.

Thus, noteworthiness is necessary: it specifies what makes a particular individual (in this case, a particular movie) noteworthy, and sets up an expectation on the part of the hearer that the noteworthy property of this individual will be discussed. The reason that *this* is infelicitous in the sentences in (76) is that the speaker does not tell the hearer what makes a particular movie noteworthy.

However, the existence of a noteworthy property is not on its own sufficient. Suppose that we simply said that in order for a sentence containing a *this*-indefinite to be felicitous, the predicate must hold of some x in the set denoted by the restrictor NP, where x has the noteworthy property φ. The problem with this proposal is that there may be multiple individuals which satisfy this condition. This is illustrated in (77).

If the speaker uses a *this*-indefinite such as *this book about the habits of the underwing moths* in (77a), nothing precludes the existence of multiple books about the habits of underwing moths – as is shown by the felicity of the continuation in (77a). Compare this to the use of a definite description in (77b): as soon as the speaker has uttered *the book about the habits of the underwing moths,* the hearer has accommodated the fact that only one such book exists, and the continuation concerning multiple such books is infelicitous.

77. a) I read this book about the habits of the underwing moths the other day. There are only three books about the habits of the underwing moths in existence.
    b) I read the book about the habits of the underwing moths the other day. #There are only three books about the habits of the underwing moths in existence.

As (77a) shows, there is no problem with the existence of multiple members of a set which share a single noteworthy property such as *x is about the habits of the underwing moths.* Suppose, moreover, that the noteworthy property of the DP's referent in (77a) is *x is one of only three books about underwing moths.* By definition, there are three individuals in the world which have this noteworthy property! But crucially, the speaker of (77a) *intends to refer* to only one of the books which have this noteworthy property.

We might of course say that the noteworthy property in fact states that *x is one of only three books about underwing moths, and I read x the other day.* This noteworthy property will most likely be held by only one individual. I say *most likely,* because we can imagine a scenario in which the speaker read two books about underwing moths the other day, but only wishes to talk about one of them. Then we have to change the noteworthy property to *x is one of only three books about underwing moths, and I read x the other day at 3pm* or *x is one of only three books about underwing moths, and I read x the other day and x really made an impression.* At some point, however, it becomes quite questionable whether this lengthy description is in fact a noteworthy property. The

noteworthiness of the book is not coming from the fact that it was read at 3pm, but, most likely, simply from the fact that it is a rare book about underwing moths.

In order to ensure that a *this*-indefinite refers to a particular individual from the set denoted by the NP, rather than to any individual from the set who happens to have the noteworthy property φ, we need to have *speaker intent to refer* be part of the lexical entry.

5.2.4. Specific indefinites, scope, and embedding

The lexical entry in (73) ensures that specific indefinites obligatorily take scope above intensional verbs and modals: since the speaker intends to refer to a particular individual in the world of utterance, this individual must exist in the world of utterance. This is a correct prediction, since we have seen that *this*-indefinites cannot scope under intensional or modal operators.

However, there is evidence that *this*-indefinites can take scope under attitude verbs. When a *this*-indefinite is embedded under an attitude verb, it can sometimes reflect the state of mind of the referent of the matrix subject, rather than that of the speaker.

Consider (78a), where the speaker is unlikely to be referring to a particular unicorn which exists in the actual world – the speaker does not even believe that unicorns exist. The belief that unicorns exist, and the intent to refer to a particular unicorn, is all on the part of Sarah.

Compare (78a) to (78b), where there is no embedding. By stating that *Sarah found this unicorn in her garden,* the speaker is committing herself to referring to a particular unicorn[16].

> 78. a) Sarah said that she found this unicorn in her garden. Unicorns don't exist, so she is either lying or crazy!
>
> b)#Sarah found this unicorn in her garden. But unicorns don't exist!

Prince (1981) has other examples, given in (79), in which a *this*-indefinite does not reflect the speaker's state of mind. In (79a), the speaker is clearly not intending to talk about any Eskimo restaurant that exists in the world of utterance – (79a) does not entail that the speaker even believes Eskimo restaurants to exist. Similarly, in (79b), the belief in the existence and importance of a particular Eskimo restaurant is on the part of John rather than the speaker (who believes John to be delusional). While (79b) does not have embedding under an attitude verb, it may be understood as having implicit embedding: the meaning expressed by (79b) is, roughly, *Poor old delusional John said that he wanted to eat in this Eskimo restaurant.* Crucially, the *this*-indefinite in (79) is neither *de re* nor *de dicto*: it neither states that a particular Eskimo restaurant exists (the *de re* reading), nor that John wants to eat in just any Eskimo restaurant (the *de dicto* reading).

> 79. a) John dreamt that he was in this Eskimo restaurant.
> b) Poor old delusional John wanted to eat in this Eskimo restaurant.
>
> (Prince 1981:241, ex. 35b, 37a)

---

[16] To the extent that (78b) is acceptable, it implies that Sarah told the speaker that "she found this unicorn in her garden" – see the discussion concerning *want* further in this section.

If *this*-indefinites are indexicals, as I have been claiming, they should always refer to individuals in the actual world and not be affected by embedding. One possible solution to the facts in this section is to treat *this*-indefinites in the framework of Schlenker (2002, 2003), who proposes that attitude verbs manipulate the context variable. Schlenker's proposal for the verb *say* is shown in (80).

80. *John says$_{c_i}$ that p* is true$_s$ iff every context c compatible with John's claim is such that *p* is true$_{s[c_i->c]}$ when uttered in c. (Schlenker 2002, ex. 28)

Under this proposal, the first sentence in (78a) would have the truth-conditions in (81).

81. *Sarah said$_{c_i}$ that she found this unicorn in her garden* expresses a proposition when the following condition is met: in every context c compatible with Sarah's claim, the speaker of c (namely, Sarah) intends to refer to exactly one individual x_c in c, and there exists a property φ that Sarah considers noteworthy in c, and x_c is a unicorn in c and x_c is φ in c. When this condition holds, *Sarah said$_{c_i}$ that she found this unicorn in her garden* is true$_s$ iff every context c compatible with Sarah's claim is such that Sarah in w_c at t_c found x_c in her garden.

It may be possible to explain cases such as (79b) under this proposal as well, if we understand there to be an implicit attitude verb in the context, as follows. The reason that I, the speaker, possess the information conveyed by (79b) is that *John said that he wanted to eat in this Eskimo restaurant* or else *I heard from a reliable source that John wanted to eat in this Eskimo restaurant* or *John thought that he wanted to eat in this Eskimo restaurant (and I have access to John's thoughts)*.

However, it is not possible to explain the *dream* cases such as (79a) under Schlenker's proposal. Suppose that we treated *dream* as an attitude verb, which, like *say* on Schlenker's proposal, manipulates the context variable. (79a) will then have the truth-conditions spelled out in (82) (I substitute *speaker of c* by *author of c*). However, (82) is rather bizarre, since it requires John *in John's dream-world* to be referring to some object in this dream-world. However, a dream isn't a conversation – in his dream, John is sitting in a restaurant, not referring to one.

82. *John dreamt that he was in this Eskimo restaurant* expresses a proposition when the following condition is met: in every context c compatible with what John dreamt, the author of c (namely, John) intends to refer to exactly one individual x_c in c, and there exists a property φ that John considers noteworthy in c, and x_c is an Eskimo restaurant in c and x_c is φ in c. When this condition holds, *John dreamt that he was in this Eskimo restaurant* is true$_s$ iff every context c compatible with John's dream is such that John in w_c at t_c was in x_c.

An alternative explanation, suggested by Irene Heim (p.c.), is that it is possible to use a *this*-indefinite to refer to fictional objects. Thus, in (79a), the matrix speaker is referring to the fictional restaurant of John's dream, with the truth-conditions given in (83).

83. *John dreamt that he was in this Eskimo restaurant* expresses a proposition only in those utterance contexts c where the following felicity condition is fulfilled: the speaker intends to refer to exactly one individual x_c in c, and there exists a property φ that the speaker considers noteworthy in c, and x_c is a (fictional) Eskimo restaurant in c and x_c is φ in c. When this condition holds, *John dreamt that he was in this Eskimo restaurant* is true$_s$ iff John in w_c at t_c dreamt that he was in x_c.

The fictional-object proposal can be extended to cover other cases besides dreams: for instance, reports such as (79a), where Sarah's story has established a fictional unicorn. This is particularly plausible when the matrix speaker possesses some information about the fictional object, as in the scenario in (84). Here, the speaker is intending to refer to a particular unicorn: the fictional unicorn of Sarah's imagination.

84. Sarah said that she found this unicorn in her garden. She named it Mabel, and feeds it on grapes and mozzarrella cheese.

This predicts that sentences such as (79a) in principle have two readings available to them[17]. On the first reading, the *this*-indefinite is evaluated with regard to the reported context: the referent of the matrix subject (in this case, Sarah) is the person who intends to refer to a particular unicorn. On the second reading, the *this*-indefinite is evaluated with regard to the original matrix context, and the matrix speaker is intending to refer to a particular (fictional) unicorn. This seems like an intuitively correct prediction, since (79a) does appear to have both the reading on which the matrix speaker considers the unicorn noteworthy, and the reading on which Sarah considers the unicorn noteworthy.

5.2.5. A note on Skolemization

As discussed previously, *this*-indefinites can never take narrow scope with respect to a higher quantifier. On the other hand, indefinites headed by *a certain* are able to take scope under a higher quantifier; at the same time, indefinites headed by *a certain* share many of the properties of *this*-indefinites, such as inability to scope under an intensional verb and implication of speaker knowledge. Thus, it may be possible to extend my analysis to indefinites headed by *a certain,* saying that they are specific, with the additional possibility of Skolemization[18].

For instance, in (85a), *a certain woman he knows* denotes for each child, a woman who stands in a particular relationship to that child (e.g., the child's mother); according to Schwarz 2001, this reading can only be derived through a contextually determined Skolemized choice function. When this function takes as its arguments an individual x

---

[17] It may also be possible to analyze the data in this section under the analysis of Abusch and Rooth (1997) for *certain* – see footnote 9. I will not pursue this here.

[18] A word of caution is needed, however. While *this*-indefinites and *a certain*-indefinites both imply speaker knowledge, the felicity conditions on the use of *this* vs. *certain* are actually slightly different. As Ken Wexler (p.c.) points out, (ia) is felicitous while (ib) is not: the speaker cannot use *a certain* to express her surprise at the appearance of a blue apple on her plate. This suggests that *certain* is not a marker of specificity in the same way that *this* is. Rather, that the semantics of *certain* make it compatible with most (but not all) contexts in which the felicity conditions on specificity (and use of *this*$_{ref}$) have been satisfied. I leave the question of exactly how the semantics of *certain* relate to specificity to further research.

    (i) a)   I found this blue apple on my plate!
         b) # I found a certain blue apple on my plate!

and the set of all women, it will return a woman from the set who stands in a particular relationship to x. The relationship (such as the mother relationship) is contextually determined.

Indefinites headed by *a certain* should, according to Schwarz, have the functional reading available to them in the absence of any scope islands. Thus, in (85b), *a certain date* denotes, for each child, the date that is the birthday of that child's mother (see also the discussion in Section 4.1.2).

85. a) Every child who hates <u>a certain woman</u> he knows will develop a serious
         complex.   (Schwarz 2001, ex. 66)
     b) Every boy forgot <u>a certain date</u> (his mother's birthday).

It should be possible to derive these functional readings under my analysis of specificity: the "contextually determined relationship" described above is very similar to the noteworthy property φ that I have been arguing for. In (85b), for instance, there is a noteworthy relationship between children and dates such that a contextually determined function maps each child to the function which chooses, from the set of all dates, the date on which that child's mother was born. In (86), I give the non-Skolemized entry for a specific determiner such as *this*$_{ref}$, and in (87), I give a Skolemized version that takes an additional type e argument, which will appear as an index bound by the higher quantifier at LF.

86. $[[sp]]^{c, i} = \lambda\alpha_{\langle s, et\rangle}$. the individual x_c, as long as the following conditions on the
     context c are fulfilled: s_c in w_c at t_c intends to refer to exactly one individual x_c
     and $\exists\varphi_{\langle s, et\rangle}$ such that s_c in w_c at t_c considers φ noteworthy, and $\alpha(w_c)(x_c)=$
     $\varphi(w_c)(x_c) = 1$

87. $[[sp]]^{c, i} = \lambda y_e. \lambda\alpha_{\langle s, et\rangle}$. the individual x_c, as long as the following conditions on the
     context c are fulfilled: $\exists f_{\langle e, \langle s, et\rangle\rangle}$ such that s_c in w_c at t_c considers $\varphi=f(y)$
     noteworthy, and for a unique individual x_c, $\alpha(w_c)(x_c)= \varphi(w_c)(x_c) = 1$

I will not look at intermediate readings of specific indefinites in this thesis and will from now on concentrate on widest-scope specific indefinites, which receive the analysis in (73).

5.3. Noteworthiness and definites

The previous discussion has concerned exclusively specific indefinites. However, the felicity conditions on specificity, spelled out in (73), can in principle be satisfied by a context which also satisfies the conditions on definiteness. As I will later show, both cross-linguistic data and L2-English data support the view that definites, just like indefinites, can be either specific or non-specific. In this section, I take a look at English definites and their relationship to specificity.

5.3.1. Specificity and the referential/attributive distinction in definites

Recall Donnellan's referential/attributive distinction, discussed in Section 4.2. Examples of a referential and an attributive definite are given in (88a) and (88b), respectively.

88.   a) I want to talk to <u>the winner of this race</u>; her name is Samantha Jones, and I've
been wanting to talk to Samantha for a long time.
b) I want to talk to <u>the winner of this race</u>, whoever that is; I'm writing a story
about this race for tomorrow's newspaper.

Now, how can we tell whether the definite DPs in (88) are specific? Definites are incompatible with the specificity marker *this*$_{ref}$ in English – I will discuss in Section 5.4.3 why that is the case. As discussed in Section 4.2.2, there is no compelling evidence to believe that English definites are ambiguous: thus, *the* never has the lexical entry in (86).

However, even though specificity is not morphologically encoded with definites, the *context* in which a definite DP is used may still satisfy, or fail to satisfy, the felicity conditions on specificity. This can be seen if we construct parallel indefinite and definite examples, as in (89) and (90). (I am using extensional rather than an intensional contexts so as to avoid narrow-scope indefinites, which are obligatorily non-specific).

The indefinite in (89a) is more likely to be specific than the indefinite in (89b): use of a *this*-indefinite is more felicitous in (89a), where the speaker is stating a possibly noteworthy property, *x is well-known for x's political opinions,* then in (89b), where the speaker is not stating anything noteworthy about a particular professor.

In (90), the indefinite from (89) is substituted with a definite. Given the parallel contexts, we can expect that the felicity condition on specificity is equally well satisfied in the (a) cases in (89) and (90) and equally not satisfied in the (b) cases. One may construe the (b) contexts as specific if the property of being *professor of history* or *chair of the history department* are considered noteworthy; however, the (a) contexts are more likely to be construed as specific, since additional information is given about the referent. In the (a) cases, unlike the (b) cases, the identity of the particular professor or chair is important for the discourse.

89.   a) William is interviewing <u>a/this professor of history tonight</u> – Professor
Mitchell, who is well-known for her political opinions; William will have quite an
interesting story!
b) William is interviewing <u>a/#this professor of history</u> tonight (whoever it is).
His next assignment is to interview someone from the physics department.

90.   a) William is interviewing <u>the chair of the history department</u> tonight – Professor
Mitchell, who is well-known for her political opinions; William will have quite an
interesting story!
b) William is interviewing <u>the chair of the history department</u> tonight (whoever
that is). His next assignment is to interview someone from the physics
department.

The point of the above comparison is to simply to show that, in principle, both definite and indefinite DPs can be used in contexts which are specific and in contexts which are non-specific. This will become relevant for the discussion of cross-linguistic data in Section 6. For now, I simply note that definite contexts may satisfy the felicity conditions on specificity; crucially, specificity is not morphologically expressed in these contexts: *this*$_{ref}$ cannot be used with definites, as I will discuss in Section 5.4.3.

5.3.2. Specificity and previous-mention definites

Finally, I address the case of previous-mention definites. The contexts in which previous-mention definites are used necessarily satisfy the conditions on specificity. In (91), for instance, the felicity condition on specific DPs is satisfied since the property *I bought x at the store* is quite noteworthy: this is the property that makes a particular book stand out from the set of books.

91. I bought a book and a magazine at the store. When I got home, I read <u>the book</u>.

In fact, previous-mention definites *must* be specific. In (91), a particular book is noteworthy precisely because it is *the book that I bought at the store*: this noteworthy property distinguishes this book from all others. The property of the referent that has been previously mentioned is precisely what sets the referent of the definite DP apart from all the other members of the set denoted by the restrictor NP.

5.3.3. Specificity marking with definites

A logical question that flows out of the above discussion is the following: if definites can be specific, why can't they ever be headed by $this_{ref}$? That they aren't is illustrated by (92), which shows that $this_{ref}$ is incompatible with DPs that are obligatorily definite.

92. a)# I talked to this mother of my friend Sam – she is really great.
    b)#I want to meet this best teacher in the school – I've heard so much about her.

One possible answer would be to say that $this_{ref}$ is incompatible with the presupposition of uniqueness: that it carries a "non-uniqueness condition". However, this proposal does not work, for the same reason that a proposal assigning a non-uniqueness condition on the use of *a* does not work (see Section 2.3.1). The sentences in (93), with $this_{ref}$, are perfectly felicitous, and yet do not presuppose the existence of multiple 20 ft. long catfish or multiple pathologically nosy neighbors of mine.

93. a) Robert caught this 20 ft. long catfish yesterday!
    b) This pathologically nosy neighbor of mine broke into the attic last night!

Thus, a different explanation for the infelicity of (92) is needed. I address this in the next section.

5.4. Definiteness and specificity in the English article system

In this section, I will look at how definiteness and specificity are encoded by English determiners.

5.4.1. Feature inventory

Given the discussion in this chapter so far, it is now possible to classify each DP as being [+definite], [+specific], or neither. The [+definite] feature has the standard Fregean lexical entry for the definite article (94), and the [+specific] feature has the lexical entry that I have proposed (95).

Consider now articles in (spoken) English, which can be treated as a three-article language – the articles being *the, a,* and $this_{ref}$. While $this_{ref}$ may not be allowed in all dialects and registers of English, I am concerned with those dialects/registers where it is allowed. I will now discuss which of the three articles is inserted in which environment.

94. $[[+\text{definite}]]^{c,\,i} = \lambda\alpha_{\langle e,\,t\rangle}.$ the unique individual x such that $[[\alpha]]^{c,\,i}(x) = 1.$

95. $[[+\text{specific}]]^{c,\,i} = \lambda\alpha_{\langle s,\,et\rangle}.$ the individual x_c, as long as the following conditions on the context c are fulfilled: s_c in w_c at t_c intends to refer to exactly one individual x_c, and $\exists\phi_{\langle s,\,et\rangle}$ such that s_c in w_c at t_c considers ϕ noteworthy, and $\alpha(w_c)(x_c) = \phi(w_c)(x_c) = 1$

The is specified as [+definite] and *this* on its referential use is specified as [+specific], as shown in (96) and (97). The indefinite determiner *a* is not specified for either definiteness or specificity, and has the simple quantificational entry in (98). (As discussed in Section 5, I differ from Fodor and Sag in not considering *a* to be ambiguous).

96. $[[\text{the}]]^{c,\,i} = \lambda\alpha_{\langle e,\,t\rangle}.$ the unique individual x such that $[[\alpha]]^{c,\,i}(x) = 1.$

97. $[[\text{this}]]^{c,\,i} = \lambda\alpha_{\langle s,\,et\rangle}.$ the individual x_c, as long as the following conditions on the context c are fulfilled: s_c in w_c at t_c intends to refer to exactly one individual x_c, and $\exists\phi_{\langle s,\,et\rangle}$ such that s_c in w_c at t_c considers ϕ noteworthy, and $\alpha(w_c)(x_c) = \phi(w_c)(x_c) = 1$

98. $[[a]]^{c,\,i} = \lambda\zeta_{\langle e,\,t\rangle}.\,\lambda f_{\langle e,\,t\rangle}.\,\exists x\,[\,[[\zeta]]^{c,\,i}(x) = 1\ \&\ f(x) = 1\,].$

I now look at what determines which English article is inserted in which environment. I will talk of environments as [+/-definite] and [+/-specific], as shorthand for "satisfying the conditions on definiteness/specificity": for instance, when I talk about an environment which is "[+definite, +specific]", I am talking about an environment in which both the conditions on definiteness and the conditions on specificity have been met. An environment which is "[+definite, -specific]" is one which satisfies the conditions on definiteness, but not the conditions on specificity. And so on.

5.4.2. The [+definite, -specific] environment

Suppose that we have an environment which is [+definite, -specific]: i.e., an environment in which the existence of a unique individual of which the predicate holds is presupposed. The lexical entry for *the* is clearly compatible with this environment, but so is the lexical entry for *a*, since the truth-conditions of *a* entail the truth-conditions of *the*. As already discussed in Section 2.3, Maximize Presupposition ensures that *the* rather than *a* is used. As for *this*ref, it is obviously incompatible with a [-specific] environment.

5.4.3. The [+definite, +specific] environment

Suppose next that we have an environment which is [+definite, +specific]. Both *the* and *this* are compatible with this environment (so is *a*, but we have already ruled out *a* with definites). However, *the* carries a presupposition while *this*ref carries only a felicity condition (see Section 5.2.2). *The* is used whenever both speaker and hearer are able to presuppose the existence of a unique individual in the restrictor set. *This*ref is used whenever the speaker wishes to convey that the DP's referent has some noteworthy property.

From the standpoint of the hearer, use of *the* is much more informative than use of *this*ref. As an illustration, consider (99). Upon hearing *the* in (99a), the hearer classifies *the cat* as [+definite] and understands that there must be a unique cat in the discourse; the logical candidate for the referent of *the cat* is the cat that was previously mentioned.

Next, consider (99b), and imagine for a moment that English does allow $this_{ref}$ with definites (ignore the demonstrative use of *this* for the purposes of this example). Upon hearing *this,* the hearer classifies *this cat* as [+specific] and understands that the speaker intends to convey something noteworthy about a particular cat. There are a number of possible candidates for noteworthiness: perhaps what's noteworthy is that *the speaker just saw x*; or perhaps it is that *x started meowing*; or perhaps it is some property that the speaker hasn't stated yet, such as *x was obviously hungry.* Suppose that the hearer will settle on one of the last two possibilities: he will then not necessarily know that the cat that started meowing is the same cat that the speaker saw – which, after all, is what the speaker was trying to convey.

99. a) I saw a cat. Suddenly, *the cat* started meowing.
    b) I saw a cat. Suddenly, *this_{ref} cat* started meowing.

Similarly, consider the case of a definite whose referent has not been previously mentioned, as in (100). Use of *the* in (100a) clearly conveys that the store has one owner. Use of $this_{ref}$ in (100b) conveys that there is something noteworthy about a particular owner (e.g., that her name is Ms. Greene) but in principle allows for the possibility of multiple owners.

100. a) I need to talk to the owner of this store – Ms. Greene.
     b) I need to talk to $this_{ref}$ owner of this store – Ms. Greene.

Thus, use of *the* in [+definite] environments is clearly more informative than use of $this_{ref}$. When the speaker uses *the* in order to establish uniqueness, she can still list various noteworthy properties of the referent. However, when she uses $this_{ref}$ in order to establish noteworthiness, she cannot easily establish uniqueness.

Given this discussion, I propose that felicity conditions, which are known only to the speaker, should not be given the same status as presuppositions, which are known to both speaker and hearer. It can be argued that from the standpoint of "Maximize Presupposition", a determiner whose lexical entry contains a felicity condition, such as $this_{ref}$, has the same status as a determiner with no felicity condition or presuppositions, such as *a.* This means that *the* will be used rather than $this_{ref}$ in all [+definite, +specific] environments. This is a desirable prediction, since $this_{ref}$ is in fact infelicitous with definites, as shown in (92). The "Maximize Presupposition" principle explains why $this_{ref}$ occurs only with indefinites in English.

5.4.4. The [-definite, +specific] environment

Next, suppose that we have a [-definite, +specific] environment. Both $this_{ref}$ and *a* are compatible with specific indefinite readings. Since neither one carries a presupposition, both should be possible in such environments. This seems like a correct prediction, since in fact $this_{ref}$ is always optional: the existence of a noteworthy property does not force the speaker to use $this_{ref}$ over *a.*

If that is the case, then why is $this_{ref}$ ever used at all? A natural answer is that a speaker uses $this_{ref}$ whenever she wants to signal the existence of a noteworthy quality that holds of a particular individual in the restrictor set. So, if the speaker says *I saw this beautiful cat yesterday!,* she is signaling that she intends to refer to a particular cat, which bears a noteworthy property; the hearer can then reasonably expect the speaker to talk about this particular cat again, perhaps explaining what the noteworthy quality is. This

expectation is apparently confirmed, given Prince's (1981) finding that individuals denoted by *this*-indefinites are nearly always referred to again later in the discourse.

An interesting question is whether the relationship between $this_{ref}$ and *a* can be derived via scalar implicature: i.e., does use of *a* imply that the felicity conditions on $this_{ref}$ have not been met? This does not seem to be the case. If the speaker says *I saw a beautiful cat yesterday,* there does not seem to be any implicature that the cat is completely not noteworthy is any way. On the other hand, there is no implicature that the cat *is* noteworthy, or that it will be referred to again later in the discourse. It is not clear whether any standard Gricean conversational maxims would be able to derive a relationship between $this_{ref}$ and *a*. For instance, consider the maxim of quantity: "Make your contribution as informative as is necessary given the purpose of the conversation!" Is signaling the existence of a noteworthy property (noteworthy for the speaker) necessarily informative for the listener? This is something to be explored. For the time being, I conclude that there is no "Maximize Felicity Conditions" principle.

5.4.5. The [-definite], [-specific] environment

Finally, consider the environment which is [-definite, -specific]. Here, neither *the* nor $this_{ref}$ are felicitous, so *a* is naturally the only option.

5.4.6. Summary

The above discussion is summarized in (101), which shows which English articles are inserted in which environments.

101. Article insertion for spoken English:
   [+definite, +specific]:   *the*
   [+definite, -specific]:   *the*
   [-definite, +specific]: $this_{ref}$ (and maybe *a*)
   [-definite, -specific]:   *a*

In the next section, I will consider other possible combinations of article features and compare them with actual article distributions in natural language.

6. Specificity and definiteness cross-linguistically

In this section, I will consider some possible ways in which languages encode definiteness and specificity. I consider hypothetical languages which have two, three, or four articles which can be used with singular DPs (I ignore plurals so as not to consider the additional [number] feature specification on articles). I do not address in detail the case of a language which has only one article, used across all singular contexts, since such an article cannot be specified for either definiteness or specificity, but must be compatible with both, as well as with absence of either.

For the purposes of this section, a language which uses an article with some types of singular DPs and licenses article omission with other types of singular DPs is considered to be a two-article language which has a null article option[19].

---

[19] Whether null articles actually exist is irrelevant. What's relevant is that a language which allows article omission with singulars in some contexts, but requires an article in other contexts, has two types of DPs: lexically-headed DPs and bare DPs.

6.1. Positive vs. negative feature specifications

I am assuming that articles have only positive feature specifications: that there are no articles marked [-definite] or [-specific]. A [-definite] article would be one that is inherently incompatible with a [+definite] context, i.e., with the presupposition of uniqueness. An article that carries a "non-uniqueness" condition would be such a possibility – see Section 2.3.1 for a summary of Heim's (1991) argument that *a* does not carry a non-uniqueness condition. Suppose that a language has a [+definite] article which presupposes uniqueness, and a [-definite] article which presupposes non-uniqueness. This language would need yet a third article for those (indefinite) contexts where neither presupposition is satisfied. This shows that an article with a non-uniqueness presupposition is not a true opposite of a definite article – it does not cover all cases of indefinites. The true opposite of the [+definite] article is an article that can be inserted in all indefinite contexts – i.e., an article that is simply not specified as [+definite].

The same reasoning can be applied to articles marking specificity. A [-specific] article would be one that is inherently incompatible with specificity: an example might be an article that carries a felicity condition of *no speaker knowledge*. This article would be felicitous only when the speaker has no knowledge of a particular member of the restrictor set, and therefore cannot intend to refer to it.

However, there are clearly cases of DPs which require neither that the speaker must intend to refer, nor that the speaker cannot possibly intend to refer. For instance, if I simply want to express the fact that I read a book last night, I neither want to attach any importance to a particular book, nor state that I lack knowledge of any particular book. Thus, an article with a *no speaker knowledge* condition is not the true opposite of the [+specific] article. The true opposite of the [+specific] article is one that can be inserted into all non-specific contexts – i.e., one that is not specified for specificity.

I am not aware of any evidence that non-uniqueness or a *no speaker knowledge* condition play a role in article systems cross-linguistically[20]. I thus assume that, as far as definiteness and specificity are concerned, articles have three possibilities open to them: being marked as [+definite], being marked as [+specific], and being underspecified.

I now discuss what article classification systems are predicted to exist given the above discussion[21].

---

[20] There is evidence that something akin to *no speaker knowledge* is encoded by forms of *some* cross-linguistically. Haspelmath (1997:47) discusses the German indefinite pronoun *irgend*, 'someone', which implies that the speaker is unaware of the referent's identity. Petrova (2003) makes a similar claim for one form of the Bulgarian adjectival 'some', *njakoj*. Related proposals have been put forth for *nanika* 'something' in Japanese (Moore 2003) and *some X or other* in English (Becker 1999).

Thus, some forms of *some* cross-linguistically are inherently incompatible with specificity. However, it is not sufficient to say that these forms are [-specific], since there are many different ways in which a lexical item could be incompatible with specificity. Rather, the semantics of these forms of *some* are incompatible with the semantics of specificity.

[21] Throughout this discussion, I am assuming that there is a category of "article" which corresponds to the D head of a DP, and that any given DP can contain only one article, which is specified for some combination of features. An alternative view (pointed out to me by David Pesetsky, p.c.) is to think of DPs as potentially having multiple positions for different features, rather than a single article: e.g., a DP might have a "definiteness" position as well as a "specificity" position inside it, either, both, or neither of which may be filled. In English, *the* fills the definiteness slot, and *this* fills the specificity slot; some constraint on English prevents the two slots from being filled at the same time. But one might imagine a language which is allowed to fill both slots at once: e.g., a language in which a [+definite, +specific] DP would look like

6.2. Two-article languages: definiteness vs. specificity distinctions

Take a language with only two articles. How might it encode definiteness and specificity? One possibility is that it will encode definiteness only – this is the case for those dialects/registers of English which do not allow the referential use of *this*. Such languages marks *the* as [+definite] and gives *a* the standard quantificational reading. Specificity will play no role in article choice – *the* will be used with both specific and non-specific definites, and *a* – with both specific and non-specific indefinites.

6.2.1. A two-article language with a specificity distinction

Consider next a hypothetical language, called English2, which has two articles: *the2*, which is [+specific], and *a2,* which is underspecified. This language has no definite or indefinite articles. What will article use in this language be like?

Since there is no definiteness marker, there will be no difference between [+definite, +specific] and [-definite, +specific] environments: both will be fully compatible with *the2*. Both will also be compatible with *a2* (cf. the discussion of *this*$_{ref}$ vs. *a* for English), but *the2* will be preferred whenever the speaker wants to signal the noteworthiness of a particular individual or group of individuals. The lack of presuppositions on *the2* means that the distinction between *the2* and *a2* will be much less rigid than the one between *the* and *a* in regular English.

In both [+definite, -specific] and [-definite, -specific] environments, on the other hand, *the2* will be outlawed and only *a2* will be possible. This is summarized in (102).

102. Article insertion for English2:
    [+definite, +specific]:   *the2 (*and maybe *a2)*
    [+definite, -specific]:   *a2*
    [-definite, +specific]:   *the2 (*and maybe *a2)*
    [-definite, -specific]:   *a2*

There is evidence that "English2" has a real-world counterpart: Samoan. I discuss this in the next section.

6.2.2. Evidence for the specificity distinction: articles in Samoan

Lyons (1999) notes that a number of languages use articles to encode *referentiality,* where "referentiality and non-referentiality are extended to embrace instances where the speaker may be in a position to identify the referent of the noun phrase but chooses to treat its identity as significant or not" (Lyons 1999:178). This is reminiscent of the specificity distinction that I have been discussing. Lyons suggests that some languages that incorporate referentiality into their determiner systems are Samoan, Hausa, Nama

---

"the this NP." While some languages do have morphologically complex determiners, I do not know of any language in which the morphological composition of the determiner corresponds to its semantics. For instance, Chung and Ladusaw 2003 discuss the Maori indefinite article *te¯tahi,* which is morphologically created from the definite article + the numeral *one.* Chung and Ladusaw show that this article behaves neither like a definite nor like a numeral, and argue that its meaning is independent of its morphological make-up.

As will be discussed in Section 6.2.3, it is not in fact possible to combine the features [+definite] and [+specific] without positing a new principle of semantic composition. I leave open the question of whether such a principle exists, and of whether DPs in any language can have multiple determiner slots corresponding to different features. I proceed on the assumption that a DP can contain only one article.

and Sissala. Of these, Samoan seems to present the clearest case of a language whose articles encode specificity.

Lyons's data on Samoan come from the work of Mosel and Hovdhaugen (1992), according to whom 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."

Consider (103). In (103a), 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, *this*$_{\text{ref}}$ could be used in this context.

In (103b), the story continues – the characters have been previously mentioned, so the DPs are definite: again, *le* is used. Mosel and Hovdhaugen give other examples of *le* use to mark definite and specific DPs (as well as generics in affirmative sentences, and predicative NPs in equative clauses). There is no plural specific article: the plural counterpart of *le* is absence of an article.

103. a)

'O	*le*	*ulugāli'i,*	*fānau*	*l=a*	*lā*	*tama*	*'o*	*le*
PRES	ART	couple	give birth	ART=Poss3.du.	child	PRES	ART

teine	*'o*	*Sina*
girl	PRES	Sina

"There was a̲ couple who had a̲ child, a̲ girl called Sina."

(Mosel and Hovdhaugen:259, ex. 6.37)

b)

Māsani	*'o*	*le*	*tamāloa*	*e*	*usua'i=ina*	*lava*	*ia.....*
used	PRES	ART	man	GENR	get up early=ES	EMPH	3sg

'ae nonofo	*'o*	*le*	*fafine*	*ma l=a=na*	*tama*	*i*
but stay(pl.)	PRES	ART	woman	and ART=POSS=3.sg	child	LD

le	*fale*
ART	house

"It was the̲ man's practice to get up early and... while the̲ woman stayed at home with her child."   (Mosel and Hovdhaugen:259, ex. 6:38)

Consider next the use of *se*: "[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). This use of *se* is illustrated in (104a), where *a coconut* is a narrow-scope indefinite. Crucially, *se* is also used in wide-scope indefinite environments, as illustrated in (104b). This sentence is about "a certain lady whose identity has not been recognized by the speaker or is not of any interest to him" (Mosel and Hovdhaugen 1992:261); this is arguably a non-specific indefinite use of *se*.

104. a)

'Au=mai	*se*	*niu!*
take=DIR	ART(nsp.sg.)	coconut

"Bring me a̲ coconut [no matter which one]!"

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

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

The specific and non-specific articles can be used in very similar contexts. As Mosel and Hovdhaugen (1992:262) note, "comparing the beginnings of stories shows that common noun phrases introducing the discourse topic are marked by the nonspecific article if its exact identity is not known or is unimportant". This is shown by the contrast in (105): the discourse topic is important in (105b), but is unimportant in (105a). This is consistent with the previous section's conclusion that the distinction between specific and non-specific articles is not a rigid one: in wide-scope environments, the speaker can use either, depending on whether she wants to draw attention to a particular individual. In the English equivalents of the sentences in (105), *this*$_{ref}$ would similarly be possible but not obligatory.

105. a) *Sa   i ai   <u>se</u>   matua=moa   ma   <u>s</u>=a=na*
PAST   exist   ART(nsp.sg.) old=hen   and   ART(nsp.sg.)=Poss=3.sg.
toloai. O   l=o   latou   aiga   o   lalo   o   le
brood. PRES ART=Poss3pl. family PRES under Poss ART
fai.
banana=tree
"Once upon a time there was <u>a</u> hen and her brood. Their home was beneath the banana tree."

b) *Sa   i ai   <u>le</u>   ulugalii   o   Papa   <u>le</u>   tane   a o*
PAST   exist   ART   couple   PRES Papa   ART   husband but PRES
Eleele   <u>le</u>   fafine   i   Manua.
Eleele   ART   woman   LD Manua

"There was <u>a</u> couple, Papa, the husband, and Eleele, the wife, who lived in Manua."

(Mosel and Hovdhaugen 1992:262, ex. 6.51, 6.52)

Most of the discussion on use of *se* concerns indefinites (*se* and its plural counterpart, *ni,* are also used in indefinite generic contexts, in negative generic sentences, and in the predicative position of negative predicative constructions). However, note that in (105a), the non-specific article *s* is used as part of the possessive determiner *her*; in English, the DP *her brood* would be considered definite. Mosel and Hovdhaugen have some other examples of use of *se* with definite DPs, given in (106). The possessives *your family* and *your father* are obligatorily definite in English. To express the meaning that *se* contributes to the possessive phrases in (106), Mosel and Hovdhaugen insert phrases like *whoever that is* in the translation. This recalls the non-specific definites in English that I have been discussing (cf. *I need to interview the winner of this race – whoever that is!*).

106. a) *Alui   se   tou   aiga   e   moe. Pe se*
     go LD ART(nsp.sg.) 2.pl. family GENR sleep. Q ART(nsp.sg.)
     tama   a   ai!
     boy   POSS   who
     "Go to your family – whoever that may be – and sleep! [I wonder] whose boy you might be!" [said to a boy who is selling necklaces at night in front of a hotel]

    b) *Tapagai   lava   ulavale   l̲=o=u   pua'a po='o*
     [term of abuse] EMPH troublesome ART=Poss=2.sg. pig   Q=PRES
     ai   s̲=o=u   tamā.[22]
     who   ART(nsp.sg.) father
     "Oh you filthy little bastard, you pig, whoever is your father."
                         (Mosel and Hovdhaugen 1992:262, ex. 6.53, 6.54)

Based only on the few examples in Mosel and Hovdhaugen, it is impossible to develop a theory of article semantics in Samoan. However, a preliminary generalization can be made: Samoan is an example of a language which uses one article in [+specific] environments and a different article in [-specific] environments[23].

---

[22] It is not clear whether *your father* in this case is predicative, and whether this makes any difference for article use.

[23] More work has been done on the article semantics of Maori, another Polynesian language. However, the semantics of Maori articles appear to be quite different from those of Samoan. As Chung and Ladusaw 2003 show, Maori has a definite article, *te,* as well as two indefinite articles, *he* and *te⁻tahi.* The distribution patterns of the two indefinite articles are quite similar: both can introduce new referents in episodic sentences, and both can take narrow scope with respect to scope operators. Chung and Ladusaw point to two differences between the two indefinite articles: first, only *te⁻tahi,* but not *he,* can take wide scope over an operator; second, only *he* but not *te⁻tahi,* can appear in existential sentences. Chung and Ladusaw provide an analysis of these facts by proposing that indefinites can be composed in two ways: via Restrict (which involves Existential Closure (EC) at no higher than event level) or via Specify (which involves a Reinhart/Winter-style choice function, with EC at any level). Both types of composition allow indefinites to take wide scope in episodic sentences, as well as to take narrow scope under an operator. Chung and Ladusaw propose that *he*-indefinites are composed via Restrict, and therefore cannot take scope over an operator (since this would involve EC above event-level). On the other hand, *te⁻tahi* indefinites are composed via Specify and therefore can take either wide or narrow scope with respect to an operator. Chung and Ladusaw also develop an account on which only Restrict can target the internal argument of the existential verb, making *te⁻tahi* indefinites incompatible with existential constructions. On Chung and Ladusaw's analysis, the article system of Maori bears no relationship to specificity. Like English, Maori has a definiteness distinction. However, it also has two indefinite (and non-specific) articles, which signal different modes of semantic composition.

Given the data from Mosel and Hovdhaugen on Samoan, it looks like the article systems in Samoan and Maori are quite different, even though both languages are part of the Polynesian language family. However, we should not necessarily expect the article systems of these two languages to be the same or similar, since Samoan and Maori fall into different subtypes of the Polynesian language family. Samoan is a Samoic-Outlier Nuclear Polynesian language, and Maori is a Tahitic Central East Nuclear Polynesian language (see http://www.ethnologue.com/show_family.asp?subid=119). (An additional difference between the article systems of the two languages is that (discounting the singular/plural distinction) Samoan has two articles while Maori has three).

6.2.3. Summary

Standard English (with no *this*_ref) and Samoan are examples of the two possible systems of article classification that are available to two-article languages under my proposal. A two-article language can either distinguish its articles by definiteness, or distinguish them by specificity. A language which marks one article as [+definite] and another as [+specific] will need a third article for contexts which are neither definite nor specific – and will thus be a three-article language.

We might imagine a two-article language which marks one article as [+definite, +specific] and leaves the second article underspecified. However, as Irene Heim (p.c.) points out, a lexical item which is both [+definite] and [+specific] would be uninterpretable, unless a compositional mechanism is proposed which would conjoin the lexical entries for [+definite] and [+specific] features. Unless such a mechanism is proposed, articles marked [+definite, +specific] are not predicted to exist. In fact, I know of no language which has such an article.

I now move on to a discussion of three-article languages.

6.3. Three- and four-article languages and specificity marking

Given the availability of both [+definite] and [+specific] features for articles cross-linguistically, a three-article language should have the following distribution of articles: one article marked [+definite], one article marked [+specific], and a third article which is underspecified. I have already discussed that spoken English is an example of such a three-article language, with *the, this*_ref, and *a*. I now discuss another example of a three-article language – Modern Hebrew. Hebrew has [+definite] and [+specific] articles as well as a null determiner which is not specified for either definiteness or specificity.

6.3.1. Definiteness and specificity in Hebrew

Hebrew has a definite article, *ha,* and no indefinite article. It does, however, have a specificity marker, *xad,* 'one', as was pointed out by Givón (1981). As Borer (2003) explains, "[i]n contrast with the usual use of *'exád',* 'one', on which it takes primary stress and occurs as a modifier of a noun which itself receives a secondary stress (cf. [(107a)]), when 'one' occurs as a specificity marker it is unstressed and phonologically reduced, essentially a clitic on the head N, which in this case bears the primary stress, as illustrated by [(107b)]."

107. a. *baxurà* 'axát
young-woman one
'one young woman'

   b. *baxurá.xƏt*
'a certain young woman' (Borer 2003, ex. 27)

Givón (2001) shows that indefinites marked with *xad* obligatorily scope over intensional operators (108a), while bare indefinites take narrow scope (108b); only bare indefinites, and not *xad*-marked indefinites, are possible in the scope of negation (109). Borer (2003) shows that *xad*-indefinites must also take scope over a higher quantifier, while bare indefinites must take narrow scope with respect to a higher quantifier.

108. a) *hi mexapeset* <u>*ish-**xad***</u> *she-hayta* *tsrixa*
     she search/IMPERF man-**one** REL-was/3FEM.SG must/FEM>SG.
     li-fgosh *kan*
     to-meet here
     "She's looking for *a man* she was supposed to meet here".

   b) *hi mexapeset* *l-a* <u>*ish,*</u> *ve-im hi timtsa*
     she search/IMPERF DAT-3SG.MASC man and-if she find/FUT
     mishehu...
     anybody
     "She's looking for *a man,* and if she finds one..."

                     (Givón 2001:447, ex. 29a-b)

109. a) *hilo* *lo* *raata* *sham* <u>*ish-**xad***</u>
     she NEG saw/3FEM.SG there man-one

   b) hi *lo* *raata* *sham* *af* <u>*ish*</u>
     she NEG saw/3FEM.SG there even man
     "She didn't see *any man* there."   (Givón 2001 :447, ex. 29c-d)

In the absence of any operators, both *xad*-indefinites and bare indefinites are possible, as shown in the following example from Givón 1981 (cited in Givón 2001):

110. a) *...az* *nixnasti* *le-xanut* *sfarim ve-kaniti* <u>*sefer-**xad**,*</u> *ve-ratsti*
     then entered-I to-store-of-books and-bought-I book-one and-ran-I
     habayta *ve-karati* **oto,** *ve-ze* *beemet haya* *sefer* *metsuyan...*
     home and-read-I it and-it truly was-it book excellent
     "...so I went into a bookstore and bought *a book,* and I ran home and read *it,* and *it* was truly a terrific book..."

   b) *...az* *nixnasti* *le-xanut* *sfarim ve-kaniti* <u>*sefer,*</u> *ve-ratsti*
     then entered-I to-store-of-books and-bought-I book and-ran-I
     habayta *ve-axalti* *aruxat* *erev* *ve-halaxti li-shon...*
     home and-ate-I meal-of evening and-went-I to-sleep

     "...so I went into a bookstore and bought *a book,* and I ran home and ate supper and went to sleep..."   (Givón 2001:456, ex. 51)

Givón (2001:456) explains the difference between (110a) and (110b) as follows: in (110a), "where one runs home and proceeds to read the book and discusses it, the specific referential identity of the book *matters,* it remains *topical* in the subsequent discourse. In [(110b)], one does some 'book-buying', then goes about one's routine. The book is never mentioned again, its specific referential identity doesn't matter". Thus, as is the case with the English *this*ref and the Samoan *le,* the Hebrew *xad* is related to discourse prominence, and therefore to noteworthiness. In (110a), the book has the noteworthy property of being terrific: the speaker intends to refer to a particular book, giving it the quality of *noteworthiness.* In contrast, in (110b), the identity of the book that the speaker bought does not matter, and there is nothing noteworthy to set this book apart from other books.

Thus, there is good reason to believe that *xad*-indefinites are specific on the definition that I gave[24].

Definites in Hebrew are incompatible with the *xad* specificity marker. The explanation, I would argue, is the same as for why *this*$_{ref}$ is incompatible with definites in English: Maximize Presupposition.

6.3.2. Four-article languages and definiteness/specificity marking

Suppose that a hypothetical language, English4, has four articles. What features might be expect those articles to have? If only positive feature specifications are possible, then the four articles would have the following features: art1 is [+definite, +specific]; art2 is [+definite]; art3 is [+specific]; and art4 is underspecified. In definite environments, art1 and art2 would compete for insertion, with art1 typically winning out when the context is specific – but not necessarily, since there is no "Maximize felicity condition". In indefinite environments, art3 and art4 would compete for insertion, with art3 typically (but not necessarily) winning out. The predictions for article insertion in English4 are laid out in (111).

111. Article insertion for English4:
[+definite, +specific]:	*art1 (*and maybe *art2)*
[+definite, -specific]:	*art2*
[-definite, +specific]:	*art3 (*and maybe *art4)*
[-definite, -specific]:	*art4*

However, as discussed in Section 6.2.3, an article which is both [+definite] and [+specific] (art1 in this case) would be uninterpretable, in the absence of a special semantic mechanism which can conjoin these two features. In the absence of such a mechanism, a language with four articles for marking definiteness/specificity distinctions, such as English4, is not in fact predicted to exist. I know of no language which has the system in (111).

6.3.3. Summary

In this section, I discussed the logical possibilities for three-article and four-article languages. While there are no known examples of four-article languages, both (spoken) English and Hebrew are examples of three-article languages. These languages mark both definiteness and specificity. In this they differ from two-article languages: Samoan, which marks specificity only, and those western European languages which mark definiteness only.

An interesting case is presented by Norwegian: while Norwegian has [+definite] and [-definite] articles, just like English, it also allows article omission with singular DPs. According to Borthen (2003), article omission with indefinites is licensed only in the

---

[24] One might wonder why non-specific (bare) indefinites in Hebrew *must* take narrowest scope in intensional contexts, given the fact that in English, non-specific indefinites are perfectly compatible with wide scope. Borer (2003) argues that bare indefinites in Hebrew must be existentially closed in the c-command domain of the VP, while English indefinites have no such requirement; Borer further supports this view by showing that bare indefinites cannot appear preverbally in Hebrew. In order for a Hebrew indefinite to be interpreted outside of the VP, it must bear lexical marking such as *xad* or *eyze* 'some'.

Note that non-specific indefinites in other languages (e.g., bare indefinites in Kannada – see Section 6.4.2) are not restricted to the narrowest-scope reading.

absence of specificity (referentiality). Borthen (1998) shows that article omission is also licensed with superlatives, again in the absence of referentiality. In both cases, Borthen argues, non-referentiality is a necessary but not a sufficient condition for article drop: some lexical verbs do not allow article omission on their argument DPs even when the DP is clearly non-referential. This suggests that the null article in Norwegian does not simply correspond to the [-specific] feature: rather, there is a particular condition on article omission, and this condition is incompatible with specificity. One such condition is suggested by Borthen (2003), who proposes that bare indefinites have type rather than token readings. It is not clear what accounts for article omission with superlatives.

6.4. Specificity, presuppositionality, and case marking

Can specificity be encoded outside of the article system? In this section, I discuss two languages in which specificity may be related to object case marking. Importantly, however, the accusative-case marker in these languages is *not* marked as [+specific]. Rather, object case marking may denote a broader semantic distinction, one which incorporates specificity. I discuss the evidence below.

6.4.1. Accusative case marking in Turkish: different types of specificity

Accusative case marking in object position in Turkish is obligatory for definites, for quantificational DPs with determiners like *every, most,* and *all,* and with partitives, as shown in (112) (see Enç 1991, Kelepir 2001, among others). Indefinites in object position, on the other hand, may appear both with and without accusative-case marking.

112. a) *Hasan baskan-ı   aradi./*...başkan-Ø*
     Hasan president-Acc called
     "Hasan called the president"

    b) *Hasan herkes-i   aradı./*...herkes-Ø aradı.*
     Hasan everybody-Acc called
     "Hasan called everybody."

    c) *Hasan aday-lar-dan   bir-i-ni aradi./*Hasan aday-lar-dan bir-i-Ø aradi.*
     Hasan candidate-pl-abl one-3poss-Acc   called
     "Hasan called one of the candidates."   (Kelepir 2001:81-82)

Enç 1991 argues that accusative case marking on Turkish indefinites marks specificity. However, Enç's view of specificity is quite different from the view of specificity that I have been discussing here. Enç defines specificity as *partitivity*: she argues that specific indefinites in Turkish (i.e., indefinites marked with accusative case) denote members of a previously mentioned set. This is illustrated in Enç's examples below. While either sentence in (114) can be uttered following the sentence in (113), the meanings conveyed are subtly different. Enç explains that (114a) "is about two girls who are included in the set of children, established by the utterance of [(113)], that entered the room" while (114b) "is about two girls who are excluded from the original set of children."

113. *Odam-a   birkaç çocuk girdi.*
    my-room-DAT   several child   entered
    "Several children entered my room."   (Enç 1991, ex. 16)

114. a) *Iki   kiz-ı   taniyordum.*
        two   girl-ACC   I-knew
        "I knew two girls."

    b) *Iki   kiz taniyordum.*
        two   girl I-knew
        "I knew two girls."   (Enç 1991, ex. 17-18)

A similar proposal was put forth by Diesing (1992), who argued that specific indefinites cross-linguistically are always presuppositional – i.e., they presuppose the existence of the set denoted by the NP. (Unlike Enç, however, Diesing does not consider specific (presuppositional) indefinites to be obligatorily partitive).

However, there is reason to believe that accusative marked Turkish indefinites do not have to be presuppositional. Kelepir (2001) shows that accusative case is used both with partitive indefinites, and with wide-scope indefinites that are not partitive. As shown in (115), accusative-case marking forces the indefinite to scope over an intensional verb, while a bare indefinite must scope under it. Kelepir makes the point that in order for (115b) to be felicitous, it is not necessary to have a salient group of interpreters in the discourse.

115. *Ahmet bugünlerde ne yapryor?*
    What is Ahmed doing these days?

    a) *Bir çevirmen-∅   ariyor.*
        an interpreter · looking.for
        "(He) is looking for an interpreter (de dicto)"

    b) *Bir çevirmen-i   ariyor.*
        an interpreter-acc   looking.for
        "(He) is looking for an interpreter (de re)"   (Kelepir 2001:91, ex. 121)

Kelepir investigates various other properties of accusative-marked indefinites (including their ability to scope under a higher quantifier and under negation) and proposes a choice-function analysis of accusative-marked indefinites in Turkish. She shows that accusative-marked indefinites are not specific either in the sense discussed here (they do not have to be referential), or in the sense discussed in Enç (1991) (they do not have to be partitive). The conditions on accusative-case marking are broader, not limited to specificity or to presuppositionality.

It is notable, however, that both specific indefinites (115b) and definites (112a) fall into the category of DPs which receive accusative case marking. This suggests that there is a broader semantic distinction, which groups together both presuppositional DPs (definites, partitive indefinites, strong quantificational phrases, etc.) and specific DPs. Specificity in and of itself may not receive morphological expression in the case system, as it does in the article system.

6.4.2. Accusative-case marking and specificity in Kannada

Lidz (1999) discusses accusative case marking in Kannada and argues that the readings of accusative-marked objects are derived through a Kratzer-style choice function mechanism[25]. He shows that bare indefinites (116a) may take either wide or narrow scope with respect to an intensional operator, while accusative indefinites (116b) must take wide scope. Thus, (116b) is true only if Hari is looking for a specific book, not if Hari is looking for some book or other. The same pattern holds with negation – accusative-marked indefinites must scope over negation, whereas bare indefinites may scope either over or under it.

116. a) *Hari   pustaka   huduk-utt-idd-aane*
    Hari   book   look.for-PPL-PROG-3SM
    'Hari is looking for a book.'

   b) *Hari   pustaka-vannu   huduk-utt-idd-aane*
    Hari   book-ACC   look.for-PPL-PROG-3SM
    'Hari is looking for a book.'   (Lidz 1999, ex. 1)

Lidz explains these facts by proposing that while bare indefinites are quantificational (and hence have both wide and narrow scope readings available to them), accusative-marked indefinites are choice-function variables which are existentially closed at the top, in a Kratzer/Matthewson-style analysis.

Next, Lidz shows what happens the sentence contains a universal quantifier, an indefinite, and an intensional verb, as in (117). In principle, the sentences in (117) might have any of the three readings in (118) available to them. On the reading in (118a), the indefinite takes narrowest scope: the sentence then says that every student is looking for some book or other – for something to read. In (118b), the indefinite takes intermediate scope below the universal quantifier but above the intensional verb: the sentence then says that for each student, there is a specific book which that student is looking for. Finally, in (118c), the indefinite takes widest scope: there is a particular book that all the students are looking for.

117. a) *pratiyobbavidyaarthi pustaka   huduk-utt-idd-aane*
    every   student   book   look.for-PPL-PROG-3SM
    'Every student is looking for a book.'

   b) *pratiyobbavidyarthi   pustakav-annu   huduk-utt-idd-aane*
    every   student   book-ACC   look.for-PPL-PROG-3SM
    'Every student is looking for a book.'   (Lidz 1999, ex. 7)

118. a) $\forall$x [student(x) $\rightarrow$ look_for(x, $^\cap\exists$y[book(y)])]
   b) $\forall$x [student(x) $\rightarrow$ $\exists$y[book(y) $^\wedge$ look_for(x, y)]]
   c) $\exists$y[book(y) $^\wedge$ $\forall$x [student(x) $\rightarrow$ look_for(x, y)]]   (Lidz 1999, ex. 8)

---

[25] Lidz shows that this analysis is applicable only to those DPs for which case marking is optional. Some classes of DPs (e.g., animates and plurals) obligatorily take accusative case marking and do not receive choice function interpretations; Lidz shows that the choice function interpretation is available for these DPs if additional emphatic morphology is added.

According to Lidz, the sentence with a bare indefinite, (117a), has the readings in (118b, c) available to it: the indefinite can take narrow or intermediate scope, but not widest scope. Lidz explains this lack of wide scope by lack of a position above the subject to which the (quantificational) indefinite may raise.

The sentence with an accusative-marked indefinite (117b) has the readings in (118a-b): the indefinite may take wide or intermediate scope. The lack of narrowest scope is explained under the Kratzer/Matthewson choice function analysis of accusative-marked indefinites. The widest-scope reading is clearly available through the choice function. As for the intermediate reading, Lidz argues that it is a result of Skolemization: the function chooses for each student the book that is related to that student in a particular way.

Unfortunately, not much is known about case marking on definites in Kannada. According to Sridhar (1990), accusative case marking with definites is optional. However, Sridhar (1990:161) also states that "the presence of accusative case marking regularly denotes definite reference" – a statement with which Lidz disagrees, showing that choice-function indefinites bear accusative case marking, as discussed above.

For now, I can conclude that, in the case of indefinites, accusative-case marking in Kannada is compatible with the view of specificity that I have been arguing for (with the additional possibility of Skolemized specific indefinites). More data are needed concerning accusative case marking in definites. It may be that, as in the case of Turkish, accusative-case making in Kannada corresponds to a broader semantic distinction that subsumes specificity.

7. The Article Choice Parameter

In the previous section, I argued that articles cross-linguistically can encode definiteness or specificity. I laid out the predictions for article specifications in languages with two, three, or four articles. I will now capture these predictions by specifying a parameter for article choice.

7.1. Lexical specifications of articles as parameter values

As discussed in the previous section, a two-article language has two options of article classification: it may divide articles on the basis of definiteness or on the basis of specificity. We can thus formulate the Article Choice Parameter as it applies to two-article languages. This parameter is given in (119). The word "parameter" as used here refers to lexical specification: there is parametric variation in how articles may be specified cross-linguistically. The Article Choice Parameter in (119) captures this parametric variation for two-article languages.

119. The Article Choice Parameter (for two-article languages):
     A language which has two articles distinguishes them as follows:
     Setting I. Articles are distinguished on the basis of specificity.
     Setting II. Articles are distinguished on the basis of definiteness.

What possible languages do the two settings in (119) predict? Languages which divide their articles on the basis of definiteness, adopting Setting II, will use one article with all definites and another article with all indefinites, as shown in (120a). This is the case for English (in the absence of *this*ref). Languages which divide articles based on specificity, adopting Setting I, will use one article with all specific DPs, definite and

indefinite, and a different article with all non-specific DPs, definite and indefinite, as shown in (120b). Samoan is an example of a Setting I language – as we saw in Section 6.2.2, both specific and non-specific articles in Samoan can be used with both definites and indefinites, depending on the context.

120. a) Article grouping by definiteness

	+definite	-definite
+specific		
-specific		

b) Article grouping by specificity

	+definite	-definite
+specific		
-specific		

One can just as easily formulate versions of the Article Choice Parameter for one-article and three-article languages. Spoken (colloquial) English and Hebrew both instantiate the three-article option in (121b).

121. a) The Article Choice Parameter (for one-article languages):
A language which has one article cannot encode it for either definiteness or specificity.

b) The Article Choice Parameter (for three-article languages):
A language which has three articles must distinguish them as follows:
One article is [+definite], one article is [+specific], and one article is underspecified.

A full version of the Article Choice Parameter would bring together all of the three options above, specifying what options are available to a language depending on the number of articles that it has. Since my focus in this thesis is on the acquisition of a two-article language (standard English), I will refer throughout to the version of the Article Choice Parameter for two-article languages (120) rather to the rather unwieldy version which includes all three possible configurations.

7.2. Article choice, presuppositionality, and genericity

The Article Choice Parameter focuses on definiteness and specificity. But might there be any other semantic distinctions which can be encoded in a language's article system?

Another semantic distinction related to DPs is *presuppositionality*, a term which refers to the presupposition that the set denoted by the restrictor NP exists. There is evidence that presuppositionality interacts with syntactic position (see Diesing 1992); it may also bear a relation to case-making (see Enç 1991). Strong quantifiers such as *every* and *most* are also presuppositional (see Diesing 1992). However, I am not aware of any language whose articles encode presuppositionality – i.e., a language in which one article would be used whenever the existence presupposition is satisfied. I thus conclude that presuppositionality does not need to be treated as an additional setting of the Article Choice Parameter. (But see Appendix 4 for some discussion on the possible role of presuppositionality in acquisition).

While articles cross-linguistically do not encode presuppositionality, they do bear a relation to *genericity*. In English, generic DPs take *the* in the singular (*The lion has a busy tail*) or appear as bare plurals (*Lions have bushy tails*). In Romance languages, plural generic DPs take the definite article (See Krifka et al. 1995 for more discussion). Article use with generics appears to be quite distinct from the definiteness / specificity

distinctions and may involve a different parameter (such as the one formulated by Chierchia 1998 concerning presence vs. absence of articles with plural generics cross-linguistically).

7.3. Article choice, scope, and licensing conditions on articles

Another possible distinction that articles may encode is scope. A proposal along with these lines was put forth by Matthewson (1998), who argued that articles in St'át'imcets distinguish between narrow-scope DPs and all others. However, Matthewson (1999) shows that scope is not the (only) relevant distinction: one determiner obligatorily takes narrow-scope, but the other determiners do not simply take wide scope – rather, they are interpreted through choice functions with topmost EC (see Section 3.2.3). I am not aware of any conclusive evidence that any language distinguishes its articles on the basis of scope, or on the basis of *assertion of existence*.

7.3.1. Article distinctions which correspond to neither definiteness nor specificity

However, St'át'imcets is an example of a language which, despite having two articles, distinguishes them neither on the basis of definiteness nor on the basis of specificity. The work of Chung and Ladusaw (2003) shows that Maori is a language which has three articles, and employs the definiteness distinction but not the specificity distinction (see footnote 23). Each of these languages has a pair of articles (or article types) which are distinguished neither on the basis of definiteness nor on the basis of specificity. According to Matthewson (1999), in St'át'imcets, indefinites headed by non-polarity articles like *ti...a* are composed via choice functions with topmost EC, while indefinites headed by *ku* are quantificational, and must, additionally, be licensed by a c-commanding operator. According to Chung and Ladusaw, in Maori, indefinites headed by *te⁻tahi* are composed through choice functions with EC at any level, and indefinites headed by *he* are interpreted through predicate restriction with VP-level EC. Where do these languages fall with respect to the Article Choice Parameter that I proposed?

7.3.2. St'át'imcets and Maori, and the Article Choice Parameter

From the standpoint of the Article Choice Parameter, St'át'imcets is a one-article language: as stated in (121a), it has one article which is neither [+definite] nor [+specific]. Additionally, however, the form of this article varies depending on whether it (a) receives a choice-function interpretation; or (b) is quantificational, and furthermore licensed by a c-commanding operator.

Maori, from this standpoint, is a two-article language which adopts the Definiteness distinction in (119). It has one [+definite] article and one underspecified article. However, the underspecified article takes different forms depending on the type of semantic composition.

Motivation for treating the relevant articles in St'át'imcets and Maori as two forms of the same article, from the standpoint of the Article Choice Parameter, is as follows. The Article Choice Parameter is concerned with discourse-based distinctions only: with whether articles in a given language signal something about the speaker's state of mind (specificity) or about the hearer's state of mind (definiteness). Articles which differ from each other on the basis of definiteness or specificity always convey different meanings, even when used in the same syntactic configuration: for instance, the specificity articles

in English, Hebrew, and Samoan all signal that the speaker attaches importance to the referent.

In contrast, articles in St'át'imcets and Maori appear to have exactly the same meanings when used in the same syntactic configuration. Matthewson (1999) does not discuss any differences in interpretation between polarity and non-polarity St'át'imcets articles in contexts where both are allowed (i.e., in the scope of a higher quantifier). Chung and Ladusaw show that the two indefinite articles in Maori are equally available in episodic sentences, and do not suggest any differences in interpretation between the two. It is possible that further investigation would show that there are indeed differences in how the two Maori indefinite articles are interpreted in episodic sentences; similar differences might be uncovered for St'át'imcets. At the moment, however, no such differences are known, so from the standpoint of the Article Choice Parameter, the two article forms in each of these languages are equivalent: both signal neither definiteness nor specificity (nor any other discourse-related distinction).

Thus, the Article Choice Parameter in principle allows the possibility that articles specified for neither definiteness nor specificity may reflect different modes of semantic composition and/or be subject to syntactic licensing conditions.

7.3.3. Different article distinctions, parameters, and acquisition

What does it mean to say that a language has "one article that takes different forms"? Clearly, *ti...a* vs. *ku* in St'át'imcets as well as *te⁻tahi* vs. *he* in Maori are different lexical items. They morphologically signal a particular mode of semantic composition and/or the satisfaction of a particular licensing condition (such as c-command by a higher operator in the case of *ku*). It is possible that a parameter (or multiple parameters) governs the morphological expression of different modes of semantic composition and/or syntactic licensing conditions on the DP. Given the present evidence, there seem to be many possible combinations for morphological encoding of these properties: for instance, one might imagine a language which has an article with the properties of the St'át'imcets *ti...a* alongside a different article with the properties of the Maori *he*[26].

How does an (L1 or L2) learner know whether articles in a given language encode a discourse-related distinction such as definiteness or specificity, or a purely grammatical distinction such as those employed by St'át'imcets and Maori articles? Suppose that a learner approaches a language with two distinct lexical items in article position. The learner needs to decide whether she is dealing with a two-article language (i.e., one of the articles is [+definite] or [+specific]), or a one-article language where the one article takes different forms depending on the licensing conditions and/or mode of semantic composition. One might hypothesize that in a language like St'át'imcets, the syntactic distribution of articles (e.g., the fact that *ku* can only occur in narrow-scope environments) signal to the learner that articles in this language encode grammatical rather than discourse distinctions. Given the very similar syntactic distribution, as well as

---

[26] Additional evidence for this comes from Hebrew. Like English, Hebrew has a [+definite] article, a [+specific] article, and an underspecified null article. However, bare indefinites in Hebrew are not identical to *a*-indefinites in English in their behavior: as Borer (2003) shows, bare indefinites in Hebrew cannot scope over an operator or appear in preverbal position; Borer argues that bare indefinites must receive VP-level EC. This is similar to Chung and Ladusaw's proposal for *he*-indefinites in Maori. Unlike Maori, however, Hebrew has only a single form of the underspecified indefinite – the bare form.

apparently identical interpretation, of the two indefinite articles in Maori, it is not clear what the learner's initial hypothesis regarding these articles would be (e.g., would the learner consider the two articles to be identical in distribution? Or would she decide that one of them encodes specificity?). In the absence of acquisition data on St'át'imcets and Maori, there is not much that can be said.

In acquiring a language such as English, however, learners hear two articles used in more or less the same environments, but in different discourse situations. Their job, then, is to learn what discourse distinction articles encode. I suggest that learners do not posit grammatical restrictions on articles unless the language (e.g., Maori or Salish) gives evidence for the existence of such restrictions. This is the most economical hypothesis: there are many different kinds of grammatical distinctions that articles can make, and learners would be faced with many different options. In contrast, if they confine their initial hypothesis to discourse-based distinctions, they have only two options: definiteness and specificity.

The data on both L1-English and L2-English acquisition of articles presented in subsequent chapters strongly suggest that learners do not make non-discourse-based distinctions between *the* and *a*: they do not, for instance, consider *the* a marker of wide scope, confine *a* to narrow-scope environments, or (to the extent that data on this are available) disallow *a* to take wide scope over an operator. Learners appear to know that English articles encode discourse distinctions rather than licensing conditions / modes of composition. Their task then is to decide *which* discourse distinction(s) English articles encode – i.e., to assign the appropriate setting to the Article Choice Parameter. I move on to this next.

8. Conclusion: articles in acquisition

In this chapter, I reviewed a number of distinctions pertaining to DP classifications cross-linguistically, focusing on *scope, definiteness,* and *referentiality/specificity,* with brief mentions of *presuppositionality* and various types of *choice-function readings.* In discussing *specificity,* I formulated a lexical entry for specific DPs which is based on Fodor and Sag's discussion of referentiality, but which incorporates the concept of *noteworthiness.*

I then argued that languages which have a discourse-based system of article classification can specify their articles as [+definite] or [+specific]. I showed that two-article languages can employ *either* the definiteness *or* the specificity distinction, while three- or four-article languages employ both. I proposed the Article Choice Parameter, and argued that it is discourse-related.

In the rest of this thesis, I will look at how L2-English learners acquire the Article Choice Parameter. I will assume that L2-English learners are acquiring a two-article language: standard English, which contains *the* and *a*. The colloquial referential use of *this* is not part of standard of English, and is not part of all dialects and registers of English[27]. It is highly unlikely that L2-learners receive enough exposure to referential *this*

---

[27] Anecdotal support for this view comes from informal responses of L1-English control participants in the studies reported in Chapters 4 through 6, whom I questioned about their acceptance of sentences with referential *this*. Younger (college-age) participants considered *this* a part of their dialects, while some older participants remarked that referential *this* was something they know younger speakers would use, but that they would not use themselves. Many participants of all ages said that referential use of *this* was "slang".

to incorporate it into their article system, and to adopt the "three-article language" setting of the Article Choice Parameter (121b). Instead, they treat English as a two-article language and therefore need to decide which setting of the parameter in (120) to adopt. In the absence of transfer, they have no reason to initially prefer one setting over another. In the rest of this thesis, I will show that L2-English learners in fact undergo fluctuation between the two settings, as predicted by the Fluctuation Hypothesis in Chapter 1.

The L2-English participants in our studies rarely went to high school or college in the U.S. (although a number of the L1-Korean participants attended graduate school here), and thus were unlikely to have much interactions with teenage and young adult populations for whom use of referential *this* is most acceptable. Given the typical ages (late 20's to early 50's) and social backgrounds (graduate students and professionals) of the L2-learners in our study, it is likely that they interacted mostly with adult native English speakers (such as teachers and co-workers) for whom referential *this* is not very socially acceptable. Finally, in classrooms, L2-learners are taught that English has two articles, *the,* and *a,* and may not even realize that *this* has an "article" use, treating all instances of *this* that they hear as demonstratives.

My proposal predicts that L2-learners who acquire a three-article language all of whose articles have equal social status should have relatively little difficulty acquiring the specifications of the three articles as [+definite], [+specific], and underspecified – the only option available to three-article languages (barring grammatical distinctions such as those in Salish and Maori). (Although learners of such a language may take some time to realize which article is [+definite] and which is [+specific], given the ambiguity of [+definite, +specific] contexts). An interesting future study (suggested by David Pesetsky, p.c.) would be to investigate whether rigorous exposure of L2-English learners to the referential use of *this,* and emphasis on *this* as an article marking specificity, would decrease errors of *the* overuse with specific indefinites among the learners; this would be predicted, if intensive exposure to referential *this* can cause L2-learners to reanalyze English as a three-article language.

Chapter 3: Articles in L1 and L2 acquisition

1. Introduction

In the previous chapter, I argued for the existence of an Article Choice Parameter, repeated in (1) (for two-article languages). I will now see how this parameter relates to L2-acquisition.

1. The Article Choice Parameter:
   A language which has two articles distinguishes them as follows:
   Setting I.   Articles are distinguished on the basis of specificity.
   Setting II. Articles are distinguished on the basis of definiteness.

It is well-known that L2-English learners misuse English articles, which suggests that they do not (at least initially) distinguish English articles on the basis of definiteness. A reasonable hypothesis, therefore, is that L2-learners erroneously think that English distinguishes articles on the basis of specificity, and that this is what causes article misuse in L2-English. I will examine this hypothesis in this chapter. This chapter brings together the Fluctuation Hypothesis of Chapter 1 and the discussion of article semantics in Chapter 2, and makes explicit predictions for the acquisition of articles. It lays the groundwork for the next four chapters, which report empirical studies of L2-English article choice.

This chapter is organized as follows. In Section 2, I propose the Fluctuation Hypothesis for L2-English article choice, and lay out its predictions. In Sections 3, 4, and 5, I review some previous studies of article choice in L1 and L2-acquisition, and look at whether their results are predicted under my proposal. Section 6 considers the potential role of transfer in the L2-English article choice of Russian and Korean speakers – speakers whose data are reported in this thesis. This section shows that no transfer effects from Russian or Korean could account for the role of specificity in L2-English article use. Section 6.2 concludes the chapter.

2. The Article Choice Parameter and L2-acquisition

This thesis investigates the role of the Article Choice Parameter in L2-acquisition. If the Article Choice Parameter is real, and languages can in fact use articles to encode either definiteness or specificity, what does this mean for L2-acquisition?

2.1. The FH for article choice in L2-acquisition

There is certainly a possibility that speakers of a language which has articles will transfer the specifications of these articles onto articles in their L2. For instance, French, like English, has a definiteness distinction; we may therefore expect French speakers to transfer the specifications of French articles onto English and correctly encode *the* and *a* as making [+definite] and [-definite] contexts, respectively. On the other hand, speakers of Samoan, which has a specificity distinction, may treat English articles as also obligatorily marking specificity.

In this thesis, I will not look at the acquisition of English by L2-learners who have articles in their L1, and will leave for further research the role of L1-transfer in article

choice. I will look instead what happens when speakers of L1's with no articles acquire an L2 such as English, which does have articles.

The acquisition of English articles in the absence of transfer provides a testing ground for the Fluctuation Hypothesis (FH) proposed in Chapter 1. The FH predicts that L2-learners should access both settings of the Article Choice Parameter. In the absence of transfer effects, there is no reason for one setting to preferred over another. The proposal for the acquisition of articles under the FH is formulated in (2).

2. **The Fluctuation Hypothesis (FH) for L2-English article choice:**
   1) L2-learners have full UG access to the two settings of the Article Choice Parameter in (1).
   2) L2-learners fluctuate between the two settings of the Article Choice Parameter until the input leads them to set this parameter to the appropriate value.

The hypothesis in (2) rests on the assumption that L2-English learners are acquiring a two-article language: standard English, which has *the* and *a* (see the discussion in Chapter 2). Their task is to decide whether these two articles are distinguished on the basis of definiteness, or on the basis of specificity.

Given the Article Choice Parameter in (1), and assuming that L2-learners have full access to both settings of this parameter, we may in principle expect individual L2-learners to follow any one of the three patterns in (3). I will now look at which of these patterns are actually predicted under the FH.

3. a) <u>The Definiteness Pattern</u> (Setting II of the Article Choice Parameter)
   L2-English learners correctly use *the* and *a* to mark [+definite] and [-definite] contexts, respectively.

   b) <u>The Fluctuation Pattern</u> (Fluctuation between settings)
   L2-English learners go back and forth between distinguishing *the* and *a* on the basis of definiteness, and distinguishing them on the basis of specificity.

   c) <u>The Specificity Pattern</u> (Setting I of the Article Choice Parameter)
   L2-English learners use *the* and *a* to mark [+specific] and [-specific] contexts, respectively.

The FH for L2-English article choice does not exclude the possibility that learners will eventually set the parameter to its appropriate value, given sufficient input – i.e., that individual learners will follow the pattern in (3a). Importantly, the FH predicts that L2-learners who have not yet succeeded in setting the Article Choice Parameter will follow the pattern in (3b). No learners should follow the pattern in (3c), since this would require them to mis-set the Article Choice Parameter to a value not present either in their L1 or in the input. In the absence of transfer, the FH predicts fluctuation between the two parameter settings. The fluctuation will cease only if learners can set the parameter to the target value for their L2 – which, in the case of English, is Setting II. As I report data from empirical studies in Chapters 5 and 6, I will examine whether individual L2-learners exhibit the patterns in (3a) and (3b), as predicted, rather than the pattern in (3c) or some random pattern.

2.2. Predictions for article use in L2-English

The hypothesis in (2) makes specific predictions for L2-English article choice. It predicts that while L2-English learners may make errors, their errors should be non-random. Errors should occur whenever L2-learners divide English articles on the basis of specificity rather than on the basis of definiteness. This means that L2-English errors of article misuse should be confined to overuse of *the* with [+specific] indefinites, as well as overuse of *a* with [-specific] definites. This is stated in (4).

4. Predictions for article use in L2-English:
    correct use of *the* in [+definite, +specific] contexts
    overuse of *a* in [+definite, -specific] contexts
    overuse of *the* in [-definite, +specific] contexts
    correct use of *a* in [-definite, -specific] contexts

The predictions in (4) concern only article use and misuse, not article omission. I make no predictions for article omission, since omission may be due to a large variety of sources, including lack of articles in the grammar (for beginner L2-learners) and retrieval difficulties. I will be concerned primarily with the patterns of article choice when L2-learners *do* use articles. However, in Chapter 7, which reports written production data, I will consider whether there are any patterns to article omission in L2-English.

2.3. The Article Choice Parameter and triggers

In Chapter 1, I suggested that L2-learners' ability to set different parameters is related to the availability of the triggers for those parameters. I suggest that the triggers related to the Article Choice Parameter are particularly difficult from the standpoint of L2-acquisition because they do not arise (at least not obviously) from the syntactic configuration. Both specific and definite articles may in principle appear in the same environments – e.g., in simple SVO sentences with no intensional or modal operators or quantifiers.

In order to determine whether *the* is [+definite] or [+specific], the L2-learner needs to evaluate the discourse situation and decide whether *the* is marking the presupposition of uniqueness (from the hearer's perspective) or the existence of a noteworthy property (from the speaker's perspective). Since definites are often specific, both hypotheses will be compatible with many situations. The learner thus also needs to pay attention to use of *a,* and note that in contexts which are [+specific] but [-definite], *the* is never used.

The discourse triggers related to the Article Choice Parameter are often ambiguous. For instance, suppose that an L2-learner hears someone use a phrase like *I talked to the doctor from next door this morning.* This phrase is compatible with the hypothesis that *the* is [+definite]: even if the learner has never heard of the speaker's next-door neighbors before, it is fairly easy to accommodate the knowledge that there is a unique, salient doctor next door. The phrase is also compatible with the hypothesis that *the* is [+specific]: the speaker may be wishing to attract attention to the identity of the doctor in question.

Suppose next that an L2-learner hears the phrase *I talked to a doctor from next door this morning.* This phrase is compatible with the hypothesis that *a* is used with indefinites: the speaker is not presupposing her listener to have knowledge of the doctor. However, this phrase is also compatible with the hypothesis that *a* is used in the absence

of specificity: the speaker is simply choosing not to attach any importance to a particular doctor's identity. The same ambiguity persists across contexts.

One might wonder then how any learner, L1 or L2, can *ever* set the Article Choice Parameter. The answer, I think, is generalization across individual instances. A single DP might be ambiguous between definite and specific, or indefinite and non-specific, readings. However, if the learner is consistently hearing *the* used only when the presupposition of uniqueness has been met (even when the speaker attaches no importance to the referent's identity), and consistently hearing *a* used only when the presupposition of uniqueness hasn't been met (even when the speaker attaches importance to the identity of the referent), the learner should generalize that *the* marks definiteness rather than specificity. Given the subtlety of the discourse triggers related to speaker and hearer knowledge, generalizing from them is likely to be a longer and more difficult process than generalizing from multiple instances of verbs following adverbs in English, or of objects preceding verbs in German, or various other syntactic triggers.

A definitive answer about the differential nature of triggers and their subsequent availability in L2-acquisition would require a discourse processing account that is beyond the scope of this thesis. However, I would like to suggest that L2-learners have less ability than L1-learners to generalize from the input triggers (see also Chapter 1), and that parameter-setting is particularly difficult for L2-learners in domains where the triggers are discourse-related and hence especially subtle and ambiguous. We will see in subsequent chapters that even many advanced L2-learners are unable to set the Article Choice Parameter and exhibit fluctuation between the two settings.

2.4. The FH and L1-acquisition

The predictions of the FH are in principle applicable to L1 as well as L2-acquisition. It is possible that child L1-learners, like adult L2-learners, take some time to set the Article Choice Parameter and undergo fluctuation between the two settings of the parameter. However, the fluctuation might not be as pronounced, or may end fairly quickly, since children are known to be quite good at parameter-setting (see Wexler 1998, among others).

It is also possible that child L1-learners set the Article Choice Parameter very early on in the course of acquisition and use articles appropriately from the start. Finally, it is possible that article errors do occur in child L1-acquisition, but are due to a different source than article errors in adult L2-acquisition: as I will discuss below, article errors in child language have frequently been given psychological rather than linguistic explanations.

In the next two sections, I will examine some studies of article choice in L1-acquisition and see whether young children's article use is predicted under the FH, or whether it stems from a different source. I will also suggest that the FH is more likely to be operative in the domain of article choice than in any syntactic domain of L1-acquisition.

3. Early studies of article choice in L1-acquisition

There have been a number of studies on L1-acquisition of articles, mostly for English (although see Karmiloff-Smith 1979 on articles in child French). The studies have involved different methodologies, including elicited production, comprehension tasks,

and truth-value judgments. While results have varied, to the extent that there is agreement on article use in L1-English, it is the following: young children tend to overuse the definite article *the* in indefinite contexts. The first major work which showed this was Maratsos's (1976) study of articles in the English of three- and four-year-old children. I will now discuss the design of Maratsos's study and the relevant results. Where relevant, I will also mention the results of a similar study with L1-French children by Karmiloff-Smith (1979).

I will then examine two different explanations that have been proposed for Maratsos's results: Maratsos's own psychological explanation and Wexler's (2003) linguistic explanation. Finally, I will look at how Maratsos's findings relate to my proposal.

3.1. Study design

Maratsos used stories to elicit definite and indefinite articles from young children[1]. Five of the stories came in two versions: definite-eliciting and indefinite-eliciting. In the definite-eliciting stories, the experimenter used the form *a X,* and the target response was of the form *the X*: thus, in (5), the experimenter mentions *a frog* and *a turtle,* and the target response is a definite: *the frog* or *the turtle.*

5. Pond Story (Maratsos 1976:52)
    A man with a wooden box goes to a pond to get an animal. He sees a frog and a turtle. So he puts one of them into his box. What did he put in?
    Target response: the frog (or: the turtle)

In the indefinite-eliciting stories, the experimenter used the plural form *Xs,* and the target response was *a X.* Thus, in (6), the experimenter mentions *squirrels* and *turtles,* and the target response is an indefinite: *a squirrel* or *a turtle.* Crucially, in all stories of this type, the experimenter introduced a set into the discourse, and the child had to then refer to a member of this set.

6. Give Story (Maratsos 1976:52)
    Pretend you have lots of turtles and lots of squirrels, say four turtles and four squirrels. Now pretend your mommy wanted one of your animals. Would you give her one? Well, what would you give her?
    Target response: a squirrel (or: a turtle)

The stories in (5) and (6), as well as three other similar stories, appeared in both definite-eliciting and indefinite-eliciting versions. The question word was always "what" or "who". Maratsos included an additional story, given in (7), which came in two versions: one with a "which one" question, designed to elicit a definite response, and another with a "what" question which, according to Maratsos, was more likely to elicit an indefinite response in this context.

7. Car-Boat story (Maratsos 1976:54)
    A man was going to a jungle. He had a car and a boat and used one of them to go to the jungle.
    Which one did he use?   Target response: the car (or: the boat)
    What did he use?   Target response: a car (or: a boat)

---

[1] All of the examples I give here are concise summaries of these stories, since the original descriptions in Maratsos are often fairly lengthy.

There were some individual stories in other types as well. Two stories, which I will not discuss here, tested article use in entailment and generic contexts. Two more stories were designed to elicit indefinites in first-mention environments, where no set had been previously introduced. In one of these stories, given in (8), there was repetitive mention of the indefinites *a lion* and *a zebra,* to make it clear that the man was not looking for any particular lion or zebra. If a lion or zebra comes running out at the man in the end of the story, it is a lion or zebra that has not been previously mentioned, and it is not a member of a previously mentioned set. Thus, an indefinite should be used. In a second indefinite-eliciting story type, some characters were introduced, and the children were then told that a new animal came running out at these characters. The children were then asked "Who came running out at them?" In this story type, the children were free to name any animal, and the DP should be indefinite.

8. Looking For story (Maratsos 1976:53)
    A man went to jungle because he wanted to see a lion or a zebra. He looked all
    over, to see if he could find a lion or a zebra. He looked for a lion or a zebra
    everywhere. He looked and looked. Who came running out at the man?
    Target response: a lion (or: a zebra)

3.2. Results

Maratsos's main findings were as follows. On the definite-eliciting stories, three-year-olds exhibited overuse of *a* while four-year-olds correctly used *the.* On the indefinite-eliciting stories, many four-year-olds overused *the.* I will now discuss the findings in turn, and briefly summarize Maratsos's explanations of these findings. A more detailed discussion follows in the next section.

3.2.1. Results: article use in definite contexts

On the definite-eliciting stories such as (5), Maratsos found that a large proportion of 3-year-olds inappropriately gave an indefinite answer (e.g., *a frog*), while the vast majority of 4-year-olds appropriately gave a definite answer. The accuracy rates for this story were only 55% for the 3-year-olds but more than 90% for the 4-year-olds. Maratsos (p. 64) suggested that 3-year-olds "had difficulty establishing and maintaining representations of unique referents well enough to give definite noun phrases consistently as answers."

A potential problem with the example in (5) is that an indefinite answer is not completely infelicitous: it is not necessarily wrong to answer "a frog" to the "what" question, meaning "the animal that he put in his box was a frog". In fact, in discussing the story in (7), Maratsos (1976:74-75) notes that "what" is more likely to elicit indefinites than "which one": children in both age groups gave significantly more definites as answers to the "which one" question than as answers to "what" question (although the four-year-olds gave more definite answers to this story overall). Thus, at least some of the overuse of *a* in definite-eliciting stories was due to the question type (see Maratsos 1976:76 for more discussion).

In a similar study with young L1-French children, Karmiloff-Smith (1979) found children as young as three correctly using definite articles in stories similar to (5). This further suggests that Maratsos's finding of *a* overuse stems from the way that the English

wh-question was asked. See the next section for similarities between L1-English and L1-French children on overuse of definite articles.

3.2.2. Results: article use in indefinite contexts

On the definite-eliciting stories such as (6), Maratsos found that many of the four-year-olds, but very few of the three-year-olds, gave a definite answer (e.g., *the squirrel*). Maratsos had divided the four-year-olds into the "4 Low" and "4 High" groups based on their performance on an imitation task. Children's placement into the Low or High group on this basis of this task had a strong association with children's responses to indefinite-eliciting stories such as (6). Members of the "4 Low" group correctly used *the* in stories like (5) 94% of the time, but were only 42% accurate in producing *a* for stories like (6). The "4 High" group performed accurately on both story types. On the other hand, all of the children were extremely accurate at producing indefinites in contexts such as (8), where no set is introduced. Overuse of *the* appears to be limited to contexts where the child needs to refer to a member of a previously mentioned set.

Maratsos (1976:63) gives a psychological explanation to overuse of *the* in stories like (6), suggesting the existence of "a developmental stage where egocentric definite responding is quite common. The children fail to take into account that even if they have established for themselves a particular boy or girl, or monkey or pig that does something, that referent is not yet uniquely specified for their listener, and must be introduced to the listener with an indefinite expression." Maratsos's explanation of the results is psychological rather than linguistic in nature: children's responses are governed by an inability to consider the hearer's state of mind. This explanation thus links directly to the idea that young children do not have a fully developed theory of mind. (But see Cziko (1986) for a reinterpretation of Maratsos's results in light of Bickerton's (1981) Bioprogram Hypothesis).

In a similar study with L1-French children, Karmiloff-Smith (1979) found that children from three to seven years of age produced more definites than indefinites in contexts like (6); only eight-year-olds and older children produced more indefinites than definites. If Maratsos's and Karmiloff-Smith's data are taken together, they suggest that overuse of *the* in indefinite contexts is more persistent in L1-acquisition than overuse of *a* in definite contexts.

3.3. Different explanations of article errors in child English

I will now discuss two different accounts of Maratsos's findings of *the* overuse in child English. I will not discuss Maratsos's findings of *a* overuse among three-year-olds, since these findings may be confounded by the fact that "what" questions were used, which are often compatible with indefinite responses.

3.3.1. A psychological explanation: egocentricity

The first explanation is the original account of Maratsos (1976): children attribute their own knowledge to the hearer, and hence overuse *the* when they are aware of a particular referent. On the surface, a psychological rather than linguistic explanation for Maratsos' (as well as Karmiloff-Smith's) findings seems plausible – as Maratsos suggests, young children are egocentric and may ignore other people's state of knowledge, focusing on their own state of knowledge. However, as Heim (2003) points

out, Maratsos's hypothesis also requires children to attach particular importance to one of the individuals in the established context (e.g., one of the four squirrels in (6)), even though the individuals have been in no way differentiated.

Consider again the story in (6). From the adult standpoint, the fact that needs to be described in the answer is "I give mommy one of my four squirrels." As Heim (2003) shows, the adult must choose between the options in (9). The adult will exclude both options (ii) and (iv), since both would induce presupposition failures. In choosing (ii), the adult would be presupposing that there is only one squirrel in the world: this presupposition does not hold. In choosing (iv), the adult would be presupposing that there is only one squirrel in the established context C; however, the assertion in (6) clearly established the existence of four squirrels in C – so again, the presupposition does not hold. The adult will therefore choose option (i) or (iii) and use *a*.

9. choices considered:
   (i) (I give her)   a squirrel
   (ii)   the squirrel
   (iii)   a C squirrel
   (iv)   the C squirrel
   where C = the set of animals that I own   (Heim 2003:5, ex. 24)

Consider next what the child would do. According to Heim, Maratsos's hypothesis, spelled out, says that a child will often choose option (iv), and assign to C a proper subset of the animals that he or she owns. If the child assigns only one squirrel to C, then the uniqueness presupposition holds, and the child uses *the*. Heim (2003) raises the question of whether this is a plausible hypothesis. Why should the child assign only one of the squirrels to C, when the four squirrels were in no way differentiated?

Given Heim's critique, Maratsos's account needs modification, as follows. Children have lively imaginations, so as soon as a child hears a story, she imagines a set of animals (e.g., squirrels), imagines their appearances and/or personalities for them, and decides that a particular animal (e.g., the really furry squirrel) is particularly relevant for the story (i.e., it is the one that the child will give to her mother). The child then assumes that her hearer shares knowledge of this furry squirrel, and uses *the*.

3.3.2. A linguistic explanation: presuppositionality

Wexler (2003) proposes an alternative explanation to Maratsos's findings, which relies on article specifications rather than on children's egocentricity. He argues that children treat *the* as having an existence presupposition, but no uniqueness/maximality presupposition. Wexler's analysis for the child's *the* (*theC*) with singulars is given in (10). Wexler shows that this proposal generalizes to plurals as well.

10. Regardless of the utterance context, [theC x] P expresses that proposition which is:
   -true at an index i, if there is an x at i, and it is P at i
   -false at an index i, if (i) there is an x at i, and (ii) there is no x such that x is P at i
   -truth-valueless at an index i, if there is no x at i   (Wexler 2003, ex. 76)

On this view, *the squirrel* essentially means the partitive *one of the squirrels* from the child's standpoint. The definite article in child English is felicitously used whenever the set denoted by the restrictor NP (in this case, the set of squirrels) is presupposed to exist;

it is not necessary that a unique referent exists. Since the context in (6) has ensured that a set of four squirrels exists, it is felicitous for the child to say *the squirrel*. Wexler's explanation of the overuse of *the* in child English is thus linguistic rather than psychological. However, it is crucially a different linguistic explanation than the one proposed under my Article Choice Parameter. Wexler's explanation relies on *presuppositionality,* rather than on *specificity.*

3.3.3. Discussion

Both Maratsos's psychological explanation and Wexler's linguistic explanation receive support from additional points of Maratsos's study. Wexler's explanation is supported by Maratsos's finding that children did not overuse *the* with first-mention indefinites in contexts such as (8). In these contexts, the indefinite is not presuppositional: no set of zebras or lions has been mentioned. While the man is looking for a zebra or a lion, he is not looking for any particular one, and it may well be that zebras and lions do not even exist (in the relevant domain). Since there is no presupposition of existence, Wexler's analysis correctly predicts that children should not overuse *the* in this context. Maratsos's explanation of children's egocentricity cannot quite account for lack of *the* overuse in this context: the child could imagine a particular lion or zebra in (8), assume hearer knowledge of this lion or zebra, and subsequently use *the*. However, this does not happen.

On the other hand, Maratsos's proposal does receive support from two other sources. The first piece of evidence concerns adult performance. Maratsos tested ten randomly chosen parents of the children in the study on one indefinite-eliciting story of the type in (6); as expected, all adults correctly gave indefinite responses. However, Maratsos (p.103) found adults overusing *the* in a different kind of context. Maratsos tested both children and their parents as follows. An opaque screen was placed between the experimenter and the participant. The participant was given a number of toys, including a car, five ducks, and four rabbits. The participant was asked to put one of the animals in the car and was subsequently asked "Who got into the car?" Crucially, only the speaker (the participant) but not the listener (the investigator) actually knew which rabbit or duck was put in the car; therefore, an indefinite response was expected. Not surprisingly, given the previous findings, young children overused *the* in this context. However, so did seven of the thirteen adults who were tested. There was in fact no statistical difference between the child and adult groups.

This finding can be interpreted as follows. Both children and adults have some tendency to focus in on a particular object, ignoring others. An adult who has been given four toy rabbits places one of them in the toy car and focuses on this rabbit, ignoring the existence of other rabbits. When asked who was put in the car, the adult will sometimes compute uniqueness with respect to the set containing the single rabbit in the car rather than all four rabbits. Children simply go one step further: while adults focus on a unique member of the set only when this individual is directly in front of them, children do so even when the individual exists only in their imagination (as in the case of singling out a particular squirrel from the imaginary set of squirrels in (6)). This would not be very surprising, given children's imaginative capacities.

Finally, Maratsos's proposal is supported by the findings that young children overuse *the* in naturalistic production. Brown (1973:353) reports errors of article misuse on the

part of the three young L1-English children in his study, and states that "[t]he result I think most significant is the large number of errors in the category: speaker specific and listener nonspecific. This is the case in which the points of view of the speaker and listener diverge." Brown's use of the term *specific* is very similar to the use that I have adopted: the "speaker specific and listener nonspecific" contexts are specific indefinite contexts. Examples of *the* overuse with specific indefinites are sentences like *The cat's dead, And the monkey hit the leopard,* and *Where's the stool?*[2] – where the adult listener is clearly unaware of the identity of the cat, monkey, leopard, or stool (Brown 1973:354)[3].

Maratsos (1976:97) similarly found that the children in his studies made "numerous egocentric errors" when interacting with the investigator or with their mothers. For example, one child said "I gave the pounding thing to Ken", when the child's listener had no knowledge of what the pounding thing was, or who Ken was. A different child, when told the Car-Boat story (7), answered the question with "So what did he use?" by saying "the reindeer". The child later clarified that he himself had reindeer at home. Maratsos gives several such examples of *the* overuse when the referent of the indefinite was known to the child, but not to the child's listener. He also cites the work of Peterson (1974), who asked children to describe week-old, real-life events to someone who had not witnessed the event. Peterson found that "even when they were talking to the naïve listener a majority of the articles used by three-year-olds were, incorrectly, definite articles" (Maratsos 1976:97, footnote 1). Peterson found that even four-year-olds overused definite articles, although not to the same extent as the three-year-olds.

These cases of *the* overuse in naturalistic production do not appear to occur in presuppositional contexts (e.g., the child who answered "the reindeer" to (7) did so even though no set of reindeer had been established to exist by the previous discourse). These cases of *the* overuse therefore cannot be easily accounted for under Wexler's proposal. On the other hand, Maratsos's proposal needs to explain why children overuse *the* in naturalistic production but not in response to the stories in (8).

3.4. Maratsos's findings and the FH

Maratsos found that young children overuse *the* with indefinites. While my proposal in (2) also predicts overuse of *the* with indefinites, it crucially ties overuse of *the* to specificity. The contexts in which children overused *the* in Maratsos's study were not specific for an adult speaker. This is illustrated by (11): it is infelicitous to use the specificity marker *this* when referring to a member of a previously established set, unless the context singles out a particular member of this set. The contexts in Maratsos's study did not single out a particular member of the set. Thus, on the surface, Maratsos's findings are incompatible with the FH – but see the discussion below.

11. I have five squirrels and three turtles. #I will give my mother this squirrel.

---

[2] As Ken Wexler (p.c.) points out, the utterance *Where's a stool?* would also be infelicitous. An adult speaker would probably use *the* and add modification, as in *Where's the stool that was supposed to be here?* Modification would narrow down the discourse domain sufficiently to enable the listener to establish uniqueness.

[3] Brown also reports some cases of *a* overuse with definites *specified by entailment*, such as *I don't like a crust [of bread]* and *Where there's a heel [of the sock].* Of all the errors in article usage reported in Brown (1974:354), specific indefinites constitute the largest category.

3.4.1. Explanations of Maratsos's findings: implications for the FH

If Maratsos is correct and young children's overuse of *the* is due to their egocentricity, then the results are irrelevant for the FH: children's article misuse is psychological rather than linguistic in nature. This would mean that children, unlike adult L2-learners, successfully set the Article Choice Parameter and treat *the* as a marker of definiteness. However, they mistakenly treat some DPs that should be indefinite as definite, as a result of their egocentricity.

On the other hand, if children's overuse of *the* is tied to presuppositionality, as proposed by Wexler (2003), then an explanation is necessary for why presuppositionality plays a role in L1-acquisition of articles while specificity plays a role in L2-acquisition of articles. This difference is not predicted under my proposal. It would mean that different linguistic processes are at work in L1 vs. L2 acquisition of articles, and would require an explanation (but see also Appendix 4 on the possible role of presuppositionality in the L2-English of some learners).

3.4.2. A possible link between Maratsos's findings and the FH

Finally, it is possible that Maratsos's results are consistent with the FH, as follows. It is possible that children, having lively imaginations, do indeed attach particular importance to a particular member of the established set (e.g., a particular squirrel). However, the reason that they go on to say *the squirrel* is not because they attribute their own knowledge to the hearer, but rather because they use *the* to mark specificity. If young children, like adult L2-learners, are subject to the FH for Article Choice, then they should sometimes divide English articles on the basis of specificity rather than definiteness. This would account for children's overuse of *the* with specific indefinites in naturalistic production as well as for overuse of *the* found in Maratsos's study (with the additional provision that children imagine a particular member of the set).

Recall that the children's overuse of *the* in the elicitation tasks was not at the level of 100%: it was 58% for the "4 Low" group (and much lower for the other groups). This optionality of article use is fully consistent with the FH.

Independent evidence in favor of the FH would need to come from the domain of non-specific definites. Maratsos's explanation does not predict any overuse of *a* with definites in child English, since there is no obvious link between overuse of *a* and egocentricity[4]. (While Maratsos did find overuse of *a* with definites among 3-year-olds, he had to give an independent explanation of this fact. As discussed above, the results may have been confounded by the form of the elicitation question). In contrast, the FH predicts overuse of *a* with non-specific definites. It would be necessary to test children on sentences such as (12), varying whether the definite DP is specific or non-specific.

12. I want to talk to <u>the owner of this toy store</u>.

A situation would need to be set up in which the child is told about a toy store (or perhaps actually shown a "toy" toy store). In the specific scenario, the child is told about who the owner of the store is – the name and/or description of the owner are given. In the

---

[4] Maratsos does not address the question of non-specific definites at all, and one might imagine a possible formulation of his proposal that would predict overuse of *a* with non-specific definites. As Maratsos's explanation stands, however, it addresses only overuse of *the* with indefinites, and does not make any predictions concerning overuse of *a* with definites.

non-specific scenario, the owner is never mentioned. In both scenarios, the child is told that she can get a toy from the toy store but that she'll need to ask someone for it. The child is then asked "Who do you want to talk to?" or "Who will you ask?" The hope is that the child will answer with a sentence of the form in (12); the prediction would be that in the specific scenario, the child will say *the owner,* and in the non-specific scenario the child will say *an owner.* A potential problem might be is that instead of *owner,* the child will use a word like *sales person*; since there can be multiple sales people in a store, *a* would be felicitous. The situation needs to be constrained enough to focus the child's attention on a referent that is obligatorily unique (e.g., *the principal of this school*; *the president of this country*; etc.).

If it is found that children indeed overuse *a* with non-specific definites, this would provide support for the specificity proposal over Maratsos's egocentricity explanation.

3.4.3. Which explanation is correct?

All three explanations discussed above – Maratsos's explanation of egocentricity, Wexler's explanation of presuppositionality, and the FH as applied to L1-acquisition – are in principle possible accounts of *the* overuse in the data. For ease of reference, I will henceforth refer to these proposals by the names given in (13).

13. Three explanations of Maratsos's findings:
   a) The egocentricity proposal (Maratsos 1976)
   b) The presuppositionality proposal (Wexler 2003)
   c) The specificity proposal (based on the FH for Article Choice)

Both the egocentricity proposal and the specificity proposal require a stipulation in order to account for Maratsos's data. The stipulation is that children's imaginations allow them to focus on a particular member of the set in contexts where adult speakers do not single out a particular member. The presuppositionality proposal has the advantage of accounting for the data without additional stipulations. However, this proposal cannot account for overuse of *the* in non-partitive contexts in naturalistic data.

As discussed in the previous section, the specificity proposal and the egocentricity proposal can be teased apart by testing article use with non-specific definites. In order to tease both of these proposals apart from the presuppositionality proposal, it would be necessary to test whether or not children's imaginations contribute to *the* overuse – i.e., whether the stipulation discussed above is warranted. One possibility might be as follows. The investigator would instruct children to make up a story about an animal. The child would be left alone for a minute (or less) and given an opportunity to do this[5]. A puppet would then come in, and ask the child to tell it the story. If children then proceed to talk about *the squirrel* (or *the cat, the fox,* etc.), this would provide evidence that children overuse *the* when talking about imaginary creatures, as long as the creature is important for the purposes of the story; it would then be possible to argue that the same power of imagination contributes to overuse of *the* in partitive contexts. On the other hand, if

---

[5] It is important to give the child a little time to make up the story, in order to allow the child to establish the importance of a particular animal. In the partitive stories in Maratsos's studies, the groups of animals were mentioned early on in the story, and the child had an opportunity to focus in on a particular animal before the investigator asked the question. In contrast, in stories like (8), no actual animal or animals were ever mentioned (only the hypothetical zebras and lions that the man was looking for). Thus, the child had no opportunity to focus her attention on any particular animal.

children do not overuse *the* in these contexts, confining their *the* overuse to presuppositional environments, this would provide strong support for the presuppositionality proposal.

Finally, it is possible that two processes are at work in L1-acquisition: that children overuse *the* both in presuppositional contexts, per Wexler 2003, and in specific indefinite contexts, per the FH (or because of egocentricity). If this is the case, then more work is needed to relate the two phenomena. For instance, it may be necessary to propose two parameters for article choice, one which relates to specificity, and another which relates to presuppositionality.

4. Later studies of article choice in L1-acquisition

More recent studies of article choice in child English have added to Maratsos's findings by examining article use in a wider variety of contexts. I discuss three relevant studies below, and look at how they relate to the FH.

4.1. Article choice with different types of indefinites in child English

The study of Schafer and de Villiers (2000) examined article choice in several different types of indefinite contexts.

4.1.1. `Study design

This study used an elicitation task with children ages three through five, testing article use in six different context types. The context in (14a) was designed to elicit *the* with previous-mention definites, while (14b) elicited definites specified by entailment (e.g., *the door of the cage*). The other contexts were indefinite. Schafer and de Villiers called the context in (14c) specific because the child, but not the child's listener, is aware of the referent; I am not using their label for the context, since I will argue that this context is in fact non-specific. The contexts in (14d) arguably elicits a narrow-scope indefinite. The context in (14e) looks similar to (14b) but requires the child to imagine the scenario. The context in (14f) is partitive, resembling the contexts tested by Maratsos.

14. a) *Definite – previous mention*
    Emily has two pets, a frog and a horse. She wanted to ride one of them, and so she put a saddle on it. Guess which?/What was it?

    b) *Definite by entailment*
    Adrienne got a pet hamster for her birthday and put it in a nice cage. It tried to escape so she quickly closed something – What did she close?

    c) *Indefinite: non-partitive, wide scope*
    I'll bet you have something hanging on the wall of your room at home. What is it?

    d) *Indefinite: non-partitive, narrow scope*
    Cindy is going to the pond. She wants to catch some fish. What will she need?

    e) *Indefinite following* have
    Think of a baseball player. Can you imagine what one looks like? What does he have?

f) *Indefinite: partitive*
Three ducks and two dogs were walking across a bridge. One of the animals fell off the bridge and said "Quack". Guess which?/What was it?

(from Schafer and de Villiers 2000:612)[6]

4.1.2. Results

The results were as follows. On items like (14a), children mainly (~60% of the time) used *the* correctly, but did exhibit relatively large (~20%) overuse of *a* (compared to adult controls, who used *the* correctly). Overuse of *a* was especially frequent if the question asked by the experimenter was "What was it?" rather than "Guess which?" This isn't too surprising – a natural answer to "What was it?" might be "It was a horse", meaning "The animal that she put a saddle on was a horse." In this predicative use, an indefinite is quite felicitous. This is the same critique that applies to Maratsos's definite-eliciting context in (5). The children were much more accurate in the *definite by entailment* context (14b), in which they used *the* around 90% of the time.

On non-partitive indefinites of the type in (14c) through (14e), children as well as adults overwhelmingly appropriately produced *a*. In (14f), on the other hand, production of *a* or *one of the* (also appropriate) was only around 50% or 60% for children of all ages, with overuse of *the* being at least 40% for all of them. Interestingly, however, adults also used *a* or *one of the* only 70% of the time, and overused *the* as much as 30% of the time.

4.1.3. Discussion: Schafer and de Villiers' results and the FH

Thus, it would appear that children overuse *the* with partitive indefinites (14f), but not with non-partitive indefinites. While Schafer and de Villiers called non-partitive scenarios like (14c) specific, these scenarios do not in fact elicit indefinites which are specific in the sense discussed in Chapter 2. While Schafer and de Villiers do not provide examples of the kinds of answers children gave to (14b), typical answers might be "It's a picture" or "I have a picture on my wall." Neither context is likely to be specific, since the exact identity of a particular picture is irrelevant for the discourse: the child is not asked to talk about a particular picture, there is no story context set up to single out a particular picture, etc. The other non-partitive indefinite contexts are even less likely to be specific.

As for overuse of *the* in partitive contexts, the discussion of partitive contexts in Maratsos's study (Section 3) is equally applicable to the partitive contexts in Schafer and de Villiers' study. Schafer and de Villiers' findings of *the* overuse in partitive contexts should be approached with caution, however, since adults also made errors on these contexts.

Finally, the fact that children correctly used *the* in *definite specified by entailment* contexts such as (14b) may have implications for my proposal: one could argue that definites such as *the door of the cage* are non-specific, since the speaker (the child) has no knowledge of the cage. Since children correctly used *the* in this context, this suggests that children do not after all associate *the* with specificity. On the other hand, one could also argue that *the door of the cage* is specific, since the property of being the door of the

---

[6] The examples reported here are taken verbatim from Schafer and de Villiers, but the labels for the contexts are changed so as to be more consistent with my overall discussion.

cage is quite noteworthy: Adrienne needs to close this door (and no other) in order to prevent the hamster's escape. It is important to find some way in which specific and non-specific definites in child language can be distinguished. In adult language, we can use denial of speaker knowledge (as in, *the owner of this store – whoever that is*), but such constructions would be quite difficult to elicit from a child. (See also footnote 3 concerning some cases of *a* overuse in *definite by entailment* contexts reported in Brown 1973).

Thus, Schafer and de Villiers' study does not either directly support or directly refute the predictions of the FH. The high use of *a* in contexts which are most likely non-specific is consistent with the FH but does not provide evidence in favor of it. The overuse of *the* with partitive indefinites may or may not be consistent with the FH, depending on how these results are interpreted. The same holds for correct use of *the* in constructions like (14b).

4.2. Child English vs. adult Salish

The work of Matthewson (1998) on Salish (see Chapter 2) inspired two studies of article acquisition in L1-English[7]. Matthewson and Schaeffer (2000), as well as Matthewson, Bryant and Roeper (2001), compared article choice in child English to article choice in adult English and in adult Salish. I will discuss these studies below, and will then show that their results can be interpreted as providing evidence for the FH (but with some caveats concerning methodology).

4.2.1. Matthewson and Schaeffer 2000: study design

In an experiment with young L1-English children, Matthewson and Schaeffer (2000) tested children ages two through five, as well as adult controls, on the four types of contexts exemplified in (15) through (18).

15. A context: *the* required for adults
    Situation: car on table
    Elmo: Hey, who is this (pointing at Donald Duck)?
    Child: Donald Duck!
    Elmo: And this (pointing at the car on the table)?
    Child: A car!
    (Donald Duck pushes the car)
    Elmo: What did Donald Duck just do?
    Child: He pushed **the** car. (target response)

16. B context: *a* required for adults
    Situation: picture of Mickey Mouse who just finished drawing a house
    Elmo: Hey, who is this (pointing at Mickey Mouse)?
    Child: Mickey Mouse!
    Elmo: And what did Mickey Mouse just do?
    Child: He drew **a** house. (target response)

---

[7] Both of these studies are based on Matthewson's (1998) proposal that Salish non-polarity articles encode *assertion of existence,* rather than Matthewson's (1999) proposal of these articles as choice functions.

17. C context: *a* required for adults; (i) incomplete object
     Situation: picture of Bert[8] painting a car (NOT finished)
     Elmo: Hey, who is this (pointing at Bert)?
     Child: Bert!
     Elmo: And what is Bert doing?
     Child: He's painting **a** car. (target response)

18. C context: *a* required for adults; (ii) non-existing object
     Big Bird: Oh, I'm so bored. I don't know what to do. Oh, you know what,
                I'm going to the forest, and I'm gonna DRAW something there.
     Elmo: What do you [think] Big Bird is gonna do in the forest?
     Child: He's gonna draw **a** tree. (target response)

                                                    (Matthewson and Schaeffer 2000:26)

4.2.2. *Matthewson and Schaeffer: results*

The results of this study were as follows. On the definite context in (15), children of all ages almost never overused *a*. This is in contrast to the findings of Maratsos (1976) and Schafer and de Villiers (2000), who found some overuse of *a* with previous-mention definites. It is possible that the physical presence of the car in Matthewson and Schaeffer's experiment contributed to correct use of *the*.

On the indefinite context in (16), the youngest children (two-year-olds) exhibited 23% overuse of *the*. Overuse of *the* was much smaller for children three years of age and older, but it should be noted that adults exhibited 4% *the* overuse in this context. In the two "type C" scenarios in (17) and (18), overuse of *the* was much lower (12% for two-year-olds).

It is important to note that the contexts in (16) and (17) both suffer from a potential confound: it is not made clear in the paper whether the listener, Elmo, can see the house or car under discussion. Thus, use of *the* may not be entirely infelicitous (although it is dispreferred, given the adult speakers' results). A better experiment would be one in which the listener clearly did not share the speaker's knowledge of the referent.

Matthewson and Schaeffer take overuse of *the* in (16) as indication that children overuse *the* in indefinite contexts when the speaker has knowledge of the referent. They point out that while English requires *a* in the response to (16), Salish requires a wide-scope (non-polarity) article. If L1-English children treat English like Salish, using *the* to mark wide scope / assertion of existence, then their overuse of *the* in (16) is expected.

However, Matthewson and Schaeffer argue that if children treat English like Salish, they should consistently overuse *the* in contexts like (16). Yet overuse of *the* was fairly low even for two-year-olds (and much lower for older children). Matthewson and Schaeffer therefore conclude that L1-English children know the semantics of *the* and *a* in English, and that overuse of *the* is due to a pragmatic failure, namely lack of the *Concept of non-shared assumptions* proposed by Schaeffer (1999). It is given in (19).

---

[8] The original context in Matthewson and Schaeffer states "picture of Elmo painting a car". However, the subsequent context makes clear that Elmo is doing the asking while Bert is doing the painting, so this was probably a typo.

19. *Concept of Non-Shared Assumptions* (pragmatic):
Speaker and hearer assumptions are always independent
(Matthewson and Schaeffer 2000:26)

Lack of the concept in (19) leads young children to sometimes attribute their own knowledge to the listener, using *the* when the child alone (i.e., the speaker) is familiar with the object. Thus, Matthewson and Schaeffer's explanation, like Maratsos's, is psychological in nature.

4.2.3. Matthewson et al. 2001

On the other hand, Matthewson et al. (2001) argue that L1-English children *do* treat English like Salish, using *the* to mark assertion of existence (wide scope) rather than definiteness. The relevant scenario is provided in (20). The adult controls replied negatively to the question in (20), adding clarifications which showed they considered the question inappropriate (e.g., for a story concerning hats, adults might say "Not THE hat, A hat" – Matthewson et al. 2001, ex. 17a).

20. *scenario from Matthewson et al. (2001), ex. (11)*
[The experimenter tries to put a necklace on Bert, but it falls off]
"Every time Bert puts his necklace on the necklace falls off. The necklace is
ᵥbroken."
[The experimenter puts a different necklace on Ernie]
"Look at Ernie. Did Ernie wear the necklace?"

On the other hand, children (ages three to seven) said "yes" to the question in (20) and other stories like it as much as 72% of the time. Matthewson et al.'s explanation is that children use *the* as if it were a Salish wide-scope article – i.e., an article that marks the speaker's assertion that the referent exists (since the referent of *the necklace* in (20) undoubtedly exists). However, this experiment is problematic, since it does not directly test children's article use but, rather, children's ability to react to presupposition failure (but see the discussion in Matthewson et al. for children's responses to presupposition failure in other contexts).

4.2.4. Discussion: relation to the FH

The results of Matthewson and Schaeffer (2000) are easily explained under my proposal. While this study compared child English to adult Salish, which encodes *assertion of existence,* the results are fully compatible with the view of *specificity* that I have been arguing for. In Matthewson and Schaeffer's study, the two-year-olds overused *the* in indefinite contexts in which the referent (the house) had particular importance to the speaker (the child). The data from this "Salish-based" study of L1-English article choice in fact provide support for the FH for Article Choice. The optional (23%) overuse of *the* that presented a difficulty to Matthewson and Schaeffer is not a problem for my analysis: under the FH, overuse of *the* is in fact predicted to be optional.

It is less clear whether the results of Matthewson et al.'s (2001) study are compatible with my proposal. One might argue that *the necklace* in (20) is likely to have a specific reading: the speaker (the experimenter) is referring to a particular necklace, the one that has the noteworthy property of being worn by Ernie. On the other hand, *this*ref is not possible in this context in adult English, which makes it doubtful that the context licenses

a specific indefinite. The results of Matthewson et al.'s study are therefore not directly compatible with the FH. On the other hand, overuse of *the* in contexts like (20) can be explained under Wexler's (2003) presuppositionality proposal (see Section 3.3.2), since the existence of a set of necklaces (which consists of two necklaces) has been established.

The results of both "Salish-based" studies should be taken with caution, however, because of the methodological difficulties discussed above. It is possible that overuse of *the* in Matthewson and Schaeffer's study was due to the fact that in the experiment, Elmo could see the house; it is also possible that the results of Matthewson et al.'s study were due to children's inability to react to presupposition failure. Thus, more evidence is necessary.

4.3. Summary: article choice in child English

The results of the L1-acquisition studies reported in this chapter leave open the question of whether specificity plays a role in L1-acquisition of English articles. Further testing is necessary, which would distinguish between specific and non-specific indefinites (e.g., by varying the importance attached to the referent) as well as between specific and non-specific definites.

If specificity does play a role in L1-acquisition of articles, what would that mean? Given my proposal, it would mean that the FH is for article choice is operative in both L1 and L2 acquisition – that both types of learners fluctuate between the possible settings of the Article Choice Parameter. In the case of L1-acquisition, this may seem a surprising finding, given the evidence from other domains that children set parameters very early on (e.g., see Wexler 1998 on early parameter setting in the domains of word order, verb movement, and null subjects, and Snyder 2002 for early parameter setting in the domain of pied-piping and preposition stranding). It may be preferable to attribute article misuse in L1-acquisition to pragmatic deficits rather than unset parameters.

However, it is in fact possible to tie pragmatic deficits and unset parameters together in the domain of article choice (thanks to Ora Matushansky, p.c., for suggesting this line of reasoning). The FH states that learners should fluctuate between parameter settings until the triggers in the input lead them to set the parameter to the target value. I have suggested that adult L2-learners have difficulty generalizing from the input triggers, and that this difficulty is especially visible in the domain of article choice, where the input is subject to ambiguity. On the other hand, children are known to be quite good at generalizing from the input and setting parameters. I would like to suggest that the Article Choice Parameter is particularly difficult for children to set because the triggers are discourse-based.

In order to set the Article Choice Parameter, the child needs to establish whether *the* encodes the state of hearer knowledge (definiteness) or the state of speaker knowledge (specificity). Each time she hears *the,* she must decide (on an unconscious level, of course) whether the speaker and listener shared knowledge of the referent, or whether the speaker along had knowledge of the referent. This is a fairly complex computation – all the more complex for young children who are egocentric and have trouble distinguishing between speaker and hearer knowledge. (In contrast, we do not expect adult L2-learners to be egocentric, since they are adults; see also the discussion in Chapter 8). Children's egocentricity may prevent them from successfully setting the parameter, and the resulting state of fluctuation causes errors of *the* overuse with specific indefinites. Thus,

egocentricity plays a role, but only in so far as it interacts with parameter setting. On this view, child L1-learners and adult L2-learners both initially fail to set the Article Choice Parameter, but for different reasons: adult L2-learners have difficulty generalizing from the input triggers, while L1-learners have difficulty evaluating the triggers themselves.

Of course, this proposal is contingent on evidence that L1-learners do in fact undergo the fluctuation predicted by the FH. Definitive evidence either for or against this does not exist at the present.

5. Previous studies of article choice in adult L2-acquisition

The previous section discussed evidence for *the* overuse in child English, and both psychological and linguistic explanations for this phenomenon. In the case of adult L2-learners, the psychological explanation for article errors is typically not applied: *a priori*, there is no reason to think that adults learning a second language should have any tendency to attribute their assumptions to their listeners. Analyses of articles in L2-English have therefore relied on linguistic rather than psychological theories.[9] I will now discuss some of the work on L2-English article choice, and then look at the relationship between the results of previous studies and my proposal.

5.1. L2-English article use in the Bickertonian framework

Many studies that have looked at L2-acquisition of definite and indefinite articles have classified articles on the basis of two features, which have to do with *existence in the world* and *hearer knowledge*; this classification system is due to Bickerton (1981:146-8). It has been used by many L2 researchers (e.g., Huebner 1983, 1985; Parrish 1987; Tarone and Parrish 1988; Thomas 1989a; Young 1996; Murphy 1997; Robertson 2000). I will use here the terminology for this system as given in Huebner (1983).

5.1.1. The Bickerton/Huebner classification system of articles

This system classifies each DP as plus or minus *specific referent* ([+/-SR]) and plus or minus *assumed known to hearer* ([+/-HK]). While the term *specific referent* is not clearly defined, it is usually taken to mean *has a referent in the actual world* (as opposed to a possible world).

In English, only [+SR, +HK] DPs obligatorily take *the*. These DPs are definites: they include previous-mention DPs (21a), DPs that have referents which are unique in all contexts (21b), DPs that are unique by entailment (21c), etc. Definites receive the specification [+SR] because their referents exist in the actual world and [+HK] because their referents are known to the hearer as well as the speaker[10].

Indefinites in this framework can be either [+SR, -HK] or [-SR, -HK]. The [+SR, -HK] context corresponds to the wide scope reading of indefinites; thus, (22a), a [+SR, -HK] context, asserts the existence of a dog in the actual world, while (22b), a [-SR, -HK]

---

[9] Work on articles in L2-acquisition has also looked at the role of instruction and explicit rule knowledge on the part of L2-learners – see, e.g., Murphy 1997. I do not discuss this approach here – but see the discussion in Chapter 8 on why article choice in L2-English cannot be accounted for by explicit strategies alone.

[10] This classification system ignores the existence of definites which do not have a referent in the actual world – see the discussion in Chapter 2.

context, does not assert the existence of a car in the actual world, since *a new car* scopes under a modal.

Huebner's framework assigns the [-SR, +HK] specification to generics, exemplified in (23). However, it is not clear why this specification should be assigned to generics, and how it can account for the fact that singular generics sometimes occur with *a* (23b) and sometimes with *the* (23c).

21. a) (Chris approached me carrying a dog) *The dog* jumped down and started barking.
    b) *The moon* will be full tomorrow.
    c) I approached his front door and rang *the bell.*

22. a) Chris approached me carrying *a dog.*
    b) I guess I should buy *a new car.*

23. a) *Fruit* flourishes in the valley.
    b) *A paper clip* comes in handy.
    c) *The Grenomian* is an excitable person.   (Thomas 1989a:337)

5.1.2. Studies of L2-English article choice in the Bickertonian framework

A number of analyses have been proposed concerning articles in L2-English in this framework. However, they have not necessarily arrived at similar conclusions. For instance, Huebner (1983), studying the development of L2-English by one adult Hmong speaker, found that this speaker initially used *the* across all environments and later associated *the* with the [+HK] feature, using it with definites and generics. Master (1987), looking at L2-English learners from five different L1 backgrounds, similarly claimed that *the* was associated with [+HK].

Huebner's and Master's findings that *the* was used in [+HK] contexts means that L2-learners appropriately used *the* with definites ([+SR, +HK]) and also used *the* with generics ([-SR, +HK]). It is not clear to what extent the latter constitutes an error, since some types of generic DPs can be used with *the.* Thomas (1989a) notes that, despite Huebner's and Master's claims that *the* was associated with the [+HK] feature, the data from both studies actually show overuse of *the* in indefinite [+SR, -HK] contexts.

While Huebner's subject originally overused *the* in all indefinite contexts, the subject starting dropping *the* from [-SR, -HK] contexts about six weeks earlier than from [+SR, -HK] contexts. In a follow-up study 20 months later, Huebner (1985) found relatively little overuse of *the* with indefinites: 7 tokens in [+SR, -HK] contexts (19% of total) and 7 tokens in [-SR, -HK] contexts (23% of total); the learner had no *a* overuse in definite contexts.

Finally, Parrish (1987), in studying the acquisition of English by an L1-Japanese speaker, found 9.4% *the* overuse in [+SR, -HK] contexts, compared to no *the* overuse in [-SR, -HK] contexts and no *a* overuse in [+SR, +HK] contexts[11].

Thomas (1989a) examined the relationship between *the* overuse and [+SR, -HK] contexts further by testing article choice among adult L2-English learners from nine L1 backgrounds (mostly Japanese and Chinese), hypothesizing that *the* overuse is linked to

---

[11] The data for both Huebner (1985) and Parrish (1987) discussed in this section are taken from Hawkins (2001:237-239).

the [+SR] feature. The learners were shown photographs of various scenes and asked to provide descriptions of the scenes. The learners' responses were then analyzed for presence of articles in obligatory contexts. I will discuss here only her results for the [-art] group, those learners whose L1's lacked articles (learners whose L1's had articles performed higher overall than the [-art] group, but showed a similar pattern of responses).

Thomas found that the L2 learners in the [-art] group had very high (81%) use of *the* in (appropriate) [+SR, +HK] contexts, and almost never used *a* in these contexts. The learners produced *a* more often in [-SR, -HK] contexts than they did in [+SR, -HK] context. While use of a null determiner was about equal in both contexts, the L2 learners inappropriately produced *the* in 16% of [+SR, -HK] contexts but only 5% of [-SR, -HK] contexts (this difference was statistically significant). Thomas suggested that the learners had initially associated *the* with [+SR], but were already in the stage of abandoning this generalization (hence only 16% overgeneralization of *the* in [+SR, -HK] contexts).

On the other hand, Murphy (1997), looking at L1-Korean and L1-Spanish adult learners of English, failed to find an association between *the* and either [+SR] or [+HK]. While the Spanish speakers, whose L1 has articles, performed better overall than the Korean speakers, whose L1 lacks articles, neither group overgeneralized *the* in any of three tasks (an oral task, a written task and a cloze test) administered to them. The biggest source of error for learners in both groups was article omission.

5.1.3. L2-English article choice and specificity

All of the studies of articles in the Bickertonian framework look at definiteness and scope, but do not look at specificity as defined in Chapter 2. Thomas's (1989a) results suggest that L2-learners have a slight tendency to overuse *the* with wide scope indefinites [+SR, -HK], compared to narrow scope indefinites [-SR, -HK] (this difference was observed in Huebner's and Master's studies as well). This could mean that the learners are in fact associating *the* with wide scope (the [+SR] setting), as Thomas suggests. Alternatively, it could mean that they are associating *the* with specificity, since specific indefinites are obligatorily wide scope. There is some indirect evidence that the latter hypothesis is correct.

Thomas notes that the participants in her picture-description study often used the *there*-construction, which she counted as [+SR, -HK]. The kinds of *there*-constructions that are likely to be elicited by picture-description tasks are probably more likely to contain non-specific than specific indefinites. If the speaker describes a picture by saying *There is a table in the room, There is a rabbit in the garden,* etc., she is probably not intending to describe a particular table or rabbit, but to simply list the objects in the picture. In fact, it would be infelicitous to describe a picture by saying *There is this table in the room* or *There is this rabbit in the garden* (see Chapter 7 for more discussion).

Thomas notes that when *there*-sentences are removed from the data, overuse of *the* in [+SR, -HK] category goes up somewhat (to as much as 25.9% for the mid-level group). She notes that overuse of *the* goes up even more if *have*-constructions are removed from the data. While Thomas does not give any examples of *there*-constructions in the data, she does give some examples of *have*-constructions (24). All of these contexts are more compatible with non-specific than with specific indefinites: the speaker is simply listing the objects that she sees in the picture, without intending to say anything else about them.

24. a) This picture have a doctor and a child.
    b) [Describing the view through a window] In outside have a tree.
    c) At room have a flower.   (Thomas 1989a:351, ex. 3)

Thus, Thomas's data suggest that overuse of *the* is fairly low in environments that are clearly non-specific (*there*-constructions and *have*-constructions). Since Thomas does not give examples of other wide scope contexts in the data, we cannot conclude that overuse of *the* occurred in contexts that were specific. However, Thomas's data indirectly suggest a relationship between overuse of *the* with indefinites and specificity.

5.2. L2-English article choice and partitivity

Kaneko (1996) investigated article use in a variety of contexts among L1-Japanese and L1-Spanish learners of English. Japanese has no articles, while the article system in Spanish is similar to the one in English. Of particular interest is Kaneko's context involving partitive indefinites, illustrated in (25); the learners' task was to fill in the missing article, *a* in this case.

25. Once there was a boy. He wanted to write a letter. He went to his mother. She
    showed him some pencils. So he took ___ pencil. And he wrote his letter.

L1-Japanese learners of English used *the* 52.7% of the time in contexts like (25), compared to 26.9% *the* use by L1-Spanish speakers. For the Spanish speakers, the error rates for this context and for the various other contexts in Kaneko's study were quite similar. In contrast, the Japanese speakers had a much greater error rate in this context than in other contexts. It appears that learners whose L1 (Japanese) has no articles often used *the* for partitive indefinites.

Kaneko's findings are problematic for my proposal. The contexts she tested were not specific, but were partitive. Like the child L1-learners in Maratsos's study, the L1-Japanese learners of English in Kaneko's study appear to be associating *the* with partitivity. This is not predicted under the FH.

In Appendix 4, I report some data from L1-Korean speakers concerning overuse of *the* in partitive contexts, and discuss possible explanations, including the influence of topic-marking in Korean and possible confounds of the testing procedure. As discussed in Appendix 4, the same explanations apply to Kaneko's findings for L1-Japanese speakers. The relationship between L2-English article use and partitive contexts requires more investigation. If L2-learners indeed associate *the* with partitivity, a new account is needed that would show how specificity and partitivity interact in L2-acquisition. I leave the issue open (but see Appendix 4 for discussion).

5.3. Overuse of *a* in L2-English

Leung (2001) looked at article choice for L1-Chinese speakers who had acquired English as an L2 and were acquiring French as an L3. She tested learners' article use in both their L2 and their L3. She used an elicitation task similar to the one used by Schafer and de Villiers in their study of L1-acquisition of articles. The questions used in Leung's elicitation study are given in (26). In the case of the indefinite contexts, I am using the terms *wide scope* and *narrow scope* where Leung used *specific* and *non-specific,* respectively, since Leung's use of the term *specificity* most closely corresponds to *wide scope.* In my terminology, (26b), which Leung called specific, is actually a non-specific

indefinite context, since the most likely answer to it is "I have a X", where the exact identity of the X is irrelevant for the discourse.

26. a) Calvin has two pets, a pig and a crocodile. He decided to sell one of them. Which one do you think it was? (definite)

   b) You probably have something on your desk in your room at home. What is it? (wide-scope indefinite)

   c) You are going to the cinema. You want to watch a movie on your own. What will you need to buy at the cinema? (narrow-scope indefinite)

Leung found that the learners were fairly accurate at using the indefinite article in both wide-scope and narrow-scope indefinite contexts, in both their L2 and their L3. On the other hand, they were highly inaccurate in definite contexts like (26a): correct use of the definite article was only around 40% in their L2, and even lower in their L3. However, it is not clear whether (26a) is a good context for eliciting definites: it seems felicitous to answer (26a) with "It was a pig", meaning "The pet that Calvin decided to sell was a pig". The control native speakers in Leung's study used *the* 88% of the time in this context, with overuse of *a* at 12%. This suggests that, at least some of the time, native English speakers consider an indefinite to be a felicitous answer to (26a).

Leung also conducted a picture elicitation task with the L2/L3 learners. The learners were extremely accurate at using the indefinite article in indefinite contexts, in both English and French. Leung does not give examples of the learners' utterances, so we do not know whether the indefinite contexts were specific or non-specific. In the case of definites, the learners were fairly accurate, using the definite article in at least 80% of all instances of article use; one exception were beginner L3 learners, who used the definite article in French in only 60% of all instances of article use with definites, while being accurate with indefinites.

5.4. Summary: L2-English article choice and the FH

None of the studies reported in this section directly tested specificity. However, some of the studies are consistent with my proposal. The findings of Huebner (1983), Master (1987) and Thomas (1989a) that *the* was overused with wide-scope indefinites is consistent with my proposal, since wide-scope indefinites may be specific. No study tested article use with non-specific definites. While Leung (2001) found overuse of *a* with previous-mention (and hence specific) definites, the methodology for eliciting definites may be problematic, as discussed above. Finally, Kaneko's (1996) findings of *the* overuse with partitive contexts are incompatible with the FH and require a separate explanation (see Appendix 4).

6. L2-English article use and transfer

In the rest of this thesis, I will argue that L2-English learners have access to both settings of the Article Choice Parameter, and that this knowledge is coming directly from UG. I will show that the [+HK/+SR] distinction made by previous studies of L2-acquisition cannot account for the distribution of articles in L2-English: the data in Chapters 4 through 6 will show that *the* is not associated with wide scope or assertion of existence, but with specificity.

This argument will be based on the findings of empirical studies with L1-Russian and L1-Korean learners of English. Before I can argue for direct UG access to the Article Choice Parameter on the part of these learners, I need to rule out any role of transfer.

Russian and Korean do not have articles. However, they may still have some ways of coding for definiteness or specificity that may cause transfer effects. I will now examine definiteness and specificity in Russian and Korean, and argue that nothing in these languages could guide the L2-learners to divide English articles on the basis of either definiteness or specificity.

6.1. Definiteness and specificity in Russian

Russian has no articles, and a bare DP can in principle be either definite or indefinite, specific or non-specific. Russian has demonstratives, which, as in English, may be used with previous-mention definites. Russian also has some lexical marking on indefinites, which I will now discuss.

6.1.1. Specificity marking in Russian

Like Hebrew (see Chapter 2), Russian has a potential specificity marker corresponding to the numeral 'one'. When the numeral *odin,* 'one', is de-stressed, it has a specific reading. In simple declarative sentences, both bare indefinites and *odin*-marked indefinites are possible. As shown in (27), use of *odin* indicates that there is something important about a particular member of the set denoted by the NP.

27. a) *Ja pročitala včera odnu knigu. Ona byla očen'*
    I   read-PST yesterday one   book-ACC she was-FEM-SG very
    interesnaja!
    interesting
    "I read a (certain) book yesterday. It was very interesting!"

   b) *Ja počitala knigu, a potom legla spat'.*
    I   read-PST book-ACC and then lay-PST sleep-INF
    "I read a book for a bit, and then went to sleep."

In intensional contexts, *odin*-marked indefinites obligatorily take wide scope while bare indefinites typically take narrow scope, as shown in (28). In this, Russian also looks like Hebrew.

28. a) *Ja xoču pročest' odnu knigu   o   babočkax.*
    I   want read-INF one   book-ACC   butterflies-PRP
    "I'd like to read a (certain) book about butterflies"
                                (*want>indefinite, indefinite>want)

   b) *Ja xoču pročest' knigu   o   babočkax.*
    I   want read-INF book-ACC about butterflies-PRP
    "I'd like to read a book about butterflies" (want>indefinite, *indefinite>want)

However, it is possible to construct contexts in which a bare indefinite scopes over an intensional verb, if the indefinite bears relative clause or other modification. It is difficult to construct an appropriate context – since Russian lacks definite as well as indefinite articles, a heavily modified noun phrase could potentially be definite rather than

indefinite. Consider, however, a scenario in which a visitor at an airport comes up to a security guard and utters (29): since there are probably multiple little girls on any flight, and since the security guard has no idea which girl is being discussed, the underlined phrase in (29) is indefinite. Despite lack of an *odin* marker, this indefinite can still take wide scope over *try*.

29. *Ja pytajus'   najti   malen'kuju   devočku   s   Vašingtonskogo   rejsa.*
I   try   find-INF little   girl-ACC from Washington   flight
Ona   moja   vnučka.
she   my   granddaughter-NOM

"I am trying to find a little girl from the Washington flight. She is my granddaughter."

Even though Russian has no definite article, *odin* is incompatible with definites. This is not predicted under my analysis so far: since there is no definite article in Russian, if *odin* is marked [+specific], there is nothing preventing it from occurring in both definite and indefinite environments. A possible solution is to say that Russian has a bare definite article (distinct from a bare indefinite article), which is inserted in all [+definite] contexts, thus preventing the insertion of *odin* via Maximize Presupposition. Alternatively, *odin* may have additional conditions on it that make it incompatible with definiteness[12]. I leave the issue open.

---

[12] *Odin* is not quite identical to *this*_ref in its distribution. For instance, *odin* is unlikely to be used in the Russian equivalent of (ia), as shown by (ib). The use of *odin* in (ib) is not entirely infelicitous, but it implies that the speaker is somehow familiar with the set of unicorns – i.e., that seeing a unicorn walking down the street is not entirely unexpected. Unlike *this*_ref , *odin* does not succeed in introducing a completely novel and unexpected entity.

(i) a) A most amazing thing happened to me yesterday: I was walking down the street, and suddenly I saw this unicorn!

b) *Včera   proizošlo   porazitel'noesobytie:*
yesterday happened   amazing   event
ja   šla   po ulice   i   vdrug   uvidela   (#odnogo)   edinoroga!
I   walked-Pst on street and suddenly saw-Pst   one   unicorn-ACC

One might conclude from this that *odin*, unlike *this*_ref, carries a presupposition of non-uniqueness which makes it incompatible with definites, and which also makes it incompatible with indefinites which denote sets that cannot be presupposed to exist (such as unicorns) (see Chapter 2 for a review of Heim's (1991) argument that *a* in English does not presuppose non-uniqueness). However, it is quite felicitous to use *odin* in sentences like (iia), even though the existence of multiple terribly unpleasant neighbors of mine cannot be presupposed.

It is also perfectly felicitous to use *odin* in narrative contexts (such as fairy tales) in order to introduce a character for the first time. In this context, use of *odin* with 'unicorn' is perfectly fine, as shown by (iib) even though there is no presupposition that a set of happy unicorns must exist.

(ii) a) *Ko mnezašel   odin   moj strašno neprijatnyj   sosed.*
to   me came-over one   my terribly unpleasant   neighbor
"This terribly unpleasant neighbor of mine came over to see me."

b) *Žil-byl   odin   veselyj edinorog...*
lived-was   was   merry   unicorn-NOM
"There was once a happy unicorn..."

If *odin* marks specificity, one might hypothesize that Russian speakers would treat *the* as the English equivalent of *odin* and therefore mark it as [+specific]. However, there is no real motivation for this hypothesis. Since *odin* is used only with indefinites, there is no reason to expect a mapping between *odin* and the definite article *the*[13].

6.1.2. Other lexical marking on indefinites in Russian

Another reason not to suppose that *odin* is mapped to *the* is that *odin* is only one of many lexical modifiers on Russian indefinites.

Two lexical items in Russian, *kakoj-to* and *kakoj-nibud'*, correspond to *some* in English. Their distribution is as follows. In extensional contexts such as (30a), only *kakoj-to* is allowed: *kakoj-nibud'* is ruled out because it has to be licensed by a c-commanding operator. In intensional contexts, both forms of 'some' are possible, but while *kakoj-to* obligatorily takes wide scope (30b), *kakoj-nibud'* obligatorily takes narrow scope (30c). Finally, both forms may scope under a universal quantifier (30d); *kakoj-to* may also take scope over the universal quantifier, but this reading is dispreferred[14].

30. a) *Lena   pročla   kakuju-to / *kakuju-nibud'   knigu.*
      Lena   read-PST   some   book-ACC
      "Lena read some book."

  b) *Lena   xočet   pročest'   kakuju-to knigu*
      Lena   wants   read-INF   some   book-ACC
      "Lena wants to read some book."   √(some>want), *(want>some)

  c) *Lena   xočet   pročest'   kakuju-nibud' knigu.*
      Lena   wants   read-INF   some   book-ACC
      "Lena wants to read some book or other."   *(some>want), √(want>some)

  d) *Každaja   devočka   pročla   kakuju-to / kakuju-nibud' knigu.*
      Every   girl-NOM read-PST some   book-ACC
      "Every girl read some book."
      kakoj-to: ??(every>some), √(some>every)
      kakoj-nibud':   *(every>some), √(some>every)

Use of *odin* seems very similar to use of *this one* in English: one would not say "I saw this one unicorn on the street" in (ia) unless seeing unicorns is a fairly regular occurrence. On the other hand, it is quite felicitous to say *this one really unpleasant neighbor of mine* in (iia) and *There once lived this one unicorn...* in (iib). It is notable that both *odin* and *this one* are derived from the numeral *one*. For both *odin* and *this one*, specificity is a necessary condition but not a sufficient one; there also seems to be some constraint on familiarity – i.e., the speaker needs to be familiar with unicorns or unpleasant neighbors as a class. I leave the semantics of *odin* and *this one*, and the question of whether/how they can be derived from the meaning of the numeral *one*, to further research.

[13] One might hypothesize that Russian speakers would associate the specificity marker *odin* with *a*, and mark only specific indefinites with *a*, leaving out articles with non-specific indefinites. As the data reported in Chapters 4 through 6 show, this is not the case.

[14] A possible source for this dis-preference is that *každyj* 'every' prefers to have a distributive reading, similarly to the English *each*. (cf.: Each boy read some book).

Thus, *kakoj-nibud'* is a narrow-scope quantifier, while *kakoj-to* must be *de re* but may scope under another quantifier. Additionally, *kakoj-to* carries an implicature that the speaker does not know the identity of the individual denoted by the DP, as shown in (31).

31. a) *Lena   pročla   kakuju-to   knigu: #Vlastelin   Kolec.*
    Lena   read-PST   some   book-ACC #Lord   Rings-GEN
    "Lena read some book – #Lord of the Rings."

  b) *Lena   pročla   kakuju-to   knigu.   Ja   ne   znaju,   kakuju.*
    Lena   read-PST   some   book-ACC   I   NEG   know   which
    "Lena read some book. I don't know what it is."

Just as one may hypothesize that Russian speakers map the specificity marker *odin* to *the*, one may also hypothesize that Russian speakers would map *kakoj-to* and *kakoj-nibud'* to *the* and *a*, respectively, and hence use *the* whenever the DP has wide scope over an intensional operator. As the data in this thesis will show, that is clearly not the case. There is no *a priori* reason to believe that Russian speakers should map any lexical modifier (*odin*, *kakoj-to*, or *kakoj-nibud'*) to a particular English article.

6.1.3. Word order, definiteness, and specificity in Russian

Finally, definiteness and specificity interact with word order in Russian to some extent. As discussed much in the literature (see Bailyn 1995, Ch. 3, and the references cited therein), the preverbal position in Russian is associated with old information. Thus, if a bare DP subject such as *koška* 'cat' is placed preverbally (32a), it is interpreted as a definite. If the subject *koška* is placed postverbally, it is interpreted as an indefinite (32b). However, if the postverbal subject is given some modification, as in (32c), it can be interpreted as either definite or indefinite, depending on the discourse scenario[15].

32. a) *Koška   vbežala   v   komnatu.*
    cat-NOM   ran   into   room
    "The cat ran into the room."

  b) *V   komnatu   vbežala   koška.*
    in   room   ran   cat-NOM
    "Into the room ran a cat."

---

[15] It is also possible for bare indefinites to appear preverbally, in narrative contexts such as (i). Typically, however, speakers have a preference for putting bare indefinites postverbally.

(i) *Ja   vošla   v   biblioteku.   Tam   bylo   dovol'no   pusto.*
    I   entered   in   library-ACC there   was-NEUT   quite   empty
    Malen'kaja   devočka   čitala   knigu,   mužčina listal   gazetu,   a
    little   girl-NOM   read   book-ACC   man   leafed   newspaper-ACC and
    bol'še   nikogo   ne   bylo.
    more   nobody   NEG   was-NEUT

"I entered the library. It was quite empty there. A little girl was reading a book, a man was leafing through a newspaper, and there was nobody else there."

c) *V komnatu vbežala koška moego soseda.*
   in room   ran   cat-NOM my   neighbor-GEN
   "Into the room ran a/the cat of my neighbor."

A bare object DP in a postverbal position can be interpreted as either definite or indefinite, (33a), depending on the context. A bare DP may also have a possessive reading, as in (33c). Obligatorily definite DPs such as *the president* in (33d) also have no problem appearing postverbally.

33. a) *Maša čitaet knigu.*
     Mary reads   book-ACC
     "Mary is reading the/a book."

    b) *Maša poterjala šapku.*
     Mary lost   hat-ACC
     "Mary lost the/a/her hat."

    c) *Lena videla prezidenta.*
     Lena saw   president-ACC
     "Lena saw the president."

Indefinites which are overt partitives appear preverbally, as shown in (34).

34. *V komnatu vošli tri devočki. Odna iz devoček*
   in room-Loc entered-pl three girls-NOM one-NOM from girls-GEN
   nesla v rukax kotenka.
   carried-PST-FEM in hands kitten-ACC
   "Three girls entered the room. One of the girls was carrying a kitten."

And finally, indefinites modified by *odin* or *kakoj-to* may appear either preverbally or postverbally, as shown in (35) and (36) (compare to the bare DP in (37)).

35. a) *Odin činovnik pozvonil mne prjamo vo vremja obeda.*
     one   official-NOM called   me-DAT right in time   dinner-GEN
     "An/this official called me right during dinner."

    b) *Mne prjamo vo vremja obeda pozvonil odin činovnik.*
     me-DAT right in time dinner-GEN called   one   official-NOM

36. a) *Kakoj-to činovnik zvonil.*
     some   official-NOM called
     "Some official called."

    b) *Zvonil kakoj-to činovnik.*
     called some   official-NOM

37. a) *Činovnik pozvonil mne prjamo vo vremja obeda.*
     official-NOM called   me-DAT right in time   dinner-GEN
     "The official called me right during dinner."

b) *Mne   prjamovo vremja obeda   pozvonil   činovnik.*
   me-DAT   right in time dinner-GEN   called   official-NOM
   "An official called me right during dinner."

To sum up, there is no one-way relationship between Russian word order and either definiteness or specificity. The only generalization that can be made based on the above data is that indefinites typically cannot appear preverbally unless they bear lexical modification (but see footnote 15). As far as transfer is concerned, one might expect that L1-Russian speakers would associate the preverbal subject position in English with definiteness. In order to avoid transfer effects, in all of our elicitation tasks testing for definiteness / specificity, the target DP was always in object position: as seen in (32) and (33), object DPs are not associated with either definiteness or indefiniteness. (See Appendix 4 for an investigation of article use in the subject position in L2-English).

6.2. Definiteness and specificity in Korean

Like Russian, Korean has no articles. It has demonstratives, which may be used with previous-mention definites, as well as topic-marking. As illustrated in (38), the topic marker gives an indefinite a contrastive interpretation.

38. *John-i   chayk-un   ilk-ess-ta*
   John-NOM   book-TOP read-PST-DEC
   'John read a book (but not a magazine).'   (Heejeong Ko, p.c.)

The topic marker is incompatible with specific indefinites (unless there is a contrastive interpretation). For instance, suppose a customer goes into a "Lost and Found" department and utters (39); in this scenario, *a green scarf* is specific, since the speaker is singling out a particular green scarf from the set of scarves. This is in fact one of the contexts for specific indefinites used in our studies.

The Korean variant of (39) is (40a). There is a topic marker on the first-person subject, which is old information (alternatively, the subject may be dropped). It is not possible to put a topic marker on the specific indefinite, as shown in (40b), since the context does not allow a contrastive interpretation (i.e., the green scarf is not being contrasted with anything; the event of *looking for a green scarf* is also not being contrasted with anything. None of the specific indefinite contexts in our study allowed for a contrastive interpretation that would be compatible with topic-marking in Korean.

39. I am looking for a green scarf.

40. a) *(Ce-nun)   noksayk   scapu-lul   chackoiss-eyo.*
   (I-TOP)   green   scarf-ACC   looking.for-DEC.
   "I am looking for a green scarf."

   b)* *Ce-nun   noksayk   scapu-nun chackoiss-eyo.*
     I-TOP   green   scarf-TOP looking.for-DEC.   (Heejeong Ko, p.c.)

Thus, topic-marking in Korean does not correspond to specificity. Nor does it correspond to definiteness: while previous-mention definites may be topic-marked, definites which are not old information, such as *the winner of this race* in (41), do not receive topic-marking.

41. a) *Na-nun*   *ikyengki-uy*   *uwsungca-lul*   *chackoiss-eyo.*
    I-TOP   this. race-of   winner-ACC   looking-DEC.
    "I am looking for the winner of this race."

    b)* *Na-nun*   *ikyengki-uy*   *uwsungca-nun*   *chackoiss-eyo.*
    I-TOP   this. race-of   winner-TOP   looking-DEC.

(Heejeong Ko, p.c.)

For the details on topic-marking in Korean, see Choe 1995 and Kim 2000, among others. Crucially, topic-marking in Korean does not map onto either definiteness or specificity; nor does Korean appear to have any other marker that would correspond unambiguously to either the English *the* or to a specificity marker such as *this*$_{ref}$. Thus, any effects of definiteness or specificity in the L2-English of L1-Korean speakers is quite unlikely to be due to transfer. (But see Appendix 4 for possible transfer effects which are independent of definiteness or specificity, and that may be related to contrastive topic marking in Korean).

6.3. Summary

As the above discussion shows, there is no reason to expect Russian or Korean speakers to treat *the* as a specificity marker as a result of transfer. Neither L1 helps its speakers to decide which setting of the Article Choice Parameter to adopt for their L2, English. Thus, if L2-learners (optionally) associate *the* with specificity, as predicted in (4), this is indicative of direct access to the Article Choice Parameter rather than to transfer of the L1 setting.

7. Conclusion

This chapter investigated the predictions of the FH in the domain of article choice, by looking at previous studies in both L1 and L2 acquisition. Some data from both L1 and L2 studies are consistent with the FH: learners do in fact overuse *the* in specific indefinite contexts. (However, no data are available on non-specific definite contexts). This evidence is generally indirect, since both L1-studies like Matthewson and Schaeffer (2000) and L2-studies like Thomas (1989a) tested for the role of *assertion of existence* rather than specificity in acquisition. However, their results provide indirect evidence in favor of the FH (but there were some methodological problems in the case of the L1 studies, as discussed above).

The findings concerning overuse of *the* in partitive contexts, in both L1 and L2 acquisition, are not predicted under the FH. These results suggest that an additional process might be at work – an association of *the* with presuppositionality. However, more data are needed to distinguish the presuppositionality explanation, the specificity explanation, and, in the case of L1-acquisition, the egocentricity explanation. I will not address presuppositionality in the body of the thesis, but see Appendix 4 for a discussion of presuppositionality and possible transfer effects in L2-acquisition.

This chapter also showed that any effects of specificity found in the L2-English of speakers of Russian and Korean are highly unlikely to be due to L1-transfer.

Chapter 4: Experiment 1

This chapter describes the first of a series of three studies of article choice in L2-English. The main goal of this study was to determine whether overuse of *the* with indefinites in L2-English is related to scope (as suggested by previous L2-studies) or to specificity (as predicted by my proposal). The results reported in this chapter show that L2-English learners overuse *the* with specific indefinites, but not with non-specific indefinites, as predicted by the Fluctuation Hypothesis for L2-English article choice formulated in Chapter 3.

1. Background and predictions

Previous studies of L2-English article choice, reported in Chapter 2, found that L2-English learners sometimes overuse *the* in indefinite contexts. However, it is not clear what this overuse is due to. As discussed in Chapter 2, one possibility that has been considered in the literature is that *the* is associated with *assertion of existence* (see Thomas 1989a for an overview). On this view, L2-learners are predicted to overuse *the* with all indefinites that take widest scope with respect to intensional verbs or modals: such wide-scope indefinites assert the existence of at least one element in the set denoted by the restrictor NP. (A similar proposal is discussed in Matthewson and Schaeffer 2000, as well as Matthewson et al. 2001, with respect to L1-acquisition of English articles).

Another possibility, the one that I will argue for, is that L2-learners associate *the* with specificity. I proposed in Chapter 2 that definiteness and specificity are two settings of the Article Choice Parameter, repeated in (1). I argued that the Article Choice Parameter is discourse-related: it governs whether articles in a given language encode the state of speaker knowledge (specificity) or hearer knowledge (definiteness). While specific DPs necessarily assert existence (i.e., take wide scope over intensional/ modal operators), not all wide-scope DPs are specific. Scope, which is a grammatical distinction, does not enter into the settings of the Article Choice Parameter.

1. The Article Choice Parameter (for two-article languages):
   A language which has two articles distinguishes them as follows:
   Setting I.   Articles are distinguished on the basis of specificity.
   Setting II.   Articles are distinguished on the basis of definiteness.

I proposed that L2-English learners have access to the Article Choice Parameter, but have not determined that English is a Setting II language. Thus, errors of article use that we see in L2-English should stem from the learners treating English as a Setting I language. Whenever L2-learners treat English as a Setting I language, they should encode *the* as [+specific] rather than a [+definite].

Thus, overuse of *the* with indefinites (an error frequently cited in L2-literature) should be limited to indefinite contexts that are specific. I discuss this in more detail below. In this chapter, I ignore the specificity distinction in definites, which will be addressed in Chapter 6. I will look only at previous-mention definites, which, as discussed in Chapter 2, are obligatorily specific.

1.1. Hypothesis

My goal in this chapter will be to show that specificity and definiteness play a role in L2-English article choice, but that scope does not, contrary to previous claims in the L2-literature (see Thomas 1989a). I therefore argue against the hypothesis in (2), and advance the hypothesis in (3). In Section 4 of this chapter, I will refine the notion of "association with specificity" and show that L2-learners' use of indefinites follows from the FH. For now, I am interested simply in whether errors with indefinites in L2-English are related to wide scope or to specificity.

2. *Hypothesis of previous L2-literature (Thomas 1989a):*
   Overuse of *the* with indefinites in L2-English is due to an association of *the* with wide scope / assertion of existence[1].

3. *Hypothesis advanced here:*
   Overuse of *the* with indefinites in L2-English is due to an association of *the* with specificity.

In order to distinguish between the predictions of (2) and (3), it is necessary to test L2-learners on indefinite contexts that clearly distinguish wide scope from specificity.

1.2. Predictions

As discussed in Chapter 2, specific indefinites obligatorily take wide scope (or, rather, give the appearance of wide scope): if the speaker intends to refer to a particular individual, this individual must exist in the actual world. Non-specific indefinites, on the other hand, may take either wide or narrow scope, like any quantifier phrases.

This chapter will therefore look at L2-learners' performance in four types of contexts, shown in (4): previous mention definites (A); specific indefinites (B); wide-scope non-specific indefinites (C); and narrow-scope non-specific indefinites (D). Also listed in (4) are the predictions for L2-English article use made by the hypotheses in (2) and (3).

4. Predictions for article use in L2-English

context code	semantic property			article used in L1-English	article that L2-learners will use if...	
	definiteness	specificity	wide scope		they associate *the* with specificity (Hypothesis 3)	they associate *the* with wide scope (Hypothesis 2)
A	+	+	+	the	the	the
B	-	+	+	a	the	the
C	-	-	+	a	a	the
D	-	-	-	a	a	a

---

[1] For the purpose of this chapter, I use the term *wide scope* as shorthand for *widest scope over intensional and modal operators*. This form of *wide scope* necessarily entails *assertion of existence*. I ignore indefinites which take narrow-scope with respect to a higher quantifier, and assert existence. The proposal of Thomas (1989a), and the Bickertonian system of article classification more generally, does not say anything about indefinites that are in the scope of a higher quantifier.

2. Methods

We tested the predictions of the two competing models in (4) in a written elicitation study with adult L1-Russian speakers. Since Russian has no articles, no direct transfer effects were expected (see Chapter 3 for arguments on why L1-transfer from Russian cannot lead to an association of *the* with specificity).

2.1. Participants

There were 31 L2-learner participants in this study. The characteristics of the participants are summarized in (5). Individual information about the participants is given in Appendix 1. Individual participant numbers mentioned in this section correspond to the participant numbers in the first table in Appendix 1.

5. Characteristics of the L2-learners:

	L1-Russian participants
Number	31 (19 female, 12 male)
Age range	20 to 55 (mean = 39; median = 41)
Age of first exposure to English	5 to 48 (mean = 18; median = 12)
Age at arrival in the US (start of intensive exposure to English)	14 to 53 (mean = 35; median = 37)
Time in the US	0;2 to 23;0 (mean = 3;11, median = 2;2)

Most of the learners (27 out of 31) 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 adolescents or adults[2].

All of the participants were native Russian speakers, with the exception of #25, a native speaker of Ukrainian who had acquired Russian at age seven and considered Russian to be his primary language. Participants #8 and #16 were fluent in Ukrainian and Georgian, respectively, but both spoke Russian as their native language.

The study was piloted with three L1-English speakers, who used *the* in contexts where *the* was expected (context A in (4)), and *a* where *a* was expected (contexts B through D in (4)). Two of the native speakers made no errors at all. One native speaker made two errors, but neither was on the relevant test items discussed here. One of the errors involved using *the* with a plural NP that was supposed to be a bare plural; the other error was omitting the article before an NP in a generic context that is not discussed here.

2.2. Standard methodology for chapters 4 through 6

In this subsection, I list some methodological points that apply to all three empirical studies reported in this thesis: the studies reported in this chapter as well as the next two chapters.

---

[2] One participant arrived at age 14. The others all arrived as late adolescents or adults, at age 17 or older.

2.2.1. Standard task format

First, the main task in each of these studies was *written elicitation*: a task that required L2-learners to choose the appropriate article for a given context. The choice of this testing format was made so that the investigators could have control over the context types. This format allowed us to design contexts which differed from each other with regard to such factors as definiteness, specificity, and scope. A less controlled production task, such as picture description (a task frequently used in studies of L2 article choice, as shown in Chapter 3), would not allow us control over the contexts: it would not be possible to tell, for instance, whether a context was specific or wide-scope non-specific. The elicitation tasks in the thee studies had slight methodological differences, which will be described in the relevant chapters.

Second, the L2-learners were always provided with a vocabulary sheet containing lexical items from the elicitation task, with corresponding translations into the learners' L1. This ensured that the learners' responses were not influenced by lack of vocabulary knowledge. The vocabulary sheet contained those items that were most likely to be unfamiliar to the learners. Whenever the lexical item *certain* (which marks specificity) was used in a task, it was included on the vocabulary sheet. (See Chapter 8 for a discussion of a possible confound introduced by the Russian translation of *certain*).

Third, the test items for each elicitation task were presented in random order. For each elicitation task, two random orders of the test items were created, with roughly half of the participants receiving each order. In the first study, reported in this chapter, the two orders of the test were truly random. In the two later studies, the ordering was pseudo-random, as follows: the test was divided into two halves, each of which contained exactly half of the items from each test category. The order of presentation between the two halves of the test was then randomized across participants. The test was always accompanied by written instructions in the learners' L1 as well as example and/or practice items.

Finally, in addition to the elicitation task, the participants always took a standardized proficiency test of English as a Second Language, the Michigan Test. Only the written portion of the Michigan test, which consists of 30 multiple-choice items, was administered to the learners. This test tests various aspects of English grammar, and comes with a standardized scoring scale: L2-learners who get scores between 23 and 30 are considered advanced; those scoring between 13 and 22 are intermediate; and those scoring 12 and under are beginners. This test was used to gauge the L2-learners' overall levels of English. Whenever another task was used in addition to the elicitation task and the Michigan test, it is described in the relevant chapter.

2.2.2. Standard procedure

Testing for all three studies took place in a laboratory or classroom environment. Participants were tested singly or in small groups (the maximum number of participants tested in one session in any study was ten). The investigator asked each participant to fill out a short questionnaire which collected such information as age of first exposure to English, type and length of exposure, etc. Then the investigator proceeded to administer the tests to the participant(s).

The Michigan test was always administered last, after the other test(s). This was done so that the Michigan test's emphasis on grammar did not force the L2-learners

into the mode of thinking about grammatical rules: the major goal of the studies was to elicit the learners' intuitions about article choice, rather than test their knowledge of explicit rules.

Testing took place in a single session (for the last two studies, participants were offered the option of coming in for two shorter session, but no participant chose this option). Participants were reimbursed monetarily at the end of the testing session. Different groups of participants were recruited for each study.

2.3. Tasks

In the first study of L2-English article choice, the participants completed three written tests of English: an article elicitation test, a relative clause comprehension test, and the Michigan test of L2-proficiency described above. The tests were untimed, and the total testing time was typically about one and a half hours.

2.3.1. Elicitation task

The primary task in the study was an article elicitation test. The test consisted of 52 items, with each item containing a context and a target sentence. The context was in Russian. The target sentence was in English, and contained a blank corresponding to a missed article. Having the context in Russian had three motivations: first, it ensured that the L2-learners would comprehend the context fully; second, it decreased the total time of the test, which was already quite long, by allowing the learners to read the contexts quickly; and finally, it allowed us to put no articles in the context (since Russian has no articles), and thus avoid biasing the learners in any way.

Participants were instructed to read through each context carefully, to read the stimulus sentence, and to write the article on a separate answer sheet, next to the number corresponding to the item number on the test. They were given the choice of writing *a, the,* or a dash if no article was needed. They were told to pay attention to the Russian context, since the context could influence their choice of article. The context was always given as a dialogue between two people, *A* and *B*.

The 52 contexts in the elicitation test broke down into 14 context types, of which ten will be discussed here. The remaining four contexts were items testing use of generics, and the corresponding controls. These categories of items are irrelevant for the discussion of definiteness and specificity, and will not be discussed here.

Each of the ten relevant contexts contained four items. The examples of various contexts are given in English here, but were in Russian on the actual test (except for the target sentence). The target sentence, which was in English in the actual test, is highlighted in bold. I will now describe each context in turn, giving one example of each. The full list of stimuli can be found in Appendix 2A.

Most of the contexts that were set up to elicit specific indefinites involved relative clause (RC) modification, since Fodor and Sag pointed out that RC-modification biases a wide scope indefinite context in favor of specificity. However, we wanted to be certain that L2-learners do not simply use a strategy of "use *the* whenever you see a relative clause." We therefore also included a narrow-scope environment with RC-modification: this environment is obligatorily non-specific, and use of *the* is not predicted. We also included RC-modification with one of the definite contexts, for balance.

The ten contexts are summarized in (6), which also gives the example numbers illustrating each context. The contexts are described in detail below.

6. Context numbers and types

Context number	Description	Target article	Example	Use of *the* predicted by Hypothesis (2)?	Use of *the* predicted by Hypothesis (3)?
I	narrow scope	a	(14)	no	no
II	narrow scope, RC	a	(15)	no	no
III	no scope[3]	a	(12)	yes	no
IV	no scope, RC	a	(13)	yes	yes
V	wide scope	a	(9)	yes	yes
VI	wide scope, RC	a	(10)	yes	yes
VII	wide scope, certain	a	(11)	yes	yes
VIII	previous mention	the	(7)	yes	yes
IX	previous mention, RC	the	(8)	yes	yes
X	bare plural	∅	(16)	no	no

Environment A in (4), previous mention definites, was exemplified by two context types. Both were designed to elicit *the*. In the first, there was no modification of the definite; the context always mentioned two referents, e.g., *novel* and *magazine* in (7), one of which was mentioned again in the stimulus sentence. The second definite-eliciting context involved RC modification of the definite. The context mentioned two individuals from the same class which shared different properties: e.g., *healthy bird* and *sick bird* in (8). One of these individuals was mentioned in the stimulus sentence, with RC modification to distinguish it from the other one.

7. *previous-mention definite, no RC*
    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 _____ magazine first.**

8. *previous-mention definite, RC*
    A: Miranda bought two birds in the pet shop yesterday; one was healthy and one was a little bit sick.
    B: What did she do when she brought the birds home?
    A: **She gave some seeds to _____ bird that was sick.**

Three contexts were set up to give the indefinite wide scope over an operator (an intensional verb or a future modal); all three contexts were designed as specific indefinite – context B in (4).

The first of these contexts contained no RC modification (9); the second involved RC modification (10); and the third contained *certain*-indefinites with RC-modification (11). These contexts were set up so that use of *the* could never be felicitous for a native

---

[3] By "no scope", I mean no scope interactions between the indefinite and another scope-bearing operator – i.e., there are no intensional verbs, modals, etc., in the sentence.

English speaker[4]: for instance in (9), *the* is infelicitous because the hearer does not share knowledge of a particular red-haired girl. However, the speaker clearly possesses knowledge of a particular girl (and can say something noteworthy about her – e.g., *she is a good friend of mine* or *I'm supposed to find her here*) – so the context is specific.

9. *wide-scope indefinite, no RC: specific*
   In an airport, in a crowd of people who are meeting arriving passengers
   A: Excuse me, do you work here?
   B: Yes.
   A: In that case, perhaps you could help me.
   A: **I am trying to find ____ girl with long red hair.**

10. *wide-scope indefinite, RC: specific*
   A: I heard that Mary was sick. Has she found medical help?
   B: Yes, she has.
   B: **She is going to see ____ doctor who went to medical school at Harvard and now lives in Brookline.**

11. *wide-scope indefinite, RC + certain: specific*
   A: Nancy went to the museum yesterday. She didn't look around at all, but went directly to the room with 18th century Dutch art.
   B: Why?
   A: **Because she wanted to see ____ certain painting that she had heard a lot about.**

Two more contexts involved no intensional or modal operators, so the indefinite took wide scope by default. One of these contexts contained an unmodified indefinite (12), and the other – an indefinite with RC modification (13). We hypothesized that the unmodified indefinite (12) was more likely to be non-specific (context C in (4)), while the modified indefinite in (13) was more likely to be specific (context B in (4)), given Fodor and Sag's (1982) discussion of relative clauses. That is, the speaker in (13) is likely to intend to be talking about a particular picture, while the speaker in (12) is simply asserting that Nick is watching a cartoon; the identity of the cartoon is irrelevant.

12. *indefinite with no scope interactions, no RC: non-specific*
   A: Nick just went to the living room. Do you know what he is doing there?
   B: He is watching ____ cartoon on television.

13. *indefinite with no scope interactions, RC: specific*
   A: Alice just came to visit her nephew Andy.
   B: Did she bring him anything?
   A: Yes, she brought him ____ picture which shows some children playing games in their garden.

Two contexts were set up to provide a narrow-scope reading of the indefinite (context type D in (4)). One of these contexts involved no relative clause (RC)

---

modification (14), and the other had RC modification (15). Since the indefinite takes narrow scope in both cases, it is obligatorily non-specific regardless of whether RC-modification is present.

14. *narrow-scope indefinite, no RC: non-specific*
    In a children's library
    A: I'd like to get something to read, but I don't know what myself.
    B: Well, what are some of your interests? We have books on any subject.
    A: Well, I like all sorts of mechanic things. Cars, trains, airplanes... Yes! I've decided.
    A: I'd like to get ____ book about airplanes.

15. *narrow-scope indefinite, RC: non-specific*
    A: Roger was in a car accident last week - he hit somebody else's car. The police claim it was his fault, but Roger says he was innocent. He doesn't want to pay the fine.
    B: So what is he going to do?
    A: He is going to find ____ lawyer who can give him good advice.

Finally, one context tested bare plurals, to ensure that the learners knew that articles can be omitted with plural NPs. An example is given in (16).

16. *bare plural*
    A: I wonder what George was doing in the toy store yesterday?
    B: He was buying ____ dolls for his daughter.

2.3.2. *The RC comprehension test*

In addition to the elicitation task and the Michigan test, the participants completed a short (10-item) comprehension task, which tested the learners' knowledge of relative clauses. This test was motivated by the presence of items with relative clauses in the elicitation test.

This test was entirely in English (except for written instructions, which were in Russian). Each item contained a short (two or three sentence) story, and a corresponding question. Participants were instructed to answer the question with one or two words (although some participants ignored the instructions and answered with a full sentence, as they had probably been taught to do in their ESL classrooms).

Six of the test items contained a relative clause within the question, as exemplified in (17). In order to answer the question successfully, the learners had to understand that the second wh-word was not asking a question, but rather signaling RC modification. Thus, in (17), the correct answer would be "an old woman." Fuller answers, such as "She met an old woman" or "Anne met an old woman", were also acceptable. An example of an incorrect answer would be "Anne" – this would indicate that the L2-learner was treating the RC as a question, answering "who was wearing a blue dress?"

17. Two little girls, Mary and Anne, were walking outside. Mary was wearing a red dress, and Anne was wearing a blue dress. Mary met a little boy during her walk. Anne met an old woman.
    Who did the girl who was wearing a blue dress meet?

In order to pass the comprehension task, the participant had to answer at least four of the six test items appropriately. The four remaining items on the task were fillers, with simple comprehension questions not involving relative clauses.

3. Results

I examined the results from all 31 learners for all three tests. On the general proficiency test (the Michigan test), two learners placed as beginners, 13 learners were intermediate, and the majority of the learners, 16, were advanced[5].

The table in (18) reports the summary data (means and standard deviations) for all ten categories, using raw numbers. Only use of *the* and *a* is reported. Since each category contained four items, the maximum score on any category for use of *the* + use of *a* + article omission is 4. For instance, on Category I in (18), the mean score for use of *the* was 0.55 and the mean score for use of *a* was 3.26. Since 4-0.55-3.26=0.19, that means that article omission was, on average, at 0.19 out of 4 for this category. For each category, the mean correct and the corresponding standard deviation highlighted: thus, for the definite categories, the highlighted cells correspond to the use of *the,* and for the indefinite singular categories – to the use of *a.* Neither article is the target for bare plurals. From now on, I will use percentages rather than raw numbers in describing the results, for ease of reference.

18. Elicitation task results: summary of means and standard deviations (N=31)

CATEGORY	TARGET ARTICLE	TYPE (from (4))	use of *the*		use of *a*	
			mean	st. dev.	mean	st. dev.
I (narrow scope)	a	D	0.55	0.85	**3.26**	**0.96**
II (narrow scope, RC)	a	D	0.48	0.68	**3.19**	**0.91**
III (no scope)	a	C	0.74	0.89	**3.06**	**0.93**
IV (no scope, RC)	a	B	2.23	1.38	**1.68**	**1.30**
V (wide scope)	a	B	1.68	1.30	**2.10**	**1.33**
VI (wide scope, RC)	a	B	2.58	1.29	**1.35**	**1.23**
VII (wide scope, *certain*)	a	B	2.13	1.36	**1.39**	**1.41**
VIII (previous mention)	the	A	**3.03**	**1.33**	0.77	1.09
IX (previous mention, RC)	the	A	**3.29**	**0.94**	0.52	0.77
X (bare plural)	ø		0.61	0.95	0.39	0.92

3.1. RC-comprehension and article choice

I now turn to the results of the RC comprehension task, described in section 2.3.2. Recall that in order to pass this task, the learner had to answer at least four of the six RC-comprehension questions correctly. Twenty-seven of the 31 participants passed this task. The most common mistake for the four learners who did not pass was treating the RC as a question, giving an answer to the embedded wh-word. We concluded that these four learners had poor knowledge of the relative clause construction, which is crucial for the elicitation task. The four learners who failed the comprehension task

---

[5] One participant who placed as "advanced" (#14) was actually taking the Michigan test for the second time; he had taken it a year previously as part of a completely unrelated study. That time, he had placed as "intermediate."

(participants #16, #23, #30, and #38 – see Appendix 1) included one beginner and three intermediate learners.

The table in (19) gives the breakdown of article use as well as omission across all 31 L2-learners in percentages. The table in (20) gives the elicitation task results for the 27 L2-learners who passed the comprehension tasks. In both tables, the percentage corresponding to the target article for each category is highlighted.

19. Elicitation task results: %article use by context (N=31).

CATEGORY	TARGET ARTICLE	TYPE (from (4))	%the	%a	%omission
I (narrow scope)	a	D	13.71%	81.45%	4.84%
II (narrow scope, RC)	a	D	12.10%	79.84%	8.06%
III (no scope)	a	C	18.55%	76.61%	4.84%
IV (no scope, RC)	a	B	55.65%	41.94%	2.42%
V (wide scope)	a	B	41.94%	52.42%	5.65%
VI (wide scope, RC)	a	B	64.52%	33.87%	1.61%
VII (wide scope, certain)	a	B	53.23%	34.68%	12.10%
VIII (previous mention)	the	A	75.81%	19.35%	4.84%
IX (previous mention, RC)	the	A	82.26%	12.90%	4.84%
X (bare plural)	∅		15.32%	9.68%	75.00%

20. Elicitation task results for the L2-learners who passed the RC-comprehension task: %article use by context (N=27).

CATEGORY	TARGET ARTICLE	TYPE (from (4))	%the	%a	%missing
I (narrow scope)	a	D	12.04%	84.26%	3.70%
II (narrow scope, RC)	a	D	12.04%	82.41%	5.56%
III (no scope)	a	C	18.52%	76.85%	4.63%
IV (no scope, RC)	a	B	52.78%	44.44%	2.78%
V (wide scope)	a	B	44.44%	52.78%	2.78%
VI (wide scope, RC)	a	B	62.96%	35.19%	1.85%
VII (wide scope, certain)	a	B	54.63%	34.26%	11.11%
VIII (previous mention)	the	A	83.33%	12.96%	3.70%
IX (previous mention, RC)	the	A	86.11%	11.11%	2.78%
X (bare plural)	∅		14.81%	5.56%	79.63%

For the rest of the paper, I will concentrate on the data of the 27 learners who passed the comprehension task; however, it should be noted that the numbers in (19) and (20) show very similar patterns. The next section considers performance in each context in detail.

3.2. Performance in individual contexts

In the three contexts that were set up as non-specific (i.e., types C and D in (4)), learners use *a* the vast majority of the time, with only slight overuse of *the*. In the four specific indefinite categories (B), which require *a* in L1-English, the L2-learners overuse *the* around 50% of the time. In the two categories where *the* is required in L1-English (A), the learners appropriately use *the* most of the time; overuse of *a* is fairly low. And in the bare plurals category, the learners appropriately use no article at all more than 70% of the time. On the other hand, article omission with singular NPs is extremely rare, under 5% for all categories with the exception of the category containing *certain*.

Thus, the L2-learners exhibit knowledge of three properties of English articles: first, they know that null determiners are allowed with plural but not with singular NPs; second, they know that *the* is required in previous-mention definite contexts; and third, they know that *a* is required for narrow-scope DPs (type D). However, the situation is more complex for wide scope DPs, categories III through VII. The learners clearly exhibit less overuse of *the* on environment III, which we considered non-specific (type C), than in any of the environments IV through VII, which are more likely to be specific (type B).

Given the similarity of performance on contexts of types C and D (environments I through III), I will now refer to them both as non-specific, without distinguishing the wide and narrow scope cases. I will refer to environments IV through VII (type B) as specific.

3.3. Statistical comparisons

I turn next to the question of whether the differences between categories are statistically significant. All of the statistical tests reported here were performed on use of *the* by category. Use of *a* was grouped together with article omission, since the latter was (for singular DPs) extremely low.

The tables in (21) and (22) report the results of t-tests (paired two-sample for means) on the use of *the* in each category: the two tables report the results of subjects analysis and items analysis, respectively. Each cell in each table reports the p-value corresponding to the statistical comparison between two categories. The categories are listed in the leftmost column as well as the topmost row: the cell corresponding to the intersection of two category numbers is the cell reporting the p-value for the comparison between these categories. Note that category X (bare plurals) is not included in this table, since it was the only category requiring neither *a* nor *the*.

First, consider the regions in italics (but not in bold) in both tables. These are the results of statistical comparisons between two categories of the same type: i.e., two categories eliciting *a* in non-specific contexts, or two categories eliciting *a* in specific contexts, or two categories eliciting *the*. Nearly all of these "same-type" comparisons failed to yield a significant difference; the one exception is a significant difference between categories V (wide scope, no RC) and VI (wide scope, RC modification) on the subjects analysis (this comparison also yields a marginally significant difference on the items analysis). Note that both specific indefinite categories that involved RC modification, IV and VI, resulted in numerically larger uses of *the* than did category V, with no RC modification. This suggests that RC-modification contributes to a context

being interpreted as specific, and hence to overuse of *the*. However, note that the L2-learners *cannot* be using a simple strategy of "use *the* when you see a relative clause": use of *the* with narrow scope (obligatorily non-specific) RC-modified NPs (category II) is extremely low, as low as use of *the* with narrow-scope indefinites which have no modification (category I).

21. Subjects analysis on use of *the*: two-tailed p-values (N=27).

category	II	III	IV	V	VI	VII	VIII	IX
I	.86	.13	<.0001	<.001	<.0001	<.0001	*<.0001*	*<.0001*
II		.34	<.0001	<.0001	<.0001	<.0001	*<.0001*	*<.0001*
III			<.001	<.01	<.0001	<.001	*<.0001*	*<.0001*
IV				.14	.13	.82	*<.001*	*<.0001*
V					<.01	.21	*<.0001*	*<.0001*
VI						.27	*<.01*	*<.01*
VII							*<.01*	*<.001*
VIII								*.50*

22. Items analysis on use of *the*: two-tailed p-values (4 items per category)

category	II	III	IV	V	VI	VII	VIII	IX
I	.76	.47	<.001	<.05	<.01	<.001	*<.001*	*<.01*
II		.43	<.01	<.05	<.001	<.01	*<.001*	*<.001*
III			<.05	<.05	<.01	<.05	*<.01*	*<.01*
IV				.20	.06	.82	*<.01*	*<.01*
V					.07	.41	*<.05*	*<.001*
VI						.38	*.06*	*<.05*
VII							*<.01*	*<.05*
VIII								*.59*

I turn next to the more exciting question of whether performance on categories of different types (non-specific vs. specific vs. definite) results in significant differences. The answer is clearly yes. The relevant region for these comparisons is in bold in both tables: regular bold for comparisons between specific and non-specific indefinite contexts, and bold-italic for comparisons between definite and all other contexts. All of the comparisons are highly significant. Given the percentages in (20) and the p-values in (21) and (22), we can summarize the results as follows:

23. a) use of *the* is significantly higher in each context where *the* is required than in each context where *a* is required; and
   b) use of *the* is significantly higher in each specific context requiring *a* than in each non-specific context requiring *a*.

3.4. Individual patterns of article choice

A final question to address is what individual results look like. This is particularly relevant for specific indefinites, since all categories of specific indefinites show nearly 50/50 use of *the* vs. *a*. This pattern could potentially come from two sources: (1)

roughly half of the learners use only *the* in specific indefinite contexts, while roughly half use only *a* in specific indefinite contexts; or (2) some, or even most, of the learners use *a* and *the* interchangeably in specific indefinite contexts.

In order to determine whether (1) or (2) is correct, I report individual results. Figure 1 shows individual article use in four specific indefinite contexts (contexts of type B, i.e., contexts IV through VII) while Figure 2 shows individual article use in the three non-specific indefinite contexts (contexts of types C and D, i.e., contexts I through III). The y-axis in each figure corresponds to the number of items (16 total for the four specific categories, 12 total for the three non-specific ones). The x-axis corresponds to the 27 individual participant results. Individual results are arranged by increasing overall proficiency score – see section 3.5 for more discussion of overall proficiency.

Figure 1: Article use in specific contexts

Figure 2: Article use in non-specific contexts
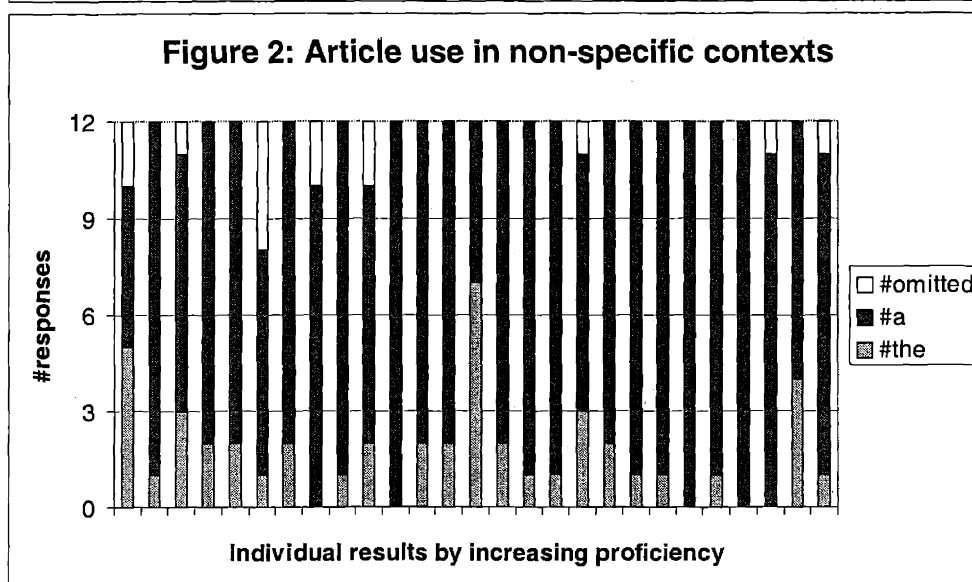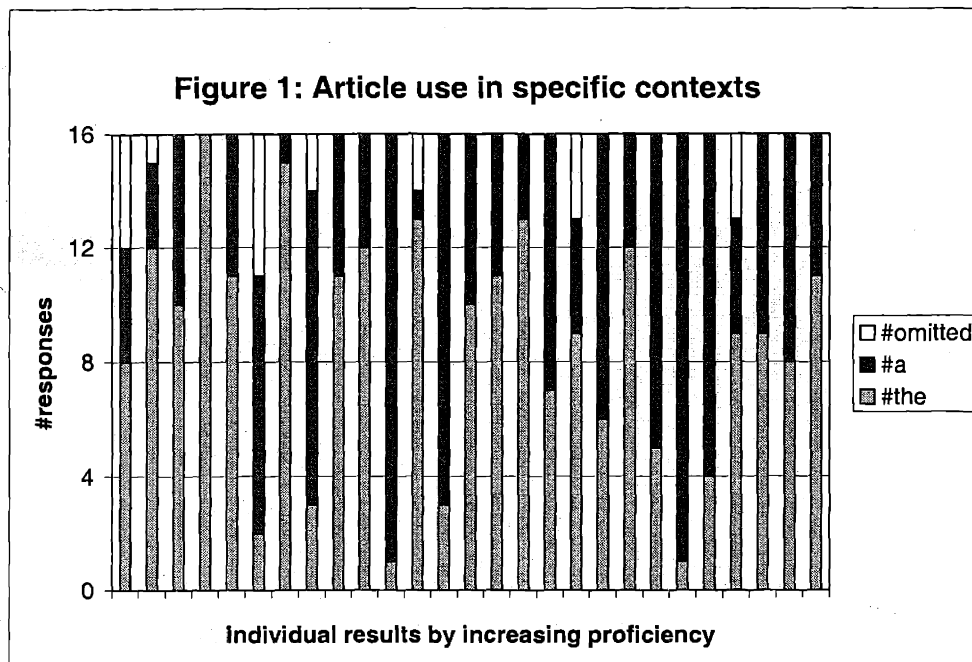
As Figure 1 shows, most of the L2-learners used both *a* and *the* in specific indefinite contexts; while a few used almost exclusively *a,* and a few used almost exclusively *the,* this was not the overall pattern. Figure 2 shows that in non-specific contexts, most subjects used *a* most of the time.

3.5. Overall proficiency and article use

I have so far talked of all the L2-learners together, ignoring individual proficiency in English. Recall that the Michigan test allowed us to gauge learners' overall proficiency in English, as well as break learners into groups of beginner, intermediate, and advanced.

A natural hypothesis would be that the more advanced a learner's overall knowledge of English, the more likely that learner is to use English articles appropriately. To test this hypothesis, we computed correlations between scores on the Michigan test and scores in each of the categories on the elicitation task. For the elicitation task score, we used the proportion of time that the *appropriate* article was used in each context: i.e., for each of the contexts I through VII, we computed the correlation between the Michigan test score and the proportion of *a* use; for contexts VIII and IX, it was the proportion of *the* use; and for context X (bare plurals), it was the proportion of null article use. The resulting coefficients are reported in (24). As can be seen from the last column, the correlations were significant in only two of the categories. Overall proficiency thus does not appear to be a good predictor of performance on the elicitation task.

24. Correlations between Michigan test scores and each article-eliciting category (N=27)

Category	Target Article	%suppliance of target article	Pearson correlation coefficient with Michigan test score	two-tailed p
I	a	84.26%	0.50	p < .01
II	a	82.41%	0.30	p = .13
III	a	76.85%	0.15	p = .46
IV	a	44.44%	0.15	p = .46
V	a	52.78%	0.32	p = .10
VI	a	35.19%	0.24	p = .22
VII	a	34.26%	0.29	p = .14
VIII	the	83.33%	0.21	p = .28
IX	the	86.11%	0.09	p = .65
X	ø	79.63%	0.55	p < .01

In order to further explore the role of proficiency in L2-English article choice, I compared the performance of intermediate L2-learners (N=10) and advanced L2-learners (N=16) (the single beginner was excluded). I ran independent two-sample t-tests for means on both uses of *the* and uses of *a* for all of the categories (with the exception of bare plurals, where neither article is appropriate). There were no significant differences between the intermediate and advanced L2-learners on either use

of *the* or use of *a* on any of the categories. While advanced L2-learners were slightly more accurate on most categories than intermediate L2-learners, this difference was not significant for any category.

To sum up, overall proficiency in English (at least insofar as it can be measured by the Michigan test) seems to have little effect on article use. Overuse of *the* in specific indefinite contexts is a mistake that persists even for learners who have mastered many other aspects of English grammar (as measured by the proficiency test). The question is what causes this persistent error. I turn to this question next.

4. Discussion

We can now come back to the table in (4), and examine the predictions made by two different models of L2-English article choice. The predictions are repeated in (25), with results from the L2-learners incorporated.

25. Predictions for article use in L2-English .

context code	semantic property			article used in L1-English	article that L2-learners will use if...		article most often used by L2-learners in this study
	definite-ness	specifi-city	wide scope		they associate *the* with specificity (3)	they associate *the* wide scope (2)	
A	+	+	+	the	the	the	the
B	-	+	+	a	the	the	the / a
C	-	-	+	a	a	the	a
D	-	-	-	a	a	a	a

4.1. Article choice and fluctuation

The results show that the L2-learners clearly do not associate *the* with wide scope: they show very little overuse of *the* with non-specific indefinites, regardless of whether these are wide or narrow scope. Second, the learners are clearly overusing *the* with specific indefinites. Overuse of *the* is high in context B (specific indefinites), but it is still lower than appropriate use of *the* in context A (previous-mention definites).

The pattern is expected under the Fluctuation Hypothesis for L2-acquisition. I argued in Chapter 1 that in the absence of transfer, L2-learners fluctuate between the possible parameter settings. In the case of the Article Choice Parameter, repeated in (26), L2-learners should fluctuate between dividing English articles on the basis of definiteness, and dividing them on the basis of specificity. The error pattern of the L2-learners in our study is explained under the proposal that the learners access *both* settings of the Article Choice Parameter in (26)[6].

26. The Article Choice Parameter (for two-article languages):
A language which has two articles distinguishes them as follows:
Setting I. Articles are distinguished on the basis of specificity.
Setting II. Articles are distinguished on the basis of definiteness.

---

[6] Thanks to Hagit Borer (p.c.) for initially suggesting this line of thought to me.

Consider the pictorial representation of the Article Choice Parameter in (27), in which the change in shading pattern corresponds to the morphological distinction in articles[7]. Thus, the column marked "Setting II" is exemplified by English (formal English, with no $this_{ref}$): all indefinites are marked by *a*, and all definites by *the*. The "Setting I" column may be exemplified by Samoan, in which the article *se* corresponds to non-specific indefinites, and the article *le* – to both specific indefinites and definites.

27. Article use crosslinguistically: possible UG options

DP type	Setting I: Distinction by specificity (Samoan)	Setting II: Distinction by definiteness (English)
Non-specific indefinites		
Specific indefinites		
Previous-mention definites		

Now consider what happens with an L2-learner who is fluctuating between the two parameter settings. Both parameter settings tell the learner that non-specific indefinites and definites must be marked by different articles: in both the Setting I and Setting II columns in (27), the cells corresponding to non-specific indefinites and to definites receive different shading. Thus, the learner will be able to assign one article (*a*) to non-specific indefinites and another (*the*) to definites. (Presumably, there is sufficient positive evidence for the learner not to perform the assignment of *a* and *the* backwards). There is no conflict between the two UG-provided options.

However, in the case of specific indefinites, there *is* a conflict. Setting II tells the learner to assimilate specific indefinites to non-specific indefinites, using *a* for both. Setting I tells the learner to assimilate specific indefinites to definites, using *the* for both. I proposed in Chapter 3 that this is a conflict that the learner cannot easily resolve – i.e., L2-learners have difficulty generalizing from the input triggers. This conflict gives rise to interchangeable use of *a* and *the* in specific indefinite contexts.

4.2. Triggers and the Article Choice parameter

In Chapter 3, I proposed that the Article Choice Parameter is difficult to set. The study reported here shows that even learners who scored very high on the Michigan test, thus showing knowledge of fairly subtle points of English grammar, nevertheless have not set the Article Choice Parameter. We saw in Section 3.5 that there was no significant correlation between overall proficiency and use of the target article for most categories, and that advanced learners did not differ significantly from intermediate learners. This is in sharp contrast to our earlier findings that in the syntactic domain of verb-raising, proficiency correlates significantly with accuracy of adverb-verb placement (r=.40) (see Ionin and Wexler 2002). While there was evidence for fluctuation in the domain of verb raising, it was not as persistent as fluctuation in the domain of article choice. This provides indirect support to my proposal (discussed in Chapter 3) that L2-learners have more difficulty generalizing from discourse-based triggers than from syntactic triggers (for more direct evidence, the same group of L2-learners would need to be tested on both syntactic and discourse-related knowledge).

---

[7] This is a simplified representation, which treats all definites as specific and ignores non-specific definites, which were not tested in this study. See Chapter 6 for modification to this representation.

5. Conclusion and remaining questions

In this chapter, I showed that L2-English article errors are not random. L2-English learners are fairly accurate at article use in previous-mention definite contexts as well as narrow scope indefinite contexts. The difficulty lies in wide-scope indefinites. I showed that wide scope alone is not, however, sufficient to account for the patterns of article misuse in L2-English: L2-learners overuse *the* in wide-scope indefinite contexts only when the context suggests the speaker's intent to refer to a particular individual in the set denoted by the NP. This distinction is easily explained under the view of specificity advocated in Chapter 2: L2-English learners overuse *the* with specific indefinites, but not with non-specific (either wide or narrow scope) indefinites. I showed that the overuse of *the* with specific indefinites, coupled with target-like performance in other contexts, is predicted under the Fluctuation Hypothesis advanced in Chapter 1: L2-learners have access to both settings of the Article Choice Parameter, and fluctuate between them.

A number issues arise from the study described in this chapter, which will be addressed in subsequent chapters. They are discussed below.

5.1. A greater variety of contexts

The first issue concerns the generalization of the findings of this study to other indefinites. Most of the specific indefinite contexts in this study involved RC-modification; we want to be certain that other types of specific indefinites similarly elicit use of *the*. Chapter 5 shows that that is indeed the case.

On a similar note, it is necessary to investigate a wider variety of non-specific wide-scope contexts. The study reported in this chapter contained only one category of non-specific wide-scope indefinites. Chapter 5 investigates article use in a vaster variety of non-specific wide-scope contexts and shows that the generalization that *the* is overused only with specific indefinites holds.

Finally, this chapter did not address article choice with specific vs. non-specific definites in L2-English. This issue is investigated in Chapter 6.

5.2. An alternative explanation

Then, there is the issue of an alternative explanation of this study's results[8]. The Fluctuation Hypothesis is not the only explanation for why use of *the* vs. *a* with specific indefinites is optional. Another possibility is that L2-learners in fact obligatorily mis-set the Article Choice Parameter to Setting I for English: they always mark *the* as [+specific]. The reason that they do not always use *the* with specific indefinites is that they do not always construe the specific indefinite contexts as specific. By definition, specificity is in the mind of the speaker: a speaker will use a specific indefinite if she wants to draw attention to an individual who possesses some noteworthy property. In evaluating the dialogues in the forced choice elicitation task, the L2-learners may sometimes decide that the speaker is using a specific indefinite (and thus fill in *the*) and sometimes decide that the speaker is using a non-specific indefinite (and thus fill in *a*). The same context – e.g., a wide-scope context with RC-modification – may be

---

[8] For a more comprehensive list of alternative explanations, and arguments against them, see Chapter 8.

construed as specific by one speaker and non-specific by another. Under the view of specificity that I have been advocating, there is no such thing as an indefinite context that *must* be interpreted as specific.

This alternative explanation cannot be ruled out on the basis of the present study alone. However, data from two subsequent studies will show this alternative explanation to be incorrect – see Chapter 8 for more discussion, as well as for arguments against other alternative explanations of the data.

5.3. The question of transfer

I have argued that the results of this study stem from a general property of L2-acquisition: when L2-learners do not have articles in their L1, they have equal access to both settings of the Article Choice Parameter when using articles in their L2. In order to truly make this claim, I need to show that the results generalize to a population of L2-learners from a completely different (but also article-less) L1. In the absence of such a generalization, there is always a possibility that some form of transfer accounts for the results. For instance, I discussed in Chapter 3 the presence of the specificity marker *odin* in Russian; it is possible, although unlikely (as discussed in Chapter 2) that L2-learners would treat *the* as the English equivalent of *odin*. In order to fully rule out transfer, however, it is necessary to look at domains where *odin*-transfer is inapplicable (i.e., definites – see Chapter 6) and, secondly, to look at different L1 groups. This is done in Chapter 5.

Chapter 5: Experiment 2

This chapter describes the second elicitation study of article choice in L2-English. This study expands upon the study reported in Chapter 4, and provides further evidence for the role of specificity in article choice. This study examines a wider variety of contexts, including singular as well as plural contexts, and shows that the Fluctuation Hypothesis for L2-English article choice holds across these contexts.

This study also compares choice in the L2-English of two groups of learners, L1-Russian speakers and L1-Korean speakers. The similar patterns of performance among the two groups provide evidence that the results are not due to L1-transfer.

1. Background and predictions

The study reported in Chapter 4 showed that L2-English learners exhibit non-random patterns of article use: they use *the* correctly with definites, correctly use *a* correctly with non-specific indefinites, and alternate between using *a* and *the* with specific indefinites. These findings supported the hypothesis that L2-English learners fluctuate between the two settings of the Article Choice Parameter, repeated in (1). The hypothesis is repeated in (2).

1. The Article Choice Parameter (for two-article languages):
   A language which has two articles distinguishes them as follows:
   Setting I.   Articles are distinguished on the basis of specificity.
   Setting II. Articles are distinguished on the basis of definiteness.

2. **The Fluctuation Hypothesis (FH) for L2-English article choice:**
   1) L2-learners have full UG access to the two settings of the Article Choice
      Parameter in (1).
   2) L2-learners fluctuate between the two settings of the Article Choice Parameter
      until the input leads them to set this parameter to the appropriate value.

The proposal in (2) makes explicit predictions for article use and misuse in L2-English. When L2-English learners adopt Setting I in (1), they use *the* with (previous mention) definites[1] and with specific indefinites, and *a* with non-specific indefinites. When they adopt Setting II in (1), they use *the* with definites, and *a* with all indefinites (specific and non-specific). Under both settings, definites get *the* and non-specific indefinites get *a*. However, the two settings are in conflict with respect to specific indefinites, so interchangeable use of *the* and *a* is predicted with specific indefinites. This is summarized pictorially in (3). The predictions of (3) were in fact supported by the data reported in Chapter 4.

3. Article use cross-linguistically: predictions for L2-English

DP type	Setting I: Distinction by specificity	Setting II: Distinction by definiteness		L2-English: fluctuation
Non-specific indefinites				
Specific indefinites				
Definites				

---
[1] This chapter continues to ignore non-specific definites, which are addressed in Chapter 6. The representation in (3) is a simplified one, since it does not take non-specific definites into account.

1.1. Predictions for L2-English article choice

In this chapter, which reports on our second study of articles in L2-English, I will provide additional evidence in favor of the proposal in (2). The main goal of this chapter is to provide further support to the proposal that L2-English learners fluctuate between the two settings of the Article Choice Parameter. The specific predictions are given in (4). Prediction 1 simply states that the pattern of article use found in the previous study for a limited number of contexts should also be seen when a greater variety of contexts is tested. Prediction 2 is concerned with generalizing the results obtained for singular contexts to plural contexts: the FH in (3) does not predict any differences between singular and plural contexts. Finally, Prediction 3 is concerned with the universality of the FH. In order to show that the results obtained in the previous study are not due to transfer from the learners' L1, Russian, but rather to universal access to the Article Choice Parameter in L2-acquisition, it is necessary to test L2-learners from different L1-backgrounds. The present study therefore examines L2-English article use among two distinct groups of L2-learners: L1-Russian speakers and L1-Korean speakers.

4. Prediction 1: The pattern of *the* overuse with specific indefinites (vs. correct article use with definites and non-specific indefinites) will be seen across a wider variety of singular contexts.

Prediction 2: The pattern of *the* overuse with specific indefinites (vs. correct article use with definites and non-specific indefinites) will be seen in plural contexts.

Prediction 3: Distinct groups of L2-learners whose L1's lack articles will show similar patterns of article use.

1.2. Predictions for individual results

Finally, the present study, which tested a much larger sample of participants than the previous study, allows me to investigate individual performance. As discussed in Chapter 3, the Article Choice Parameter as formulated in (1) allows for three types of patterns among individual learners, given in (5). The FH predicts that individual L2-English learners should either be in a state of fluctuation, exhibiting the pattern in (5b), or should succeed in setting the Article Choice Parameter, exhibiting the pattern in (5a). The FH predicts that L2-English learners should not exhibit the pattern in (5c), since the input should lead them to abandon fluctuation in favor of the definiteness setting rather than the specificity pattern. In this chapter, I will examine the results of individual L2-learners to see whether the predictions of the FH are supported at the individual level.

5. a) The Definiteness Pattern (Setting II of the Article Choice Parameter)
   L2-English learners correctly use *the* and *a* to mark [+definite] and [-definite] contexts, respectively.

   b) The Fluctuation Pattern (Fluctuation between settings)
   L2-English learners go back and forth between distinguishing *the* and *a* on the basis of definiteness, and distinguishing them on the basis of specificity.

   c) The Specificity Pattern (Setting I of the Article Choice Parameter)
   L2-English learners use *the* and *a* to mark [+specific] and [-specific] contexts, respectively.

2. Methods

We tested the above predictions in a study with L1-Russian and L1-Korean adult learners of English[2].

2.1. Participants

The participants in this study were 88 adult L2-English learners: 50 L1-Russian speakers and 38 L1-Korean speakers. The results of three additional L1-Russian participants were excluded from the final analysis: two did not meet the age requirement (were past 60 years of age)[3], and one failed to complete the Michigan test.

The characteristics of the two groups of participants are given in (6). Characteristics of individual participants are given in Appendix 1. The participant numbers given in this section correspond to the participant numbers in the second table in Appendix 1. The numbers of L1-Russian participants start with an "r", and the numbers of the L1-Korean participants start with a "k".

6. Characteristics of the L2-learners:

	L1-Russian participants	L1-Korean participants
Number	50 (27 female, 23 male)	38 (18 female, 20 male)
Age range	17;11 to 56;9 (mean = 38; median = 42)	17;1 to 38;7 (mean = 28; median = 29)
Age of first exposure to English	6 to 50 (mean = 13; median = 11)	9 to 20 (mean = 12; median = 13)
Age at arrival in the US (start of intensive exposure to English)	16;11 to 55;11 (mean = 36; median = 37)	15;8 to 36;7 (mean = 26, median = 27)
Time in the US	<0;1 to 13;8 (mean = 3;2, median = 1;0)	<0;1 to 7;1 (mean = 1;10, median = 1;4)

Most of the learners (45 of the L1-Russian speakers and all 38 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.

All of the Russian-speaking participants spoke Russian as their primary language but some were also native in other languages: four participants were native in Ukrainian, and one was native in Judeo-Tat (a language of the Jews of the Caucasus Mountains)[4]; all had spoken Russian as their primary language since early childhood. Other speakers were fluent (but non-native) in languages other than Russian. The full details are described in the "special notes" section in the second table in Appendix 1. All of the Korean-speaking participants spoke Korean as their first and primary language.

---

[2] The studies described in this and the following chapters were conducted in collaboration with Heejeong Ko. Heejeong Ko collected the L1-Korean speakers' data.

[3] Following Eubank et al. (1997), we set the age limit for participation in our studies at 60 in order to avoid the possible confound of aging effects on L2-acquisition.

[4] The participant called his language simply "Jewish". Judeo-Tat is considered to be the official name of this language – see http://www.ethnologue.com/show_language.asp?code=TAT

The L1-Russian participants included immigrants, foreign workers who are here on visas, and spouses or children of 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 and through ESL classes. The L1-Korean participants were primarily international students and their spouses, as well as a few foreign workers and their spouses. Most of the participants (#k8 through #k38) resided in Gainesville, FL at the time of the study and were recruited through the University of Florida. The other seven participants resided in the greater Boston area and were recruited at schools for international students as well as through advertisement in the Korean community.

This study was piloted with seven adult L1-English speakers, who performed as expected, supplying the target article in all of the contexts enumerated below[5].

2.2. Tasks

The participants completed three tasks: a forced-choice elicitation task, a written translation task, and the Michigan test of L2-proficiency. This chapter focuses on the results of the elicitation task. See Chapter 4 for a discussion of the Michigan test and of the standard methodological and procedural points used in our studies.

The translation task was designed to test L2-learners' implicit knowledge of English articles: unlike the elicitation task, the translation task did not reveal to the learners that they were being tested on articles. However, this task encountered a number of methodological problems. This task, and the corresponding results, are discussed in Appendix 3 and are not given here. See Chapter 7 for a discussion of a more successful methodology (written production data) used for testing L2-learners' implicit knowledge of English articles.

2.2.1. Task design for the forced-choice elicitation task

The main task in this study was a forced choice elicitation test consisting of 56 short dialogues. The format of this task was largely the same as that in the previous study, reported in Chapter 4, since this format proved to work well in eliciting articles in a variety of contexts.

As in the previous study, each dialogue in the test was in the learner's L1 (Russian or Korean), with only the target sentence and any sentence(s) following it given in English. The participants' task was to choose the appropriate article for the target sentence in each dialogue.

The main change from the previous study was that this time, instead of filling in the blank, the learner was given three options of possible articles and had to choose between them. The learner had to choose between *a, the,* and the null article (--) for singulars, and between *some, the,* and -- for plurals. As the previous study found, the learners almost never allowed article omission in singular contexts (the only context type tested in the previous study). We were concerned that the learners felt that they had to use an article because they had to fill in a blank "with something". In order to avoid biasing the learners into supplying an article, we provided "no article" as an explicit alternative to *a* or *the* in the current study.

2.2.2. Item types: summary

There were 14 contexts types of four items each in the task. Of these, 10 context types are relevant to the discussion of specificity and are discussed here. The four remaining types are not

---

[5] An additional participant was excluded from the control group because he was bilingual in Spanish and English, and it was not clear whether English was his primary language. This participant made three errors, in different contexts.

directly concerned with definiteness or specificity, but rather, with word order and partitivity. These types, and the corresponding results, are reported in Appendix 4.

The table in (7) summarizes the 10 relevant context types and lays out the predictions for L2-English article use. Note that all of the specific indefinite categories are predicted to take both *the* and *a* (or, in the case of plurals, both *the* and *some / nothing*), under the hypothesis that L2-English learners fluctuate between the two settings of the Article Choice Parameter.

7. Summary of context types and predictions (sg=singular, pl=plural)

context type	example number	article used in L1-English	article predicted for L2-English under (2)
sg specific wide scope indefinite	(8)	*a*	*the / a*
sg specific indefinite with *certain*	(9)	*a*	*the / a*
sg specific indefinite, no scope[6]	(10)	*a*	*the / a*
sg non-specific narrow scope indefinite	(11)	*a*	*a*
sg non-specific indefinite, no scope	(12)	*a*	*a*
sg non-specific indefinite, denial of knowledge	(13)	*a*	*a*
pl specific wide scope indefinite	(14)	*some or --*	*the / some or --*
pl non-specific narrow scope indefinite	(15)	*some or --*	*some or --*
sg definite	(16)	*the*	*the*
pl definite	(17)	*the*	*the*

The ten relevant contexts are exemplified in (8) through (17), one example per context type. The full list of items can be found in Appendix 2B. The contexts are given in English here but were in the learners' L1 in the actual test. The sentences which were in English on the actual test are highlighted in bold, and the target article for each context is underlined (for indefinite plurals, both *some* and '--' are underlined, since both are in principle acceptable). The contexts are all dialogues between two people. The target DP is always in object position.

2.2.3. Singular specific indefinite contexts

Three of the context types aimed to elicit singular specific indefinites. The context type exemplified in (8) was designed to give the indefinite wide scope over an operator; speaker knowledge of the referent was explicitly stated, making it likely that the speaker intends to refer to a particular individual (in this case, a particular scarf). The context type in (9) involved wide scope as well as use of *certain,* a marker of specificity (see Chapter 2). The context type in (10) involved no scope interactions, but explicitly stated speaker knowledge of the referent. This variety of contexts was used to ensure that overuse of *the* was not tied to any one particular context type. Unlike the previous study, this study contained no items with RC-modification, in order to make sure that overuse of *the* was not a reflex of the presence of RC-modification.

All of the specific indefinite contexts are contexts which are compatible with use of referential *this* in L1-English (except the context in (9), since *this* and *certain* cannot be used together in English).

---

[6] The shorthand "no scope" means no scope interactions with an intensional or modal operator; the indefinite in the "no scope" contexts takes wide scope by default.

8. **specific indefinite, singular (wide scope)**
   In a "Lost and Found"
   Clerk: Can I help you? Are you looking for something you lost?
   Customer: Yes... I realize you have a lot of things here, but maybe you have what I need.
   You see, **I am looking for (a, the, --) green scarf. I think that I lost it here last week.**

9. **specific indefinite, singular (wide scope, *certain* modification)**
   In a school
   Becky: Tom seemed very nervous to me. I think he is having problems in class. Do
   you know why he is so nervous right now?
   Ben: **He is going to meet with (a, the, --) certain professor – someone that Tom
   is really afraid of.**

10. **specific indefinite, singular (no scope interactions, explicit speaker knowledge)**
    At a university
    Rob: Hi Katie - can you help me? I need to talk to Professor Christina Jones, but I
    haven't been able to find her. Do you know if she is here this week?
    Katie: Well, I know she was here yesterday. **She met with (a, the, --) student - he is in
    my physics class.**

2.2.4. Singular non-specific indefinite contexts

Three context types were designed to elicit singular non-specific indefinites. In the context
type exemplified in (11), the indefinite took narrow scope under an operator. The context type in
(12) had no scope interactions and did not either state or deny speaker knowledge; the indefinite
in this context type is non-specific since the speaker is not singling out a particular individual (in
this case, a particular student) as having some noteworthy property. Finally, the singular non-
specific indefinite context in (13) contained no scope interactions, and involved explicit denial of
speaker knowledge.

All of the non-specific contexts are contexts in which use of referential *this* in English is
infelicitous, since no importance is attached to a particular individual (e.g., a particular hat or
student).

11. **non-specific indefinite, singular (narrow scope)**
    In a clothing store
    Clerk: May I help you?
    Customer: Yes, please! I've rummaged through every stall, without any success. **I am
    looking for (a, the, --) warm hat. It's getting rather cold outside.**

12. **non-specific indefinite, singular**
    (no scope interactions, no explicit statement of speaker knowledge)
    Visitor: Excuse me - can you help me? I'm looking for Professor James Smith.
    Secretary: I'm afraid he's not here right now.
    Visitor: Is he out today?
    Secretary: No, he was here this morning. **He met with (a, the, --) student... but I don't
    know where Professor Smith is right now.**

13. **non-specific indefinite, singular (no scope interactions, denial of speaker knowledge)**
   At a university
   Professor Clark: I'm looking for Professor Anne Peterson.
   Secretary: I'm afraid she is out right now.
   Professor Clark: Do you know if she is meeting somebody?
   Secretary: I am not sure. **This afternoon, she met with (a, the, --) student – but I don't know which one.**

2.2.5. *Plural indefinite contexts*

Two context types tested plural indefinites. The wide-scope environment in (14) aims to elicit a specific indefinite, and the narrow-scope environment in (15) aims to elicit a non-specific indefinite. In both environments, L1-English controls allowed both *some* and article omission.

14. **specific indefinite, plural (wide scope)**
   Phone conversation
   Jeweler: Hello, this is Robertson's Jewelry. What can I do for you, ma'am? Are you looking for a piece of jewelry? Or are you interested in selling?
   Client: Yes, selling is right. **I would like to sell you (some, the, --) beautiful necklaces. They are very valuable.**

15. **non-specific indefinite, plural (narrow-scope)**
   Phone conversation
   Salesperson: Hello, Erik's Grocery Deliveries. What can I do for you?
   Customer: Well, I have a rather exotic order.
   Salesperson: We may be able to help you.
   Customer: **I would like to buy (some, the, --) green tomatoes. I'm making a special Mexican sauce.**

The plural contexts were included in order to test whether specificity plays the same role with plural as with singular indefinites (Prediction 2 in (4)).

2.2.6. *Definite contexts*

Finally, two context types were designed to elicit definite DPs in previous-mention environments. One context type was singular (16) and one plural (17).

16. **definite, singular**
   Richard: I visited my friend Kelly yesterday. Kelly really likes animals - she has two cats and one dog. Kelly was busy last night - she was studying for an exam. So I helped her out with her animals.
   Maryanne: What did you do?
   Richard: **I took (a, the, --) dog for a walk.**

17. **definite, plural**
   Rosalyn: My cousin started school yesterday. He took one notebook and two new books with him to school, and he was very excited. He was so proud of having his own school things! But he came home really sad.
   Jane: What made him so sad? Did he lose any of his things?
   Rosalyn: Yes! **He lost (some, the, --) books.**

2.3. Procedure

This study employed the standard procedure described in Chapter 4. There were two major differences with the procedure for the previous study. First, the translation task (discussed in Appendix 3) was always given before the elicitation task, so that the learners did not realize, while taking the translation task, that they were being tested on article use. The forced choice task, which made it clear that articles were being tested, was given second, followed by the Michigan test.

The other change was that the tests were timed. Participants were given one hour for the translation task, one hour for the forced choice task, and 20 minutes for the Michigan test. The total possible testing time was thus 2 hours 20 minutes, but the vast majority of participants finished in two hours or less.

3. Results

In this section, I will examine the L2-learners' performance on the forced choice elicitation task. First, I make a brief note about the learners' L2-proficiency.

The results of the Michigan proficiency test grouped the learners as follows. The L1-Korean group contained 1 beginner, 12 intermediate, and 25 advanced L2-learners. The L1-Russian group contained 13 beginner, 15 intermediate, and 22 advanced L2-learners. In order to make comparisons between L1-Korean and L1-Russian speakers, I concentrate on the results of intermediate and advanced L2-learners, excluding beginners (but see Section 3.7 for a summary of the beginners' results). There were 37 L2-learners in each intermediate/advanced group; the average Michigan test score was 23.65 out of 30 for the L1-Korean intermediate/advanced group, and 22.59 for the L1-Russian intermediate/advanced group. The difference in proficiency between the two L1 intermediate/advanced groups was not significant (two-sample t-test, p = .31)[7].

3.1. Overall results and proficiency levels

Within each L1 group, there were almost no differences between intermediate and advanced L2-learners, as measured by independent two-sample t-tests. For L1-Russian speakers, the only significant differences were as follows. Advanced learners were significantly less likely than intermediate learners to overuse *the* in one of the categories of non-specific singular indefinites (13), as well as on the category of non-specific plural indefinites (15). They were also more likely to (appropriately) use *the* with singular definites (16). In the case of the L1-Korean speakers, advanced learners were significantly less likely than intermediate L2-learners to overuse *the* on specific indefinites containing *certain* (9) and more likely to (appropriately) use *the* with singular definites (16).

Thus, there were very few significant differences between advanced and intermediate L2-learners. This suggests that patterns of article use do not change significantly with proficiency. In the rest of this section, I will therefore group the intermediate and advanced L2-learners in each L1-group together.

I now turn to the overall results of the forced choice elicitation task. The tables in (18) and (19) summarize the means and standard deviations for each category in (7) for the L1-Russian

---

[7] One of the L1-Russian participants who scored as "intermediate" (#r19) was actually taking the Michigan test for the second time; she had taken it more than a year previously as part of a completely unrelated study. That time, she had also placed as "intermediate", with a slightly lower score.

and L1-Korean speakers, respectively. The mean raw scores are reported here – as in the last study the maximum score in each category was 4. Uses of *the* and *a / some* are reported, excluding omission. The means and standard deviations corresponding to the target article for each context are highlighted in bold.

In the rest of this chapter, I will report percentages only, for ease of reference.

18. Elicitation task results: summary of means and standard deviations: L1-Russian
intermediate/advanced L2-learners (N=37)

context type	Target	example	use of *the*		use of *a / some*	
			mean	st. dev.	mean	st. dev.
sg specific wide scope indefinite	*a*	(8)	2.19	1.33	**1.49**	**1.30**
sg specific indefinite, *certain*	*a*	(9)	1.62	1.40	**1.62**	**1.55**
sg specific indefinite, no scope	*a*	(10)	1.46	1.35	**2.38**	**1.34**
sg non-specific narrow scope indefinite	*a*	(11)	0.35	0.72	**3.46**	**0.90**
sg non-specific indefinite, no scope	*a*	(12)	0.54	0.93	**3.41**	**1.07**
sg non-specific indefinite, denial of knowledge	*a*	(13)	0.41	0.76	**3.54**	**0.87**
pl specific wide scope indefinite	*some /* --	(14)	1.08	1.16	**2.11**	**0.94**
pl non-specific narrow scope indefinite	*some /* --	(15)	0.46	0.73	**2.65**	**0.98**
sg definite	*the*	(16)	**3.41**	**0.90**	0.57	0.90
pl definite	*the*	(17)	**2.11**	**1.49**	0.65	0.82

19. Elicitation task results: summary of means and standard deviations: L1-Korean
intermediate/advanced L2-learners (N=37)

context type	Target	example	use of *the*		use of *a / some*	
			mean	st. dev.	mean	st. dev.
sg specific wide scope indefinite	*a*	(8)	1.11	1.05	**2.81**	**1.05**
sg specific indefinite, *certain*	*a*	(9)	1.30	1.61	**2.38**	**1.72**
sg specific indefinite, no scope	*a*	(10)	0.51	0.84	**3.35**	**0.86**
sg non-specific narrow scope indefinite	*a*	(11)	0.27	0.56	**3.59**	**0.64**
sg non-specific indefinite, no scope	*a*	(12)	0.08	0.28	**3.78**	**0.63**
sg non-specific indefinite, denial of knowledge	*a*	(13)	0.51	0.56	**3.32**	**0.63**
pl specific wide scope indefinite	*some /* --	(14)	0.76	1.01	**2.11**	**1.22**
pl non-specific narrow scope indefinite	*some /* --	(15)	0.08	0.36	**3.14**	**0.98**
sg definite	*the*	(16)	**3.32**	**1.00**	0.57	0.90
pl definite	*the*	(17)	**2.49**	**1.33**	0.68	0.82

3.2. Article choice with singular DPs

In this section, I examine the learners' performance in all singular contexts. The prediction concerning article choice in singular contexts is repeated in (20).

20. <u>Prediction 1:</u> The pattern of *the* overuse with specific indefinites (vs. correct article use with definites and non-specific indefinites) will be seen across a wider variety of singular contexts.

The overall results for article use in singular contexts (for intermediate/advanced L2-learners) are given in Figure 1. The three specific indefinite contexts are grouped together, and the three non-specific indefinite contexts are grouped together. The figures shows percentages of *the* use across contexts. The error bars show standard deviations.

Figure 1: % Use of 'the' in all singular contexts
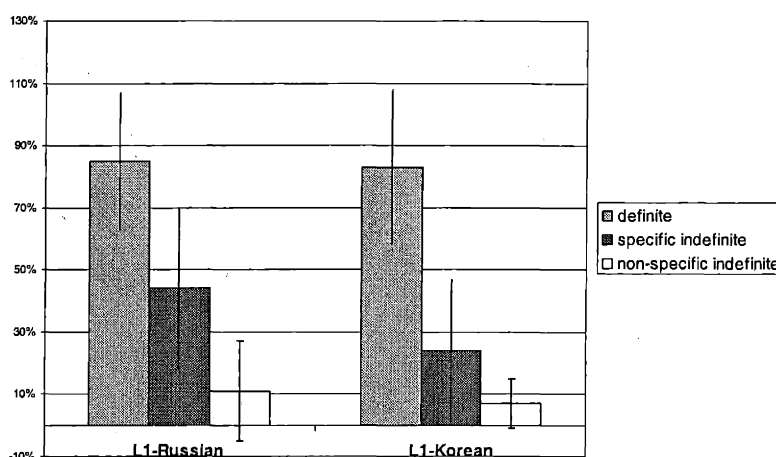
As Figure 1 shows, use of *the* was higher with definites than with specific indefinites, and higher with specific than with non-specific indefinites. These differences were significant for both L1 groups, for both singulars and plurals (paired two-sample t-tests for means, p < .01).

The results for singular contexts are reported numerically in (21), with (22) reporting performance on each singular indefinite context type separately. As these tables show, there was very little article omission in singular contexts. The patterns of more *the* overuse with specific than non-specific indefinites hold up across nearly all of the contexts in (22). Prediction 1 is supported: the interchangeable use of *the* and *a* is restricted to specific indefinite contexts, while article use is quite accurate with definites and with non-specific indefinites.

For the L1-Russian speakers, there was significantly more overuse of *the* in each specific indefinite category than in each non-specific indefinite category (p < .001). For the L1-Korean speakers, there was also significantly more overuse of *the* in each specific indefinite category than in each non-specific indefinite category (p < .01), with two exceptions: the degree of *the* overuse on the specific indefinite category exemplified by (10) did not differ significantly from the degree of *the* overuse on the non-specific indefinite categories exemplified by (11) and (13).

There is some variation in proportions of *the* overuse between the different categories of singular specific indefinites. For the L1-Korean group, there was significantly less overuse of *the* on the category exemplified in (10) than in either of the two other specific indefinite categories (p < .01). For the L1-Russian group, there was significantly more overuse of *the* on the category exemplified by (8) than in either of the other two specific indefinite categories (p < .05).

21. Article use and omission in singular contexts: intermediate/advanced L2-learners

CATEGORY	Target	L1-Russian participants (N=37)			L1-Korean participants (N=37)		
		%the	%a	%omission	%the	%a	%omission
definites	the	85%	14%	1%	83%	14%	3%
specific indefinites	a	44%	46%	10%	24%	71%	5%
non-specific indefinites	a	11%	80%	9%	7%	89%	4%

22. Performance in singular indefinite contexts: detailed breakdown by context type

CATEGORY	L1-Russian participants (N=37)			L1-Korean participants (N=37)		
	%the	%a	%omission	%the	%a	%omission
specific indefinites						
exemplified in (8)	55%	37%	7%	28%	70%	2%
exemplified in (9)	41%	41%	19%	32%	59%	8%
exemplified in (10)	36%	59%	4%	13%	84%	3%
non-specific indefinites						
exemplified in (11)	9%	86%	5%	7%	90%	3%
exemplified in (12)	14%	85%	1%	2%	95%	3%
exemplified in (13)	10%	89%	1%	13%	83%	4%

3.3. Article choice and number

Next, I examine article choice in plural contexts. Prediction 2, which concerns number, is repeated in (23).

23. Prediction 2: The pattern of *the* overuse with specific indefinites (vs. correct article use with definites and non-specific indefinites) will be seen in plural contexts.

Figure 2 reports use of *the* in plural contexts for the two groups of intermediate/advanced L2-English learners. The figure shows percentages of *the* use across contexts. The error bars show standard deviations.

The table in (24) reports the numerical results for plurals. As seen from this table, omission was fairly high across all plural contexts, definite as well as indefinite. The patterns of article use, however, are consistent between the singular and plural contexts[8].

---

[8] It should be noted that control L1-English speakers almost uniformly chose *some* in wide-scope plural indefinite contexts, while allowing both *some* and article omission in narrow-scope plural indefinite contexts. In contrast, L2-English learners allowed article omission in both context types. See Section 4.2 for a discussion of why L2-English learners allow much article omission in plural contexts across the board.

Figure 2: % Use of 'the' across all plural contexts
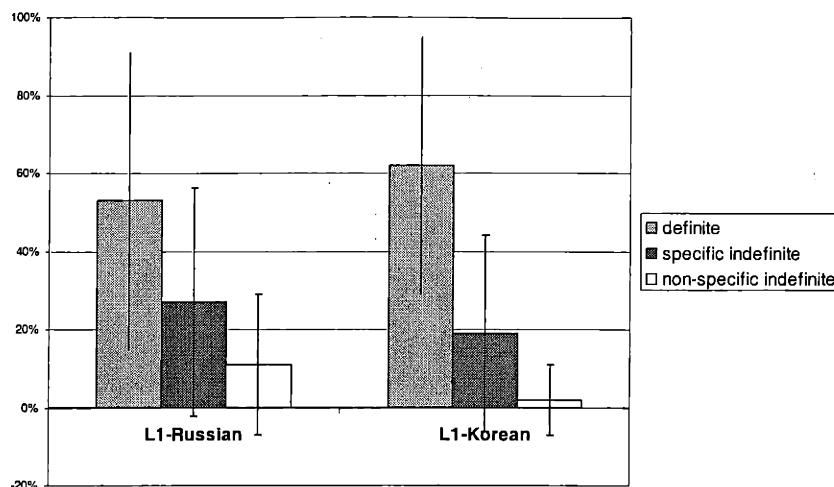
24. Article use and omission in plural contexts: intermediate/advanced L2-learners

CATEGORY	TARGET	L1-Russian participants (N=37)			L1-Korean participants (N=37)		
		%the	%some	%omission	%the	%some	%omission
definites	the	53%	16%	31%	62%	17%	21%
specific indefinites	some / ø	27%	53%	20%	19%	53%	28%
non-specific indefinites	some / ø	11%	66%	22%	2%	78%	20%

Prediction 2 in (23) does not predict a difference between singular and plural DPs. We see from (24) that, in fact, the pattern of overusing *the* with specific indefinites holds in plural as well as singular contexts. However, there is less *the* overuse, and less article use in general, with singular than with plural DPs. I now investigate this in more detail.

In order to examine the effects of number vs. specificity on article choice, I performed a repeated-measures ANOVA on use of *the* with two independent variables: scope and number. The relevant categories for the comparison were wide scope singular indefinites (8), narrow scope singular indefinites (11), wide scope plural indefinites (14) and narrow scope plural indefinites (15). The learners' L1 (Russian vs. Korean) was the between-subject variable. The results are reported in (25).

The significant interaction of scope and number reported in (25) is a result of use of *the* being significantly higher with wide scope singular indefinites than with wide scope plural indefinites; performance on the two narrow scope indefinite categories is not significantly different.

As (25) further shows, there was a significant effect of L1 on use of *the,* and a significant interaction between L1, scope, and number. The source of this was that the L1-Korean speakers were significantly less likely than the L1-Russian speakers to overuse *the* in the categories of narrow scope plural indefinites ($p < .01$) and wide scope singular indefinites ($p < .001$), but not in the other categories.

25. Effects of scope, number, and L1: ANOVA comparisons on use of *the* in wide scope (specific) vs. narrow scope (non-specific), singular vs. plural indefinite contexts.

	F
Effect of scope	$F(1, 72) = 85***$
Effect of number	$F(1, 72) = 20***$
Scope x number	$F(1, 72) = 21***$
Scope x L1	$F(1, 72) = 4.8*$
Number x L1	$F(1, 72) = 1.74$
Scope x number x L1	$F(1, 72) = 12**$
Between subjects effect: L1	$F(1, 72) = 10.55**$

* $p < .05$ ** $p < .01$ *** $p < .001$

In order to further examine the interaction with L1, I look at the relationship between scope and number for each language group, comparing performance on wide scope vs. narrow scope, plural vs. singular indefinites separately for each group. The results of repeated-measures ANOVAs on use of *the* are reported in (26).

Both groups of L2-learners overuse *the* more with specific (wide-scope) than with non-specific (narrow scope) indefinites, and both use *the* more with singular than with plural DPs. Only for L1-Russian speakers is there an interaction: these learners overuse *the* significantly more (p < .001) with wide-scope singular than with wide-scope plural indefinites, while showing no significant differences between the two categories of narrow-scope indefinites.

26. Effects of scope and number: ANOVA comparisons on use of *the* in wide scope (specific) vs. narrow scope (non-specific), singular vs. plural indefinite contexts.

	L1-Russian speakers	L1-Korean speakers
Effect of scope	$F(1, 36) = 51.92***$	$F(1, 36) = 32.72***$
Effect of number	$F(1, 36) = 14.8***$	$F(1, 36) = 5.47*$
Scope x number	$F(1, 36) = 28.86***$	$F(1, 36) = 0.42$

* $p < .05$ *** $p < .001$

At least part of the reason for lowered overuse of *the* with plurals may be the fact that article use overall is lowered with plurals – both indefinite plurals, which allow article omission, and definite plurals, which require *the*. In order to determine whether use of *the* was lowered with both definite and indefinite plurals, I computed repeated-measures ANOVAs with definiteness and number as two independent variables. The comparison categories were wide-scope singular specific indefinites (8), wide scope plural specific indefinites (14), singular definites (16) and plural definites (17). The results for each language group are given in (27).

As (27) shows, both definiteness and number contributed to use of *the*, with no significant interaction. We can conclude that, regardless of definiteness, L2-learners are more likely to use *the* in singular than in plural contexts.

27. Effects of definiteness and number: ANOVA comparisons on use of *the* in definite vs. wide scope (specific) indefinite, singular vs. plural contexts.

	L1-Russian speakers	L1-Korean speakers
Effect of definiteness	$F(1, 36) = 25.9$***	$F(1, 36) = 79.8$***
Effect of number	$F(1, 36) = 35.86$***	$F(1, 36) = 10.97$**
Definiteness x number	$F(1, 36) = .44$	$F(1, 36) = 3.31$

*** $p < .001$

3.4. Article choice and effects of L1

I now turn to Prediction 3, repeated in (28), which concerns performance among L2-learners from different L1's.

28. <u>Prediction 3</u>: Distinct groups of L2-learners whose L1's lack articles will show similar patterns of article use.

The data reported in the previous sections show that Korean speakers were more accurate than Russian speakers on most categories. In fact, the differences in *the* overuse between the two groups were statistically significant on all three categories of singular specific indefinites, on one category of singular non-specific indefinites (12), and on the category of plural non-specific indefinites. There were no significant differences in *the* use on the other categories, either definite or indefinite.

At the same time, the two groups show similar patterns of article use. While the L1-Korean speakers are more accurate than the L1-Russian speakers, both groups exhibit overuse of *the* with specific indefinites. For both groups, overuse of *the* was significantly higher with specific than with non-specific indefinites, as discussed earlier.

3.5. Effects of other variables on article choice

Finally, I briefly consider the effects of various factors (proficiency, age, age of L2 exposure, length of L2 exposure) on the L2-learners' performance. All of these factors showed small and non-significant correlations with *the* overuse in specific indefinite contexts. The (non-significant) correlations are reported in (29).

29. Correlations between %*the* overuse with specific indefinites and proficiency/age variables.

Variable	L1-Russian	L1-Korean
Overall proficiency (Michigan test score)	$r = -.15$	$r = -.23$
Age	$r = .18$	$r = -.20$
Age of first L2 exposure	$r = .15$	$r = -.10$
Length of U.S. residence	$r = -.21$	$r = -.10$
Age at U.S. arrival	$r = .25$	$r = -.19$

3.6. Individual results of intermediate/advanced L2-learners

I now come back to the discussion of individual patterns of L2-learners. In (5), I outlined three possible patterns of article choice in the framework of the Article Choice Parameter, and argued that only two of them – the Definiteness pattern and the Fluctuation Pattern – are predicted to be shown by L2-English learners. In this section, I look at how many individual L2-learners in our study actually show these patterns, as opposed to the Specificity Pattern or various random / miscellaneous patterns.

The patterns of performance shown by the L2-learners in this study in fact fall into five patterns, described in (30) through (34).

30. **The Definiteness Pattern**: *adopting Setting II of the Article Choice Parameter*
    (target-like grammar)
    high use of *the* with definites only
    little or no overuse of *the* with indefinites

31. **The Fluctuation Pattern**
    high use of *the* with definites
    optional use of *the* with specific indefinites
    little or no overuse of *the* with non-specific indefinites

32. **The Specificity Pattern**: *adopting Setting I of the Article Choice Parameter*
    high use of *the* with both definites and specific indefinites
    little or no overuse of *the* with non-specific indefinites

33. **Miscellaneous Pattern 1**
    optional (or high) use of *the* with all indefinites, specific and non-specific

34. **Miscellaneous Pattern 2**
    optional use of *the* with definites

The above terms are defined as follows: for each category (definite, specific indefinite, non-specific indefinite) "high use of *the*" refers to a 75% or higher use of *the* across obligatory contexts; "optional use of *the*" refers to use of *the* in at least 25% but less than 75% of obligatory contexts; and "little or no use of *the*" refers to use of *the* that is below 25%. These proportions were calculated across all singular contexts. Plural contexts were excluded, since the high rates of article omission may obscure the effects of *the* overuse with plurals.

Figure 3 reports the number of L2-learners in each pattern. As this figure shows, the vast majority of the L2-learners fall into one of two groups: 26 L2-learners exhibit the Definiteness Pattern, and 25 L2-learners exhibit the Fluctuation Pattern. Thus, most of the L2-learners are either fluctuating between the two possible settings of the Article Choice Parameter, or have converged on target-like grammar. Very few (5) L2-learners exhibit the Specificity Pattern (Setting 1 of the Article Choice Parameter), which is what we would expect, since the input should lead L2-learners to abandon fluctuation in favor of Setting 2, the correct setting for English.

148

Figure 3: Number of intermediate/advanced L2-learners exhibiting each
pattern
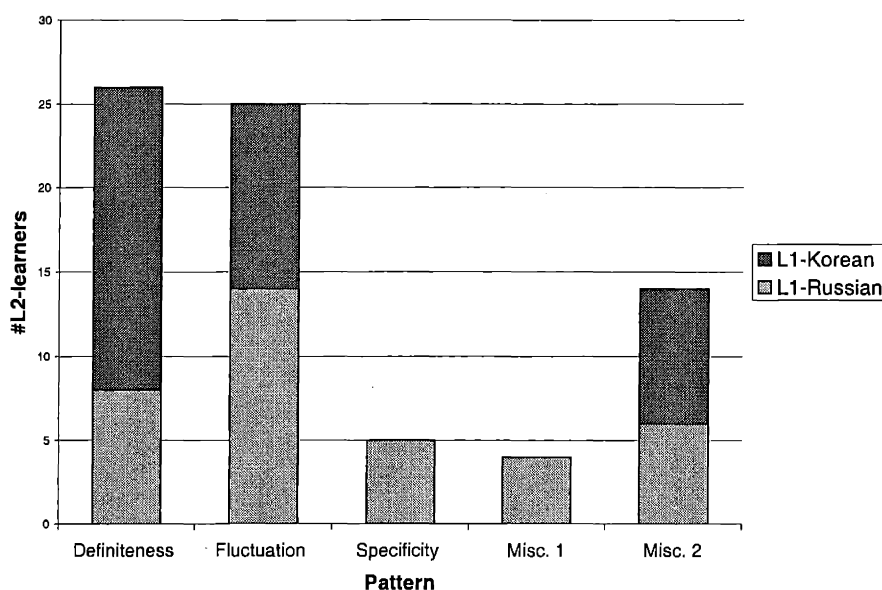
The table in (35) breaks down the L2-learners in each pattern group by proficiency. It shows that L2-learners who exhibit one of the two Miscellaneous patterns are mostly intermediate[9]. There does not seem to be a difference in proficiency, however, between those L2-learners who adopt the Definiteness Pattern and those who exhibit Fluctuation. As already noted, overall proficiency does not seem to be a very good predictor of article use in L2-English.

35. Patterns and proficiency:

subjects	Definiteness Pattern	Fluctuation Pattern	Specificity Pattern	Misc. 1	Misc. 2
L1-Russian: avg. Michigan score (max:30)	**24.50**	**24.00**	**23.00**	**18.75**	**19.00**
L1-Russian: #participants in each proficiency category	5 adv. 3 intr.	11 adv. 3 intr.	4 adv. 1 intr.	1 adv. 3 intr.	1 adv. 5 intr.
L1-Korean: avg. Michigan score (max:30)	**24.94**	**23.36**	--	--	**20.88**
L1-Korean: #participants in each proficiency category	16 adv. 2 intr.	5 adv. 6 intr.	--	--	3 adv. 5 intr.

[9] The four L1-Russian speakers in Miscellaneous Pattern I do not show random performance. The pattern is: use of *the* with definites > use of *the* with specific indefinites ≥ use of *the* with non-specific indefinites: i.e., a pattern resembling Fluctuation, but with unexpectedly high *the* use with non-specific indefinite. In Miscellaneous Pattern 2, three L1-Russian speakers and four L1-Korean speakers also show a pattern resembling Fluctuation, but with overall lowered use of *the*: they use *the* with definites around 50% of the time, overuse *the* with specific indefinites between 25% and 50% of the time, and do not overuse *the* with non-specific indefinites. The remaining three L1-Korean speakers and three L1-Russian speakers show truly random performance. With one exception, these "random" performers are all low intermediate in proficiency.

3.7. Results of beginner L2-learners

In (36) and (37), I report the group results of the 14 beginner L2-learners, grouping together the 13 L1-Russian beginners and the single L1-Korean beginner. Overall, these learners exhibit more article omission across contexts than intermediate/advanced L2-learners, and show lower accuracy of article use. The individual patterns of these learners were as follows: seven of the beginner learners used *the* more with definites than with indefinites, and had a tendency to overuse *the* with specific indefinites (i.e., exhibited a pattern resembling the Fluctuation Pattern, but with more noise). The other seven beginner learners exhibit fairly random performance.

36. Performance in singular contexts: all beginners (N=14)

CATEGORY	TARGET	%the	%a	%omission
definites	the	57%	34%	9%
specific indefinites	a	34%	42%	24%
non-specific indefinites	a	23%	61%	16%

37. Performance in plural contexts: all beginners (N=14)

CATEGORY	TARGET	%the	%some	%omission
definites	the	36%	41%	21%
specific indefinites	some / ø	30%	46%	23%
non-specific indefinites	some / ø	20%	66%	14%

4. Discussion

The data reported in the previous section for intermediate and advanced L2-English learners largely support the predictions in (3). We have observed the patterns of results summarized in (38): L2-learners are quite accurate on the categories of both definites and non-specific indefinites, but show interchangeable use of *a* and *the* on the category of specific indefinites. This study further supported the results of the previous study in showing that overuse of *the* is tied to specificity rather than to wide scope: the L2-learners in this study were tested on two wide-scope non-specific categories (exemplified in (12) and (13)) and did not show significant *the* overuse in these contexts.

38. Summary of article use in L2-English

definiteness	specificity	article used in L1-English	article used in L2-English
+	+	the	the
-	+	a	the / a
-	-	a	a

This pattern of results reported in (38) supports the Fluctuation Hypothesis: L2-English learners fluctuate between the two settings of the Article Choice Parameter, which results in interchangeable article use with specific indefinites, as illustrated in (3).

I will now discuss the three specific predictions in (4) in turn.

4.1. Singular contexts and the degree of specificity

The first prediction, repeated in (39), concerns article use with singular DPs.

39. <u>Prediction 1:</u> The pattern of *the* overuse with specific indefinites (vs. correct article use with definites and non-specific indefinites) will be seen across a wider variety of singular contexts.

Our results largely supported this prediction: overuse of *the* was indeed prevalent in specific indefinite contexts.

However, the L2-learners showed somewhat different patterns of overuse of *the* across different contexts of singular specific indefinites. For instance, both groups of L2-learners overused *the* more in the context in (8), repeated in (40), than in the context in (10), repeated in (41).

40. **specific indefinite, singular (wide scope)**
In a "Lost and Found"
Clerk: Can I help you? Are you looking for something you lost?
Customer: Yes... I realize you have a lot of things here, but maybe you have what I need.
You see, **I am looking for (a̲, the, --) green scarf. I think that I lost it here last week.**

41. **specific indefinite, singular (no scope interactions, explicit speaker knowledge)**
At a university
Rob: Hi Katie - can you help me? I need to talk to Professor Christina Jones, but I haven't been able to find her. Do you know if she is here this week?
Katie: Well, I know she was here yesterday. **She met with (a̲, the, --) student - he is in my physics class.**

Under my hypothesis, this difference in performance suggests that the L2-learners were more likely to consider the indefinite in (40) than the indefinite in (41) specific. This is consistent with the discussion in Chapter 2, since *noteworthiness* is more clearly expressed in (40) (and other contexts of this type) than in (41) (and other contexts of this type). The property *I lost x here last week* in (40) is quite noteworthy: it conveys to the hearer why the speaker considers a particular scarf relevant for the discourse. On the other hand, the property *x is in my physics class* in (41) is not particularly noteworthy; the identity of the particular student that Professor Jones had met with is not important for the immediate discourse. Thus, the differences in overuse of *the* between individual contexts of (singular) indefinites correlate with the degree to which these contexts meet the felicity condition on the use of a specific indefinite.

There is yet another possibility (suggested by David Pesetsky, p.c.). It may be that L2-learners are more likely to use *the* when an indefinite takes scope over an operator. The logic would be as follows. In examining the context in (40), the learners are making a choice: should they treat the indefinite as specific, and use *the*, or should they treat it as non-specific, and use *a*? If they use *a*, the sentence will in principle be ambiguous – it can mean either that the speaker is looking for a particular green scarf (wide-scope reading), or that he is looking for any green scarf (narrow-scope reading). The sentence would not unambiguously convey that the speaker is looking for a particular green scarf (although this reading is strongly preferred, given the context). On the other hand, if the L2-learner uses *the*, marking the context as specific, the context is unambiguous: the speaker is necessarily looking for a particular green scarf. Thus, treating the context as specific (and using *the*) allows the L2-learners to disambiguate the context. In contrast, sentences like (41), with no operators, are unambiguous on both the specific

and the non-specific reading of the indefinite. Whether the L2-learner uses *the* or *a,* the sentence will mean that Dr. Jones talked to a particular student; the only difference is in whether the speaker attaches importance to the student's identity. It is possible that L2-learners are somewhat more likely to mark specificity in an indefinite if this allows them to disambiguate the context in favor of the desired reading.

Finally, both groups of learners consistently used *a* with singular non-specific indefinites, as expected. One exception was the "denial of speaker knowledge" context (13), in which L1-Korean speakers showed unexpectedly high (13%) overuse of *the.* I hypothesize that this effect may have been due to the presuppositionality effect induced by the D-linked phrase *which one* in this context type. The role of presuppositionality in the English of L1-Korean speakers, and its possible explanation of transfer, are discussed in Appendix 4 and are beyond the scope of this chapter.

4.2. Performance in plural contexts

In this section, I consider Prediction 2, repeated in (42), which concerns plural DPs.

42. <u>Prediction 2:</u> The pattern of *the* overuse with specific indefinites (vs. correct article use with definites and non-specific indefinites) will be seen in plural contexts.

This prediction was largely supported: the L2-learners showed similar patterns of performance in singular and plural contexts. However, as shown in Section 3.3, L2-learners omit articles to a much greater degree in plural than in singular contexts (additionally, for the L1-Russian speakers, overuse of *the* was much higher with singular than with plural specific indefinites). There is no obvious linguistic reason for why this should be the case.

I suggest that this effect is in fact an artifact of the test, as follows. L2-learners above beginner level typically know, through explicit instruction, that singular DPs in English must have an article: obligatory use of articles with singular countable nouns in English is one of the aspects of article use consistently emphasized in textbooks of L2-English (see the review of instructions concerning articles in Chapter 8). In contrast, the learners know that plurals may appear without an article in English (when they are indefinite).

When an L2-learner taking the elicitation task sees a target sentence containing a plural indefinite, she may simply decide that no article is required, without bothering to read the preceding context (which, in the case of previous-mention definites, would indicate that *the* is required). In contrast, when the learner sees a context with a singular DP, she knows that an article is required, and therefore reads the context in order to find out which article (*a* or *the*) is appropriate. Crucially, this strategy of "omit articles with plurals" isn't all-determining, since we see more use of *the* with plural definites than with plural indefinites, and more *the* overuse with specific than with non-specific plural indefinites.

If this explanation is correct, and article omission with definite plurals is an artifact of the task, then we predict there to be no differences in *the* use between singular and plural definites in naturalistic production. This prediction will receive support from the production data reported in Chapter 7.

4.3. Group differences

Finally, I turn to Prediction 3, repeated in (43), which concerns article use between the two different L1 groups.

43. <u>Prediction 3:</u> Distinct groups of L2-learners whose L1's lack articles will show similar patterns of article use.

Prediction 3 was partially confirmed: the two groups of L2-learners did indeed show very similar patterns of article use where definiteness and specificity were concerned (but see Appendix 4 for some differences between the two groups that are not related to definiteness or specificity). However, the L1-Korean speakers were more accurate in their article use in most contexts than the L1-Russian speakers – a pattern that was not expected.

I suggest that this effect is traceable to the different type of L2-exposure received by the two groups, rather than to any effect of the learners' L1. While the two L1 groups were balanced for L2-English proficiency, they differed in the type of exposure to the L2 that they had received. The L1-Korean participants were predominantly international students receiving intensive English instruction, while the L1-Russian speakers came from a wide variety of backgrounds. Student status is known to have a positive effect on L2-acquisition (cf. Flege and Liu 2001)[10], and the Michigan test may not have been a very accurate measure of L2-proficiency.

Crucially, the patterns of performance of the two groups are very similar, despite the quantitative differences. This suggests that the quantitative differences are not due to differences in the learners' L1's. Specificity contributes to L2-English article choice among speakers of L1's as different as Russian and Korean. This strongly suggests that transfer cannot account for overuse of *the* with specific indefinites. (For aspects of article choice which may be affected by transfer, see Appendix 4).

4.4. Individual results

Finally, I consider individuals' learners' performance. This study showed that the majority of the L2-learners fell either into the Definiteness Pattern (44a) or the Fluctuation Pattern (44b): most learners either had target-like grammar or were undergoing fluctuation between the settings of the Article Choice Parameter. This was predicted by the FH: learners should start out fluctuating between parameter settings, and eventually set the parameter to its target value (i.e., exhibit the Definiteness Pattern).

44. a) <u>The Definiteness Pattern</u> (Setting II of the Article Choice Parameter)
       L2-English learners correctly use *the* and *a* to mark [+definite] and [-definite] contexts, respectively.

    b) <u>The Fluctuation Pattern</u> (Fluctuation between settings)
       L2-English learners go back and forth between distinguishing *the* and *a* on the basis of definiteness, and distinguishing them on the basis of specificity.

    c) <u>The Specificity Pattern</u> (Setting I of the Article Choice Parameter)
       L2-English learners use *the* and *a* to mark [+specific] and [-specific] contexts, respectively.

Only five L2-learners appeared to incorrectly choose Setting I of the Article Choice Parameter (the pattern in(44c)). That very few L2-learners exhibited the Specificity Pattern is

---

[10] Flege and Liu found that increased length of residence in the U.S. (i.e., increased exposure to the L2) had a positive effect on performance for student L2-English learners, but did not have an effect for non-student L2-English learners. While Flege and Liu's results are not directly comparable to the results of our study, their findings do suggest that there are differences between how students and non-students acquire their L2.

expected: the FH predicts that L2-learners should be in the process of fluctuation until the input leads them to set the parameter to the value appropriate for the target language (in this case, Setting II of the Article Choice Parameter). The five L2-learners who exhibited the Specificity Pattern may have misinterpreted the input. It should be noted that three of these five learners showed higher use of *the* with definites than with specific indefinites, although the difference was very small (e.g., 100% use of *the* with definites vs. 90% use of *the* with specific indefinites); these learners may therefore be fluctuating between parameter setting, with strong leaning in the direction of Setting I.

Another group of learners showed fairly random behavior not predicted by the FH – but see footnote 9 concerning patterns in the behavior of learners who fall under one of the Miscellaneous patterns.

In principle, these individual results do not allow me to claim that L2-learners *start out* by fluctuating between the settings of the Article Choice Parameter. This claim would require evidence from learners at the very start of acquisition, whereas the data considered in this chapter come primarily from intermediate and advanced L2-learners. It is conceivable that learners in the earliest stage of acquisition in fact adopt the Specificity Pattern, and more advanced learners undergo fluctuation as they reset the parameter to the Definiteness setting.

However, the results of the beginner learners suggest that this is not the case. The main characteristics of the beginner data are high article omission and random performance. Where patterns begin to emerge, they resemble Fluctuation rather than Specificity. This suggests that Specificity does not serve as the default setting of the Article Choice Parameter.

5. Conclusion

To sum up, the study reported in this chapter provided additional evidence in favor of the Fluctuation Hypothesis in (2): both L1-Russian and L1-Korean learners of English correctly used *the* with definites, correctly used *a* with singular non-specific indefinites, and overused *the* with specific indefinites. This pattern was upheld in plural contexts: L2-learners were more likely to overuse *the* with specific than with non-specific plural indefinites. The fluctuation pattern was observed at both group and individual level.

The fact that L1-Russian and L1-Korean speakers exhibited qualitatively (if not quantitatively) similar patterns of article use and misuse suggests that the association of *the* with specificity in L2-English is not a result of L1-transfer, but, rather, of direct access to the Article Choice Parameter in L2-acquisition. An examination of individual results suggested that at least some L2-learners are able to set this parameter to the target value, although the ability to set the Article Choice Parameter did not strongly relate to L2-proficiency.

Given the evidence that L2-English learners from different L1-backgrounds draw a specificity distinction in indefinites, the next logical question to ask is whether they also draw a specificity distinction for definites. This question is addressed in the next chapter.

Chapter 6: Experiment 3

This chapter describes the third elicitation study of article choice in L2-English. While the previous studies concentrated on the specificity distinction in indefinites, the study described here addresses the role of specificity with both indefinites and definites, and provides further support for the Fluctuation Hypothesis for L2-English article choice.

1. Background and predictions

The two studies described in Chapters 4 and 5 provided evidence for the reality of the Article Choice Parameter, repeated in (1). I have shown that L2-English learners are in fact sensitive to both definiteness and specificity distinctions: they correctly use *the* with previous-mention definites, correctly use *a* with non-specific indefinites, and overuse *the* with specific indefinites.

1. The Article Choice Parameter (for two-article languages):
   A language which has two articles distinguishes them as follows:
   Setting I.   Articles are distinguished on the basis of specificity.
   Setting II. Articles are distinguished on the basis of definiteness.

The pattern of L2-English article use supports the Fluctuation Hypothesis for L2-English article choice, repeated in (2).

2. **The Fluctuation Hypothesis (FH) for L2-English article choice:**
   1) L2-learners have full UG access to the two settings of the Article Choice Parameter in (1).
   2) L2-learners fluctuate between the two settings of the Article Choice Parameter until the input leads them to set this parameter to the appropriate value.

The study reported in this chapter tests the predictions that the FH makes for definites, and provides additional support for my proposal. It also examines specificity in indefinites in more detail.

1.1. The Fluctuation Hypothesis and definites

The two previous studies examined the role of specificity with indefinites only: the only definites that were tested in those two studies were previous-mention definites, which are obligatorily specific. However, as discussed in Chapter 2, the specificity distinction cross-cuts the definiteness distinction: not all definites are specific. In (3a), for instance, the underlined definite DP is likely to be non-specific: the speaker cannot state any noteworthy property that would hold of a particular individual who is also the owner of this restaurant. Compare (3a) to (3b), where the definite is specific: the speaker can state something noteworthy about the owner – e.g., that she is an old friend. Of course, if one considers the property *owner of this restaurant* to be noteworthy in and of itself, then both contexts in (3) are specific. However, the context in (3b) is in general more likely to be construed as specific than the context in (3a), since the speaker knows more about the referent than what is expressed by the definite description.

3. a) I want to talk to the owner of this restaurant – whoever that is.
   b) I want to talk to the owner of this restaurant – she is an old friend.

If the specificity distinction indeed cross-cuts the definiteness distinction, then languages which distinguish their articles on the basis of specificity (Setting I of the Article Choice Parameter) should use one article with all specific DPs (both definites and indefinites) and a different article with all non-specific DPs (both definites and indefinites). I suggested that an example of such a language is Samoan.

The Article Choice Parameter thus predicts two possible patterns of article choice cross-linguistically: article grouping by definiteness, as in English (4a) and article grouping by specificity, as in Samoan (4b). The picture in (4) represents a complete four-way definiteness/specificity distinction, whereas the representations in Chapters 4 and 5 included only three-way distinctions.

4. a) Article grouping by definiteness   b) Article grouping by specificity

	+definite	-definite
+specific		
-specific		

	+definite	-definite
+specific		
-specific		

The FH predicts that L2-English learners should fluctuate between the two possibilities in (4): some of the time, they should divide articles on the basis of specificity, as in (4a), and some of the time, they should divide articles on the basis of definiteness, as in (4b). With sufficient input, the learners may succeed in setting the Article Choice Parameter to the appropriate setting for English, and divide articles on the basis of definiteness only, as in (4b). Here, I am concerned with what L2-English article choice looks like *before* the learners have set the parameter.

Given the picture in (4), the FH makes explicit predictions for L2-English article use when L2-learners fluctuate between the settings in (4a) and (4b). Both settings predict that one article (e.g., *the*) should be used with specific definites, and a different article (e.g., *a*) should be used with non-specific indefinites: whichever setting the L2-learners adopt, specific definites and non-specific indefinites receive different lexical specifications.

However, the two settings in (4) differ as to how they group specific indefinites and non-specific definites. While (4a) groups specific indefinites with non-specific indefinites, and groups non-specific definites with specific definites, (4b) does exactly the opposite. This is shown in (5): the two cells with striped shading are where the two settings of the Article Choice Parameter are in conflict.

5. The two possible article groupings together

	+definite	-definite
+specific		
-specific		

L2-learners who are in the process of fluctuation should go back and forth between using *the* and *a* on the striped shadings in (5), while being accurate in their article use on the two solidly shaded categories. The predictions for L2-English article choice are spelled out in (6).

6. Predictions for article choice in L2-English: the specificity distinction with definites and indefinites

	+definite (Target: *the*)	-definite (Target: *a*)
+specific	correct use of *the*	overuse of *the*
-specific	overuse of *a*	correct use of *a*

Chapters 4 and 5 provided evidence that L2-learners do indeed overuse *the* on the category [-definite, +specific], while correctly using *the* in [+definite, +specific] contexts and correctly using *a* in [-definite, -specific] contexts. The fourth cell of the table, [+definite, -specific], has not so far been tested. Testing article choice in [+definite, -specific] contexts was the main goal of the study reported in this chapter.

1.2. Indefinites and the degree of specificity

The second goal of this study was to examine in more detail what makes an indefinite context specific. The previous studies showed that such factors as relative clause modification of a wide-scope indefinite and explicit statement of speaker knowledge contribute to overuse of *the* in L2-English, i.e., to the indefinite being treated as specific. This was predicted, since these factors also contribute to the licensing of the specificity marker *this*$_{ref}$ in L1-English.

However, it was noted in Chapter 2 that native English speakers found use of *this*$_{ref}$ felicitous even when speaker knowledge is denied, as long as some noteworthy property is stated in the context. Thus, a contrast was noted between (7a), where the speaker denies all knowledge of the movie in question, and (7b), where the speaker knows something noteworthy about the movie, namely, that *Mary has been excited about seeing it for weeks*. In (7b), the indefinite is more likely to be specific, and use of *this*$_{ref}$ is possible.

The indefinite can also be specific in (7c), where it is embedded under an attitude verb. In principle, (7c) may have two readings available to it. On one reading, the speaker is intending to refer to a particular movie x, where x has the noteworthy property *Mary has been excited about seeing x for weeks*. On the second reading, the speaker is reporting *Mary's* attitude – it is Mary who intends to refer to a particular movie that she wants to see. As discussed in Chapter 2, this reading is derivable under Schlenker's (2002, 2003) proposal that attitude verbs manipulate the context variable. This reading may also be available for (7b) if we take this sentence to contain an implicit attitude operator (see the discussion in Chapter 2).

7. a)# Mary wants to see this new movie; I don't know which movie it is.
   b) Mary wants to see this new movie; I don't know which movie it is, but she's been all excited about seeing it for weeks now.
   c) Mary said that she wants to see this new movie; I don't know which movie it is, but she's been all excited about seeing it for weeks now.

The outcome of the above discussion is as follows. Specific indefinites may be used not only when the speaker has direct knowledge of the referent, but also when the speaker has only second-hand knowledge of the referent, but knows something noteworthy about the referent (7b), or when the speaker is reporting on somebody else's attitude, and the "somebody else" has knowledge of the referent (7c). Crucially, specific indefinites are

infelicitous when the no noteworthy property of the referent is stated, and all speaker knowledge is denied, as in (7a): here, there is no indication of why either the speaker or Mary considers a particular movie noteworthy.

Given these facts, what can we expect in L2-acquisition? Since contexts such as (7b) and (7c) license specific (*this*ref) indefinites, we should expect overuse of *the* in these contexts as well: i.e., we should find overuse of *the* in indefinite contexts with explicit denial of speaker knowledge, as long as these contexts meet the felicity condition on specificity. These predictions are spelled out in (8).

8. Predictions for article choice in L2-English: overuse of *the* and its relation to the degree of speaker knowledge

context type	Predictions:
speaker knowledge is completely denied (7a)	no overuse of *the*
speaker knowledge is denied, but some importance is attached to the referent's identity (7b), (7c)	overuse of *the*

Next, consider the contexts in (9). In the *explicit speaker knowledge* context in (9a) the speaker has acquaintance with the individual denoted by *a book*. The felicity condition on specificity is easily met because the speaker is likely to be referring to a particular book, which has the noteworthy property of being *a book that my brother recommended*. I argued above that contexts with embedding, such as (9b), can also satisfy the felicity condition on specificity: either because the speaker possesses some minimal knowledge of the book in question (i.e., the book has the noteworthy property *Mary liked it a lot*), and therefore intends to refer to it; or because the speaker is reporting Mary's attitude, and Mary intends to refer to a particular book. In contrast, the indefinite in (9c) is clearly non-specific.

9. a) I read <u>a book</u> yesterday – it's the one that my brother recommended; I didn't find it very interesting.
   b) Mary said that she read <u>a book</u>. I don't know which book it was, but she liked it a lot.
   c) Mary read <u>a book</u> yesterday. I don't know which book it was.

An interesting question is whether L2-learners will overuse *the* to the same extent in (9a) and (9b). I hypothesize that L2-learners should show more *the* overuse in (9a) than in (9b). In (9a), the context unambiguously tells the learners that the indefinite is specific: there is explicit statement of speaker knowledge, indication of speaker intent to refer, statement of a noteworthy property. In (9b), however, the context is not equally clear: on the one hand, the speaker is denying knowledge of the book in question, which suggests that the indefinite is non-specific; on the other hand, the speaker is indicating some second-hand knowledge of the referent, so the indefinite may be specific, from either Mary's or the speaker's perspective.

The context in (9a) makes it more obvious than the context in (9b) that the speaker is intending to refer to a particular individual – that importance is attached to a particular book. While we should see overuse of *the* in both (9a) and (9b), there is likely to be more overuse of *the* in contexts with *explicit speaker knowledge* (9a) than in contexts with embedding and *minimal speaker knowledge* (9b). And there should be no overuse of *the* in the *no speaker knowledge* context (9c). These predictions are spelled out in (10).

10. Predictions for article choice in L2-English: overuse of *the* and its relation to the degree of speaker knowledge

context type	Predictions:
speaker knowledge is completely denied (9c)	no overuse of *the*
speaker knowledge is denied, but some importance is attached to the referent's identity (9b)	overuse of *the*
speaker knowledge clearly stated (9a)	highest overuse of *the*

In (9b), the L2-learner may potentially compute specificity from either the matrix speaker's or the embedded speaker's perspective. I will call contexts like (9b) *minimal speaker knowledge* contexts, since the knowledge of the matrix speaker is minimal and second-hand. This does not of course preclude full knowledge on the part of the embedded speaker. I am using the term *minimal speaker knowledge* to distinguish these contexts from contexts like (9a), in which the matrix speaker's knowledge of the referent is explicitly stated.

1.3. The role of *certain*

Finally, this study examined the details of article choice with *certain*-indefinites. The previous two studies showed that L2-English learners consistently overuse *the* with *certain*-indefinites, which suggested that they treat *certain*-indefinites as specific. Researchers such as Fodor and Sag 1982, Kratzer 1998, and Schwarz 2001 in fact viewed use of *certain* as a characteristic of specific or choice-function readings of indefinites.

If *a certain*-indefinites are obligatorily specific, and if L2-learners understand the semantics of *certain*, then we should expect L2-learners to overuse *the* with *certain* regardless of the context. For instance, consider (11). In (11a), speaker knowledge is explicitly stated; in (11b), with embedding, speaker knowledge is "second-hand" and *certain* may be anchoring on Mary's attitude rather than on the attitude of the speaker (see the discussion in Abusch and Rooth 1997). If L2-English learners understand the semantics of *certain*, then they should overuse *the* to the same degree in both contexts in (11), since the indefinite is specific in both[1].

11. a) I read a certain book yesterday – it's the one that my brother recommended; I didn't find it very interesting.
   b) Mary said that she read a certain book. I don't know which book it was, but she liked it a lot.

On the other hand, the learners should differentiate between the two contexts in (9). As discussed in the previous section, L2-learners are likely to overuse *the* more in (9a) than in (9b), since in the former, there is clearer indication that the speaker is intending to refer to a particular book. These predictions are laid out in (12). It is predicted that L2-learners will attend both to *degree of speaker knowledge* and to the presence of *certain*, and that these two factors will interact: the degree of speaker knowledge will play a role with *a*-indefinites but not for *a certain*-indefinites.

---

[1] Irene Heim (p.c.) pointed out to me that since the combination *this certain* is disallowed in L1-English, we might expect *the certain* to be similarly disallowed in L2-English. I do not have an explanation for why *this certain* is disallowed in L1-English. The data of this and previous studies suggest that *the* is compatible with *certain* in L2-English.

12. Predictions for article choice with indefinites in L2-English: the role of *certain*

	no embedding	embedding
a indefinites	ex. (9a) high overuse of *the*	ex. (9b) lower overuse of *the*
a certain indefinites	ex. (11a) high overuse of *the*	ex. (11b) high overuse of *the*

1.4. Summary

Thus, the study described in this chapter set out to test the three hypotheses summarized in (13). The hypothesis in (13a) is crucial for my proposal that the specificity distinction cuts across both definites and indefinites. The hypothesis in (13b) stems from the finding that the felicity condition on specificity may be satisfied even when speaker knowledge is denied, as long as the speaker has some knowledge of the referent.

Finally, the hypothesis in (13c) tests whether L2-English learners treat *certain* as a marker of specificity, which it appears to be in L1-English.

13. a) L2-English learners distinguish both definites and indefinites on the basis of specificity.

b) L2-English learners should construe an indefinite as specific as long as the speaker (or the matrix subject) has some knowledge of the referent.

c) L2-English learners should obligatorily treat indefinites headed by *a certain* as specific, while their evaluation of *a*-indefinites as specific vs. non-specific should depend on the degree of explicitly stated speaker knowledge.

Additionally, this study tested L2-English article choice with previous-mention definites (which are obligatorily specific) and first-mention indefinites (which, in the absence of any stated speaker knowledge, are non-specific) in order to replicate the results of the previous studies.

1.5. Predictions for individual results

Finally, this study allowed for more examination of individual results. The possible patterns of article use in the framework of the Article Choice Parameter are repeated in (14). As discussed in Chapter 5, the FH predicts that L2-English learners will follow the pattern in (14a) or (14b), but not the pattern of (14c) or a random pattern. The individual results reported in Chapter 5 showed that most of the learners did in fact adopt either the pattern in (14a) or the pattern of (14b).

14. a) The Definiteness Pattern (Setting II of the Article Choice Parameter)
L2-English learners correctly use *the* and *a* to mark [+definite] and [-definite] contexts, respectively.

b) The Fluctuation Pattern (Fluctuation between settings)
L2-English learners go back and forth between distinguishing *the* and *a* on the basis of definiteness, and distinguishing them on the basis of specificity.

c) The Specificity Pattern (Setting I of the Article Choice Parameter)
L2-English learners use *the* and *a* to mark [+specific] and [-specific] contexts, respectively.

The present chapter investigates whether the same patterns of individual results hold up as predicted when the specificity distinction is examined with both definites and indefinites.

2. Methods

The predictions in (13) were examined in a forced-choice elicitation study[2].

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 (15). Characteristics of individual participants are given in Appendix 1. Individual participant numbers refer to those in the third table of Appendix 1.

15. Characteristics of L2-learners:

	L1-Russian participants	L1-Korean participants
Number	30 (18 female, 12 male)[3]	40 (22 female, 18 male)
Age range	19;2 to 56;7 (mean = 38; median = 35)	19;11 to 40;0 (mean = 31; median = 31)
Age of first exposure to English	8 to 53 (mean = 14; median = 11)	9 to 14 (mean = 12; median = 12)
Age at arrival in the US (start of intensive exposure to English)	19;1 to 55;10 (mean = 36; median = 33)	16;0 to 35;1 (mean = 28, median = 29)
Time in the US	<0;1 to 10;4 (mean = 1;9, median = 1;2)	0;1 to 8;7 (mean = 2;6, median = 1;8)

2.1.1. Language background

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.

All of the Russian-speaking participants spoke Russian as their primary language but some were also fluent, or even native, in other languages: two speakers were native in Buriat (a Mongolian language) and Tatar, respectively, but had been Russian-dominant since age four; six were bilingual in Russian and another language (Ukrainian, Azeri, Turkmen, or Armenian) and were Russian-dominant; others were fluent, but non-native, in one of the various languages spoken in the former U.S.S.R. The full details are described in the "special notes" section in Table 3 of Appendix 1.

All of the Korean-speaking participants spoke Korean as their first and primary language.

---

[2] This study, like the previous one, was conducted in collaboration with Heejeong Ko.

[3] Two of the participants (#r20 and #r21) had also participated in the previous study, described in Chapter 5. Their performance did not appear to improve noticeably between Study 2 and Study 3: #r20 scored as advanced on the Michigan test both times, and #r21 scored as intermediate both times. Both participants made errors in article choice both times they were tested, with no noticeable improvement.

2.1.2. Recruitment procedure

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 the University of Florida.

2.1.3. Control participants

This study was also administered to 14 native English speaking controls. An additional participant was excluded from the control group because she turned out to be a native speaker of Chinese rather than English. Interestingly, this participant, while fully fluent in English, made three errors of *the* overuse with indefinites. The results of the control participants are reported in Section 3.1.

2.2. Tasks

The participants completed three tasks: a forced-choice elicitation task, a written production task, and the written portion of the Michigan test of L2-proficiency. The Michigan test was described in Chapter 4. The written production task will be described, and the results reported, in Chapter 7. In this chapter, I focus on the forced choice elicitation task.

2.2.1. The format of the elicitation task

The forced choice elicitation task consisted of 76 short dialogues. The format of the forced choice elicitation task was chosen since it had proven to work well in the previous two studies. However, one modification was made: this time, the entire dialogue was given in English, while in the previous studies, all but the target sentence had been given in the learners' L1. The change was put in place to ensure that the results of the previous studies had not been affected by code-switching.

The target sentence in each dialogue was missing an article: the learner had to choose between *a, the,* and the null article (--), basing his or her response on the preceding context. All of the contexts were singular (see Chapter 6 for testing of article choice in plural contexts). The target item was always in object position, to avoid possible transfer effects from Russian (see Chapter 2 for a discussion of Russian word order and definiteness).

There were four contexts per item type. The number of indefinite contexts equaled the number of definite contexts: this was a change from the previous studies, in which the number of indefinite contexts had outweighed the number of definite contexts. We wanted to be certain that the results were not influenced by this skewed distribution, so in the study reported here, the number of indefinite and definite contexts was the same.

I now describe each context type in turn, and repeat the hypotheses from Section 1, indicating the relevant item types which test these hypotheses. The target article in each item is underlined.

See Appendix 2C for the full list of test items.

2.2.2. Specificity with definites and indefinites

In order to test the predictions in (6), four definite and four indefinite context types were included in the test.

First, four definite contexts were used to test for the role of specificity with definites. Two contexts, exemplified in (16) and (17), contained specific definites; the only difference between these contexts is that in (16), the definite takes scope over an intensional verb, while (17) contains no scope interactions. In both types of contexts, the target DP was followed by a description of the individual under discussion; the felicity conditions on the use of a specific DP are satisfied, since the speaker clearly considers something about the individual murderer or creator noteworthy.

16. **[+definite, +specific]: wide scope**
    Conversation between two police officers
    Police officer Clark: I haven't seen you in a long time. You must be very busy.
    Police officer Smith: Yes. Did you hear about Miss Sarah Andrews, a famous
    lawyer who was murdered several weeks ago? **We are trying to find (a, <u>the</u>, --)**
    murderer of Miss Andrews – his name is Roger Williams, and he is a well-
    known criminal.

17. **[+definite, +specific]: no scope interactions, explicit speaker knowledge**
    Kathy: My daughter Jeannie loves that new comic strip about Super Mouse.
    Elise: Well, she is in luck! **Tomorrow, I'm having lunch with (a, <u>the</u>, --)**
    creator of this comic strip – he is an old friend of mine. So I can get his
    autograph for Jeannie!

Two more definite contexts, (18) and (19), contained non-specific definites. The difference between the two contexts again is scope – in (18), the definite has narrow scope with respect to the intensional verb, while in (19), there are no scope interactions. Both context types involved denial of speaker knowledge of the referent.

18. **[+definite, -specific]: narrow scope**
    Conversation between a police officer and a reporter:
    Reporter: Several days ago, Mr. James Peterson, a famous politician, was
    murdered! Are you investigating his murder?
    Police officer: Yes. **We are trying to find (a, <u>the</u>, --) murderer of Mr.**
    Peterson – but we still don't know who he is.

19. **[+definite, -specific]: no scope interactions, denial of speaker knowledge**
    Bill: I'm looking for Erik. Is he home?
    Rick: Yes, but he's on the phone. It's an important business matter. **He is**
    talking to (a, <u>the</u>, --) owner of his company! I don't know who that person is
    – but I know that this conversation is important to Erik.

Four indefinite contexts served as counterparts to the four definite contexts described above. Two contexts, exemplified in (20) and (21), contained specific indefinites; in (20), the indefinite takes wide scope over an intensional verb, while in (21), there are no scope interactions. The felicity condition on the use of specific DPs are met in both cases.

20. **[-definite, +specific]: wide scope**
phone conversation
Jeweler: Hello, this is Robertson's Jewelry. What can I do for you, ma'am? Are you looking for some new jewelry?
Client: Not quite – I heard that you also buy back people's old jewelry.
Jeweler: That is correct.
Client: **In that case, I would like to sell you (a̱, the, --) beautiful silver necklace. It is very valuable – it has been in my family for 100 years!**

21. **[-definite, +specific]: no scope interactions, explicit speaker knowledge**
Meeting on a street
Roberta: Hi, William! It's nice to see you again. I didn't know that you were in Boston.
William: I am here for a week. **I am visiting (a̱, the, --) friend from college – his name is Sam Bolton, and he lives in Cambridge now.**

The two indefinite contexts exemplified in (22) and (23) contained non-specific indefinites. The context in (22) contained a narrow-scope indefinite, and the context in (23) contained denial of speaker knowledge. The indefinite in this context was unmodified, to make clear that the speaker has no knowledge of the referent: modification by a PP or a relative clause (e.g., *a friend from New York*) may indicate some speaker knowledge.

22. **[-definite, -specific]: narrow scope**
In a school
Student: I am new in this school. This is my first day.
Teacher: Welcome! Are you going to be at the school party tonight?
Student: Yes. I'd like to get to know my classmates. **I am planning to find (a̱, the, --) new good friend!** I don't like being all alone.

23. **[-definite, -specific]: no scope interactions, denial of speaker knowledge**
Chris: I need to find your roommate Jonathan right away.
Clara: He is not here – he went to New York.
Chris: Really? In what part of New York is he staying?
Clara: I don't really know. **He is staying with (a̱, the, --) friend – but I have no idea who that is.** Jonathan didn't leave me any phone number or address.

The above eight contexts allowed us to test the predictions in (6) concerning the role of specificity with both definites and indefinites. The predictions are repeated in (24), with the example item numbers incorporated.

24. Predictions for article choice in L2-English: the specificity distinction with definites and indefinites

	[+definite] (Target: *the*)	[-definite] (Target: *a*)
[+specific]	categories in (16) and (17) **correct use of *the***	categories in (20) and (21) **overuse of *the***
[-specific]	categories in (18) and (19) **overuse of *a***	categories in (22) and (23) **correct use of *a***

2.2.3. Specificity and the degree of speaker knowledge in indefinites

The next item type was designed in order to examine the relationship between L2-English article use and *the degree of specificity,* testing the predictions in (10).

Consider the example item in (25): here, the speaker (the second clerk) is not personally acquainted with the necklace; however, the speaker does know that the necklace has the property of being *the necklace that the lady lost here last night, and that was very valuable.* Thus, the indefinite *a gold necklace* may be construed as specific. Alternatively, the indefinite may be construed as specific if specificity is evaluated from the standpoint of the embedded speaker (the lady) rather than the matrix speaker.

However, the indefinite here is less likely to be construed as specific than, for instance, the indefinite in (20), where the matrix speaker is attaching great importance to the identity of a particular silver necklace. Thus, the context in (25) involves *minimal speaker knowledge,* while (20) involves *explicit speaker knowledge.*

25. **[-definite, +specific]: minimal speaker knowledge**
    In a "Lost and Found"
    Clerk1: That lady you were talking with looked very upset. What was the matter?
    Clerk2: She was upset because I couldn't help her. **She said that she is looking for (a, the, --) gold necklace.** She said that she lost it here last night, and that it's really valuable; unfortunately, I couldn't find it.

In all four contexts of the type in (25), the indefinite was embedded under a verb of saying. This was done in order to make the nature of the speaker knowledge clear: (25) makes it clear that the matrix speaker (the second clerk) is not personally acquainted with the necklace in question, but received all of his information from the lady who came into the "Lost and Found." The only knowledge possessed by the speaker is that *the lady said that the necklace is really valuable.*

The contexts exemplified by (25), as well as those exemplified by (20), give indefinites wide scope over an intensional operator. However, (20) is more likely than (25) to be construed as specific. In both (21) and (23), the indefinite takes wide scope in the absence of any intensional operator; (21) is specific and (23) is non-specific, with complete denial of speaker knowledge. Thus, I predict the pattern of overuse of *the* in (26): *the* should be overused in the *explicit speaker knowledge* contexts in (20) and (21); *the* should also be overused, but to a somewhat lesser extent, in the *minimal speaker knowledge* context in (25); and *the* should not be overused in the clearly non-specific *no speaker knowledge* context in (23).

26. Predictions for article choice in L2-English: overuse of *the* and its relation to the degree of speaker knowledge

	scopes over an intensional operator	no scope interactions
explicit speaker knowledge	category in (20) **high overuse of *the***	category in (21) **high overuse of *the***
minimal or no speaker knowledge	category in (25): minimal speaker knowledge **some overuse of *the***	category in (23): no speaker knowledge **no overuse of *the***

2.2.4. *The role of* certain

Two indefinite contexts contained indefinites headed by *a certain*, in order to allow us to test the predictions in (12). The context exemplified by (27) involves *explicit speaker knowledge*: Cathy in (27) is acquainted with the colleague in question. The context exemplified in (28), on the other hand, contains only *minimal speaker knowledge*: the only thing that Ben in (28) knows about the professor in question is that *Tom is really afraid of this professor*. In all of the *minimal speaker knowledge* contexts, the indefinite is embedded under an attitude verb.

27. **[-definite, +specific]:** *certain*-**indefinite, explicit speaker knowledge**
    Robert: Hi, Cathy. Do you have time to talk?
    Cathy: Sorry, not right now – I am about to leave. **I am planning to have coffee with (a̲, the, --) certain colleague; she is very punctual, so I should be on time!**

28. **[-definite, +specific]:** *certain*-**indefinite, minimal speaker knowledge**
    In a school
    Becky: Tom seemed very nervous to me. I think he is having problems in class. He looked really nervous just now!
    Ben: I am not surprised. **He said that he is going to meet with (a̲, the, --) certain professor; I don't know who it is, but Tom is really afraid of this person!**

The above contexts containing *a certain*-indefinites, along with their *a*-indefinite counterparts, allow us to test the predictions in (12). The predictions, along with the corresponding example numbers, are laid out in (29).

29. Effect of speaker knowledge vs. presence of certain: predictions

	explicit speaker knowledge (widest scope)	minimal speaker knowledge (embedding)
indefinite headed by *a*	category in (20) **high overuse of *the***	category in (25) **lower overuse of *the***
indefinite headed by *a certain*	category in (27) **high overuse of *the***	category in (28) **high overuse of *the***

2.2.5. *Other contexts*

Finally, there was a context with simple first-mention indefinites (no scope interactions, statement or denial of knowledge, etc.), shown in (30), and a context with simple previous-mention definites, shown in (31). Based on previous discussion and findings, (30) and (31) were predicted to contain non-specific indefinites and specific definites, respectively. They were included for comparison with other indefinite and definite contexts: previous-mention definites were expected to pattern with other specific definites, and first-mention indefinites were expected to pattern with other non-specific indefinites. One more indefinite context contained a covert partitive (32). I will not discuss results of the covert partitive context in this chapter, but see Appendix 4.

30. **[-definite, -specific]: first-mention indefinite**

Mary: I heard that it was your son Roger's birthday last week. Did he have a good celebration?

Roger: Yes! It was great. He got lots of gifts – books, toys. And best of all – **he got (a, the, --) puppy!**

31. **[+definite, +specific]: previous-mention definite**

Molly: How is your grandpa Sam's farm doing?

Tom: All right, thanks. Last summer, grandpa needed some new animals, so he went to an animal market.

Molly: Did he find any?

Tom: Yes – he found a big cow and a small, friendly horse. But he didn't have enough money for both. **In the end, he bought (a, the, --) horse.**

32. **partitive indefinite**

Rudolph: My niece Janet likes animals a lot. Last week, she decided to get a pet and went to a pet shop.

Lisa: Did she find any pets that she liked?

Rudolph: Yes – she saw three beautiful puppies and six lovely kittens. She couldn't decide! **Finally, she bought (a, the, --) kitten.**

This brings the total number of contexts to nine indefinite contexts and five definite contexts. To balance the number of definites, four filler definite contexts were added. These four contexts were universal definites (33), geographic names which require *the* (34), proper names which require *the* (35), and generics (36).

33. **filler, universal definite**

Laura: I'd like to go for a walk. Is it nice outside?

Jenny: I think so – **I can see (a, the, --) sun!**

34. **filler, geographic name definite**

Tom: You know my uncle Ed? He is a doctor, and once, he went on an expedition!

Louis: Where did he go?

Tom: **He went to (a, the, --) South Pole! He spent a year down there!**

35. **filler, proper name definite**

Louise: I just saw a movie about a ship that was hit by an iceberg, a long time ago. But I can't remember what this ship was called!

Betsy: It **was called (a, the, --) Titanic. It was very famous!**

36. **filler, definite generic**

Peter: Is Sally home? I need to talk to her right away.

Kim: You'll have to wait a few minutes. **She is talking on (a, the, --) telephone. I'll tell her you are here.**

An additional filler context contained possessive DPs, as in (37). Since no article is allowed with possessive pronouns in English, the target her was --. These contexts were included in order to give the L2-learners an indication that articles could be omitted in some contexts.

37. **filler, possessive**

Julie: What did you do last night?

Peter: Not much. **I just worked on (a, the, --) my physics homework.**

2.2.6. *Superlatives and ordinals*

Additional item types containing superlatives and ordinals were administered to a subset of the L2-learners in this study. We included four categories of superlatives/ordinals in the second version of the test, replacing the four categories of definite filler items described in (33) through (37). The superlative/ordinal items were included in order to test the role of specificity with a greater range of definites – i.e., to further test the predictions in (7).

The four contexts below are specific superlatives (38), non-specific superlatives (39), specific ordinals (40) and non-specific ordinals (41). Superlatives were balanced for synthetic vs. analytic forms.

The two *ordinal* categories in fact each contained three items with ordinals plus one item with *only* (as in *the only professor*).

38. **[+definite, +specific]: superlative**

Jim: What are you going to do this summer?

Rose: I'm taking a trip to Paris! I am going to go to the Louvre. **I want to see (a, the, --) most famous painting in Europe – the Mona Lisa! I studied it in my art history class.**

39. **[+definite, -specific]: superlative**

George: My mother needs to have an operation.

Anne: Are you worried?

George: A little. But I'm doing something about that! **I am trying to find (a, the, --) best doctor in Boston – I don't know who that is, but I will find out!**

40. **[+definite, +specific]: ordinal**

At a supermarket

Lesley: It's almost our turn in line.

Sarah: Not yet. I'll be right back.

Lesley: Where are you going?

Sarah: **I am going to talk to (a, the, --) fourth person in this line – it's my friend Peter!**

41. **[+definite, -specific]: ordinal**

In a line for movie tickets

George: That clerk just made an announcement. Did you hear what he said?

Ella: He said that there are only five tickets left! We are lucky, since we are next! And there are so many people behind us.

George: Yeah, we are lucky. **You know, I feel sorry for (a, the, --) sixth person in this line – even though I have no idea who that is.**

The predictions for article use in the superlative and ordinal contexts are given in (42). These predictions are derived from the predictions in (6).

42. Predictions for article use with superlatives/ordinals

	superlatives	ordinals
[+specific]	category in (38) **correct use of** *the*	category in (40) **correct use of** *the*
[-specific]	category in (39) **overuse of** *a*	category in (41) **overuse of** *a*

2.2.7. Test orders

Two test versions were created: one which contained the filler definite items, and one which contained the superlative/ordinal items. The version with fillers was administered to the first 15 L1-Russian participants (r1 through r15) and to all of the L1-Korean participants. The version with superlatives/ordinals was administered to the last 15 L1-Russian participants (r16 through r30); this version was also administered to the last three of the 14 control L1-English participants.

In both versions, the test contained 76 items. These items were divided into two halves of 38 items each; each half contained two items from each category. The two halves were counterbalanced for order of presentation.

2.3. Procedure

The procedure employed in this study was the standard procedure described in Chapter 4. Order of presentation of the naturalistic production task and the forced choice elicitation task was counterbalanced. As in the previous studies, the Michigan test was always given last.

The forced choice task was, as in the previous studies, accompanied by a translation sheet. No translation sheets were provided for the production task or the Michigan test.

The participants were given 90 minutes to complete the forced choice task, but the vast majority finished in 60 minutes or less. The two other tests were untimed. Most participants completed the entire testing session in about two hours, though the more proficient L2-learners finished in 90 minutes or less.

3. Results

This section describes the data from the forced choice elicitation task. I briefly describe the results of the control participants before moving on to a detailed discussion of the L2-learners' performance. Then the L2-learners' results are reported, with the corresponding hypotheses.

3.1. Results of the control participants

The forced choice elicitation task was administered to 14 native English speakers. Eleven of these speakers performed as expected, supplying the target article in each context[4]. I will now discuss the performance of the other three control participants.

---

[4] One of these 11 participants actually omitted *the* before "United Kingdom" in one of the filler items. This is irrelevant for our study. This participant was trilingual in English, Hindi, and Marathi, and had been exposed to English since age 3. If this participant is excluded from the control group, as a result of not having been exposed to English since birth, we are still left with 13 L1-English participants, ten of whom performed as expected.

One participant gave four unexpected responses (as well as leaving one item unanswered). Two of the unexpected responses involved use of *a* in previous-mention definite contexts. I will discuss these contexts, and performance of the control participants, in more detail in Section 3.3.6. This participant also put *the* in two specific indefinite contexts. One was a wide-scope specific indefinite context, given in (43). The other was a context with embedding of the specific indefinite.

43. **[-definite, +specific]: wide scope over an operator**
    In an airport, in a crowd of people who are meeting arriving passengers
    Man: Excuse me, do you work here?
    Security guard: Yes.
    Man: In that case, perhaps you could help me. **I am trying to find (a, the, --) girl with long red hair; I think that she flew in on Flight 239.**

A second control participant gave a single unexpected answer, which also involved putting *the* in the specific indefinite context in (43). At this point, testing of L2-learners had not yet begun, so the item in (43) was changed before further testing: *girl with long red hair* was substituted by *red-haired girl,* since it has been noted that additional modification facilitates a definite reading. Once the item had been changed in this manner, no L1-English speaking control put *the* in this context.

The third control participant who did not perform as expected had received the version of the test with superlatives and ordinals. This participant's unexpected answers were as follows. First, he put *a* in one previous-mention definite context (this will be discussed in Section 3.3.6). Second, he put *the* in one simple indefinite context as well as in one embedded indefinite context. Third, he circled both *a* and *the* as possible answers in one partitive indefinite context.

Finally, two of this participant's unexpected answers involved contexts with superlatives/ordinals, given below. In (44), a *specific superlative* context, the participant put *a* and wrote in "one of the". In (45), a *non-specific ordinal* context, this participant circled both *a* and *the*. As will be seen, L2-English learners were nearly perfect at using *the* with both *specific superlatives* and *non-specific ordinals*, so the unexpected answers given by the control participant to the superlative/ordinal items were in fact almost never attested in L2-English.

44. **[+definite, +specific]: superlative**
    Betsy: What are you going to study when you go to college?
    Kendra: I will study Italian and Spanish films. **I especially want to study (a, the, --) most wonderful director in Italy – Federico Fellini.**

45. **[+definite, -specific]: ordinal**
    Julie: How did you spend your summer vacation?
    Ruth: I read a lot! I started reading a really interesting trilogy. **Now I am trying to find (a, the, --) second book of this trilogy – I don't know what its name is, or what it's about, unfortunately...** But I really want to know what happens next, so I have to find it!

3.2. L2-learners: summary of results

I now move on to the performance of the L2-learners. The distribution of the Michigan test scores was as follows: among the 30 L1-Russian speakers, there were 4

beginner, 11 intermediate, and 15 advanced L2-learners. Among the 40 L1-Korean speakers, there were 1 beginner, 6 intermediate, and 33 advanced L2-learners. For the remainder of this chapter, I focus on the results of the intermediate/advanced L2-learners. The results of the beginners are reported in Section 3.6.

The average Michigan test score was 23.38 for the L1-Russian intermediate/advanced L2-learners, and 25.51 for the L1-Korean intermediate/advanced L2-learners. The difference in proficiency between the two groups was significant ($p < .05$). The two groups can thus not be compared directly. However, we should expect to see the predicted patterns in both groups, since, as the previous studies showed, L2-learners at both intermediate and advanced levels of proficiency exhibited similar patterns.

Of the 15 L1-Russian speakers who received the version of the test with superlatives/ordinals, 14 were intermediate/ advanced.

3.2.1. Overall results by L1-group

The tables in (46) and (47) summarize the means and standard deviations for use of *the* and *a* in each category for the intermediate/advanced L1-Russian and L1-Korean speakers, respectively. The maximum score in each category is 4. The Korean speakers, who are more advanced in proficiency, also exhibit higher accuracy across contexts.

For each category, the mean and standard deviation corresponding to the target article for that category are highlighted.

46. Elicitation task results: summary of means and standard deviations: L1-Russian intermediate/advanced L2-learners (N=26; for superlatives/ordinals, N=14)

context type	target	example	use of *the*		use of *a*	
			mean	st. dev.	mean	st. dev.
definite, specific, wide scope	the	(16)	3.46	0.95	0.23	0.65
definite, specific, no scope interactions	the	(17)	2.85	1.29	0.42	0.64
definite, non-specific, narrow scope	the	(18)	2.31	1.52	1.38	1.55
definite, non-specific, no scope interactions	the	(19)	2.23	1.31	1.23	1.37
indefinite, specific, wide scope	a	(20)	1.42	1.27	2.08	1.41
indefinite, specific, no scope interactions	a	(21)	1.50	1.33	2.23	1.56
indefinite, non-specific, narrow scope	a	(22)	0.23	0.51	3.58	0.76
indefinite, non-specific, no scope interactions	a	(23)	0.35	0.69	3.12	0.95
indefinite, with embedding	a	(25)	0.96	0.72	2.73	0.83
indefinite, *certain*, wide scope	a	(27)	1.54	1.56	1.69	1.62
indefinite, *certain*, with embedding	a	(28)	0.73	1.25	2.42	1.45
simple indefinite	a	(30)	0.62	0.94	2.77	1.21
previous-mention definite	the	(31)	2.88	1.40	0.81	1.20
superlative, specific	the	(38)	3.86	0.53	0.07	0.27
superlative, non-specific	the	(39)	3.50	0.76	0.50	0.76
ordinal, specific	the	(40)	3.64	0.50	0.07	0.27
ordinal, non-specific	the	(41)	3.57	0.65	0.21	0.58

47. Elicitation task results: summary of means and standard deviations: L1-Korean intermediate/advanced L2-learners (N=39)

context type	target	example	use of *the*		use of *a*	
			mean	st. dev.	mean	st. dev.
definite, specific, wide scope	the	(16)	3.79	0.52	0.13	0.41
definite, specific, no scope interactions	the	(17)	3.21	0.98	0.18	0.56
definite, non-specific, narrow scope	the	(18)	3.49	0.82	0.41	0.72
definite, non-specific, no scope interactions	the	(19)	2.92	1.09	0.69	0.89
indefinite, specific, wide scope	a	(20)	1.05	1.26	2.82	1.34
indefinite, specific, no scope interactions	a	(21)	0.69	1.13	3.31	1.13
indefinite, non-specific, narrow scope	a	(22)	0.26	0.64	3.62	0.75
indefinite, non-specific, no scope interactions	a	(23)	0.08	0.27	3.82	0.39
indefinite, with embedding	a	(25)	0.95	1.07	3.00	1.10
indefinite, *certain,* wide scope	a	(27)	0.72	1.02	2.92	1.29
indefinite, *certain,* with embedding	a	(28)	0.21	0.52	3.38	1.21
simple indefinite	a	(30)	0.23	0.43	3.64	0.49
previous-mention definite	the	(31)	3.26	0.94	0.69	0.89

From now on, throughout this chapter, I will report the percentages of use of *the* and *a.* The remaining percentage points out of 100 correspond to article omission, which was fairly low.

3.2.2. Overall results, L1, and proficiency

Within each L1 group, all participants are grouped together. In the L1-Russian group, eight out of the 26 intermediate/advanced L2-learners were bilingual in Russian and some other language, such as Ukrainian or Azeri (none of the beginners were bilingual). Bilingualism did not appear to affect performance: there were no significant differences in use of *the* or *a* on any category between the 18 Russian-primary speakers and the 8 bilingual speakers. The bilingual speakers had, on average, slightly higher L2-proficiency than the Russian-primary speakers, but this difference also was not significant. Thus, for the remainder of this chapter, I group all of the intermediate/advanced L1-Russian speakers together, regardless of whether they were Russian-primary or bilingual.

There were some differences in article use between the intermediate and advanced learners in either L1-group. In the case of the L1-Russian speakers, advanced L2-learners outperformed intermediate L2-learners on several categories, as follows. The two groups differed significantly on use of *the* (p < .05) in the specific definite category in (16); on both use of *the* (p < .05) and use of *a* (p < .05) in the non-specific definite category in (18); and on use of *the* (p < .05) in the previous-mention definite category in (31). The two groups also differed marginally on both use of *the* (p = .065) and use of *a* (p = .059) in the specific indefinite category in (21). In all cases, the advanced L2-learners were more accurate in their article use than the intermediate L2-learners. There were no other

significant differences between the two groups. Thus, there is some improvement in accuracy with proficiency, but it is not visible in all categories.

In the case of the L1-Korean speakers, there were even fewer significant differences between intermediate and advanced L2-learners. The two groups differed significantly on use of *the* (p < .05) and marginally on use of *a* (p = .069) in the category of embedded *certain*-indefinites (28); the two groups also differed marginally on use of *a* with both categories of specific definites: the category in (16) (p = .057) and the category in (17) (p = .051). Surprisingly, all of the significant and marginally significant differences in the L1-Korean group went in the opposite direction than expected: intermediate L2-learners were more accurate than advanced L2-learners. For instance, intermediate L2-learners did not make a single error of *a* overuse with [+specific] definites, while advanced L2-learners had a few such errors. This surprising result probably does not have any deep meaning, however. There were only six intermediate learners in the entire group of 39 L1-Korean speakers, and these intermediate learners were nearly all high intermediate: the cut-off between intermediate and advanced groups is the score of 23 on the Michigan test, and four of the six intermediate learners scored 21 or 22. The sample of L1-Korean speakers in this study did not have a sufficiently great diversity of proficiency scores to inform us about proficiency effects on L2-acquisition. Recall that in the previous study (Chapter 5), which contained a greater number of L1-Korean intermediate L2-learners, intermediate learners were less accurate than advanced learners.

Since there were few differences between intermediate and advanced L2-learners in each L1-group, I will group the intermediate and advanced learners in each group together. However, I will also look at effects of proficiency as I discuss the results.

3.3. Specificity with definites and indefinites: the results

In this section, I look at whether L2-learners make the specificity distinction across both definites and indefinites. The predictions examined in this section are repeated in (48).

48. Predictions for article choice in L2-English: the specificity distinction with definites and indefinites

	[+definite] (Target: *the*)	[-definite] (Target: *a*)
[+specific]	categories in (16) and (17) correct use of *the*	categories in (20) and (21) overuse of *the*
[-specific]	categories in (18) and (19) overuse of *a*	categories in (22) and (23) correct use of *a*

3.3.1. The role of specificity for definites vs. indefinites: individual contexts

First, I report the results for the effects of definiteness and specificity in contexts involving intensional operators (ex. (16), (18), (20), and (22)). The results are reported in (49) and (50) for L1-Russian and L1-Korean participants, respectively. The statistical significance between two definite or two indefinite categories is marked in the cell corresponding to the [-specific] context (so for instance, in (49), the three stars next to "58%the" indicate that the difference in *the* use between specific and non-specific

definites was highly significant)[5]. See the next section for more detailed statistical analyses.

49. Definiteness vs. specificity: intensional contexts.
   L1-Russian speakers (N=26)

	[+definite]		[-definite]	
[+specific] (wide scope)	87%the	6%a	36%the	52%a
[-specific] (narrow scope)	58%the***	35%a***	6%the***	89%a***

50. Definiteness vs. specificity: intensional contexts.
   L1-Korean speakers (N=39)

	[+definite]		[-definite]	
[+specific] (wide scope)	95%the	3%a	26%the	71%a
[-specific] (narrow scope)	87%the**	10%a**	6%the***	90%a***

Next, I report the results for the effects of definiteness and specificity in extensional contexts (ex. (17), (19), (21), and (23)). These are reported in (51) and (52) for the L1-Russian and L1-Korean participants, respectively. Again, statistical comparisons between each two definite and each two indefinite contexts are reported.

51. Definiteness vs. specificity: extensional contexts.
   L1-Russian speakers (N=26)

	[+definite]		[-definite]	
[+specific]	71%the	11%a	37%the	55%a
[-specific]	56%the╪	31%a**	9%the***	78%a**

52. Definiteness vs. specificity: extensional contexts.
   L1-Korean speakers (N=39)

	[+definite]		[-definite]	
[+specific]	80%the	4%a	17%the	83%a
[-specific]	73%the	17%a**	2%the**	96%a**

---

[5] The statistical significances reported here are for paired two-sample t-tests for means, subjects analysis, where: ***p<.001, **p<.01, *p<.05, ╪p<.07 (marginal).

On the items analysis, significant differences between [+specific] and [-specific] categories were yielded for all context types except the following. For L1-Russian speakers, there were no significant differences in *the* use between the two extensional definite contexts or the two extensional indefinite contexts on the items analysis (but the differences in *a* use for these categories were significant, as were the differences for all intensional contexts). For the L1-Korean speakers, the differences in *the* use were non-significant for the two extensional definite contexts, and marginally significant (p=.10) for the two intensional definite contexts; the differences in *a* use were marginally significant both for extensional definite contexts (p=.11) and intensional definite contexts (p=.08). All of the differences for indefinite contexts were significant.

3.3.2. The role of specificity for definites and indefinites: totals

Next, I combine the intensional and extensional contexts above together. The data are given in (53) and (54) for the L1-Russian and L1-Korean speakers, respectively[6].

53. Definiteness vs. specificity: total in non-superlative/ordinal contexts
   L1-Russian speakers (N=26)

	[+definite]		[-definite]	
[+specific]	79%the	8%a	36%the	54%a
[-specific]	57%the	33%a	7%the	84%a

54. Definiteness vs. specificity: total in non-superlative/ordinal contexts
   L1-Korean speakers (N=39)

	[+definite]		[-definite]	
[+specific]	88%the	4%a	22%the	77%a
[-specific]	80%the	14%a	4%the	93%a

As shown by the above tables, both L1-Russian and L1-Korean speakers overused *the* more with specific than with non-specific indefinites, and overused *a* more with non-specific than with specific definites. In order to determine the significance of the contribution of definiteness and specificity to use of *the* vs. *a*, I performed repeated-measures ANOVAs on the use of *the* as well as on the use of *a* by category, for each language group. The results are summarized in (55).

55. Effects of definiteness and specificity: results of repeated-measures ANOVAs

L1-Russian speakers	use of *the*	use of *a*
Definiteness	$F(1, 25) = 61^{***}$	$F(1, 25) = 57^{***}$
Specificity	$F(1, 25) = 21^{***}$	$F(1, 25) = 25^{***}$
Definiteness * Specificity	$F(1, 25) = 1.66$	$F(1, 25) = 1.04$
L1-Korean speakers		
Definiteness	$F(1, 38) = 406^{***}$	$F(1, 38) = 501^{***}$
Specificity	$F(1, 38) = 29^{***}$	$F(1, 38) = 27^{***}$
Definiteness * Specificity	$F(1, 38) = 4.9^{*}$	$F(1, 38) = 1.95$

   $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .001$

---

[6] As the tables in Section 3.3.1 show, there was more use of *the* in intensional definite than in extensional contexts, regardless of specificity. The difference in *the* use between intensional and extensional definite contexts was significant in specific contexts for both groups, and in non-specific contexts for the L1-Korean group. However, this was not replicated on use of *a*, where the differences between extensional and intensional definite contexts were non-significant (although the difference in *a* use between the two non-specific definite contexts was marginal (p = .06) for the L1-Korean group). The lowered use of articles in extensional contexts may have stemmed from two items in extensional contexts in which article omission was very high: items in which the DP was of the form *mother/father of X*. These items are discussed in more detail in Section 3.3.6.

For comparisons between individual indefinite contexts, see Section 3.4. In the present section, I combine extensional and intensional contexts together, and focus on the effects of definiteness vs. specificity.

As (55) shows, both definiteness and specificity had significant effects on article use for both groups, whether use of *the* or use of *a* was measured. There was no interaction between these two factors except in the case of *the* use among L1-Korean speakers. This interaction stems from the fact that the difference in *the* use between specific and non-specific definites was smaller than that between specific and non-specific indefinites; however, both differences were significant (p < .001) (as were the differences in *the* use between each definite category and each indefinite category). There may be a ceiling effect with definites, since Korean speakers were overall quite accurate in their article use. Notably, the interaction was not replicated when use of *a* was measured.

For the L1-Russian speakers, there was a further interesting effect: there was a significant inverse correlation (r = -.47, p < .05) between use of *the* with non-specific definites and use of *the* with specific indefinites; a significant inverse correlation was also present for use of *a* with non-specific definites vs. specific indefinites (r = -.53, p < .01). That is, learners who overused *the* with specific indefinites were also quite likely to overuse *a* with non-specific definites. This further shows that the two effects overusing *the* with indefinites and overusing *a* with definites – are related, both stemming (as was predicted) from the role that specificity plays in L2-grammar.

No such inverse correlation was found for L1-Korean speakers, possibly because of the much higher accuracy rates within that group.

3.3.3. Effects of proficiency level

Next, I looked at the effect of proficiency level (intermediate vs. advanced) on article choice within each language group. The results of repeated-measures ANOVAs with proficiency level as a between-subject variable are given in (56).

56. Effects of definiteness, specificity, and proficiency level: results of repeated-measures ANOVAs

L1-Russian speakers	use of *the*	use of *a*
Definiteness	F(1, 24) = 68.19***	F(1, 24) = 63.48***
Definiteness * level	F(1, 24) = 7.71*	F(1, 24) = 7.5*
Specificity	F(1, 24) = 22.98***	F(1, 24) = 30.8***
Specificity * level	F(1, 24) = 1.9	F(1, 24) = 4.05╪
Definiteness * Specificity	F(1, 24) = 1.47	F(1, 24) = 0.71
Definiteness * Specificity * level	F(1, 24) = 0.05	F(1, 24) = 1.13
L1-Korean speakers		
Definiteness	F(1, 37) = 188.2***	F(1, 37) = 257.15***
Definiteness * level	F(1, 37) = 3.32	F(1, 37) = 0.02
Specificity	F(1, 37) = 19.13***	F(1, 37) = 17.09***
Specificity * level	F(1, 37) = 0.58	F(1, 37) = 0.35
Definiteness * Specificity	F(1, 37) = 2.95	F(1, 37) = 0.73
Definiteness * Specificity * level	F(1, 37) = 0.04	F(1, 37) = 0.04

╪p < .06 (marginal)   * p < .05   ** p < .01   ***p < .001

As shown in (56), the effects of definiteness and specificity remain highly significant when level is taken into account; moreover, the interaction between these two factors for use of *the* among the Korean speakers is no longer significant (now, p = .09).

For the Russian speakers, proficiency level interacts with definiteness whether use of *the* or *a* is measured. This is due to the fact that intermediate L2-learners are significantly more likely to use *the* (p < .05) and less likely to use *a* (p < .05) with indefinites than advanced L2-learners; conversely, advanced L2-learners are (marginally) more likely to use *the* (p = .07) and less likely to use *a* (p = .09) with definites than intermediate L2-learners. Thus, the advanced L2-learners are overall more accurate with both definites and indefinites than intermediate L2-learners.

There is also a marginal interaction between proficiency and specificity when use of *a* is measured for L1-Russian speakers. This stems from the fact that advanced L2-learners are marginally more likely (p = .11) to use *a* with specific DPs than intermediate L2-learners. There is in fact a very clear developmental effect for L1-Russian speakers. Advanced and intermediate L2-learners differ significantly from each other on article use with non-specific definites and with specific indefinites. On non-specific definites, advanced L2-learners are (marginally) more likely to use *the* (p = .057) and less likely to use *a* (p = .052). On specific indefinites, advanced L2-learners are more likely to use *a* (p < .05) and (marginally) less likely to use *the* (p = .07). Thus, on the two categories where optionality of article use was predicted (the [+definite, -specific] and the [-definite, +specific] contexts), intermediate L2-learners show greater optionality than advanced L2-learners[7]; the latter use articles more accurately. On the other hand, in the [+definite, +specific] categories and [-definite, -specific] categories, where no fluctuation was predicted, the groups did not significantly differ from each other: both used articles appropriately.

With the L1-Korean speakers, no such clear developmental effect was observed – not surprisingly given that the group consisted overwhelmingly of advanced L2-learners (33 out of 39). (See Section 0 on the finding that intermediate L2-learners were actually *more* accurate than advanced L2-learners). There was, however, a marginal interaction of level and definiteness for L1-Korean speakers on use of *the* (p = .08), which stems from higher use of *the* with definites among advanced than among intermediate L2-learners (this difference is non-significant, p = .2).

3.3.4. Effects of L1

We have already seen that L1-Korean speakers were more accurate than L1-Russian speakers. I next consider the effect of language on article choice in more detail. I performed repeated measures ANOVAs on the results of both groups taken together, with the learners' L1 as a between-subjects variable. (The additional factor of proficiency level was not included, since other ANOVAs found no significant interactions between L1 and proficiency level; see the previous section for a discussion of the effects of proficiency).

Repeated-measures ANOVAs for all participants together largely replicated the findings reported in the previous sections, as shown in (57). Both definiteness and specificity had significant effects on article use, whether use of *the* or use of *a* was measured. A significant interaction was present only on use of *the,* as was the case with the L1-Korean group (55). (The interaction between definiteness and specificity was

---

[7] This effect was observed on some but not all of the individual [+definite, -specific] and [-definite, +specific] contexts – see Section 0. Thus, while L2-learners tend to exhibit less optionality as they become more proficient, this effect is not seen for all categories tested.

marginal (p = .10) in the case of *a* use). Again, this interaction was due to less difference in *the* use between the two definite contexts than between the two indefinite contexts; however, the differences in *the* use between the two definite contexts, as well as between the two indefinite contexts, were highly significant (p < .001), as were differences between each definite and each indefinite context.

57. Effects of definiteness and specificity, and L1: results of repeated-measures ANOVAs

	use of *the*	use of *a*
Definiteness	F(1, 63) = 328.3***	F(1, 63) = 348.15***
Definiteness * L1	F(1, 63) = 14.96***	F(1, 63) = 17.23***
Specificity	F(1, 63) = 49.89***	F(1, 63) = 55.87***
Specificity * L1	F(1, 63) = 5.97*	F(1, 63) = 6.76*
Definiteness * Specificity	F(1, 63) = 5.75*	F(1, 63) = 2.76
Definiteness * Specificity * L1	F(1, 63) = 0.18	F(1, 63) = 0.03

    * p < .05   ** p < .01   ***p < .001

I will now move on to the interactions between L1 and definiteness as well as L1 and specificity: as shown in (57), these interactions were significant whether *the* or *a* use was measured. The interaction of L1 and definiteness stemmed from the fact that Korean speakers were more likely than Russian speakers to use *the* (p < .01) and less likely to use *a* (p < .05) with definites, and, conversely, that Korean speakers were less likely to use *the* (p < .05) and more likely to use *a* (p < .01) with indefinites: i.e., Korean speakers were more accurate on both categories than Russian speakers.

The interaction of L1 and specificity is due to the fact that Korean speakers used *the* more with non-specific DPs (p < .05) and used *a* more with specific DPs(p < .05).

3.3.5. Superlatives and ordinals

The previous sections showed that L2-English learners overuse *the* with specific indefinites and overuse *a* with non-specific definites – an effect that was especially pronounced for L1-Russian speakers. In order to find out whether this effect would uphold across all categories of definites, we tested a subset of the L1-Russian speakers on superlatives and ordinals (the latter category including *only*). The predictions for article use with superlatives and ordinals are repeated in (58).

58. Predictions for article use with superlatives/ordinals

	superlatives	**ordinals**
[+specific]	category in (38) **correct use of** *the*	category in (40) **correct use of** *the*
[-specific]	category in (39) **overuse of** *a*	category in (41) **overuse of** *a*

Overall, the L2-learners were much better at using *the* with superlatives and ordinals (regardless of specificity) than with other definite DPs. The results are given in (59). The small difference in article use between specific and non-specific superlatives is significant when use of *a* is measured (p < .05) and nearly significant when use of *the* is

measured (p = .055). No significant difference was found between specific and non-specific ordinals.

59. Results for superlatives/ordinals: L1-Russian speakers (N=14)

	superlatives		ordinals	
[+specific]	96% the	2%a	91% the	2%a
[-specific]	88% the	13%a	89% the	5%a

3.3.6. Previous mention definites and uniqueness

In looking at definites so far, I have concentrated on definites that were clearly specific (a noteworthy property exists that picks out a single individual from a set) vs. clearly non-specific (no such noteworthy property exists, as evidenced by denial of speaker knowledge).

In Chapter 2, I argued that previous-mention definites are necessarily specific, since the noteworthy property of their referents has been previously mentioned. I will now take a look at L2-learners' performance on previous mention definites to see if previous-mention definites in fact pattern with other specific definites. The table in (60) compares performance on different types of definites.

60. Comparison of article use on different types of definites

	L1-Russian speakers		L1-Korean speakers	
specific definites (16, 18)	79% the	8%a	88% the	4%a
non-specific definites (17, 19)	57% the	33%a	80% the	14%a
previous-mention definites (31)	72% the	20%a	81% the	17%a

The previous-mention definites present a curious case: they don't quite pattern with either specific or non-specific definites. In the case of L1-Russian speakers, previous-mention definites are closer to specific definites, but with noticeably higher overuse of *a*; for the L1-Korean speakers, previous-mention definites pattern more with non-specific definites in terms of *a* overuse.

However, these results may be confounded: use of *a* with previous-mention definites does not necessarily indicate that L2-learners interpret the context as non-specific. Rather, L2-learners may have interpreted the context as *indefinite*, not computing the uniqueness presupposition from the previous context.

I argue that, indeed, some of the previous-mention definite contexts in the test could be construed as indefinite. As mentioned in Section 3.1, two of the 14 control L1-English participants sometimes put *a* rather than *the* in previous-mention definite contexts. The context in (61) elicited *the* from one control participant, and the context in (62) – from two. This is probably the case because it is possible to interpret both *horse* and *film* below as referring to a novel entity – grandpa Sam could have bought a horse other than the one he had seen, and Robin could have watched any film on the TV, not necessarily the German one.

61. [+definite, +specific]: previous-mention context

Molly: How is your grandpa Sam's farm doing?

Tom: All right, thanks. Last summer, grandpa needed some new animals, so he went to an animal market.

Molly: Did he find any?

Sam: Yes – he found a big cow and a small, friendly horse. But he didn't have enough money for both. **In the end, he bought (a, the, --) horse.**

62. [+definite, +specific]: previous-mention context

Alice: What did you do last night?

Robin: I watched TV.

Alice: What did you watch?

Robin: Well, on one channel, I found an interesting German film. On another channel, I found an exciting news program. Finally, **I watched (a, the, --) film.**

In contrast, no control participant put *a* in either of the contexts in (63) or (64), where it is unlikely that a book or a girl other than the one just mentioned is being considered.

63. [+definite, +specific]: previous-mention context

Vicky: Where were you yesterday? I tried to call you, but you weren't home.

Rachel: I went to a bookstore yesterday.

Vicky: Oh, what did you get?

Rachel: I got lots of things – several magazines, two red pens, and an interesting new book. After I came home, **I read (a, the, --) book.**

64. [+definite, +specific]: previous-mention context

Sarah: Yesterday, I took my granddaughter Becky for a walk in the park.

Claudia: How did she like it?

Sarah: She had a good time. She saw one little girl and two little boys in the park. Claudia is a little shy. **But finally, she talked to (a, the, --) girl.**

Let's now consider how L2-learners did on the items in (61-62) vs. (63-64). The results shown in (65) are rather striking: both groups of L2-learners are clearly overusing *a* much more on items where uniqueness is less likely to be computed than on items where it is more likely to be computed. The apparently high use of *a* with previous-mention definites is thus most likely due to a confound in the test.

65. Performance on individual previous-mention items: all L2-English learners

L2-learner group	Items (61-62): control participants allowed *a*		items (63-64): control participants always put *the*	
L1-Russian speakers	63%the	29%a	81%the	12%a
L1-Korean speakers	74%the	26%a	88%the	9%a

The logical question is whether failure to compute uniqueness can account for L2-learners' overuse of *a* with other kinds of definites: the categories of specific and non-specific definites exemplified in Section 2.2.2. Crucially, no native English speaker ever allowed *a* on any item in these categories. All of the contexts were set up so that the definite could be assumed to have only a single referent: definite descriptions such as *the murderer of x, the owner of x, the president of x,* etc., which have clearly unique

referents, were used. There was only one item, a non-specific definite, which contained a definite description with a potentially non-unique referent: *the manager of his office*. This item is given in (66). (Control participants always put *the* in this context).

66. **[+definite, -specific]: no scope interactions, denial of speaker knowledge**
Rose: Let's go out to dinner with your brother Samuel tonight.
Alex: No, he is busy. **He is having dinner with (a, the, --) manager of his office** – I don't know who that is, but I'm sure that Samuel can't cancel this dinner.

If the L2-learners assumed offices to have multiple managers, they could treat the definite DP in (66) as an indefinite and hence overuse *a*. However, if we exclude this item from the count, we see that both groups of L2-learners still show high overuse of *a* with non-specific indefinites, as shown in (67) and (68).

67. Definiteness vs. specificity: total in non-superlative/ordinal contexts, excluding the 'manager' item: L1-Russian speakers (N=26)

	[+definite]		[-definite]	
[+specific]	79% the	8% a	36% the	54% a
[-specific]	60% the	30% a	7% the	84% a

68. Definiteness vs. specificity: total in non-superlative/ordinal contexts, excluding the 'manager' item: L1-Korean speakers (N=39)

	[+definite]		[-definite]	
[+specific]	88% the	4% a	22% the	77% a
[-specific]	82% the	12% a	4% the	93% a

Importantly, the L2-learners showed rather high overuse of *a* on categories where uniqueness is absolutely obligatory: these are the categories in (69) (specific definite) and (70) (non-specific definite).

69. **[+definite, +specific]: no scope interactions**
Meeting in a park
Andrew: Hi, Nora. What are you doing here in Chicago? Are you here for work?
Nora: No, for family reasons. **I am visiting (the, a, --) father of my fiancé** – **he is really nice, and he is paying for our wedding!**

70. **[+definite, -specific]: no scope interactions**
Phone conversation
Mathilda: Hi, Sam. Is your roommate Lewis there?
Sam: No, he went to San Francisco for this week-end.
Mathilda: I see. I really need to talk to him – how can I reach him in San Francisco?
Sam: I don't know. **He is staying with (a, the, --) mother of his best friend** – **I'm afraid I don't know who she is, and I don't have her phone number.**

It is a safe assumption that L2-learners realize that a person can have only one father and only one mother, and would therefore not treat *the father of my fiancé* or *the mother*

of his best friend as indefinites. Nevertheless, the learners often used *a* in the above contexts, as shown in (71). (Note that there is much article omission in these contexts. The L2-learners apparently thought that the words *mother* and *father* do not require an article, unlike other nouns).

Both groups of learners overused *a* to a rather large extent with the non-specific definite *mother of his best friend*. Surprisingly, the L1-Russian speakers also overused *a* with the specific definite *father of my fiancé*. Since this context was designed as [+specific], overuse of *a* was not predicted for this item. In fact, for the Russian speakers, there was more *a* overuse on this item than on any other item in the categories of specific definites. One L1-Russian participant, when asked for informal feedback after the test, explained that *father of my fiancé* just didn't seem definite: the speaker isn't giving the father's name or saying much about him, just stating that he's paying for the wedding. This participant had apparently decided that *father of my fiancé* does not require *the* in this context because there's not enough information given about the father to make him noteworthy – i.e., because the definite is not specific.

71. Performance on definite items where uniqueness is obligatory

L2-learners group	(69): specific definite		(70): non-specific definite	
L1-Russian speakers	54% the	23% a	46% the	19% a
L1-Korean speakers	57% the	5% a	53% the	18% a

To sum up, it looks like L2-English learners are quite good at using with *the* with definites which are specific, including previous-mention definites, but overuse *a* with definites that they consider non-specific.

3.4. Performance on different indefinite contexts

The previous section showed that L2-English learners distinguish both definites and indefinites on the basis of specificity. In this section, I will look at specificity in indefinites in more detail, examining article use with different types of indefinites.

3.4.1. Specificity in indefinites and the degree of speaker knowledge

First, I looked at the relationship between *the* overuse and the degree of speaker knowledge. The relevant predictions are repeated in (72). The category in (25) contained an indefinite embedded under a verb of saying; while the indefinite in this context could be construed as specific, it was less likely to be specific than the indefinites in (20) and (21), so less overuse of *the* was expected. Crucially, I predicted more overuse of *the* in (25), which may be construed as specific, than in (23), which is obligatorily non-specific.

72. Predictions for article choice in L2-English: overuse of *the* and its relation to the degree of speaker knowledge

	scopes over an operator	no scope interactions
explicit speaker knowledge	category in (20) **high overuse of *the***	category in (21) **high overuse of *the***
minimal or no speaker knowledge	category in (25): minimal speaker knowledge **some overuse of *the***	category in (23): no speaker knowledge **no overuse of *the***

The corresponding results are given in (73) and (74) for the L1-Russian and L1-Korean speakers, respectively.

73. The degree of speaker knowledge: L1-Russian speakers (N=26)

	scopes over an intensional operator		no scope interactions	
explicit speaker knowledge	36%the	52%a	37%the	55%a
minimal or no speaker knowledge	24%the	68%a	9%the	78%a

74. The degree of speaker knowledge: L1-Korean speakers (N=39)

	scopes over an intensional operator		no scope interactions	
explicit speaker knowledge	26%the	71%a	17%the	83%a
minimal or no speaker knowledge	24%the	75%a	2%the	96%a

In order to examine the significance of the above results, I computed a repeated-measures ANOVA on uses of *the* and *a* in the relevant categories for both groups.

75. Effects of wide scope over an operator vs. speaker knowledge: results of repeated-measures ANOVAs

L1-Russian speakers	use of *the*	use of *a*
Wide scope over an operator	$F(1, 25) = 1.88$	$F(1, 25) = 1.65$
Speaker knowledge	$F(1, 25) = 19.24***$	$F(1, 25) = 14.18**$
Scope over an operator * speaker knowledge	$F(1, 25) = 7.16*$	$F(1, 25) = 0.41$
L1-Korean speakers		
Wide scope over an operator	$F(1, 38) = 21***$	$F(1, 38) = 19.75***$
Speaker knowledge	$F(1, 38) = 8.89**$	$F(1, 38) = 6.81*$
Scope over an operator * speaker knowledge	$F(1, 38) = 6.11*$	$F(1, 38) = 3.01$

* $p < .05$ ** $p < .01$ *** $p < .001$

Turning first to the Russian speakers, we see that the presence of explicit speaker knowledge contributed significantly to article use: there was more overuse of *the* and less use of *a* on the categories in (20) and (21), where speaker knowledge was explicitly stated, than on the two categories where it was not. There was no significant effect of the presence of an intensional operator, as predicted – since the indefinites in (20) (scope over an operator) and (21) (no scope interactions) are both specific. There was a significant interaction of speaker knowledge and the presence of an intensional operator in the case of *the* overuse: this stemmed from the fact that there was significantly more overuse of *the* on the category in (25), which could be construed as specific, than on the category in (23), which was clearly non-specific; on the other hand, there was no

significant difference in *the* overuse between the specific indefinite contexts in (20) and (21).

Turning next to the L1-Korean speakers, we see that the presence of an intensional operator had a significant effect on both use of *the* and use of *a*: the L1-Korean speakers exhibited higher overuse of *the* when the indefinite scoped over an intensional operator, as in (20), than when there were no scope interactions, as in (21). There was also a significant effect of speaker knowledge, as predicted. In the case of *the* overuse, there was a significant interaction of the two factors (the interaction was marginal in the case of *a* use, p = .09). This stemmed from the fact that there was a significant difference in *the* overuse between the two categories with no scope interactions ((21) and (25), p = .001), and no difference between the two categories where the indefinite took wide scope over an operator.

3.4.2. The role of certain

I now examine the predictions concerning the presence of *certain,* repeated in (76). According to these predictions, *certain* should be viewed by L2-learners as an obligatory marker of specificity, and presence of *certain* should therefore elicit overuse of *the* in all contexts.

76. Predictions on the effect of speaker knowledge vs. presence of *certain*

	explicit speaker knowledge (widest scope)	minimal speaker knowledge (embedding)
indefinite headed by *a*	category in (20) **high overuse of *the***	category in (25) **lower overuse of *the***
indefinite headed by *a* certain	category in (27) **high overuse of *the***	category in (28) **high overuse of *the***

The corresponding results are given in (77) and (78) for L1-Russian and L1-Korean speakers, respectively. As these results show, embedding lowered overuse of *the* with *certain*-headed indefinites, contrary to the predictions.

77. Effect of speaker knowledge vs. presence of *certain*: L1-Russian speakers

	explicit speaker knowledge (widest scope)		minimal speaker knowledge (embedding)	
indefinite headed by *a*	36%the	52%a	24%the	68%a
indefinite headed by *a* certain	38%the	42%a	18%the	61%a

78. Effect of speaker knowledge vs. presence of *certain*: L1-Korean speakers

	explicit speaker knowledge (widest scope)		minimal speaker knowledge (embedding)	
indefinite headed by *a*	26%the	71%a	24%the	75%a
indefinite headed by *a* certain	18%the	73%a	5%the	85%a

To determine the significance of the above results, I computed repeated-measures ANOVAs for each group of L2-learners, both on use of *the* and on use of *a*. The results are reported in (79). The variable "embedding" corresponds to the distinction between widest scope indefinites (with explicit speaker knowledge) vs. embedded indefinites (with minimal speaker knowledge).

79. Effects of embedding and certain: results of repeated-measures ANOVAs

L1-Russian speakers	use of *the*	use of *a*
Embedding	$F(1, 25) = 16.59$***	$F(1, 25) = 16.36$***
Presence of *certain*	$F(1, 25) = .07$	$F(1, 25) = 2.85$
Embedding * presence of *certain*	$F(1, 25) = 1.05$	$F(1, 25) = .06$
L1-Korean speakers		
Embedding	$F(1, 38) = 7.88$**	$F(1, 38) = 8.35$**
Presence of *certain*	$F(1, 38) = 19.36$***	$F(1, 38) = 1.09$
Embedding * presence of *certain*	$F(1, 38) = 4.85$*	$F(1, 38) = 1.91$

* $p < .05$　　　　** $p < .01$　　　　***$p < .001$

For the L1-Russian speakers, the pattern is very straightforward. These speakers overuse *the* more (and use *a* less) in wide-scope contexts which clearly contain specific indefinites than in embedded contexts where the indefinite is less likely to be specific. Presence of *certain* makes no difference.

For the L1-Korean speakers, the results are more complex. In the case of *a* use, only embedding plays a role – i.e., there is significantly more use of *a* in embedded contexts than in wide-scope contexts, just as was the case for the L1-Russian speakers. However, in the case of *the* use, there is also a significant effect of the presence of *certain*, as well as a significant interaction. The effect of *certain* is due to less overuse of *the* with *certain*-indefinites than with *a*-indefinites. The source of the interaction is the fact that there is no significant difference in *the* overuse between the two *a*-indefinite contexts, but there is a significant difference (p < .01) between the two *a certain*-indefinite contexts.

3.4.3. Comparisons of individual categories of indefinites

In this section, I will look at performance on all indefinite contexts in order to see how performance on various indefinite contexts classified as [+specific] vs. [-specific] is related to the simple "first-mention indefinite" context (ex. (30)). The table in (80) is organized as follows. The first line reports the percentages of *the* and *a* use for simple first-mention indefinites. The rest of the table reports percentages of *the* and *a* use for all other categories of indefinites (which have already been discussed in the previous section), and notes whether uses of *the* and *a* in these categories differ significantly from uses of *the* and *a* with first-mention indefinites.

As can be seen from the table, first-mention indefinites pattern more closely with non-specific indefinites than with specific indefinites. Overall, however, the data in (80) show that the indefinite contexts in the test varied as to their likelihood of being interpreted as specific. Contexts where some noteworthy property was explicitly stated elicited the most use of *the*; contexts where there was no speaker knowledge at all, or which contained a narrow-scope indefinite, had the lowest use of *the*. Potentially

ambiguous contexts – simple indefinites and indefinites in embedded environments – pattern in between.

80. Comparisons of performance on indefinite contexts

	L1-Russian speakers		L1-Korean speakers	
Context (example #)	%the	%a	%the	%a
first-mention indefinites (30)	15%	69%	6%	91%
specific indefinite, wide scope (20)	36%**	52%*	26%***	71%***
specific indefinite, no scope interactions (21)	37%**	55%	17%**	83%‡
specific *certain*-indefinite, wide scope (27)	38%*	42%**	18%**	73%**
a-indefinite with embedding (25)	24%	68%	24%***	75%**
certain-indefinite, embedding (28)	18%	61%	5%	85%
non-specific indefinite, narrow scope (22)	6%	89%**	6%	90%
non-specific indefinite, no scope interactions (23)	9%	78%‡	2%‡	96%‡

‡p < .06 (marginal)   * p < .05   ** p < .01   ***p < .001

3.5. Individual results: intermediate/advanced L2-learners

Next, I report individual results, in order to ensure that the specificity distinction with definites and indefinites occurs at the level of individual L2-learners. I focus on the individual results on the four context types in (48), since performance on these contexts indicates whether individual learners undergo the fluctuation predicted by the FH. The predictions of the FH for L2-English article choice are that fluctuation between *the* and *a* should occur for the same learners in both [+definite, -specific] contexts and [-definite, +specific] contexts. It would be unexpected to find learners fluctuating only in one of these context types but not in the other.

Learners who do not exhibit fluctuation at all are predicted to have adopted the Definiteness setting of the Article Choice Pattern. It is predicted that learners should not adopt the Specificity Setting, which is incorrect for English.

3.5.1. Patterns of individual results

In evaluating the individual results, I divided L2-learners into six groups, with various subgroups. The relevant data for the division were the L2-learners' use of *the* out of all instances of article use in the four context types from (48)[8].

First, I isolated those learners who correctly (at least 75% of the time) used *the* in [+definite, +specific] contexts and who rarely (no more than 25% of the time) used *the* in [-definite, -specific] contexts. The majority of the L2-learners fell into this group. Next I

---

[8] The choice to compute the proportion of *the* use across all instances of article use (excluding omission) was done in order to control for the fact that there was more article omission with definites than with indefinites, largely due to high rates of omission with definites containing *mother* and *father* (see Section 3.3.6).

computed the number of L2-learners who showed the *Definiteness Pattern* of article use – i.e., those who had converged or nearly converged on the target grammar. As shown in (81), there were two subtypes of this pattern. Learners who fall under Type 1 have truly converged on the target grammar: the specificity distinction plays no role with either definites or indefinites. Learners who fall under Type 2 still make a specificity distinction with either definites or indefinites, but to a very small extent. For instance, learners who use *the* 100% of the time with all definites, overuse *the* 12% of the time with specific indefinites, and never overuse *the* with non-specific indefinites, would fall into this group.

81. The Definiteness Pattern (Setting II of the Article Choice Parameter)
L2-learners correctly use *the* and *a* to mark [+definite] and [-definite] contexts, respectively.

a) Definiteness Type 1: equally high *the* use with both specific and non-specific definites; equally[9] low *the* use with both specific and non-specific indefinites.

b) Definiteness Type 2: small difference (less than 25%) in proportions of *the* use between the two categories of definites OR between the two categories of indefinites, but not both.

Next, I looked at how many L2-learners were in a state of fluctuation. In an ideal state of fluctuation, learners who show, a, for instance, 30% difference in *the* use between specific vs. non-specific definites should show exactly the same difference in *the* use between the two types of indefinites. It is not predicted that learners will show a 30% difference in *the* use between the two kinds of definites, and a 50% or a 10% difference between the two kinds of indefinites. However, actual data are unlikely to reflect this ideal state. A learner who uses *the* to mark specificity with 30% of definites may well use *the* to mark specificity with 20% or 40% of indefinites. What would be much more unexpected is a 20% difference with definites vs. an 80% difference with indefinites. Thus, in evaluating the learners' performance, I set certain (relatively arbitrary) cut-off points between patterns. These cut-off points allow learners to be placed in the Fluctuation Pattern even if they make the specificity distinction to a greater extent with indefinites than with definites (or vice-versa), as long as the difference is not too great. For instance, a learner is placed in the Fluctuation Pattern if her difference in *the* use is 15% between the two kinds of definites and 35% between the two kinds of indefinites; she is similarly placed in this pattern if her difference is 30% between the two kinds of definites and 65% between the two kinds of indefinites. However, if the two differences are 15% vs. 60%, she will not be placed in the Fluctuation Pattern, but in one of the additional patterns discussed below.

I now move on to a discussion of *The Fluctuation Pattern*, which is divided into four subtypes, as shown in (82). The learners falling under Type 1 make a very small distinction between both definites and indefinites. While these learners are showing fluctuation, they are leaning strongly in the direction of resetting the parameter in favor of the Definiteness distinction. The learners who fall under Type 2 show a mixed

---

[9] Use of *equally* needs to be qualified. I treat use of *the* as being equal on two categories if the difference is on the level of 1% or 2% (e.g., 12% vs. 14% use of *the*).

performance: they make a larger specificity distinction with definites than with indefinites, or vice versa. These are the learners who are probably the furthest away from the ideal state of fluctuation.

The learners in Type 3 give the clearest evidence of fluctuation: these learners make a strong specificity distinction with both definites and indefinites. Finally, the learners who fall under Type 4 are leaning towards the Specificity setting of the parameter, which is inappropriate for English. These learners cannot be said to have adopted the Specificity setting completely, however, because they still differentiate between definites and indefinites: for instance, this category might include a learner who uses *the* 100% of the time with specific definites, 75% of the time with specific indefinites, and never with non-specific DPs.

82. The Fluctuation Pattern (Fluctuation between settings)
   L2-learners go back and forth between distinguishing *the* and *a* on the basis of definiteness, and distinguishing them on the basis of specificity.

   a) Fluctuation Type 1: small specificity distinction with both definites and indefinites; the difference in proportion of *the* use between specific and non-specific DPs is under 25% for both definites and indefinites.

   b) Fluctuation Type 2: slightly different patterns in the specificity distinctions with definites vs. indefinites; proportion of *the* use between specific and non-specific DPs is small (less than 25%) for one of the definite/indefinite categories, and higher (between 25% and 50%) for the other.

   c) Fluctuation Type 3: strong specificity distinction with both definites and indefinites; the difference in proportion of *the* use between specific and non-specific DPs is between 25% and 75% for both definites and indefinites.

   d) Fluctuation Type 4: very high specificity distinction with both definites and indefinites; the difference in proportion of *the* use between specific and non-specific DPs is more than 75% for both definites and indefinites. However, there is still more *the* use with specific definites than with specific indefinites.

Next, I looked for L2-learners showing the *Specificity Pattern*. As shown in (83), these are learners who treat specific definites and indefinites the same, using *the* with both, while hardly ever using *the* with non-specific DPs.

83. The Specificity Pattern (Setting I of the Article Choice Parameter)
   Equally high use of *the* with both specific definites and specific indefinites; low (<25%) use of *the* with all non-specific DPs.

Then come the various unpredicted patterns. One such pattern is what I will call the *Strange Pattern*, a pattern that the FH cannot account for. Learners who fall under this pattern, stated in (84) come in several types. Some make a large (more than 25%) specificity distinction with definites, but no specificity distinction with indefinites. Others do exactly the reverse. Finally, some make strong specificity distinctions with both definites and indefinites, but exhibit patterns of article use that the FH cannot account for (that do not fit under any of the patterns in (82) or (83)).

84. The Strange Pattern: unaccounted for by the FH

large discrepancy in the specificity distinction with definites vs. with indefinites.

Next come two miscellaneous patterns, described in (85). The first of these patterns include learners who unexpectedly show higher use of *the* on non-specific categories than on the corresponding specific categories. The second miscellaneous pattern includes L2-learners who were initially excluded from the analysis of individual results, as discussed in the beginning of this section: those whose use of *the* with specific definites is unexpectedly low, or whose use of *the* with non-specific indefinites is unexpectedly high.

85. a) Miscellaneous Pattern1: use of *the* on at least one of the non-specific
categories exceeds use of *the* on the corresponding specific category.

b) Miscellaneous Pattern 2: low (under 75%) use of *the* with specific definites
and/or high (more than 25%) use of *the* with non-specific indefinites.

3.5.2. Results: number of L2-learners in different patterns

In Figure 1, I give the numbers of L1-Russian speakers and L1-Korean speakers falling into each pattern. The two types of definiteness patterns are grouped together, as are the four types of fluctuation patterns. As shown in Figure 1, most L2-learners fall into either the Definiteness Pattern or the Fluctuation Pattern, as expected. Only two learners adopt the Specificity Pattern and relatively few learners (9) adopt the unexpected "Strange" patterns.

Figure 1: Number of L2-learners showing each pattern
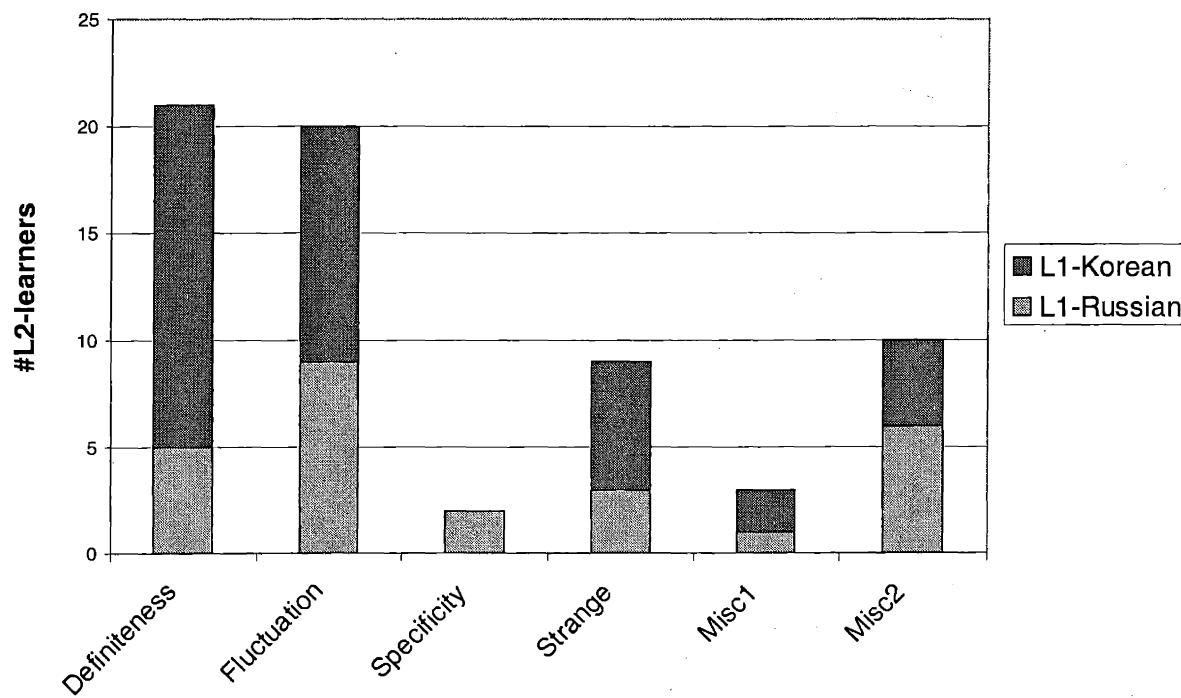
Figure 2 gives the breakdown of L2-learners in the individual Definiteness and Fluctuation patterns discussed in (81) and (82). As this graph shows, most learners who show the Definiteness Pattern have truly converged on the target grammar (Definiteness Type 1) but some still show slight fluctuation (Type 2). Most learners who are showing the Fluctuation Pattern either show true fluctuation (Type 1) or a mixed pattern (Type 2). Very few learners who are in a state of fluctuation lean strongly towards either the Definiteness Setting (Type 3) or the Specificity Setting (Type 4).

Figure 2: Number of L2-learners in individual Definiteness / Fluctuation Patterns
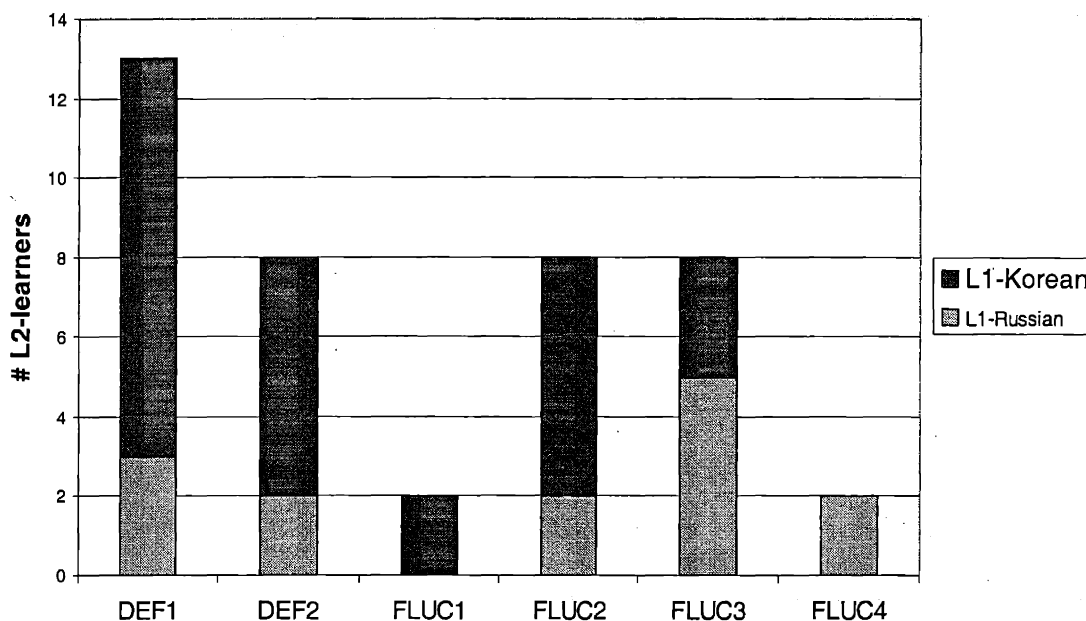
Finally, the table in (86) reports the average proficiency scores of the L2-learners in each pattern. For the definiteness and fluctuation patterns, there is a slight developmental effect. In both L1-groups, learners who exhibit the Definiteness Pattern (i.e., who have converged on the target grammar) are on average advanced. Learners who fall into the first two fluctuation patterns – i.e., who show relatively little fluctuation and lean towards the target Definiteness setting of the parameter – are also advanced, on average. On the other hand, learners who fall into the last two fluctuation patterns – i.e., who show quite strong fluctuation, and lean more towards the Specificity setting of the parameter – are on average intermediate. However, the developmental pattern does not hold for the other categories: for instance, learners who show the Specificity pattern are on average advanced, as are those who show the "Strange" pattern.

86. Individual patterns and proficiency scores

PATTERN		L1-Russian		L1-Korean	
		#learners	avg MI score	#learners	avg MI score
Definiteness	Definiteness1	3	29	10	25.2
	Definiteness2	2	25	6	25.3
Fluctuation	Fluctuation1	0	n/a	2	25.5
	Fluctuation2	2	25	6	26
	Fluctuation3	5	21.8	3	21
	Fluctuation4	2	20	0	n/a
Specificity		2	23.5	0	n/a
Strange[10]		3	24.7	6	26.5
Miscelanneous1[11]		1	28	2	27
Miscelanneous2[12]		6	20.5	4	27

---

[10]In the "Strange" pattern, two L1-Russian and three L1-Korean participants use *the* with specific definites 100% of the time and correctly never use *the* with non-specific indefinites. They also show very high use of *the* both with non-specific definites and with specific indefinites: e.g., 100% use of *the* on one and 70% use of *the* on the other, or vice-versa. There is thus some degree of fluctuation, but the FH cannot account for why these learners simultaneously use *the* to such a high degree both with non-specific definites and with specific indefinites.

Also in this pattern, one L1-Russian participant and two L1-Korean participants hardly ever use *the* with either kind of indefinite, but make a clear specificity distinction with definites; the remaining L1-Korean speaker does the reverse, correctly using *the* with both kinds of definites, but making a specificity distinction with indefinites. The FH cannot account for why learners should draw a specificity distinction only with definites or only with indefinites, rather than with both.

[11] All three learners in the "Miscellaneous1" pattern actually exhibit behavior very similar to the Definiteness Pattern, so it is not surprising that they are, on average, quite advanced. These learners exhibit high (more than 75%) *the* use with both kinds of definites, and low (less than 25%) *the* use with both kinds of indefinites. Their errors involve either slightly more *the* use with non-specific than with specific definites; or, slightly more *the* use with non-specific than with specific indefinites.

[12] The ten learners in the "Miscellaneous2" pattern behave as follows. Two L1-Russian speakers exhibit a pattern somewhat resembling the Definiteness Pattern, with much higher use of *the* with definites than with indefinites; their error is in allowing fairly high use of *the* with non-specific indefinites. Two L1-Russian speakers as well as one L1-Korean speakers show clear evidence of fluctuation. Their error is in either using *the* too much with non-specific indefinites, or in not using it enough with specific definites. One L1-Korean speaker shows the "Strange" pattern, with 100% of *the* use across all categories except non-specific indefinites; in addition, this learner has high *the* use with non-specific indefinites. The remaining two L1-Korean speakers, as well as one L1-Russian speakers, show semi-random behavior: they use *the* more with definites than with indefinites, but their use of *the* is unexpectedly low on specific definites, and unexpectedly high with non-specific definites. Finally, one L1-Russian speaker almost never uses *the* at all, on any category.

3.6. Results of the beginner learners

In this section, I briefly report the results of the beginner learners. I group together the 4 L1-Russian beginners and the single L1-Korean beginners. These learners showed low accuracy in article use across all contexts. In the first-mention simple indefinite category (30), their overuse of *the* was 25%, and their use of *a* was 55%. In the previous-mention definite category (31), their use of *the* was 45%, and their overuse of *a* was 25%.

The tables in (87) and (88) give the beginners' results in the definite and indefinite contexts related to the predictions in (48), for intensional and extensional contexts, respectively. As the results show, beginner L2-learners resembled intermediate/advanced L2-learners in overusing *the* more with specific than with non-specific indefinites, and overusing *a* more with non-specific than with specific definites. They exhibited especially great overuse of *the* with specific indefinites.

87. Definiteness vs. specificity: intensional contexts. All beginners (N=5)

	[+definite]		[-definite]	
[+specific] (wide scope)	65%the	20%a	70%the	20%a
[-specific] (narrow scope)	50%the	45%a	35%the	35%a

88. Definiteness vs. specificity: extensional contexts. All beginners (N=5)

	[+definite]		[-definite]	
[+specific]	65%the	30%a	80%the	20%a
[-specific]	55%the	35%a	15%the	70%a

Next, the table in (89) reports on the effects of different degrees of speaker knowledge, testing the predictions in (72). The beginners follow the predicted pattern of high *the* overuse with explicit speaker knowledge, lower *the* overuse with minimal speaker knowledge, and almost no *the* overuse with no speaker knowledge

89. The degree of speaker knowledge: All beginners (N=5)

	scopes over an intensional operator		no scope interactions	
explicit speaker knowledge	70%the	20%a	80%the	20%a
minimal or no speaker knowledge	45%the	45%a	15%the	70%a

Finally, the table in (90) reports the results concerning the effect of *certain*, testing the predictions in (76). The results show that beginners show less overuse of *the* with *certain*-indefinites than with *a*-indefinites, and that embedding lowers overuse of *the* for both categories. The results of the beginner learners follow the same pattern as those of the intermediate/advanced L1-Korean learners of English.

90. Effect of speaker knowledge vs. presence of *certain*: All beginners (N=5)

	explicit speaker knowledge (widest scope)		minimal speaker knowledge (embedding)	
a-indefinite	70%the	20%a	45%the	45%a
a certain indefinite	40%the	40%a	35%the	55%a

As far as individual results are concerned, one beginner learner exhibits the Specificity Pattern; two learners exhibit a pattern resembling Fluctuation, but with high use of *the* across all categories, including non-specific indefinites; one learner exhibits the "Strange" pattern; and one shows completely random behavior.

3.7. Effects of other variables on article choice

Finally, I consider whether various variables related to age and proficiency had an effect on L2-English article choice. I therefore computed correlations between proportion of *the* use and each of the following, for the intermediate/advanced L2-learners: age at the time of the study; age of first exposure to English; age of arrival into the U.S.; time of residence in the U.S.; and proficiency score on the Michigan test. I report only the significant and marginally significant correlations. As the results below show, there were relatively few significant correlations between various additional variables and article use in L2-English. None of the age-related or exposure-related variables appear to be good predictors of L2 article choice. Proficiency is a slightly better predictor, but even so, there are no significant correlations between proficiency and *the* use on most categories.

3.7.1. Effects of age

For the L1-Russian group, age at the time of the study had positive correlations ($r=.39$, $p < .05$; $r = .35$, $p = .08$) with *the* overuse on two categories of specific indefinites (20 and 21), and a positive correlation of .36 ($p = .07$) with *the* overuse with simple first-mention indefinites. Thus, older L2-learners were more likely to make errors of *the* overuse on some indefinite categories.

As for the L1-Korean group, age had positive correlations ($r = .45$, $p < .01$; $r = .35$, $p < .05$) with *the* use on two categories of definites (16 and 19). Older learners exhibited greater accuracy of article use on these two categories.

3.7.2. Effects of first age of exposure to English

For the L1-Russian group, age of first exposure to English had a positive correlation of .49 ($p < .05$) with overuse of *the* with simple first-mention indefinites. On this category, L2-learners who had started acquiring English later in life were more likely to make errors.

For the L1-Korean group, there were no significant correlations between age of first exposure and use of *the*. This is not very surprising, since age of first exposure was nearly identical for most of the learners in this group.

3.7.3. Effects of age of arrival to the U.S.

For the L1-Russian group, age of arrival in the U.S. (start of intensive exposure to English) had positive correlations ($r = .40$, $p < .05$; $r = .39$, $p = .051$) with overuse of *the* on two categories of specific indefinites (20 and 21). Learners who had arrived in the U.S. at a later age were more likely to make errors on these categories.

For the L1-Korean group, age of arrival had a negative correlation of -.27 ($p = .09$) with *the* overuse on one category of specific indefinites (21). Learners who had arrived in the U.S. at a later age were more accurate on this category.

3.7.4. Effects of time of residence in the U.S.

For the L1-Russian group, length of U.S. residence had a negative correlation of -.62 (p < .001) with *the* use on one category of specific definites (16), and a negative correlation of -.38 (p = .054) with *the* use on one category of non-specific definites (18). Time of residence also had a positive correlation of .42 (p<.05) with *the* overuse with simple first-mention indefinites (20).

For the L1-Korean group, time of residence had a positive correlation of .40 (p < .05) with *the* overuse in the category of widest-scope *certain* indefinites.

Thus, in both L1-groups, learners who had lived in the U.S. longer actually made *more* errors of article use on some categories, contrary to what might be expected. On most categories, there was no significant relationship between time of residence and performance.

3.7.5. Effects of overall proficiency

For the L1-Russian group, proficiency scores had positive correlations (r=.61, p < .001; r=.55, p < .01) with *the* use in both categories of specific definites. Proficiency had a negative correlation with *the* overuse (r = -.35, p = .077) on one category of non-specific definites (23). It also had a negative correlation of -.43 (p < .03) with *the* overuse with simple first-mention indefinites, and a positive correlation of .66 (p < .001) with *the* use on previous-mention definites. Thus, proficiency positively affected performance on some categories.

For the L1-Korean group, proficiency correlated positively (r = .27, p = .10) with *the* use on one category of non-specific definites (19). Again, advanced learners were more accurate.

4. Discussion

I will now discuss each hypothesis concerning article choice in L2-English and the corresponding results. I focus on the results of the intermediate/advanced L2-learners. The patterns of article choice among beginner L2-learners resemble those of the intermediate/advanced learners, but with higher error rates.

4.1. Specificity and definiteness

First, I look at the predictions concerning the role of specificity with definites vs. with indefinites. The predictions are repeated in (91), and the actual results are summarized in (92).

91. Definiteness vs. specificity: predictions for L2-English

	[+definite] (Target: *the*)	[-definite] (Target: *a*)
[+specific]	categories in (16) and (17) correct use of *the*	categories in (20) and (21) overuse of *the*
[-specific]	categories in (18) and (19) overuse of *a*	categories in (22) and (23) correct use of *a*

92. Definiteness vs. specificity: summary of results from both L1-Russian and L1-Korean speakers

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

4.1.1. Performance on the main test items

As the summary in (92) shows, the predictions were supported: L2-learners differentiated between specific and non-specific DPs with both definites and indefinites. This difference was more pronounced for the L1-Russian speakers than for the L1-Korean speakers. This is not very surprising, since the L1-Korean speakers were, on average, more proficient in English (and we saw in Chapter 5 that Korean speakers outperformed Russian speakers even when the two groups were balanced for proficiency). Crucially, despite the quantitative differences between the two L1 groups, the groups showed qualitatively similar patterns of performance: both groups overused *the* more with [+specific] than [-specific] indefinites, and overused *a* more with [-specific] than with [+specific] definites. Both groups also used *the* appropriately with previous-mention definites, which are obligatorily specific (overuse of *a* with some previous-mention definites was shown to result from a confound of the test).

4.1.2. Performance on superlatives/ordinals

Turning next to superlatives and ordinals, we saw that the subset of L1-Russian speakers who were tested on these item types overwhelmingly used *the* correctly with both specific and non-specific superlatives and ordinals. The high accuracy rate with superlatives and ordinals may be a result of explicit instructions: there is evidence that L2-English learners are instructed on the use of *the* with superlatives and ordinals. For instance, in an L2-English textbook, Maclin (1987:60) lists a few rules for where *the* is used, among them "Use *the* before superlatives and ordinal numbers", and gives corresponding examples. Similarly, Raimes (1990:59), in another textbook of L2-English, states that "When we use a superlative (*the best, the most powerful,* etc.), we always use *the*. The use of the superlative distinguishes the noun phrase and makes it actual and specific." In fact, when one of the L2-learners in our study was asked for feedback concerning why she put *the* in superlative contexts, her answer was that she had learned that superlatives obligatorily take *the*. It is thus possible that L2-learners automatically put *the* any time they saw a superlative or ordinal as a result of an explicitly taught strategy, without paying much attention to the context. (See Chapter 8 for a discussion of why explicit strategies cannot account for overuse of *the* with specific indefinites, or overuse of *a* with non-specific definites).

Crucially, in the case of superlatives (though not in the case of ordinals), there was still a small but significant difference in degree of *the* use with specific vs. non-specific DPs. This suggests that the L2-learners' intuitions concerning specificity to some degree override explicit teaching instruction.

4.1.3. Individual performance

The specificity distinction with both definites and indefinites was observed at the level of individual L2-learners. As shown in Section 3.5, nearly a third (20) of the L2-

learners made a four-way distinction between specific definites, non-specific definites, specific indefinites, and non-specific indefinites: these learners showed fluctuation between the two settings of the Article Choice Parameter. Another large group (21) of the L2-learners had converged or nearly converged on the Definiteness Setting of the parameter. Only two learners incorrectly converged on the Specificity setting of the Article Choice Parameter – this is expected, since the input should lead L2-learners to choose the Definiteness rather than the Specificity setting. The two learners presumably mis-analyzed the input. Since these two learners had quite high proficiency (high intermediate and advanced), one cannot easily argue from their data that the Specificity setting is a "default" setting employed at the start of acquisition (see also the discussion in Chapter 5).

Among the learners who showed fluctuation, only two showed strong leaning in favor of the Specificity setting, with the rest either showing true fluctuation or leaning towards Definiteness. Learners who fell in the Miscellaneous patterns typically showed patterns resembling Definiteness or Fluctuation, consistent with the predictions.

The FH cannot account for the behavior of the eight learners who showed the "Strange" pattern – i.e., who made a much greater specificity distinction with indefinites than with definites, or vice-versa. For the three of these learners who made the specificity distinction only with indefinites, one can speculate that they considered definites like *the murderer of X* to be specific even when there was denial of speaker knowledge (since *being a murderer* can be construed as noteworthy in and of itself). However, there is no independent evidence for this explanation, or for why some other learners in the "Strange" pattern made the specificity distinction with definites only. The performance of these eight learners remains a puzzle.

Finally, this chapter showed that there was a relationship between proficiency and the ability to set the Article Choice Parameter: advanced learners tended to be more accurate than intermediate learners in both specific indefinite and non-specific definite contexts, and learners who exhibited the Definiteness Pattern, or who were fluctuating in the direction of Definiteness, were on average more advanced than learners in a true state of fluctuation. This evidence suggests that as proficiency increases, L2-learners are able to set the Article Choice Parameter (although many advanced learners still show fluctuation). In contrast, such variables as age of first exposure and time of residence were not good predictors of performance.

4.2. Degree of speaker knowledge

I turn next to the predictions concerning overuse of *the* with different types of indefinites. These predictions are repeated in (93) and the results are summarized in (94) and (95) for the L1-Russian and L1-Korean speakers, respectively.

As shown by (94), the data from the L1-Russian speakers support the predictions in (93). These speakers draw a three-way distinction between indefinites that are clearly specific ((20), (21)), indefinites that may be construed as specific (25), and indefinites that are clearly non-specific (23). The higher the likelihood that an indefinite is specific, the greater the overuse of *the* in these speakers' L2-English.

93. Predictions for article use in L2-English: overuse of *the* and its relation to the
degree of speaker knowledge

	scopes over an intensional operator	no scope interactions
explicit speaker knowledge	category in (20) **high overuse of *the***	category in (21) **high overuse of *the***
minimal or no speaker knowledge	category in (25): minimal speaker knowledge **some overuse of *the***	category in (23): no speaker knowledge **no overuse of *the***

94. Overuse of *the* with different types of indefinites: results from L1-Russian
speakers

	scopes over an intensional operator	no scope interactions
explicit speaker knowledge	high overuse of *the*	high overuse of *the*
minimal or no speaker knowledge	some overuse of *the*	almost no overuse of *the*

95. Overuse of *the* with different types of indefinites: results from L1-Korean
speakers

	scopes over an intensional operator	no scope interactions
explicit speaker knowledge	high overuse of *the*	lower overuse of *the*
minimal or no speaker knowledge	high overuse of *the*	almost no overuse of *the*

The results of the L1-Korean speakers in (95) support the predictions in (93) only partially. As predicted, these learners show significantly less overuse of *the* with indefinites that are clearly non-specific (23) than with indefinites that may be construed as specific ((20), (21), and (25)). However, these learners show an unexpected pattern of *the* overuse between the three specific indefinite categories: they draw no distinction between the two categories where the indefinite scopes over an operator ((20) and (25)), overusing *the* to an equally high degree in both, regardless of the degree of speaker knowledge. On the other hand, they overuse *the* significantly less in the specific indefinite category with no scope interactions (21).

The fact that the L1-Korean speakers overuse *the* more when the indefinite scopes over an operator (20) than when it takes wide scope by default (21) parallels a similar finding in Study2 (see Chapter 5): in that study, both L1-Russian and L1-Korean speakers overused *the* with specific indefinites more often when an indefinite took scope over an operator than when there were no scope interactions. I suggested that L2-learners may be more likely to mark an indefinite as specific (i.e., overuse *the*) when this allows them to disambiguate the (intensional) context in favor of a wide-scope reading of the indefinite. The same explanation holds for the results of the current study, although there is no explanation for why this effect holds only with L1-Korean speakers in the present study

(especially since it held with both L1-groups in the previous study). Crucially, both groups distinguish between specific and non-specific indefinites: even in the case of no scope interactions, both L1-Russian and L1-Korean speakers overuse *the* more in the specific indefinite category (21) than on the non-specific indefinite category (23).

The next question is why L1-Korean speakers do not draw a distinction between unembedded indefinites (20) and embedded indefinites (25), while L1-Russian speakers do draw such a distinction. I hypothesized in Section 1.2 that in the cases of embedding, the indefinite is less likely to be construed as specific because the referent is not directly known to the speaker. However, the context *may* be construed as specific, either because the speaker has indirect knowledge of the referent, or because the speaker is reporting the speech of the referent of the matrix subject, and the indefinite is specific from the point of view of this referent. It is possible that L1-Korean speakers are more likely than L1-Russian speakers to evaluate specificity from the point of view of the referent of the matrix subject (rather than that of the speaker). However, this is unlikely given the results with *certain*-indefinites, discussed in the next section. The differential performance of the L1-Russian and L1-Korean speakers thus remains a puzzle.

To conclude, various categories of indefinites differ as to their likelihood of being interpreted as specific. While specificity is a binary distinction, whether a given DP is classified as [+specific] vs. [-specific] is dependent on various discourse factors, such as the speaker's degree of knowledge and the speaker's desire to single out a particular individual. As discussed in Chapter 2, specificity markers such as *this*$_{ref}$ in English or *le* in Samoan appear to be largely optional, with their use determined by the speaker's state of mind. Faced with the task of deciding whether an indefinite context in the forced choice elicitation task is specific, L2-English learners have to decide whether the speaker is intending to single out a particular individual via some noteworthy property. There is some variation in how individual L2-learners make this decision. L1-Russian and L1-Korean speakers appear to attend to slightly different properties of the context in evaluating the context as specific vs. non-specific; this remains a puzzle. However, the general pattern for both L1-groups is that greater expression of speaker knowledge in the context causes greater overuse of *the* on the part of the L2-learners: neither group overuses *the* with indefinites that are clearly non-specific.

The data in Section 3.4.3 further show that different types of indefinites have different likelihoods of being considered specific: L2-English learners show highest overuse of *the* with indefinites which involve clearly stated speaker knowledge, and lowest overuse of *the* with narrow-scope indefinites as well as indefinites which involve no speaker knowledge. First-mention indefinites which involve neither explicit statements nor explicit denials of speaker knowledge pattern in between, with significantly lower *the* overuse than specific indefinites and slightly (but not significantly) higher *the* overuse than non-specific indefinites. This shows that in the absence of any explicit statement of speaker knowledge, an indefinite is more likely to be construed as non-specific than as specific. This is consistent with the fact that *this*-indefinites in English (which are obligatorily specific) are infelicitous in contexts involving no statement of speaker knowledge.

4.3. The role of *certain*

I next consider the effect of the presence of *certain* in an indefinite. The predictions, which treat *certain* as an obligatory marker of specificity, are given in (96), and the summary of results are given in (96) and (97) for the L1-Russian and L1-Korean speakers, respectively.

96. Effect of embedding vs. presence of *certain*: predictions

	explicit speaker knowledge (widest scope)	minimal speaker knowledge (embedding)
indefinite headed by *a*	category in (20) **high overuse of** *the*	category in (25) **lower overuse of** *the*
indefinite headed by *a* **certain**	category in (27) **high overuse of** *the*	category in (28) **high overuse of** *the*

97. Effect of embedding vs. presence of *certain*: summary of results of the L1-Russian speakers

	explicit speaker knowledge (widest scope)	minimal speaker knowledge (embedding)
indefinite headed by *a*	high overuse of *the*	lower overuse of *the*
indefinite headed by *a* **certain**	high overuse of *the*	lower overuse of *the*

98. Effect of embedding vs. presence of *certain*: summary of results of the L1-Korean speakers

	explicit speaker knowledge (widest scope)	minimal speaker knowledge (embedding)
indefinite headed by *a*	high overuse of *the*	lower overuse of *the*
indefinite headed by *a* **certain**	lower overuse of *the*	almost no overuse of *the*

The results do not support the predictions. Both L1-Russian and L1-Korean speakers distinguish between those contexts containing *certain* which involve explicit speaker knowledge (27) and those that do not (28). The L1-Russian speakers treat *certain*-indefinites exactly like *a*-indefinites: regardless of the presence of *certain*, these learners overuse *the* to a high degree if speaker knowledge is explicit and to a lower degree if the indefinite is embedded under an attitude verb and speaker knowledge as minimal.

As for the L1-Korean speakers, they actually overuse *the* less with *certain*-indefinites than with *a*-indefinites. As already discussed in the previous section, these learners do not draw a distinction between the *a*-indefinites in (20) vs. (25): whether or not the indefinite is embedded under an attitude verb, as long as some speaker knowledge is present, the indefinite is treated as specific, and *the* is overused. On the other hand, these speakers do draw a distinction between embedded vs. unembedded cases of *certain*-indefinites.

It should be noted that the L1-Korean speakers' performance may to some degree be influenced by explicit instruction: L1-Korean speakers learners of English are typically

instructed that *certain* should be obligatorily used with *a* (Heejeong Ko, p.c.; see Chapter 8 for more discussion). Russian speakers are not known to receive such instruction. This instruction may account in part for the lower overuse of *the* with *certain*-indefinites than with *a*-indefinites on the part of the Korean speakers. Interestingly, despite this instruction, these learners still overuse *the* with *certain*-indefinites in unembedded environments (27), suggesting that L2-learners' intuitions concerning specificity to some degree override explicit instruction.

Crucially, neither group of learners treats *certain* as an obligatory marker of specificity: both groups distinguish between *certain* indefinites in widest-scope environments (27), where speaker knowledge is explicitly stated, and embedded environments (28), where speaker knowledge is minimal. This suggests that while *certain* is compatible with specific indefinites in L2-English, it does not obligatorily signal specificity: in evaluating an indefinite as specific or non-specific, L2-learners pay attention to such contextual cues as the degree of speaker knowledge rather than to the presence of *certain*.

This has interesting implications for the discussion of what determines specificity. As discussed earlier, Abusch and Rooth (1997) showed that when an indefinite headed by *a certain* is embedded under an attitude verb, it can be anchored to the attitude of the matrix clause: for instance, in (99a), *a certain* is anchored to story report; in (99b), of *a certain* is anchored to Claude's beliefs.

99. a) There was a story in Spy about Solange. According to the story, she has moved to a certain remote island in the Pacific. I don't know which one, it was some exotic-sound place.
b) Claude evidently believes that Solange is involved with a certain ballet dancer. I have no way of telling who this dancer is supposed to be.

(Abusch and Rooth 1997, ex. 74-75)

If *certain* is an obligatory marker of specificity, then Abusch and Rooth's findings need to be incorporated into the lexical entry for specific DPs: specific indefinites, such as those headed by *a certain,* can reflect the attitude of the matrix clause rather than that of the actual speaker. (As discussed in Chapter 2, a possible way of reconciling these facts with the discussion of specificity is to adopt Schlenker's (2002, 2003) proposal for attitude verbs as manipulators of context variables).

The data from the L2-English learners suggests that specificity is preferentially anchored on the actual speaker: both groups of L2-learners overuse *the* more when *certain*-indefinites are anchored on the attitude of the speaker (27) than when they are embedded and potentially anchored on the attitude of the matrix clause (28); L1-Russian speakers also differentiate between embedded and unembedded *a*-indefinites ((20) vs. (25)). This suggests that in L2-English, specific DPs are much more likely to reflect the attitude of the speaker rather than the attitude of the referent of the matrix subject. An interesting question is whether similar patterns are observed cross-linguistically. For instance, it is likely that *this*-indefinites in L1-English corpora anchor on the attitude of the matrix speaker more often than on the attitude of the embedded speaker. It would be interesting to see whether the proportions of anchoring on matrix speaker vs. embedded speaker are similar between L1-English and L2-English.

5. Conclusion

The main contribution of the study described in this chapter was to show that the specificity distinction is present for both definites and indefinites in L2-English. The Fluctuation Hypothesis has received further support. L2-English learners exhibit optionality in use of *the* vs. *a* precisely in those contexts in which the definiteness and specificity parameter settings are in conflict: the [+definite, -specific] and [-definite, +specific] contexts.

The specificity distinction in indefinites is generally more accepted in the literature than the specificity distinction in definites. Much literature, reviewed in Chapter 2, has been devoted to the different readings of indefinites in English. Languages such as English and Hebrew have specificity markers for indefinites only (see Chapter 2 for more discussion).

Not as much evidence exists for the specificity distinction with definites. As discussed in Chapter 2, there is no clear motivation for positing an ambiguity of English definites, and there is not much cross-linguistic evidence in favor of the specificity distinction in definites (but see Chapter 2 for some suggestive evidence from Samoan). The L2-English data discussed in this chapter provide evidence that the specificity distinction does in fact play a role with definites as well as indefinites. As shown in this chapter, L2-English learners fluctuate between the two systems of article groupings represented in (100).

100. a) Article grouping by definiteness   b) Article grouping by specificity

	+definite	-definite
+specific		
-specific		

	+definite	-definite
+specific		
-specific		

The pattern of article choice in L2-data is clearly not random. This pattern cannot be accounted for by transfer (see the discussion in Chapter 3). On the other hand, it can be accounted for if we take both patterns in (100) to be possible UG options of article groupings.

The grouping in (100a) is clearly a possible UG option, one instantiated in a variety of languages, including English. While not much evidence exists for the cross-linguistic reality of (100b), this article grouping is a logical possibility for natural language: as discussed in Chapter 2, both definite and indefinite DPs can be specific (when the felicity condition of noteworthiness is satisfied) or non-specific. A language which has no marker of definiteness but does have a marker for specificity should in principle be able to use the specificity marker with both definites and indefinites – but only as long as they are specific. I suggested in Chapter 2 that Samoan is an example of such a language. The L2-English data provide additional support for the view that the article grouping in (100b) is a possible UG option.

Chapter 7: Production data

This study describes the corpus of written production data that was collected from adult L2-English learners as part of the third study of article choice in L2-English. The data are intended to supplement the findings of the elicitation studies described in the previous chapters. They provide further support to my proposal that specificity plays a role in L2-English article choice.

1. Background and predictions

The three studies described in Chapters 4 through 6 provide evidence that L2-English learners fluctuate between dividing English articles on the basis of definiteness vs. on the basis of specificity. These conclusions were based on the results of three elicitation tasks. It is important to know whether the patterns seen in the elicitation tasks are also found in the actual production data of L2-English learners.

As part of the third study, reported in Chapter 6, we therefore also collected samples of written production data from L2-English learners and analyzed article choice in the data.

1.1. Predictions for L2-English article use in production

The prediction is that the patterns of article choice in production will be similar to article choice in elicitation tasks. The Fluctuation Hypothesis for article choice that I have been arguing for predicts that we should see overuse of *the* with specific indefinites and overuse of *a* with non-specific definites in production as well as elicitation.

Given this hypothesis, as well as the findings of the elicitation tasks reported in Chapters 4 through 6, I can therefore make the predictions in (1); here, the symbol '--' refers to article omission, which is appropriate with indefinite plural and mass nouns.

1. Prediction A for article choice in production in L2-English

	+definite	-definite
	Target: *the*	Target: *a* (sg), -- (pl, ms)
+specific	correct use of *the*	overuse of *the*
-specific	overuse of *a*, --	correct use of *a*, --

sg = singular count noun   pl = plural count noun   ms = mass noun

The main advantage of production data is that it gives us a better indication than the elicitation studies of how L2-English learners use articles in daily life. While the elicitation tasks focus the learners' attention on article use, collection of production 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.

1.2. Written production data and article omission

Production data also allow us to examine article omission and see whether it corresponds to a particular semantic distinction. We saw the L2-learners almost never omitted articles with singular DPs in the elicitation studies. However, it is well-known, both anecdotally and from previous studies of L2-English (e.g., Huebner 1983), that L2-

English learners do omit articles to a fairly large extent in production. Thus, collecting production data would allow us to examine both article use and article omission in L2-English, while the elicitation studies largely focused on article use.

Article omission may occur in production 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 me is whether article omission will follow any particular linguistic pattern: whether articles are omitted more in some semantic environments than in others.

A possible prediction concerning article omission is expressed in (2). The rationale for this hypothesis is that the [+definite] article is more informative than the [-definite] article, since it carries the maximality presupposition. 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 (2) is inapplicable to plurals, since article omission with indefinite plurals is licensed in English.

2. Prediction B:
   L2-English learners omit articles more in singular [-definite] than in singular [+definite] environments in production.

The hypothesis in (2) assumes that L2-learners at least sometimes correctly divide English articles on the basis of definiteness: as seen in the preceding chapters, this was indeed the case. On the other hand, we have seen that learners also often divide articles on the basis of specificity. 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, any results concerning omission are at best approximate.

A final hypothesis regarding article omission concerns the difference between singular and plural DPs. As reported in Chapter 5, L2-learners exhibited more article omission with plural than with singular definites in elicitation. I proposed that this was an artifact of the test, and that no such pattern should occur in production, as formulated in (3). (This prediction is inapplicable to indefinites, since article omission is licensed with plural indefinites in English).

3. Prediction C:
   There will be no difference between singular and plural definites in article omission in production.

1.3. Determining specificity in production: specific predictions

The main disadvantage of production data collection 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 indefinite 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 production data as specific or non-specific: as discussed

previously, 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 non-specific, there are some cues that can be used. Scope is one of them: indefinites and definites which take narrow-scope under an intensional verb or modal are obligatorily non-specific, as discussed in the previous chapters. Wide-scope DPs may be specific or non-specific, on the other hand.

Certain uses of definites, such as definites on their *previous-mention* use, are obligatorily specific, as previously discussed. In Section 3.4, I will talk in more detail about other types of definites and their relation to specificity.

Finally, consider indefinites in *there* and *have* constructions. As illustrated in (4a) vs. (4b), specific indefinites headed by $this_{ref}$ are incompatible with *there*-constructions that simply list a series of descriptions. Of course, 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 (4c). The same is shown for indefinites in *have*-constructions in (5).

4. 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!

5. 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 (4c) and (5c) are quite felicitous, it is not possible to argue that specific indefinites are inherently incompatible with *there* or *have* constructions. However, the production task administered in our study aimed at eliciting *there* and *have* constructions of the type in (4a) and (5a), 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 chapter, I will treat *there* and *have* constructions as being biased in favor of containing non-specific indefinites, and will analyze them separately from all other wide-scope indefinite constructions[1].

The above discussion leads to specific predictions for article use in the L2-English production data, which are laid out in (6).

6. *Prediction for article use in production:*
   a) Definites in *previous-mention* contexts are specific: no overuse of *a* is expected.
   b) Wide scope indefinites may be specific or non-specific: some overuse of *the* is expected.

---

[1] A few of the *there* and *have* constructions in the data contained indefinites that could be construed as specific – e.g., when the speaker said *I have an X* and then followed this up with a description of the referent. Since there was no sure way of separating "potentially specific" from "most likely non-specific" indefinites in the data, and since the majority of indefinites in *there* and *have* constructions occurred with no follow-up information about the referent, I counted all indefinites in *there* and *have* constructions as "most likely non-specific".

c) Narrow-scope indefinites are non-specific: no overuse of *the* is expected,

d) Indefinites in *there* and *have* constructions are most likely to be non-specific: little or no overuse of *the* is expected.

e) Narrow-scope definites are non-specific: overuse of *a* is expected

The rest of this chapter is organized as follows. Section 2 describes the methodology of the production data collection and analysis. Section 3 explains how the different contexts in the data were classified, and discusses in more detail which contexts are specific, which are non-specific, and which are ambiguous. The actual results are presented in Section 4. Section 5 discusses the data, and Section 6 concludes the chapter.

2. Methods

The collection of the written production data was part of Study 3, which is described in Chapter 6. The participants were the same 30 L1-Russian speakers and 40 L1-Korean speakers who participated in Study 3. The characteristics of the participants and the testing procedure are described in the Methodology section of Chapter 6. Here, I describe the methodology that is specific to the collection of the production data.

2.1. The 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 (7), 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 production task tested their article use.

7. 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 (7a) was aimed at eliciting specific indefinites, since the learners would write about an object that was well-known and important to them. The question in (7c) was aimed at eliciting indefinites in *there* and *have* constructions, i.e., indefinites which are typically non-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 (7e) 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 (7b) and (7d) 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 (7), with the possible exception of (7e), were likely to elicit definites in previous-mention contexts.

2.2. Coding procedure

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

2.2.1. Coding procedure: the general principle

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 put in an article. The coders were then asked to fill the blank with the article that was most appropriate for the next.

The rationale for this 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 (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.

In many English sentences, both *the* and *a* would be appropriate, depending on the exact situation. Consider the sentences in (8). In each case, both *a* and *the* might be used felicitously. For instance, in (8a), the choice of *the* vs. *a* depends on the number of windows in the speaker's room. In (8b), *a salesperson* is fine because the speaker does not presuppose that the hearer shares her knowledge of a unique salesperson. However, *the* is also fine because the hearer can accommodate there being only one (salient) salesperson. Finally, in (8c), both *a bus* and *the bus* are perfectly felicitous, with the definite phrase having a possibly generic reading.

8. a) I put my bed next to a/the window.
   b) I went into a store. A/The salesperson asked if she could help me.
   c) I took a/the bus to New York.

If we ask L1-English speakers to fill in the article in sentences such as those in (8), some coders will probably put *the* and some will put *a*. We can then exclude these contexts from the count as ambiguous, since they do not inform us about whether L2-English article use is correct or erroneous. If an L2-learner has put *the* in a context such as (8a), for instance, we do not know whether the learner evaluated the context as definite, and therefore used *the* correctly, or evaluated the context as specific indefinite, and therefore used *the* incorrectly. On the other hand, if an L2-learner put *the* in a context which L1-English coders treated as indefinite (i.e., in which all coders put *a*), then we can treat the L2-learner's response as an error and evaluate it accordingly.

I will now discuss in more detail exactly how the L2-learners' narratives were prepared for coding by native English speakers.

2.2.2. *Preparation of the narratives for coding*

The coders' versions of the narratives were prepared as follows. Every article used by an L2-learner was replaced a blank. The only exception 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 *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 being a determiner like *many*, that, like *many*, 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: a period was inserted at the end of the last sentence in each narrative if it had not been there originally. There was one exception: in three instances, L1-Russian speakers used a comma to set off a relative clause that was clearly a restrictive relative. The examples are given in (9), with the DP containing the relative clause underlined. In all three cases, the most probable reading of the relative clause is restrictive; a reading of the relative clause as non-restrictive is highly infelicitous. In fact, on the restrictive reading, use of *the* in (9b-c) is felicitous, whereas on the restrictive reading, it is infelicitous. Treating the relative clauses in these examples as non-restrictive may therefore artificially inflate the number of incorrect *the* uses in L2-English.

9. a) I got very sudden gift on this Christmas from <u>person, who was my cousin's husband's friend</u>, that was a beautiful ring, made of gold and aquamarine, as a stone.

  b) Moreover, I was afraid not meeting <u>the person, who was responsible for us</u> – students from Russia.

c) I would be very happy to receive this sum, because exactly now I need to find an additional source of money to treat my poor lovely tooth, which I broke accidently. I would prefer to apply <u>the doctor, whom I met at the morning</u>.

If the sentences in (9) are meant to contain restrictive relatives, why did the L1-Russian speakers put commas before the relative clause? The answer is quite simple: Russian rules of punctuation require commas before *all* relative clauses, restrictive as well as non-restrictive. Thus, it is not surprising that L1-Russian speakers writing in English would apply the same rule to English restrictive relatives. To avoid artificial inflation of error counts, I took commas out of the three underlined DPs in (9) for the version that was given to the coders[2].

As an illustration of the changes that an L2-learner's narrative underwent before being given to a coder, consider (10). The passage in (10a) is an L1-Russian speaker's response to the question in (7b). The passage in (10b) 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*.

10. 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.

2.2.3. *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

---

[2] Of course, it is possible that in this way, the counts of correct uses of *the* were artificially inflated. However, it is preferable to inflate the number of correct uses rather than the number of errors, since the errors in this case (overuse of *the* with indefinites) are predicted by my hypothesis. It is important that results which provide support to the hypotheses are not artificially inflated.

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 add an *s* at the end of the head noun and treat the entire NP as a plural. This instruction was especially important for the coding of the L1-Korean speakers' narratives: since Korean does not have plurality marking, these speakers frequently omitted –*s* in clearly plural contexts.

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 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[3].

Crucially, the coders had no idea which article the L2-learners had actually put in any of the contexts. The coders' responses were thus completely unbiased, reflecting their own intuitions of article use.

3. Classification scheme

After the responses of all the coders had been collected and entered into a computer file, a post-coding procedure took place.

3.1. The post-coding procedure

Each context was classified as definite vs. indefinite, and, furthermore, as count noun singular vs. count noun plural vs. mass noun. This was done as follows.

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*

---

[3] 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. Notably, the non-native coders sometimes put *the* in the contexts in which all native English-speaking coders put *a*, but in which the original L2-English learners had put *the*. This suggests that the tendency to overuse *the* with specific indefinites (noted for L1-Russian and L1-Korean speakers) exists also for speakers of Hindi/Urdu, even ones who are otherwise completely fluent in English.

plural if each of the four coders put in -- ("no article") or *some.* The distinction between plurals 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 simply based on 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.

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.

All of the clear definite and indefinite contexts were then evaluated for type, as discussed in the next two sections.

3.2. Types of indefinite singular DPs

The contexts which native English speakers unambiguously coded with *a* (i.e., singular indefinite contexts) were classified into the eight types in (11). Detailed explanations and illustrations follow; illustrations are contexts in which L2-learners either correctly put *a* or omitted the article. For each indefinite context, I will also discuss whether it is more likely to be specific or non-specific.

> 11. **Types of singular indefinite DPs**
>   wide scope
>   *a certain*
>   narrow scope
>   *there* construction
>   object of have
>   predicative position
>   generic
>   *such/as* constructions

Some indefinites occurred in constructions that are likely to be formulaic, such as those shown in (12). This category also included indefinites which were exact copies of indefinites used in the original questions. Formulaic indefinites were excluded from the rest of the analysis.

12. a) It is <u>a pity</u>, but I cannot tell you something interesting about my recent vacations.

    b) They stayed at my home for <u>a week</u>.

3.2.1. *Wide scope indefinites*

Given the discussion in Section 1.3, I am separating indefinites in *there* and *have* constructions, which are likely to be non-specific, from all other wide scope indefinites, which may be specific. The context that I am calling *wide scope indefinite* therefore includes all wide-scope indefinites that are not contained in *there* and *have* constructions.

Most of the DPs classified as *wide scope indefinites* occurred in sentences with no intensional operators, such as (13a-c). There were a few cases of indefinites clearly scoping over an intensional verb or modal (13d-e). A few cases in which the indefinite was ambiguous between a wide-scope and a narrow-scope reading were excluded from the analysis.

13. a) I live in <u>a small room</u> in <u>a big private house in Brookline</u>.

    b) On last Christmas, I got <u>a new leather coat</u> from my relatives in Korea.

    c) In the airport one old lady with <u>a dog</u> greeted us with <u>a real warm smile</u>.

    d) Recently, we need $1000 to fill out $13,000 to buy <u>a used car that we have negociated</u>.

    e) Next day I visited UT at Austin and fortunately could meet <u>an american student who helped me all day</u>. I still miss her so far.

Wide-scope indefinites may be specific or non-specific. In some cases, such as (13b), (13d), and (13e), 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 (13e), the speaker may have in mind the noteworthy property *x gave me a lot of help and I still miss her.* In other cases, such as (13a) or (13c), no importance is attached to a particular member of the set denoted by the restrictor NP, be it the set of rooms, houses, or dogs.

Since it is impossible to tell for certain whether a given wide-scope indefinite context is specific or non-specific, I will treat all of the instances of this category as potentially, but not necessarily, specific.

3.2.2. *Indefinites headed by* a certain

A special subcategory of wide-scope indefinites was created for indefinites headed by *a certain.* There were very few instances of *a certain*-headed indefinites in production, and all of them were produced by a single L1-Russian participant. Examples are given in (14).

14. I think object became considerable and important for you, if it has <u>a certain value</u> for you or <u>certain meaning</u>.

Indefinites headed by *a certain* are obligatorily specific (but see Chapter 6 for how L2-learners treat indefinites headed by *a certain* in different types of contexts).

3.2.3. Narrow-scope indefinites

The category of *narrow scope indefinites* consists of indefinites take scope under an operator. Most of these occurred in answer to the question in (7e), which asked the L2-learners to speculate what they would do with a thousand dollars. Some answers are given in (15a-c). Narrow scope indefinites also occurred in other contexts (15d-e). In most cases, the narrow scope indefinites had *de dicto* readings, as a result of scoping under an intensional verb or modal. There were also a few cases of indefinites taking narrow scope with respect to a higher quantifier, as in (15e).

15. a) My answer will be very simple – I will by a car. In America you can't do anything without a car, even to do a shoping.
    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.

As previously discussed, narrow-scope indefinites are obligatorily non-specific.

3.2.4. Indefinites in there-*constructions*

The next category are indefinites in *there*-constructions. Most *there*-constructions occurred in descriptions of the narrator's room in answer to (7c) – an example is (16a). Some *there*-constructions occurred in other contexts, as in (16b). Also included in this category were a few indefinites that occurred locative copular constructions (16c) or as part of a sentence fragment that followed a *there*-construction in a previous sentence (16d).

16. 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) In my room are a big bad, a table, TV, and three cheas.
    d) 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.

As I argued in Section 1.3, indefinites in *there*-constructions of the type in (16a-b) are very likely to be non-specific. The same holds for indefinites in locative constructions such as (16c-d): for instance, in (16c), no particular importance is attached to a particular bed, table, or TV.

3.2.5. Indefinites as objects of have

Indefinites in *have* constructions also typically occurred in answer to (7c), as in (17a-b). Also included in this category were a few cases of indefinites in the *consist of* construction, since, like the *have*-construction, it denotes possession (17c).

17. a) My room is not big but not small. It has a table for studying and a chair.
    b) 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.

c) My bedroom is consist of <u>queen size bed</u>, <u>hanger</u>, <u>closet</u>, <u>dresser</u> and <u>big mirro</u>.

As discussed in Section 1.3, indefinite objects of *have* in "listing" contexts are likely to be non-specific.

3.2.6. Indefinites in predicative position

Predicative indefinites were indefinites in the predicative position of copular constructions. Since it is impossible to tell with certainty whether a given copular construction is predicative or equative, no distinction was made between the two; all post-copular indefinites were included in this category. Some examples are given in (18): while (18a-b) are more likely to be predicative copula constructions, (18c) may be a case of an equative copula.

18. a) As I am <u>a great traveler</u>, I will spend more than half of this amount to discover a new place in the U.S…
    b) Last Christmas, I visited parents' house with my husband. It's <u>a white Christmas</u> as everyone always wishes.
    c) During my last vacation I got a chance to travel to Florida. That was <u>a place I heard a lot about</u>.

Indefinites in predicative copular constructions (18a) are necessarily non-specific while indefinites in equative copular constructions (18c) may be specific: the speaker may, for instance 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*. Since it is often hard to tell whether a copular construction is predicative or equative, it is similarly hard to tell whether an indefinite in the post-copular position is specific or non-specific.

3.2.7. Indefinite generics

There were very few indefinites in the *indefinite generic* category. These indefinites typically occurred in what Kriffka et al. 1997 call *characterizing* sentences. Thus, (19a) characterizes Brio wooden trains, and (19b) characterizes people's rooms.

19. a) For example, <u>a Brio wooden train</u> costs almost half of $1000.
    b) I think room of <u>a person</u> reflects his personality in some way.

Indefinite generics are arguably non-specific, since they do not denote a particular individual: there is no particular Brio wooden train or person under discussion in the sentences in (19). On the other hand, there is evidence from Mosel and Hovdhaugen 1992 that Samoan groups generics in affirmative sentences together with specific DPs. It is not clear exactly how generics fit into the specificity discussion.

3.2.8. Indefinites in such a/ as a constructions

A number of indefinites occurred in *such a X* and *as a X* constructions, illustrated in (20). Since *such a* and *as a* could be treated as unanalyzed chunks by L2-learners, indefinites in these constructions were grouped into a separate category. Indefinites in these constructions do not bear any obvious relation to specificity.

20. a) I first arrived in the U.S 5 years ago, in September, 97 as <u>a tourist</u>.

b) I did not know such <u>a long flight</u> could become boring and exhausting.

3.3. Types of plural/mass indefinites

Plural and mass indefinites were divided into the six types in (21). These largely correspond to the types for singular indefinites, except that there were no mass or plural indefinites in predicative position, or with *certain*. There was also the category of *general concepts*, which will be discussed below. 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 used in categories other than *general concepts*.

21. **Types of singular indefinite DPs:**
   wide scope
   narrow scope
   there-construction
   object of have
   generic
   general concepts

Some formulaic uses of plural/mass indefinite DPs were excluded from the rest of the analysis.

3.3.1. Wide-scope plural and mass indefinites

As was the case with singulars, wide-scope plural/mass indefinites encompassed all cases of wide-scope indefinites that are not part of *there* or *have* constructions. Examples are given in (22).

22. a) From airport (here in Boston) I was taken by <u>representatives of my University</u> and driven to a hotel.

b) The object that is of value for me is my father's gift: <u>beautiful diamond earings</u>.

Just like singular wide-scope indefinites, plural/mass wide-scope indefinites may be specific or non-specific.

3.3.2. Narrow-scope plural/mass indefinites

Narrow-scope plural/mass indefinites, which are obligatorily non-specific, always took scope under an intensional operator, modal, or negation, as in (23).

23. a) For the rest of the money, I will buy <u>souvenirs of the visited places</u> and <u>gifts for my Mom and nephews</u>.

b) There's a few dogs for inspection to look for <u>weapons</u> or <u>drugs</u>.

c) [description of a skiing experience] Friend of mine show(n) me how to deal with speed, how to manage yourself without <u>injury</u>.

3.3.3. Plural/mass indefinites in there and have constructions

Examples of plural/mass indefinites in *there*-constructions and *have*-constructions are given in (24) and (25), respectively.

24. a) There only bed, iron table, some chairs, mirror and <u>closets</u>.
    b) On the floor, there are always <u>toys</u> here and there.

25. a) The room where I moved recently is very nice. It has <u>hardwood floors</u>, closet, <u>large windows</u> in it.
    b) We went to several springs, and had <u>BBQ parties</u> and <u>fishing</u> there.

As already discussed, indefinites in *there* and *have* constructions are likely to be non-specific.

3.3.4. Plural/mass indefinite generics

The category of *generic* plural/mass indefinites contained kind-referring and property-denoting NPs, as in (26). These do not directly relate to the specificity distinction.

26. a) I'm resting there, wath <u>TV</u> and <u>VIDEO's</u>.
    b) On Christmas, my husband and I played <u>golf</u>.

3.3.5. General concepts

The category of *general concepts* contained most cases of indefinite mass nouns – DPs that were impossible or difficult to classify under any other category. Examples are given in (27). Most DPs that fell into this category denoted abstract concepts, as in (27a). *General concept* indefinites are by definition non-specific, since they denote whole concepts or classes rather than individual (and noteworthy) objects.

27. a) That day, my husband and I were really excited and our hearts were filled with <u>joy</u> and <u>hope</u>.
    b) In my room are three windows thats why it is full of <u>light</u>.

3.4. Types of definites

Definite DPs (singular, plural, and mass) were classified into the ten types given in (28). 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-learner put in the target article, *the*.

28. **Types of definite DPs:**
    anaphoric
    unique by entailment
    associative use
    obligatorily unique
    narrow-scope
    temporal
    superlative
    ordinal
    special adjective
    generic

Not included in any of the above categories are proper names which take definite articles, such as *the United States*. These were not included in the analysis.

3.4.1. *Anaphoric uses of definites*

The *anaphoric* use of a definite is the basic previous-mention use – any case where the referent of the DP was explicitly mentioned in the previous discourse. Examples are given in (29).

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

As discussed previously, anaphoric definites are obligatorily specific: in (29a), 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*.

3.4.2. *Definites that are unique by entailment*

The *unique by entailment* definites are characterized by the presence of an NP complement. 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 maximality presupposition of the definite is satisfied because the NP complement narrows down the domain of the discourse sufficiently. Thus, in (30a), the speaker refers to the maximal set of doctors at a particular hospital; in (30b), the speaker refers to the fame of a particular city; and so on for the other cases in (30).

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

Definites which are unique by entailment may be specific or non-specific, depending on whether or not the speaker has in mind some noteworthy property. In such cases as (30a), 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 (30), 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, such as the contexts that we used in the last elicitation task (cf. *We are searching for the murderer of Smith, whoever that is*). However, such contexts are extremely unlikely to be produced in L2-data; in fact, I did not find a single definite (or, for that matter, indefinite) context in the data which involved explicit denial of speaker knowledge.

3.4.3. Associative use of definites

The *associative use* type 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 maximality presupposition: thus, in (31a), the reader is led to understand that the relevant sunset view is the one at Sarasota; similarly, in (31b), 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 31c-d). The reader typically understands that there is a unique and contextually salient airport, beach, etc., under discussion.

31. 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 non-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 felicitous. For instance, in (31a), the noteworthy property might be *x is the view at Saratoga,* and in (31c) it might be *x is the beach in Georgia where I traveled to.*

3.4.4. Definites with obligatorily unique referents

Definites which fall into the type *obligatorily unique* are, as their name suggests, definites with referents that are obligatorily unique in the actual world. As shown in (32), the internet was included in this category. Definites in this class are necessarily specific, since the narrator is necessarily referring to a very particular world or sun: the only relevant one in existence, from the perspective of Earth's inhabitants.

32. a) There so many places in the US and accross 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.

3.4.5. Narrow-scope definites

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 (33a), 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 (33b), 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 non-specific.

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

3.4.6. Temporal definites

The next four categories of definite descriptions are tied to particular lexical items. *Temporal* definite descriptions have head nouns which refer to particular time periods, such as *afternoon, evening, year*, etc., as shown in (34). These expressions were grouped into a special category, since they may be learned as unanalyzed chunks by L2-English learners. For instance, an ESL textbook by Raimes (1990:66) calls expressions like *in the afternoon* and *in the morning* examples of "idiomatic usage." It is not clear in what relation temporal definites stand to specificity.

34. a) The christmas season was the busiest time of <u>the year</u>.
    b) We came by airplane to New York in <u>the evening</u>...
    c) It was early in <u>the morning</u> and I had to do everything very fast as everybody in my family at that moment.

3.4.7. Superlatives and ordinals

The categories of *superlatives* and *ordinals* are self-explanatory, and are illustrated in (35) and (36), respectively. Superlatives and ordinals were grouped separately from other definites since they represent categories of definites on which L2-English learners often get explicit instruction.

35. a) For example, I have been in Artek – <u>the best summer camp in FSU</u> back in my school time.
    b) On July, 2001, I arrived at airport of Atlanta. Atlanta airport is one of <u>the biggest airports in U.S</u>.

36. a) It happened on <u>the third day of my arrival to this country</u>.
    b) I occupy <u>the 2nd floor</u> of a family house for 500$.

Superlatives and ordinals may be specific or non-specific. As shown by the results of the elicitation task in Chapter 6, L2-learners are very accurate at using *the* with superlatives and ordinals regardless of specificity, probably as a result of explicit instruction.

3.4.8. Special adjective modifiers on definites

The *special adjective* category of definite descriptions includes DPs modified by *next, last, left, right, front, back, exact, only,* and *other.* It also includes adjectival uses of *very*

(as in *the very beginning*), nominal uses of *other* (as in *the others*), and locative uses of *right, left, back,* and *front,* (as in *on the right*). Some examples are given in (37). These definites were grouped together since the adjectival modifier contributes to the maximality presupposition: we don't usually talk of *a left side* or *a next month.*

37. a) My room is quite cosy there is a queen-size bed on the right. On the left there is a desk.

b) It was a fabulous experience at least because I got a suntan in the very month of January!

c) The only thing I remember is that I was very frustrated with missing the plane to Gainesville because of the delay of Korean Airlines to Atlanta. I thought I had to stay somewhere in Atlanta. But fortunately, I could make it to aboard the next plane to Gainesville and I remember that I said sorry to my friend who met me at the airport.

Like superlatives and ordinals, definites in the *special adjectives* category may be specific or non-specific. However, as with superlatives and ordinals, the effects of specificity may be obscured by explicit instruction: L2-learners may have memorized combinations like *the next, the last, the right,* etc.

3.4.9. ` Definite generics

Finally, there was a category of *definite generics*. There were very few examples in this category, and it is questionable whether all of them are true kind-denoting DPs. Definite descriptions from the narratives were classified as definite generics if the author was referring not to a particular individual described by the definite, but rather to a class of individuals. Some examples from the narratives are given in (38). In these cases, the reader understands that the narrator is not talking about a particular piano or a particular beach. In (38a), the speaker is stating that she used to play a particular type (or class) of musical instruments. In (38b), the speaker is stating that his family wants to spend Thanksgiving week-end at some location that belongs to the class of beaches.

38. a) I used to play the piano before I came here.

b) We plan to go to the beach on next Thanksgiving weekend.

Definite generics do not fall neatly into the category of either specific or non-specific definites.

3.5. Summary and predictions

As the above discussion shows, most types of definites and indefinites are in principle ambiguous between specific and non-specific readings. Narrow scope with respect to an intensional verb or modal is a property of non-specific DPs, while certain types of definites (e.g., *anaphoric use* definites) are obligatorily specific.

In (39) and (40), I list the predictions concerning specificity and article use for indefinites and definites, respectively. I leave out generics and other exceptional contexts. The predictions in (39) and (40) expand on the original prediction in (1). The overuse of *a* discussed in (40) is relevant only for singular definites. The counterpart for plural and mass-denoting definite descriptions would be article omission.

39. *Prediction for article use and misuse with indefinites:*

    a) Wide scope indefinites may be specific or non-specific: some overuse of *the* is expected.

    b) Narrow-scope indefinites are non-specific: no overuse of *the* is expected,

    c) Indefinites in *there* and *have* constructions are most likely to be non-specific: little or no overuse of *the* is expected.

    d) Indefinites in the post-copular position are non-specific, except in equative copular constructions. Overuse of *the* is expected as long as it is tied to equative copular constructions.

40. *Predictions for article use and misuse with definites:*

    a) Definites on the anaphoric use are specific: no overuse of *a* is expected.

    b) Definites that are unique by entailment may be specific or non-specific; in the absence of a denial of speaker knowledge, these definites are likely to be specific, so not much overuse of *a* is expected.

    c) Definites on the associative use are most likely specific: little or no overuse of *a* is expected.

    d) Definites with obligatorily unique referents are specific: no overuse of *a* is expected.

    e) Narrow-scope definites are non-specific: overuse of *a* is expected.

    f) Superlatives, ordinals, and definites in the *special adjectives* category may be specific or non-specific: some overuse of *a* is expected.

4. Results

This section reports the data on article use and omission in all of the categories described in Section 3. 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.

4.1. Articles in indefinite contexts

First, I will look at article use across all indefinite contexts in the data. These are contexts that L1-English coders considered indefinite – i.e., singular contexts in which all coders put *a* and plural contexts in which all coders put no article or *some*.

The numbers are reported in (41) and (42) for L1-Russian and L1-Korean speakers, respectively.

41. Article use across all indefinite contexts: L1-Russian speakers

article used by L2-learners:	Indefinite context, by type; target: *a*			
	singular	plural	mass	total
the	19	4	4	27
a	113	3	5	121
-- (null)	109	65	28	202
%*the* use	7.88%	5.56%	10.81%	7.71%
%*a* use	46.89%	4.17%	13.51%	34.57%
%omission	45.23%	90.28%	75.68%	57.71%

42. Article use across all indefinite contexts: L1-Korean speakers

article used by L2-learners:	Indefinite context, by type; target: -- (null)			
	singular	plural	mass	total
the	16	5	5	26
a	192	3	3	198
-- (null)	114	27	28	169
%*the* use	4.97%	14.29%	13.89%	6.62%
%*a* use	59.63%	8.57%	8.33%	50.38%
%omission	35.40%	77.14%	77.78%	43.00%

As these tables show, the L2-learners either correctly used *a* or omitted the article in most singular contexts. In plural/mass contexts, article omission (which is appropriate in those contexts) was prevalent. Crucially, there was also some overuse of *the* across all types of indefinite contexts, for both groups. I now discuss whether there was a relationship between *the* overuse and specificity.

4.1.1. Article use in indefinite contexts: a detailed breakdown

The tables in (43) and (44) report article use for singular and plural/mass indefinites, respectively, in the main types of indefinite contexts discussed in sections 3.2 and 3.3.

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

article used by L2-learners	wide scope	narrow scope	*certain*	object of *have*	*there-* construc- tion	pred	generic	*such/ as*
L1-Russian								
the	13	2	0	0	3	1	0	0
a	*17*	*18*	*2*	*32*	*17*	*18*	*2*	*7*
null	16	4	2	50	22	13	1	1
L1-Korean								
the	8	4		0	0	1	3	0
a	*37*	*35*		*42*	*42*	*28*	*2*	*6*
null	15	27		27	22	20	2	1

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

article used by L2-learners	wide scope	narrow scope	object of *have*	*there-* construction	general concept	generic
L1-Russian						
the	4	1	0	1	1	1
a	*1*	*3*	*2*	*0*	*1*	*1*
null	34	16	13	5	17	8
L1-Korean						
the	3	4	0	0	0	3
a	*2*	*1*	*2*	*0*	*1*	*0*
null	3	20	4	2	23	3

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.

4.1.2. Overuse of the *with wide-scope indefinites*

Examples of *the* overuse in wide-scope indefinite contexts, singular as well as plural, are given in (45) and (46) for L1-Russian and L1-Korean speakers, respectively; (45g) also includes some of the rare cases of *the* overuse in *there*-constructions. In all cases, there is no previous mention of the underlined DP's referent.

In a few cases of *the* overuse, use of *the* does not appear to be completely infelicitous: in (45f), for instance, *the custom* (meaning *the customs*) seems fine even though L1-English coders uniformly put no article in that context; in (46g), where L1-English speakers also uniformly put no article, *the packages* might be accommodated if one imagines packages as being an obligatory item in an airport.

However, in most cases of *the* overuse in the data, use of *the* is truly infelicitous: for instance, in (45a), there is no way that the reader can presuppose the existence of a unique kitten. In (45b), the narrator has multiple healthy teeth and has undoubtedly eaten multiple sweets. In (46b), the reader has no knowledge of the romantic memo in question. In (46c), the reader cannot identify a unique ball or a unique baseball player. And so on. The contexts are obviously indefinite. Crucially, many of them are also specific: in most cases, the author of the text can single out a particular individual by naming something noteworthy about it. For instance, in (45a), the kitten was the author's much-loved pet; in (45d), the author has a very particular secret place in his back yard in mind; in (45f), the author knows exactly what the problems were; and so on. Contexts which are less likely to be specific include (45e) and (45g), where the identity of a particular hotel room, or a particular set of books, are quite irrelevant.

45. *uses of* the *in wide-scope environments: L1-Russian participants*
> a) When I was living in Ulan-Ude yet unmarried my friends presented me <u>the small seamese kitten</u>.
> b) I lost <u>the health tooth</u>, and I have realized after some time how it was valuable for me. It happened unexpectedly – I bit off <u>the solid sweet</u> 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 <u>the small party</u>.
> 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 <u>the secret place</u> in our yard, and then at home.
> e) On Thanksgiving week-end we went to NY for the first time. We took <u>the room</u> 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 <u>the long line</u> in order to get through the custom.
> g) I have beautiful room. There are <u>the big sofa</u>, <u>the table</u>, <u>the chairs</u>, TV and shelves with <u>the books</u>.

46. *uses of* the *in wide-scope environments: L1-Korean participants*
    a) I received <u>the frame with picture</u> for Chritmas from my roommate.
    b) [discussion of how the author's husband gave her a bible] When he gave me the bible, he attached <u>the memo which was written about the his love about me</u>.
    c) The most valuable object that I have received is <u>the ball</u> and the signature of <u>the famous baseball player</u> is signed on it.
    d) I went to <u>the cruise to Bahamas</u> with my mother.
    e) In New Orleans, we visited a couple of well-known places and had a traditional cusine at <u>the restaurant in French quarter</u>.
    f) We have been looking for a really good used car. Eventually we got the information, but <u>the problems</u> came out.
    g) Through the window from airplanes, I saw African-American men working with <u>the packages</u>.

While overuse of *the* in the *there*-construction in (45g) is unexpected, it is not completely infelicitous: the narrator may be attaching great importance to the particular sofa, table and chairs that she owns. However, this is unlikely, since the context is a "listing" context which does not favor specificity; an L1-English speaker would probably be unlikely to use *this*ref in this context, compared to the other contexts in (45) and (46). In fact, the vast majority of indefinites in *there*-constructions in the data, unlike (45g), did not appear with *the*.

4.1.3. Overuse of the *with narrow-scope indefinites*

Next, I examine overuse of *the* in contexts where it is completely unexpected – narrow-scope environments, which are obligatorily non-specific. The complete lists of the instances of *the* use in narrow-scope contexts for L1-Russian and L1-Korean speakers are given in (47) and (48), respectively. Most of the answers are the L2-learners' descriptions of how they would spend a thousand dollars.

In some cases, even though the context looks like a narrow-scope indefinite context, it may not be: for instance, in (48c), 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 (48f), the narrator may need to buy his child a particular set of clothes (e.g., a school uniform). In (48b), 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 (47b) and (48d), contain what look like possible definite generics, referring to the *class* of helicopters or planes. Only a few cases ((47a, c), (48a, e, g)) are unambiguously narrow-scope and hence non-specific.

47. *uses of* the *in narrow-scope environments: L1-Russian participants*
    a) If it is happen I'll spend money for <u>the trip to California or Florida</u>. I'm tired for winter this year.
    b) I would go for example to the Grand Canyon and use the chance to fly over it in <u>the helicopter</u>.
    c) I'd spend some money for travelling anywhere. And another part I'd spend on <u>the gifts for my sons and daughter-in-law</u>.

48. *uses of* the *in narrow-scope environments: L1-Korean participants*

    a) First of all, I will donate $100 as an offering to my church. With the rest of it, I will take a trip to <u>the famous place or city</u> and some historical cites, which especially were cited in the Bible.

    b) I really want to visit <u>the famous city in a European country which has old history</u>. The exact place will be depended on the time which I have or the flight price.

    c) And I will spend the rest of money ($200) for my daughter. like buying her <u>the bike</u> and dolls.

    d) I remeber that first day I was very tired. it's first time I spent over twenty hours in <u>the plane</u>.

    e) I wanna give my church $1000 as a donation for the missionaries. I hope the money would be spent to <u>the valuable work or people</u>.

    f) After get that money, I'll buy a new car and buy <u>the clothes for my child</u>.

    g) First of all, I will donate a tenth of $1000 to the church. Then, I will buy <u>the gifts for my son and wife</u>.

    h) It feels always fresh to walk around <u>the new places</u>.

4.1.4. Overuse of the in other indefinite contexts

Instances of *the* overuse in other contexts were fairly rare[4]. There were two instances of *the* overuse in predicative position. These are given in (49a) and (49b), from an L1-Russian speaker and an L1-Korean speaker, respectively. Both may be cases of specific indefinites in equative copular constructions: in (49a), 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 (49b), 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.

49. *uses of* the *in predicative position:*

    a) On the last summer I went to the Prague. It is <u>the very beautiful, ancient city</u>.

    b) On my last birthday, I receive a special gift from one of my best friends. It was <u>the watch made by Swiss army</u>, which I always wanted to have but haven't because of its high price.

Next, I consider indefinite generics which involved overuse of *the*. Most of the examples of *the* overuse in this category come from L1-Korean speakers; they are given in (50) and (51) for singular and plural/mass indefinites, respectively. The singular generics in (50) may actually be compatible with *the* if they are construed as definite kind-denoting generics. No particular pattern is observed among the plural/mass generics in (51).

50. *uses of* the *with indefinite singular generics: L1-Korean speakers*

    a) One of the desks has the shape of "L" and the others have the shape of <u>the square</u>.

---

[4] For discussion of one instance of *the* overuse with a numeral, see Appendix 4. This is instance is not included in the counts reported in this chapter.

    b) I understand that airport should be well controlled, however, it looked like that the first visitor [*meaning:* the first-time visitor] can feel uncomfortable to some degree.

    c) I heard that the round trip to Europe cost $1000.

51. *uses of* the *with indefinite plural/mass generics: L1-Korean speakers*

    a) I'm living in a studio type of the apartment[5].

    b) X-box video game console was always on the top of my wish lists. It can support the high definition TV, in-game dolby digital surround, and also can play the DVD if you buy the separate accessory for that[6].

    c) I definitely remember the day when I arrived on here. I have seen oversea in the TV, but I've never had experience to live or visit foriene country.

The single case of *the* overuse among L1-Russian speakers in a context classified as generic is given in (52). It is not in fact clear whether this context should be treated as generic (kind-denoting) or as wide-scope indefinite – whether the narrator is intending to refer to the kind *other people* or to some particular other people (e.g., the narrator's friends and relatives).

52. *use of* the *in a potentially generic context: L1-Russian speaker*

    I've lost my value to the other people when I'd come to America.

I leave generics aside for the remainder of this chapter. A study of article choice with generics in L2-English would be an interesting topic for a separate investigation.

4.1.5. Statistical comparisons of article use with singular indefinites

In this section, I will look at whether *the* was overused significantly more often with specific than with non-specific indefinites. I will treat wide-scope indefinites as (potentially) specific, and narrow-scope indefinites, as well as indefinites in *there* and *have* constructions, as non-specific. The exceptional cases of predicative and generic indefinites are not included.

First, I looked at singular indefinites and computed the χ^2 statistic[7] on use of *the* vs. use of *a*. The tables in (53) report 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 significant for both L1 groups: both groups used *the* more in wide-scope contexts than in *there/have* constructions.

The table in (54) reports use of *the* vs. *a* for wide scope vs. narrow scope contexts. This distribution was significant only for the L1-Russian speakers.

---

[5] The noun *apartment,* which is normally a count singular noun, was classified as a mass noun in this particular case, since the L1-English coders uniformly omitted the article before it. On this generic use, *apartment* refers to the class of apartments rather than to a particular one.

[6] The L1-English coders treated *DVD* as a plural noun, omitting articles before it and sometimes adding the plural suffix *-s.*

[7] 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.

53. 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

54. 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

4.1.6. Statistical comparisons of article use with all indefinites

Next, I 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. The tables in (55) compare use of *the* vs. use of *a*/omission for wide scope vs. *there/have* contexts; for both L1 groups, the distribution is significant.

55. 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

The tables in (56) compare wide scope and narrow scope contexts; the distribution is marginally significant for L1-Russian speakers and non-significant for L1-Korean speakers.

56. 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

4.2. Article use in definite contexts

This section looks at article use with definites. Only those contexts which L1-English coders considered definite – i.e., where all coders put *the* – are considered.

First, I group all of the definite categories described in Section 3.4 together. The data are given in (57) and (58) for L1-Russian and L1-Korean speakers, respectively.

57. Article use across all definite contexts: L1-Russian speakers

article used by L2-learners:	Definite context, by type			
	singular	plural	mass	total
the	112	7	5	124
a	3	0	0	3
-- (null)	26	3	1	30
%*the* use	79.43%	70.00%	83.33%	78.48%
%*a* use	2.13%	0.00%	0.00%	1.90%
%omission	18.44%	30.00%	16.67%	19.11%

58. Article use across all definite contexts: L1-Korean speakers

article used by L2-learners:	Definite context, by type			
	singular	plural	mass	total
the	163	19	22	204
a	8	0	0	8
-- (null)	34	3	5	42
%*the* use	79.51%	86.36%	81.48%	80.31%
%*a* use	3.90%	0.00%	0.00%	3.15%
%omission	16.59%	13.64%	18.52%	16.54%

As the above tables show, L2-learners appropriately used *the* in the vast majority of all definite contexts. Notably, there are no noticeable differences in the rates of article omission between singular, plural, and mass definite contexts. There is very little overuse of *a*; as predicted, *a* overuse occurs in singular contexts only.

4.2.1. Article use with different types of definites

The table in (59) reports article use in singular contexts for the ten categories of definites discussed in Section 3.4, and the table in (60) reports article use in plural/mass contexts. Since *a* was never used with plural definites, only use of *the* vs. article omission are examined for plurals.

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

article used by L2-learners	anaphoric	entailment	associative use	obligatory uniqueness	narrow scope	temporal	superlative	ordinal	special adjective	generic
L1-Russian										
the	18	18	28	2	2	5	12	10	17	
a		*1*					*1*			*1*
null	6	3	12	1		3			1	
L1-Korean										
the	52	23	33	4	4	6	12	10	16	3
a	*1*	*1*	5		*1*					
null	6	4	8	1			1	4	7	3

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

article used by L2-learners	anaphoric	entailment	associative use	obligatory uniqueness	narrow scope	temporal	superlative	ordinal	special adjective	generic
L1-Russian										
the	3	2	5				1		1	
null		**1**	**2**						**1**	
L1-Korean										
the	9	9	13	3	2	0	2		3	
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.

4.2.2. Overuse of a with definites

I will now look at how article use with definites relates to specificity. The cases of *a* overuse for L1-Russian and L1-Korean speakers are given in (61) and (62), 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 non-specific is the narrow-scope definite in (62h): 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 superlative in (61c) may also be non-specific. The previous-mention definite in (62a) is clearly specific. The other contexts below may be specific or non-specific – we cannot know one way or the other without knowing exactly what the speaker had in mind. In (62c), it is likely that the author did not consider the nurse's identity importance; on the other hand, in (62f), 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 (62b-d), in particular.

61. Cases of *a* use with definites for L1-Russian speakers:

a) *entailment context: may be non-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.

b) *superlative context: may be non-specific*
    In this case I will try to help with this money all my frends and relatives. But part of this money I would spend for Russian church. I think this is a best way to spend this money.

c) *generic context*

I didn't feel good until I began understand conversation on a train, at <u>a street</u>, in the class.

62. Cases of *a* use with definites for L1-Korean speakers:

a) *anaphoric context: must be specifict*

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 <u>a baby</u>.

b) *entailment context: may be non-specific*

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

associative contexts: may be non-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 <u>a nurse</u> 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 <u>a dormitory</u>.

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

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

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

h) *narrow scope context: must be non-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 <u>a full collection that he needs</u>.

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 the table in (59), 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*). In the Korean speakers' data, one of the superlatives and one of the "special adjective" definites also had narrow-scope, and both were used with *the*. None of the items in these two categories were narrow-scope in the Russian speakers' data.

4.2.3. Definiteness and article omission

Finally, I will take a look at article use vs. omission in singular contexts. I computed the χ^2 statistic for article use (*the+a*) vs. article omission with definites vs. indefinites in all singular contexts. As shown in (63), 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.

63. Article use vs. omission with definites vs. indefinites in singular contexts

L1-Russian speakers		
	definite	indefinite
the+a	115	132
omission	26	109

$\chi^2 = 27.94$***

L1-Korean speakers		
	definite	indefinite
the+a	171	208
omission	34	114

$\chi^2 = 21.96$***

***$p<.001$

5. Discussion

I will now examine the predictions made in Section 1, looking at whether they are supported by the results of the production data.

5.1. Article use and misuse in production

In this section, I look at whether L2-English article use in production followed the patterns established by the elicitation studies.

5.1.1. Article use with indefinites in production

The specific predictions for article use with indefinites are repeated in (64).

64. *Prediction for article use and misuse with indefinites:*
a) Wide scope indefinites may be specific or non-specific: some overuse of *the* is expected.
b) Narrow-scope indefinites are non-specific: no overuse of *the* is expected,
c) Indefinites in *there* and *have* constructions are most likely to be non-specific: little no overuse of *the* is expected.
d) Indefinites in the post-copular position are non-specific, except in equative copular constructions. Overuse of *the* is expected as long as it is tied to equative copular constructions.

These predictions were largely supported. In accordance with (64a), there was some overuse of *the* with wide-scope indefinites. There was practically no overuse of *the* in *there* and *have* constructions, as predicted by (64c). While some overuse of *the* occurred with narrow-scope indefinites, contrary to (64b), 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 (see Section 4.1.3). As for (64d), we did

indeed see a couple of cases of *the* overuse in the predicative position; in both cases, the indefinite was likely to have a specific reading.

Overuse of *the* was higher with wide-scope indefinites than with narrow-scope indefinites or with indefinites in the *there* and *have* constructions, whether use of *the* vs. *a* or use of *the* vs. use of *a* + omission was measured. These differences were significant (or at least marginally significant) in all cases with the exception of the wide scope / narrow scope comparison for the L1-Korean speakers. In this last case, the differences went numerically in the direction that was predicted, but did not reach significance.

Thus, in the case of article use with indefinites, the conclusions of the elicitation studies are further supported by the findings from the production data: L2-English learners do indeed overuse *the* with indefinites, and overuse of *the* occurs most frequently with those indefinites that are likely to be construed as specific.

5.1.2. Article use with definites in production

The specific predictions concerning article use with definites are repeated in (65).

65. *Predictions for article use and misuse with definites:*
   a) Definites on the anaphoric use are specific: no overuse of *a* is expected.
   b) Definites that are unique by entailment may be specific or non-specific; in the absence of a denial of speaker knowledge, these definites are likely to be specific, so not much overuse of *a* is expected.
   c) Definites on the associative use are most likely specific: little or no overuse of *a* is expected.
   d) Definites with obligatorily unique referents are specific: no overuse of *a* is expected.
   e) Narrow-scope definites are non-specific: overuse of *a* is expected.
   f) Superlatives, ordinals, and definites in the *special adjectives* category may be specific or non-specific: some overuse of *a* is expected.

As shown by the results, L2-English learners overwhelmingly used *the* across all definite contexts in production. This was predicted for previous-mention definites (65a), definites on their associative use (65c) and definites with obligatorily unique referents (65d). Definites which are unique by entailment were also overwhelmingly used with *the*. As stated in (65b), these definites were quite likely to be specific: the speaker typically had knowledge of the definite DP's referent and could state something noteworthy about it.

Narrow-scope definites were the only category on which overuse of *a* was predicted (65e); however, narrow-scope definites were extremely rare in the data, so the prediction cannot be tested.

As for superlatives, ordinals, and definites in the *special adjective* category, L2-English learners overwhelmingly used *the* with them. Since nearly all of the items in this category were wide-scope definites, we have no way of knowing whether they were considered specific or non-specific by the learners.

Moreover, as discussed in Chapter 6, L2-English learners often receive explicit instruction on using *the* with superlatives and ordinals; the learners may have similarly learned that words like *next, right,* and *left* obligatorily take *the,* regardless of the context. Thus, performance on these contexts is not necessarily indicative of the L2-learners' intuitions concerning definiteness and specificity.

5.1.3. Summary: L2-English article use and Prediction A

The main prediction concerning article choice in production is repeated in (66).

66. Prediction A for article choice in production in L2-English

	+definite	-definite
	Target: *the*	Target: *a* (sg), -- (pl, ms)
+specific	correct use of *the*	overuse of *the*
-specific	overuse of *a*, --	correct use of *a*, --

sg = singular count noun   pl = plural count noun   ms = mass noun

We see that Prediction A was partially confirmed by the production data. L2-English learners did indeed overuse *the* in [-definite, +specific] context to a much greater degree than in [-definite, -specific] contexts. They also used *the* correctly in [+definite, +specific] contexts.

Unfortunately, the last cell in the table, [+definite, -specific] contexts, could not really be tested. If we disregard superlatives and ordinals (on which L2-English learners may receive explicit instruction), two kinds of definites are necessarily classified as non-specific: narrow-scope definites and wide-scope definites involving explicit denial of speaker knowledge. There were very few instances of the former in the data, and none of the latter.

This makes an interesting point. As we saw in the section on indefinites, there were plenty of non-specific indefinite contexts in the data – both narrow-scope indefinites and indefinites in *there/have* constructions. In contrast, there were very few non-specific definite contexts. This probably stems from the nature of the task, since the learners were asked to talk about their personal experience. When talking about their experience, L2-learners are probably more likely to use specific DPs than non-specific DPs, since specificity is related to speaker knowledge. Non-specific indefinites occurred in abundance because (1) they occur in *there/have* constructions when L2-learners describe their room; and (2) they occur in narrow-scope contexts, when L2-learners write about how they would spend $1000. Neither context is designed to elicit non-specific definites. In fact, it is far from obvious what kind of a scenario would induce L2-learners to write about objects whose uniqueness they can establish but to whose identity they attach no importance: i.e., objects that would be denoted by [+definite, -specific] DPs.

5.2. Article omission in production

Next, I examine the predictions concerning article omission in L2-English, and the corresponding results.

5.2.1. Prediction B: article omission and definiteness

Prediction B, which concerns article omission, is repeated in (67).

67. Prediction B:
    L2-English learners omit articles more in singular [-definite] than in singular [+definite] environments in production.

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 production than in a controlled 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 information. Thus, article omission, like article use, shows that L2-English learners are aware of article semantics.

5.2.2. *Prediction C: article omission and number*

Prediction C, which concerns article omission with definites, is repeated in (68).

68. Prediction C:
     There will be no difference between singular and plural definites in article omission in production.

This prediction was supported: the L2-learners omitted articles to roughly the same extent across singular and plural definite contexts (although there were not many instances of plural definite contexts altogether). This suggests that the differential rates of omission with singular and plural definites reported in Chapter 5 were in fact an artifact of the format of the elicitation task.

6. Conclusion

The production data reported in this chapter largely supported the conclusions drawn on the basis of the elicitation task. The data showed that overuse of *the* with specific indefinites holds in production as well as elicitation, and was not an artifact of the elicitation studies. While overuse of *a* with definites was not replicated in the production data, this quite likely stemmed from the fact that non-specific definite contexts are rare in production.

In general, production studies of article choice in L2-English (e.g., Huebner 1983, Thomas 1989) have found overuse of *the* with indefinites much more often than overuse of *a* with definites. This is not surprising: the data from our study suggests that L2-English are much more likely to use specific indefinites (talking about objects with which they are well familiar) than to use non-specific definites (which require very specialized scenarios with denial of speaker knowledge). Non-specific definites can, however, be elicited in more controlled settings, such as the elicitation study reported in Chapter 6. While production data can give us an indication of how L2-English learners use articles in daily life, such data do not provide information on the full range of contexts that can be tested in a more controlled study.

Chapter 8: Conclusion

1. Summary

In this thesis, I have examined articles in the elicitation and production data of adult L1-Russian and L1-Korean learners of English, and have made the following points:

1. L2-English article use is not random but reflects access to the semantic features of *definiteness* and *specificity*, which play a role in article semantics cross-linguistically.
2. The role of specificity in L2-English article choice cannot be attributed to either L1-transfer or L2-input. L2-learners' access to this semantic feature provides evidence for direct UG-access in L2-acquisition.
3. L2-learners fluctuate between the two settings of the *Article Choice Parameter*, sometimes distinguishing articles on the basis of definiteness, and sometimes – on the basis of specificity. This fluctuation is seen at both group level and individual level, and provides evidence in favor of the Fluctuation Hypothesis: L2-learners fluctuate between possible parameter settings until the input leads them to converge on the target setting. There is no evidence for a "default" setting of the Article Choice Parameter.
4. There is some evidence that fluctuation decreases with proficiency. A number of (mostly advanced) L2-learners are able to set the Article Choice Parameter to the target value.

The elicitation studies reported in Chapters 4 and 5 showed that L2-learners show optionality in article use with specific indefinites, while being more accurate in their article use with non-specific indefinites and with (specific) previous-mention definites. The elicitation study reported in Chapter 6 showed that L2-learners draw the specificity distinction with definites as well as indefinites, as predicted by the combination of the Article Choice Parameter and the Fluctuation Hypothesis. The production data reported in Chapter 7 supplemented elicitation studies by showing that overuse of *the* with specific indefinites occurs in production as well as elicitation (but there were not enough contexts in the data for testing article use with non-specific definites).

The proposal that I have been arguing for is falsifiable. If learners had shown a random pattern of article use, or a pattern that does not exist in any natural language, this would have been problematic for my proposal. If the learners had made the specificity distinction only with indefinites and not with definites (or vice-versa), this would have undermined the predictions of the Article Choice Parameter. As it was, very few learners in the third study showed what I termed the "Strange" pattern of making the specificity distinction only with indefinites or only with definites (see Chapter 6).

The Fluctuation Hypothesis is similarly falsifiable. One could imagine a different pattern of acquisition data, one consistent with the Article Choice Parameter but not with the FH: e.g., if L2-learners initially adopted the Specificity setting of the Article Choice Parameter and then switched abruptly to the Definiteness pattern. We would then expect to see some learners showing the Specificity pattern and others showing the

Definiteness Pattern. However, as reported in Chapters 5 and 6, very few learners showed the Specificity pattern. While about a third of the learners in both studies had converged (or almost converged) on the Definiteness setting of the parameter, approximately as many showed fluctuation between the two parameter setting (the remainder of the learners showed a variety of patterns, some of them resembling fluctuation but with more noise).

While the data reported in the previous chapters support my proposal, it is also conceivable that these data can be given a different explanation. In the next section, I will go through a list of possible alternative explanations, and give arguments for why these explanations do not successfully account for the data.

2. Alternative explanations

One alternative explanation, that of transfer, was already considered in Chapter 3. That chapter provided evidence that nothing in Russian or Korean can guide L2-learners to associate *the* with specificity. I now consider several other alternative explanations, and argue against them.

2.1. Permanent parameter mis-setting

One alternative explanation, which I considered briefly at the end of Chapter 4, concerns the reality of fluctuation between two parameter settings. The *parameter mis-setting* explanation would say that my proposal is on the right track, and that L2-English learners do in fact divide English articles on the basis of specificity. Under this explanation, however, they *always* do so: L2-English learners always use *the* with [+specific] DPs and always use *a* (or omit articles) with [-specific] DPs.

2.1.1. The proposal

This proposal is summarized in (1).

1. The Parameter Mis-setting Explanation:
    L2-English obligatorily *the* is [+specific]

    a) previous-mention definites, definites with clearly stated speaker knowledge
       obligatorily [+specific] → correct use of *the*

    b) indefinites which involve some speaker knowledge:
       sometimes treated as [+specific] → overuse of *the*
       sometimes treated as [-specific] → correct use of *a*

    c) indefinites with no speaker knowledge, or narrow-scope readings:
       obligatorily [-specific] → correct use of *a*

    d) definites with denial of speaker knowledge
       sometimes treated as [+specific], depending on the DP → correct use of *the*
       sometimes treated as [-specific] → overuse of *a*

On this view, L2-learners correctly use *the* with previous-mention definites because these definites are obligatorily [+specific] (1a). In contrast, they alternate between *the*

and *a* use with indefinites because there is no such thing as an obligatorily specific indefinite. After all, specificity is determined by the speaker's state of mind. Whether the speaker considers an indefinite specific or non-specific depends on whether she wishes to draw attention to a particular individual. As we saw in Chapter 6, even the modifier *certain* is not an obligatory marker of specificity in L2-English. Thus, the greater the likelihood that an indefinite is considered specific, the greater the overuse of *the*. L2-learners often correctly use *a* with indefinites that I have classified as [+specific] because, in fact, the learners often consider these indefinites to be [-specific]. This is summarized in (1b). Indefinites that are obligatorily [-specific] – e.g., narrow-scope indefinites – obligatorily take *a* (1c). Finally, in the case of definites that involved denial of speaker knowledge, L2-learners may sometimes consider them [+specific], contrary to what I have claimed. For instance, on seeing a phrase such as *the murderer of Smith – whoever that is*, they may decide that the property of being a murderer is sufficiently noteworthy to warrant use of *the*. Other times, they may decide that there is not enough information given about who the murderer is, treat the DP as [-specific], and use *a*. This is summarized in (1d).

2.1.2. Problem: development

This explanation has a problem. It does not allow for the possibility of gradual parameter re-setting: it predicts that L2-learners always treat *the* as being [+specific]. However, we saw that some (mostly advanced) L2-learners use articles in a target-like manner. The parameter mis-setting proposal would have to say that these advanced learners have successfully reset the parameter and classified *the* as [+definite].

Thus, on this view, there are learners who treat *the* as [+specific], following the pattern in (1), and other learners who treat *the* as [+definite]. Under the *parameter mis-setting* explanation, any learner who shows even the smallest degree of optionality in article use must have set the Article Choice Parameter to the Specificity setting: this setting allows the learner to sometimes consider a given context specific (and use *the*) and sometimes consider it non-specific (and use *a*). In contrast, the Definiteness setting allows no optionality: a learner who has set the parameter to this setting will always use *the* with definites, and with definites only. The *parameter mis-setting* explanation distinguishes between learners who show target-like article use (the Definiteness Setting), and those who show some optionality of article use (the Specificity setting). Crucially, there is no intermediate stage – no learners who are fluctuating between the two parameter settings. The switch from one parameter setting to the other is abrupt rather than gradual.

However, there is evidence from the empirical data that the switch is in fact gradual. This evidence comes from developmental effects. As we saw in Chapter 6, among the L1-Russian speakers, advanced L2-learners were less likely than intermediate L2-learners to overuse *the* with specific indefinites, as well as to overuse *a* with non-specific definites. Yet very few L2-learners showed absolutely target-like performance (i.e., use of *the* with definites only, and use of *a* with indefinites only).

When individual patterns were considered in Chapter 6 (for both L1-Russian and L1-Korean speakers), we saw that learners who showed relatively little fluctuation were more advanced, on average, than learners who showed a high degree of fluctuation. Yet, from the *parameter mis-setting* perspective, learners who show even the smallest

degree of optionality in article choice at stuck at the Specificity setting of the parameter. Then why should the optionality of article choice decrease with proficiency? This view would be forced to say that as L2-learners become more proficient, they continue to treat *the* as a marker of specificity, but their tendency to evaluate a context as [+specific] changes. Moreover, it changes in a very peculiar way: they become *more* likely to evaluate definite contexts as [+specific] and *less* likely to evaluate indefinite contexts as [+specific]. (Recall that the third study, reported in Chapter 6, found an inverse correlation between use of *the* with non-specific definites and use of *the* with specific indefinites for L1-Russian speakers). This is tantamount to saying that as a learner's proficiency increases, she starts considering properties like *murderer of x, mother of x,* and *owner of x* to be very noteworthy, and at the same time starts considering properties like *red-haired girl from New York, green scarf that I lost,* and *friend from college* as not particularly noteworthy. It is very difficult to see why proficiency should have any effect on which properties a learner considers noteworthy.

Thus, the *parameter mis-setting* explanation cannot provide a logical explanation for why overuse of *the* with specific indefinites, as well as overuse of *a* with non-specific definites, should decrease with proficiency. It has especial difficulty explaining why use of *the* increases with non-specific definites at the same time that it decreases with specific indefinites. On the other hand, these facts are easily captured by the proposal of gradual parameter re-setting.

2.2. Disregard of hearer knowledge

An explanation that takes the opposite stance of the *parameter mis-setting* explanation is one that says that specificity plays no role in L2-English article choice: L2-learners obligatorily treat *the* as [+definite]. The reason they sometimes overuse *the* with indefinites is that they treat these indefinites as definites.

2.2.1. The proposal

This explanation would run as follows[1]: L2-learners are under a great deal of pressure as they learn their second language. One way in which they might minimize this pressure is by focusing on their own state of knowledge and ignoring the state of knowledge of their listener.

Thus, whenever the referent is known to the speaker (the L2-learner herself), she uses *the,* regardless of whether or not the referent is also known to the hearer; she assumes that the hearer shares her knowledge. This is similar to the explanation frequently given for *the* overuse in L1-English (see Maratsos 1976, Matthewson and Schaeffer 2000, among others). When this explanation is given for *the* overuse in child L1-English, it makes reference to young children's egocentricity. Adult L2-learners are not expected to be any more egocentric than adult L1-English speakers. However, we might think that under the stress and the processing demands of speaking a foreign language, L2-learners would neglect to always carefully evaluate their listener's state of knowledge.

---

[1] The possibility of such an alternative explanation for *the* overuse with indefinites was first brought to my attention by David Pesetsky.

2.2.2. *Problem #1: elicitation data*

While this *disregard of hearer knowledge* explanation can explain overuse of *the* in production, it cannot easily explain overuse of *the* in forced choice elicitation. Recall that in the elicitation tasks, L2-English learners had to evaluate a conversation between two individuals, e.g., A and B; one of these individuals, e.g., A, utters the target sentence, and the learner has to fill in the appropriate article in A's utterance.

The *disregard of hearer knowledge* explanation would have to say that the learners focus on A's state of knowledge, disregarding B's state of knowledge. It is no longer a case of the learner ignoring her listener's knowledge and focusing on her own. Rather, the learner has to identify with a particular participant in the discourse, such as A, and disregard the state of knowledge of the other participant, B. It is not clear why stress and processing difficulties should cause the L2-learner to identify with a particular participant in the discourse.

2.2.3. *Problem #2: non-specific definites*

More importantly, the *disregard of hearer knowledge* explanation cannot easily account for overuse of *a* with definites. Since this explanation assumes that L2-English learners obligatorily treat *the* as [+definite], it predicts that learners should have no difficulty using *the* with all types of definites. Consider the conversation in (2), and suppose that the L2-learner identifies with speaker B, disregarding the state of knowledge of speaker A. Well, speaker B clearly knows that the restaurant has only one owner, and wants to complain to that owner – whoever that is. The underlined DP is unambiguously definite, and the L2-learner should correctly use *the*. However, we saw in Chapter 6 that L2-learners overuse *a* in precisely such [+definite, -specific] contexts as (2).

2. A: What's wrong?

   B: This food is spoiled. I want to complain to <u>the owner of this restaurant</u> – whoever that is.

Thus, while the *disregard of hearer knowledge* explanation can explain overuse of *the* with indefinites, it cannot explain overuse of *a* with definites.

2.3. Computation of uniqueness

Another alternative explanation (suggested to me by Mike Tannenhaus, p.c.) is related to the nature of the uniqueness/maximality presupposition. As discussed in Chapter 2, the domain in which uniqueness is computed is determined by the discourse. The *computation of uniqueness* explanation would say that L2-English learners obligatorily treat *the* as [+definite], i.e., as marking uniqueness, but that they are not as good as native English speakers at computing uniqueness.

2.3.1. *The proposal*

Take the conversation in (3). Speaker B can felicitously utter *the cat* even though there is clearly more than one cat in the world, and no cat has been mentioned in the immediately preceding discourse. What licenses the use of the definite?

3. A: Where are you going?

   B: I just heard <u>the cat</u> come in. I better go feed him.

One scenario in which use of the definite is licensed is if A and B have shared knowledge: for instance, they are members of the same family, and this family has only one cat, Shaw. A will naturally take B's utterance to be about Shaw.

Alternatively, A might be a friend or neighbor who just popped in for a visit, and who doesn't know that B owns a cat. Yet, it's still fairly felicitous for B to utter *the cat* in (3); B can count on A accommodating the fact that B has a unique cat and is talking about that particular cat.

Take, on the other hand, the conversation in (4). Here, it is only felicitous for B to use the definite expression *the cat* if A and B have shared knowledge of some unique, salient cat – e.g., the one that's always hanging out in the park. If A has never heard of any particular cat in the park before, B's utterance is infelicitous, and A's reaction will be along the lines of "What cat?" B cannot easily count on A to accommodate B's acquaintance with a particular cat.

4. A: What did you see in the park today?
   B: I saw the cat running after some pigeons!

The difference between (3) and (4) is a fairly subtle one. In each case, the speaker has to decide what the contextually relevant domain is in which uniqueness is computed, and whether the hearer shares knowledge of the unique referent in this domain, or can be reasonably expected to accommodate this knowledge. While in (3), B can expect A to accommodate the fact that B owns a unique cat, in (4), B cannot expect A to accommodate the existence of a unique and salient cat in the park.

The computation of uniqueness is a fairly complex process, and one might expect that L2-English learners are just not as good at it as native English speakers. Not having articles in their L1, speakers of Russian and Korean might not know exactly how the contextually salient domain for computing the uniqueness presupposition on *the* is determined.

Take an airport scenario like (5), an example of a [-definite, +specific] context from our elicitation studies. A native English speaker will not use *the* with the underlined DP. The security guard has no knowledge of the particular red-haired girl that the visitor is talking about, and the relevant domain – the airport – is likely to include more than one red-haired girl.

5. Security guard: How can I help you?
   Visitor: I am looking for a red-haired girl. She flew in from New York.

Suppose that the L2-English learner has difficulty establishing the relevant domain in which uniqueness is computed. Suppose that the L2-learner mistakenly takes the relevant domain to be the set of people at the airport that the visitor in (5) is looking for. Since the visitor is clearly looking for a unique and salient red-haired girl, the referent of the underlined DP will be unique in this domain. The L2-learner will therefore use *the*.

2.3.2. Problem: non-specific definites

Like the *disregard of hearer knowledge* explanation in the previous section, this *computation of uniqueness* explanation can account for overuse of *the* with indefinites, but not for overuse of *a* with definites. A non-specific definite such as the one in (2) has an obligatorily unique referent. The L2-learner should have no problem computing

uniqueness with non-specific definites, and should obligatorily use *the* with them. However, we saw in Chapter 6 that L2-learners overused *a* with non-specific definites. Overuse of *a* occurred (to a high extent) even with such definites as *the mother of X*, which have indisputably unique referents.

2.4. Explicit instruction and strategies

An often advanced non-linguistic explanation for article errors in L2-English is concerned with explicit instruction. Since L2-English learners receive some classroom instruction on article use, they may formulate explicit strategies for article use – strategies that may lead them to make errors. For instance, if instruction emphasizes *the* as a marker of referents that are well-known to the speaker, L2-English learners may adopt the strategy of overusing *the* with specific indefinites. The learners' errors on this view are not traceable to UG access or mis-set parameters, but to explicit, non-UG-based strategies.

I will now discuss what instruction L2-English learners get concerning English articles, and will argue that strategy-based explanations cannot easily account for the patterns that we see in the data.

2.4.1. Textbook instruction on article choice

Textbooks of English as a Second Language (ESL) generally emphasize that use of *the* requires hearer/reader knowledge as well as speaker/writer knowledge. For instance, a textbook by Azar (1993:122) advises ESL students that "A noun is definite when *both the speaker and the listener* are thinking about the same specific thing"[2] – and gives examples of *the* use in this context; on the other hand, "Indefinite nouns are actual things (not symbols), but they are not specifically identified" (Azar:122).

Similarly, an L2-English textbook by Raimes (1990:58) advises that "When you are trying to decide whether to use *a* or *an, the,* or no article at all, one of the crucial distinctions to make is whether the noun phrase refers to something actual and specific *for both you and your reader* either inside the text in front of you or outside it." On the other hand, "When we refer to someone or something that will not be actual and specific *for the reader and listener*, we do not use *the*" (Raimes 1990:60). In a book for L2-English learners writing in English, Raimes (1992:86) also says that "A *specific* reference is *known by the writer and by the reader* as something unique, specific, familiar, or previously identified to the reader" and advises use of *the* in these contexts. Thus, both speaker and hearer knowledge are emphasized in the instruction on how to use *the*.

It should be noted that ESL textbooks generally devote very small sections (just a few pages) to a review of English articles, stating a few general rules for article usage and a few examples, without extensive discussion. A Russian-language textbook by Leventhal (2000), which is aimed at Russian speakers who are learning English in the U.S., similarly spends only five pages on articles. Most of these pages are devoted to such specific rules concerning articles as putting *the* before plural last names *(the Joneses)*. The main distinction between *a* and *the* is discussed in just a couple of lines, as follows:

---

[2] The emphasis in this and other quotations in this section is mine.

6. a) "A (AN) – neopredelennyj; stavit v rjad s emu podobnymi; kakoj-to odin, nekotoryj, ljuboj; vpervye upomjanutyвj; odin iz gruppy podobnyx.

   "THE - opredelennyj; individualiziruet; konkretnyj, izvestnyj, vot étot; ne v pervyj raz upomjanutyj; edinstvennyj v svoem rode."

   (Leventhal 2000:35)

   b) *Translation*
   "A (AN) - indefinite; places [the referent] in a row with others like it; an unspecified one, some, any; mentioned for the first time; one out of a group of its kind.

   "THE - definite; individualizes; specific, known, this one; previously mentioned; the only one of its kind."

Note that Leventhal's directives are somewhat confusing: should *the* be used when the referent has been previously mentioned, when it is known to the speaker, or both? Such terms as *specific* or *known* can be understood in a variety of ways, including *known to both speaker and listener* (as in all of the English-language textbooks cited above) and *known to the speaker only*. If the L2-learners are adopting the latter use of *specific*, using the strategy that *specific to me = the*, this could lead them to overuse *the* in specific indefinite contexts. The textbooks' emphasis on familiarity and (in the case of Leventhal) speaker knowledge might also conceivably account for overuse of *a* with definites – if L2-learners adopt the strategy of using *the* only with referents that they are familiar with.

2.4.2. General problems with strategy-based explanations

The problem with this strategy-based explanation is that it has to explain why L2-learners prefer one strategy over others. There are in fact quite a few different strategies concerning article use suggested to L2-English learners.

For instance, Maclin (1987:59) suggests that *a(n)* is used "before an unidentified singular countable noun that is one example of its class" (as well as in some other attributive and generic contexts), while *the* is used as "reference backward to a noun already mentioned." Maclin (1987:59) points out that *the* is also used with a noun that shows "reference forward to an identification soon to be made, often by modifiers following the noun", giving examples such as *The dog that has been barking all day has finally stopped barking.* Note that if our L2-learners are following this advice, they would put *the* whenever they see RC or other modification. This would in fact explain why L2-English learners overwhelmingly used *the* with wide-scope RC-modified indefinites in our first study (see Chapter 4).

However, a strategy associating *the* with modification would have resulted in overuse of *the* in narrow-scope environments involving RC modification also. Yet we saw in Chapter 4 that this was not the case – the learners overwhelmingly put *a* in this context. A strategy-based explanation would have to somehow rank strategies, saying that the strategy "use *the* with specific referents" overrides the strategy "use *the* with RC modification." And of course, there is always the strategy that *the* should be used in the "*context known to both writer and reader*" (Maclin 1987:60). The idea that one's hearer/reader must be familiar with the referent in order for *the* to be licensed is emphasized across textbooks with more consistency than the idea that *the* is used to

refer to specific objects. If L2-learners relied on textbook-taught strategies only, we would expect them to use *the* primarily in previous-mention, definite contexts. There is no *a priori* reason to think that a strategy "use *the* when the object is familiar to me" should play an important role in L2-acquisition.

Jane Dunphy, director of the English Language Studies program at the Massachusetts Institute of Technology, notes (p.c.) that L2-learners typically get little instruction in article usage except at very advanced levels, and that they are usually instructed on the importance of *previous mention* as well as of specificity and post-modification.

On a final note, recall one context in which L2-learners do get explicit instruction – article use with superlatives. In Chapter 6, I provided references for explicit instruction concerning *the* use with superlatives and ordinals, and suggested that the higher use of *the* with superlatives/ordinals than with other definites is traceable to the L2-learners having received explicit instruction.

Instruction on article use with superlatives/ordinals is much clearer than instruction on any other aspect of article use: it provides learners with one clear directive (use *the* when you see a particular word form) rather than a lot of fairly vague directives (pay attention to what your listener/reader knows, and also to what is specific for you, and also to whether there is a reference forward to an identification, and also to whether the referent has been previously mentioned, etc.). Thus, superlatives/ordinals may be a domain where instruction plays a greater role in the learners' article choice. However, even so, there was a small but significant difference in use of *the* with specific vs. non-specific superlatives (though this was not found for ordinals). Thus, even when a strategy is very clear and explicit, L2-learners' intuitions override it to some extent: despite instruction that superlatives must take *the*, learners still differentiate them (albeit to a small extent) on the basis of specificity.

2.4.3. Specific problem with a strategy-based explanation: plurals

Even if a strategy-based explanation could be motivated, and could account for overuse of *the* with singular indefinites, it would have a problem accounting for L2-learners' article use with plural indefinites. (Thanks to Ora Matushansky, p.c., for pointing this out).

In discussing articles, textbooks focus on singular contexts, with hardly any mention of plurals. Usually, all that is said about plurals is that they take no article in contexts where their singular counterparts would take *a*. Examples on when *the* is used are usually confined to singular definites, although some textbooks (e.g., Raimes 1992), also give examples of *the* with plurals. Importantly, textbooks often emphasize *uniqueness* when talking about use of *the*: for instance, Raimes (1992:86), states that *the* is used with *specific reference* where "[a] specific reference is known by the writer and by the reader as something *unique,* specific, familiar, or previously identified to the reader." Leventhal (2000) makes an even stronger (and factually incorrect) statement, that *the* is used when the referent is *the only one of its kind* (see (6)). The concept of *uniqueness* is clearly inapplicable to plurals, and textbooks do not discuss any concept resembling *maximality,* which is applicable to plurals.

When discussing bare plurals, textbooks pay particular attention to generic uses. For instance, the only example of a bare plural in the "Articles" section of Raimes' (1990)

textbook is *Bicycles are popular in China* (p. 61). Hardly any discussion in the textbooks is devoted to a discussion of indefinite vs. definite uses of plurals in non-generic contexts.

Thus, textbooks provide L2-learners with much fewer strategies for article use with plurals than with singulars. In the case of singulars, an L2-learner might mistakenly decide that *the* = "unique for me" (rather than "unique for my listener"). With plurals, no such strategy is possible. L2-learners might adopt a strategy of article omission for plurals, since the textbooks emphasize the impermissibility of *a* with plurals. In the study reported in Chapter 5, we did indeed see that L2-learners frequently omitted articles with both definite and indefinite plurals (but this was not replicated for definites in production). However, the learners also showed the pattern of overuse of *the* with specific indefinites (compared to non-specific indefinites) with plurals, as well as with singulars. A strategy-based explanation cannot easily account for this fact.

2.4.4. Summary

To sum up, the burden of proof is on strategy-based explanations to spell out how and why the various strategies interact the way they do – and why the same strategies are adopted for both singulars and plurals, even though much more attention is paid to singulars in instruction.

Strategy-based explanations would also have to account for the finding that L2-learners who have received ESL instruction in a wide variety of institutions, in Russia, Korea, and the U.S., all appear to adopt very similar strategies – a surprising finding, considering how little space textbooks devote to English articles. As the following quote from Leventhal's (2000) textbook suggests, textbook instruction alone is insufficient for mastery of article use (Leventhal's quote is followed by the basic article usage rules given in (6)).

7. a) "Možno vypisat' massu pravil otnositel'no togo, kakoj imenno artikl' sleduet upotrebljat' v tom ili inom slučae, stol'ko zhe bydet isklučenij i ogovorok, no ot vsex somnenij vse ravno ne izbavit'sja. Čtoby ne ošibit'sja, nužhen očhen' bol'šoj opyt."

(Leventhal 2000:35)

b) *Translation*
"It is possible to write out a mass of rules concerning which article, exactly, it is necessary to use in any given case, there will be as many exceptions and stipulations, but it will still not be possible to get rid of all doubts. In order to avoid making mistakes, it is necessary to have a great deal of experience."

In the absence of any strategy-based explanation that can account for why certain strategies are preferred over others, I will assume that strategies are not responsible for the pattern of data that we find with L2-English learners. (See also the next section for more discussion of strategy-based explanations).

2.5. Strategies and use of *certain*

In this section, I will examine a particular strategy that might be employed by L2-English learners in their use of articles, and that may be related to transfer.

2.5.1. Proposal: a confusion of certain and definite

A particular strategy that L1-Russian speakers may employ in their article use was suggested to me by Katherine Crosswhite (p.c.) and concerns the word *certain*. One of the Russian translations for this word is *opredelennyj* – a word which can also be translated as *definite,* as in *opredelennyj artikl',* 'the definite article'. Suppose that L1-Russian learners of English evaluate each context in the elicitation task as being *opredelennyj* 'certain/definite' or *neopredelennyj* 'uncertain/indefinite'. If the context seems *opredelennyj,* they use the definite article *the*; otherwise, they use *a.* A context may seem *opredelennyj* if it is compatible with the actual use of the word *opredelennyj.* Since this word can mean 'certain', it is more or less compatible with specific indefinite contexts. Hence, overuse of *the* in the English of L1-Russian speakers is due to the confusion created by the fact that the same Russian word corresponds to both 'certain' and 'definite' in English.

2.5.2. Problem #1: differential performance on contexts with certain

The above explanation is most straightforwardly applied to contexts that actually contain the word *certain.* We saw that Russian speakers in fact overused *the* with *certain* to a fairly high extent. It is quite plausible that the Russian speakers' overuse of *the* in this context is due to direct translation. On the translation sheet given to Russian speakers in the elicitation studies, one of the translations provided for the word *certain* was in fact *opredelennyj.* If the learners viewed this word as meaning "definite", it is not surprising that they overused *the* in this context.

However, recall that in the last study, reported in Chapter 6, L1-Russian speakers actually differentiated between different contexts containing *certain* depending on the degree of speaker knowledge that was present in the context. The strategy of pairing *the* with *certain* is thus not all-determining. Some other processes, such as determining the degree of speaker knowledge, enter into the L2-learners' decision concerning article choice.

More importantly, the translation of *certain* as *opredelennyj* cannot directly influence performance in those specific indefinite contexts which did not contain the word *certain* (i.e., the majority of specific indefinite contexts in the studies). I turn next to whether these contexts could be indirectly influenced by the confusion of translation.

2.5.3. Problem #2: production data

The proposal that L1-Russian speakers associate *the* with *certain* may extend to contexts in which *certain* is not explicitly present. Take once again the airport scenario, repeated in (8). The L1-Russian speaker may reason as follows: "the visitor in (8) is looking for *opredelennaja devočka* 'a certain/specific girl'. Since the word *opredelennaja* is compatible with this context, I better use the *opredelennyj* 'definite' article – *the*."

8. Security guard: How can I help you?
    Visitor: I am looking for <u>a red-haired girl</u>. She flew in from New York.

This explanation might explain overuse of *the* in the elicitation tasks, where the learners have to consciously evaluate each context as definite or indefinite. However, the explanation is harder to apply to production: it would be forced to say that every

time that the learner plans to write a DP in a narrative, she pauses and considers whether this DP would, in Russian, be compatible with use of the word *opredelennyj*. It seems questionable that this actually takes place, especially since *opredelennyj* is a very formal lexical item which is not very compatible with the informal style of the L2-learners narratives.

2.5.4. Problem #3: non-specific definites

The confusion of *certain* and *definite* also cannot easily account for overuse of *a* with non-specific definites. Recall that L1-Russian speakers overused *a* to a fairly high extent with non-specific definites such as *the mother of X* in (9). Under the explanation in this section, the Russian speakers would be reasoning as follows: the speaker in (9) does not know anything about the mother of Sam's best friend; this mother is therefore *neopredelennaja* 'uncertain/non-specific'; therefore, I should use the *neopredelennyj* 'indefinite' article – *a*.

9. Sam is staying with <u>the mother of his best friend</u>. I have no idea who she is.

However, what does it mean to be a "non-specific mother"? While in Russian it is perfectly felicitous to talk about *opredelennaja devočka* 'a certain girl' in (8), it is rather infelicitous to talk about *neopredelennaja mat'* 'an uncertain (non-specific) mother'. The confusion of *certain* and *definite* cannot very easily account for overuse of *a* with definites[3].

2.5.5. L1-Korean speakers

The strongest argument against the '*certain* = *definite*' explanation is the data from the L1-Korean speakers. Like the L1-Russian speakers, the L1-Korean speakers overused *the* with specific indefinites and overused *a* with non-specific definites. However, Korean, unlike Russian, does not have a word that would correspond to both *certain* and *definite* (Heejeong Ko, p.c.). With no confusion stemming from the actual form of the word, Korean speakers would not be expected to adopt the strategy "use *the* whenever the word *certain* is felicitous." However, Korean speakers overused *the* with specific indefinites, including indefinites containing the word *certain*.

If anything, we could have expected the Korean speakers to adopt the opposite strategy of "use *a* with *certain*", since use of *a* with *certain* is emphasized in the L2-English instruction of Korean speakers (Heejeong Ko. p.c.)[4]. However, Korean speakers still overused *the* with *certain*. This strongly suggests that simple strategies associating *the* or *a* with a particular word cannot successfully account for article choice in L2-English.

---

[3] I cannot entirely exclude the possibility that the L1-Russian speakers may somehow be evaluating contexts like *mother of X* as indefinite by deciding that these contexts are incompatible with the meaning of *opredelennyj*. However, they clearly cannot do so through simply testing the felicity of the strings *opredelennaja mat'* "definite/certain mother" vs. *neopredelennaja mat'* "indefinite/uncertain mother", since neither string is felicitous. The learners would have to rely on their understanding of the meaning of word *opredelennyj* instead.

[4] According to Heejeong Ko, middle/high-school textbooks of English in Korea instruct students to use *a* with *certain*. This information was corroborated by several informants in Korea. Unfortunately, no textbook reference for this could be found.

2.6. Frequency

The last alternative explanation that I consider concerns the frequency of different types of DPs in the input that L2-learners receive.

2.6.1. Proposal: specific indefinites are infrequent

This explanation goes as follows. The reason that L2-English learners make errors with specific indefinites and non-specific definites is that these types of DPs are infrequent in the input. In contrast, specific definites and non-specific indefinites are quite frequent. L2-learners don't know which article to use with specific indefinites or non-specific definites because they simply haven't heard these types of contexts sufficiently often to make any conclusions about them.

2.6.2. Problem: where does the learners' knowledge come from?

While it may well be the case that specific indefinites and non-specific definites are not frequent in the input, the frequency explanation simply cannot work. This explanation requires that L2-learners have some way of differentiating between different types of indefinites, as well as between different types of definites. Upon hearing an indefinite context, for instance, the learner has to compute whether it is specific or non-specific. Having heard many non-specific indefinites that are used with *a*, the learner will correctly use *a* in this category. Not having heard many specific indefinites, the learner will not know which article is appropriate.

Given that specific and non-specific indefinites headed by *a* look exactly the same, how does the learner differentiate between them? The learner must have some linguistic knowledge, such as access to the Article Choice Parameter, in order to make this distinction. And once we allow learners to have this linguistic knowledge, the frequency explanation becomes superfluous.

3. Conclusion and unanswered questions

In the previous section, I argued that the data on L2-English article choice cannot be accounted by such explanations as disregard of hearer knowledge, explicit (non-UG-based) strategies, and frequency, among others. The data are best explained by the proposal that L2-learners have UG-access and fluctuate between the two settings of the Article Choice Parameter.

3.1. Implications for L2-acquisition

The data reported here provide support for the Full Access view of L2-acquisition. Various studies of syntactic aspects of L2-acquisition, such as CP-direction, control verbs, and functional categories, have argued that L2-learners have direct UG access to syntactic structure. The studies reported here add to the discussion by providing evidence that L2-learners also have access to universal semantic distinctions.

The studies also support previous findings (Broselow and Finer 1986, Finer and Broselow 1991, and Eubank et al. 1997, among others) which show L2-learners exhibiting linguistic knowledge that can be due to neither L1-transfer nor L2-input – but that is nevertheless consistent with the possibilities provided by UG. In the domain of article choice, as in the domains of reflexive binding and verb-raising, L2-learners

access UG parameter settings that are not instantiated in either their L1 or their L2. This provides additional evidence for direct access to UG, unmediated by the L1.

Finally, the fluctuation pattern found in the domain of L2-English article choice suggests that optionality in L2-acquisition may (at least in some domains) be due to learners' inability to choose between parameter settings. The FH provides a principled way of looking at optionality in L2-acquisition: it shows that optionality is in principle fully compatible with UG-access, since it may result from learners accessing different UG-provided possibilities.

L2-learners' use of optional variants in the same contexts has sometimes been taken as an indication that L2-learners' linguistic knowledge is permanently impaired in some way (see e.g., Eubank et al. 1997, Beck 1998, on impairment in the domain of verb-raising; see also Sorace 2000 for an overview). Under my proposal, optionality does not mean permanent impairment, however. L2-learners fluctuate between parameter settings until the input leads them to set the parameter to the target value. I have suggested that L2-learners have difficulty with parameter setting, especially so in the domain of discourse-related parameters; as a result, fluctuation may persist for a long time. While fluctuation may go on indefinitely for some L2-learners, parameter setting is possible, even in the notoriously difficult domain of article choice. About a third of the learners in our last two studies exhibited target-like use of English articles.

3.2. Implications for linguistic theory

The L2-studies reported in this thesis make a contribution to the theory of article semantics by providing evidence that the specificity distinction exists for definites as well as indefinites.

As discussed in Chapter 2, there is no compelling evidence that English definites are ambiguous between specific and non-specific readings (see Heim 1991). Since English does not morphologically encode specificity with definites (as it does with indefinites, via *this*ref), it is necessary to look at other languages for the relevant data. An example of a language which appears to make the specificity distinction with both definites and indefinites is Samoan, as discussed in Chapter 2; however, articles in Samoan have not been extensively studied. L2-English provides an additional source of data for examining specificity with definites as well as indefinites. The data reported in this work provide evidence that the specificity distinction cross-cuts the definiteness distinction. The data contribute to a parametric view of articles cross-linguistically.

3.3. Unanswered questions: suggestions for future research

The studies reported here suggest a number of directions for further research on aspects that were not addressed or could not be explained in the present work.

3.3.1. Article choice and transfer

First, the role of transfer in L2-English article choice needs to be examined. What would happen when L1-Spanish or L1-French speakers acquire English? One possibility is that these learners would transfer the definiteness distinction from their L1 to their L2, and correctly treat *the* as [+definite]. Another possibility is that transfer would not play a role, and that these learners would fluctuate between the two parameter settings just like L1-Russian and L1-Korean speakers. Both alternatives are

possible: as discussed in Chapter 1, transfer plays a role in some domains (e.g., VP-headedness and the V2-constraint) but not others (e.g., control verbs and verb-raising). The match between the parameter settings in the L1 and L2 does not necessarily lead to transfer. Fluctuation is attested both in domains where transfer takes place (e.g., V2 word order – Robertson and Sorace 1999) and in domains where it does not (e.g., verb raising – Eubank et al. 1997). It is important to see whether transfer is operative in the domain of article choice interacts as well, or whether fluctuation is an inherent property of the acquisition of articles, regardless of the L1 parameter setting.

3.3.2. Article choice in L1-acquisition

Second, article choice in L1-acquisition needs to be studied in more depth. As discussed in Chapter 3, different accounts of article use in L1-acquisition make different predictions that need to be tested. If children, like L2-adults, fluctuate between the two settings of the Article Choice Parameter, then they are predicted to overuse *a* with non-specific definites as well as overuse *the* with specific indefinites. While overuse of *the* with indefinites has been attested for L1-acquisition (see the overview in Chapter 3), more work is needed to determine whether it is tied to specificity or to a different factor, such as presuppositionality (see Wexler 2003). Article use with non-specific definites has not previously been tested in L1-acquisition, so work is needed in this domain as well.

Finally, if different factors are found to lie behind L1- and L2-acquisition of articles, then an account of this difference needs to be developed.

3.3.3. Other semantic factors in the acquisition of articles

The role of presuppositionality in acquisition needs to be examined. Presuppositionality may play a role in L1-acquisition (see Chapter 3). It may, under certain circumstances related to transfer, play a role in L2-acquisition (although the results may be due to a testing confound – see Appendix 4 for more discussion). The role of presuppositionality needs to be studied in more detail. If presuppositionality is found to play a role in L2-acquisition, it is necessary to see how it relates to the role of specificity – e.g., whether it is necessary to posit two parameters related to article choice, or whether both specificity and presuppositionality can be derived from the same source. It would also be necessary to explain why presuppositionality in L2-acquisition is related to transfer (see Appendix 4) while specificity plays a role in L2-acquisition independently of transfer.

Another semantic distinction that has not been discussed in this thesis is *genericity*. Article use in generic contexts presents an interesting case for acquisition, since languages differ in how they use articles in generic contexts (see Chierchia 1998, among others). If article use with generics is parameterized, as suggested by Chierchia 1998, then we can study parameter setting, fluctuation, and transfer in this domain with both L1- and L2-learners.

3.3.4. More cross-linguistic data

On a different track, it would be productive to look for more cross-linguistic evidence related to specificity. For instance, a study of article semantics in Samoan would provide information about whether specificity does indeed play a role with both

definites and indefinites in this language, as I have suggested, and whether the use of the specific article in Samoan resembles the use of *this*ref in English.

It would also be interesting to study the conditions on article omission in Norwegian, since Borthen (1998, 2003) suggests that article omission with both indefinites and superlatives in this language is licensed only in non-referential (non-specific) contexts.

Finally, if the Article Choice Parameter is real, it should ideally incorporate article distinctions not related to specificity, such as those in Salish and Maori (see Chapter 2). It is also necessary to explore in more details why the discourse-related distinctions of definiteness and specificity, but not the many possible grammatical distinctions that articles can make, are relevant for acquisition.

3.3.5. *Fluctuation in other domains*

Finally, it is important to look in more detail at whether fluctuation between parameter settings is operative in other domains. Reflexive-binding provides an interesting case study, since, as discussed in Chapter 1, there is evidence that L2-learners adopt parameter settings for the Governing Category Parameter that come from neither the L1 nor the L2. It would be important to examine whether L2-learners fluctuate between the different settings of the GCP: the FH predicts that learners should not show 100% adherence to a setting that is coming from neither their L1 nor their L2; rather, they should fluctuate between such a setting and the target L2-setting.

It is also important to consider in more depth why some parameters are more subject to fluctuation than others. It still remains a puzzle why transfer is operative in some domains but not others, and why fluctuation between parameter settings occurs in some domains but not others.

Appendix 1: Descriptions of individual L2-learners

This appendix gives the characteristics of individual participants in all three studies.

1. Participants in Experiment 1 (Chapter 4)

Table 1: L1-Russian participants in Experiment 1

subject number[1]	age	sex	age of first exposure to English	length of residence in the US (years; months)	approximate age of arrival in the US	Michigan test score (max = 30)
8	51	F	47	3;11	47	14
9	34	M	6	0;4	34	29
10	36	M	12	0;4	36	13
11	54	F	41	~14y	41	28
12	41	F	40	0;9	40	6
13	29	F	7	1;1	28	26
14	43	M	10	1;9	41	28
15	40	F	12	3;4	37	20
16	41	M	11	3;6	38	13
17	51	F	40	0;11	50	18
18	37	M	13	2;5	35	20
19	25	F	18	1;4	24	28
20	29	M	10	1;4	28	25
21	44	F	20	~23y	21	30
22	32	F	8	1;8	30	29
23	49	F	37	11;10	37	19
24	21	F	17	4;5	17	24
25	47	M	18	0;11	46	24
26	20	F	9	5;11	14	28
27	28	M	11	2;2	26	18
28	20	F	11	2;4	18	29
29	21	F	11	2;5	19	26
30	53	F	23	~7y	46	19
31	52	F	17	0;2	52	13
32	53	M	12	0;3	53	18
33	37	F	5	1;10	35	21
34	42	M	8	8;8	33	28
35	40	M	7.5	3;6	37	23
36	55	F	10	8;8	46	20
37	43	F	8	0;7	43	26
38	48	M	48	0;7	48	12

[1] The first seven participants took part in a pilot study, hence the numbering beginning with #8.

2. Participants in Experiment 2 (Chapter 5 + Appendix 3)

Table 2A: L1-Russian participants in Experiment 2

subject number	age	sex	age of first exposure to English	Length of U.S. residence (years)	Age at U.S. arrival	Michigan test score (max = 30)	Special notes
r1	20	F	9	3.44	17	28	
r2	21	F	12	0.09	21	20	native in Ukrainian; Russian primary since early childhood
r3	32	M	17	3.96	28	22	
r4	35	F	13	4.54^2	31	19	
$r6^3$	50	F	10	1.00	50	17	
r7	50	M	13	1.00	49	13	
r8	51	M	11	13.72	37	25	
r9	26	F	7	8.55	17	30	
r10	31	M	12	3.00	28	26	
r11	46	F	20	0.81	45	12	lived in Israel for 11 years prior to arrival in the U.S.
r12	47	F	11	10.41	36	30	
r13	50	F	49	0.37	49	16	lived in Israel for 12 years prior to arrival in the U.S.
r14	17	M	6	0.16	17	23	
r15	27	M	6	0.16	27	26	
r16	25	F	7	0.16	25	29	
r17	55	F	12	0.16	55	17	
r19	42	F	7	2.07	40	19^4	native in Ukrainian; Russian primary since early childhood
r20	21	M	20	0.15	21	17	
r21	22	F	22	0.07	22	5	
r22	34	M	12	5.15	29	25	
r23	42	F	10	5.15	37	14	

---

[2] Participant # r4 did not begin exposure to English until 9 months after arriving in the U.S. During the first 9 months, she did not attend English classes, work, or have any substantial contact with native English speakers.

[3] Participants #r6 and #r7 had participated in another study involving L2-English article choice: they were pilot subjects in Study 1 (Chapter 4). This did not appear to influence their subsequent performance.

[4] Participant # r19 was taking the Michigan test for the second time, having taken it a year and a half previously as part of a completely unrelated study. The previous time, she had also scored as intermediate, with a slightly lower score.

Table 2A (continued)

subject number	age	sex	age of first exposure to English	Length of U.S. residence	Age at U.S. arrival	Michigan test score (max = 30)	Special notes
r24	26	M	14	0.01	26	23	lived in Israel for 10 years, from age 14; fluent in Hebrew; now lives in Germany, in the U.S. on a visit
r25	56	F	11	0.55	56	11	
r26	35	F	10	8.91	26	28	
r27	42	M	12	10.75	31	24	
r28	56	M	8	0.24	55	24	
r29	55	M	45	10.44	45	10	
r30	27	M	10	0.33	26	12	
r32	56	M	12	9.26	47	26	
r33	42	F	7	0.16	42	16	native in Ukrainian; Russian primary since early childhood
r34	19	M	14	0.17	19	10	
r35	42	F	11	0.18	42	11	
r36	43	M	11	0.18	43	9	
r37	21	M	7	0.15	21	11	
r38	21	M	12	0.17	21	23	
r39	50	M	12	8.54	41	10	bilingual in Russian and Judeo-Tat
r40	50	F	12	8.54	42	21	fluent (non-native) in Judeo-Tat
r41	55	M	10	9.04	46	25	
r42	57	F	10	8.54	48	23	
r43	53	F	12	0.20	53	21	fluent (non-native) in Ukrainian
r44	36	F	11	0.21	36	23	native in Ukrainian; Russian primary since early childhood
r45	37	M	12	0.21	37	12	
r46	42	F	10	3.15	39	23	
r47	42	M	10	2.89	39	26	
r48	26	F	10	0.24	26	2	
r49	51	F	50	0.23	51	11	
r50	39	F	14	4.07	35	26	
r51	41	M	18	4.16	37	22	
r52	18	F	7	1.16	17	27	fluent (non-native) in Ukrainian
r53	42	F	12	1.16	41	19	

Table 2B. L1-Korean participants in Experiment 2

subject number	age	sex	age of first exposure to English	Length of U.S. residence (years)	Age at U.S. arrival	Michigan test score (max = 30)
k1	27	F	12	1.89	26	30
k2	27	F	13	0.98	26	28
k3	17	M	12	1.39	15	26
k4	29	F	13	2.31	27	16
k5	23	F	13	0.41	23	28
k6	28	F	9	0.82	28	23
k7	28	M	12	2.09	26	29
k8	19	F	12	0.59	18	30
k9	29	F	13	0.15	29	18
k10	24	F	14	0.11	24	14
k11	26	M	10	2.15	24	22
k12	20	F	13	0.23	20	13
k13	32	M	15	0.15	32	26
k14	20	M	12	1.32	19	20
k15	36	M	12	7.07	29	25
k16	30	M	13	1.06	29	23
k17	33	F	13	4.15	29	8
k18	38	M	13	2.06	36	25
k19	38	F	13	2.06	36	29
k20	36	M	13	4.15	31	22
k21	30	M	12	2.15	28	23
k22	29	M	13	3.07[5]	26	26
k23	29	M	20	1.06	28	25
k24	29	M	13	1.06	28	30
k25	22	F	13	2.65	20	23
k26	20	F	12	4.23	16	28
k27	29	F	12	0.15	29	19
k28	25	M	12	0.15	25	20
k29	27	M	12	0.64	27	25
k30	27	F	12	0.06	27	26
k31	23	M	10	1.06	22	21
k32	30	M	13	3.15[6]	27	24
k33	29	F	13	3.15	26	23
k34	29	F	13	1.06	28	20
k35	32	M	13	3.15	29	29
k36	29	M	13	4.15	25	26
k37	31	M	13	1.15	30	24
k38	27	F	13	3.15	24	16

[5] Participant # k22 had also spent some time in the U.S. before settling here permanently, during two separate year-long visits, and had then gone back to Korea.

[6] Participant # k32 had also spent four months in the U.S. during a visit, and had gone back to Korea, before coming back to the U.S.

3. Participants in Experiment 3 (Chapters 6 and 7)

Table 3A. L1-Russian participants in Experiment 3

subject number	age	sex	age of first exposure to English	Length of U.S. residence (years)	Age at U.S. arrival	Michigan test score (max = 30)	Special notes
r1	46	F	12	0.08	46.55	9	
r2	19	F	10	0.08	19.13	19	fluent (non-native) in Romanian
r3	48	M	12	0.08	48.49	2	
r4	34	F	10	2.10	32.02	26	native in Buriat; Russian primary since age 4
r5	42	F	10	10.38	31.67	16	bilingual in Russian and Ukrainian
r6	39	F	11	2.79	36.49	23	
r7	48	M	38	6.56	41.66	8	
r8	54	M	15	1.73	52.32	21	lived in Israel for 10 years before coming to the U.S.
r9	21	F	10	1.38	20.48	25	
r10	35	F	12	1.75	33.53	27	fluent (non-native) in Azeri
r11	34	F	19	1.75	32.48	28	bilingual in Russian and Azeri
r12	39	F	11	0.99	38.92	15	fluent (non-native) in Ukrainian
r13	49	M	11	0.91	48.28	17	fluent (non-native) in Ukrainian
r14	30	M	12	1.50	28.62	28	
r15	27	F	10	1.25	26.26	20	
r16	36	F	10	0.68	35.99	22	bilingual in Turkmen and Russian; Russian primary
r17	33	M	33	0.59	32.48	20	
r18	26	F	8	1.53	25.10	25	native in Uzbek; Russian primary since age 4
r19	22	M	11	1.51	20.64	29	bilingual in Russian and Azeri

Table 3A (continued)

subject number	age	sex	age of first exposure to English	Length of U.S. residence (years)	Age at U.S. arrival	Michigan test score (max = 30)	Special notes
r20	55	M	12	0.77	55.37	27	
r21	56	M	10	0.77	55.88	22	
r22	29	M	11	0.51	29.45	30	bilingual in Russian and Armenian; Russian primary
r23	54	F	10	0.93	53.55	27	
r24	24	F	14	1.44	23.23	26	
r25	33	M	12	1.09	32.81	23	fluent in Uzbek since age 16; Russian primary
r26	34	F	10	0.76	33.70	20	somewhat fluent in Tatar
r27	24	F	10	0.60	24.39	28	bilingual in Russian and Armenian
r28	55	F	53	2.60	52.85	20	
r29	26	M	10	0.77	25.61	24	
r30	53	F	11	5.59	48.38	10	

Table 3B. L1-Korean participants in Experiment 3

subject number	age	sex	age of first exposure to English	Length of U.S. residence (years)	Age at U.S. arrival	Michigan test score (max = 30)
k1	30	M	12	0.45	30.50	27
k2	29	F	13	0.45	29.44	23
k3	27	F	10	0.53	27.07	29
k4	27	M	10	0.53	27.37	26
k5	31	M	13	1.50	30.35	27
k6	34	F	13	8.60	26.32	27
k7	31	M	12	1.53	29.54	28
k8	34	F	13	7.45	26.62	18
k9	27	F	12	0.94	26.96	5
k10	34	M	14	6.92	27.57	27
k11	31	M	13	1.46	30.51	27
k12	33	M	12	3.63	29.93	25
k13	32	M	13	3.46	28.79	26
k14	34	M	12	3.46	31.10	25
k15	33	F	10	3.13	29.93	27
k16	29	M	10	1.71	27.36	26
k17	28	F	9	2.46	25.85	26
k18	34	M	13	3.55	30.72	30
k19	27	F	12	0.13	27.39	28
k20	31	F	13	3.55	27.84	25
k21	27	M	13	1.47	26.51	19
k22	33	M	13	1.55	32.08	24
k23	38	F	14	3.47	35.08	23
k24	33	F	13	0.88	33.02	23
k25	34	M	12	5.64	28.64	26
k26	30	M	14	1.55	28.58	28
k27	32	F	13	3.55	29.11	30
k28	31	F	14	0.47	31.03	22
k29	31	F	10	1.63	29.48	28
k30	28	M	11	1.38	27.27	24
k31	32	F	13	0.63	31.40	21
k32	33	F	13	1.52	31.49	29
k33	40	M	13	7.72	32.31	22
k34	19	F	12	3.89	36.08	28
k35	26	F	12	1.72	24.50	27
k36	30	M	12	1.05	29.40	26
k37	26	F	12	0.55	25.47	24
k38	29	F	12	2.05	27.50	26
k39	28	F	10	2.47	26.44	22
k40	32	F	12	2.55	30.39	26

Appendix 2: Elicitation task stimuli

This appendix reports the items from the three elicitation tasks reported in Chapters 4 through 6.

Appendix 2A: Items from Elicitation Study 1

This appendix contains all of the test items from the first elicitation study, reported in Chapter 4. In this study, all of the contexts, as well as instructions, were given to the L2-learners in Russian. Only the last sentence was in English. The English versions of instructions and contexts are given here. The stimulus sentence, that was originally in English, is offset from the rest of the dialogue by a blank line.

Instructions given to participants (English version)

This test consists of a number of short dialogues between two people, A and B. The last sentence in each dialogue is missing an article. Your job is to decide what article is appropriate for the given sentence. There are three possible answers: *the*, *a*, and no article. The choice of article may depend on the context, so please carefully read the entire dialogue.

Please write your answer on the answer sheet provided for you. Please do not write on the test itself.

There is no time limit. Please complete the items in the order given. Please do not go back to or change your earlier answers.

Practice Items

EXAMPLE1
A: I haven't seen Lisa in a long time. Do you know where she is?

B: She is in California visiting _____ Susie.

EXAMPLE2
A: Do you remember how you introduced me to your neighbor Robert yesterday? What does he do for a living?

B: Roger is _____ musician.

EXAMPLE3
A: Tell me, what does Erica want to be when she grows up?
B: She wants to be an explorer.

B: She wants to go to _____ North Pole.

Item types – arranged by category

Non-specific indefinites, narrow scope, no RC-modification

1.

In a children's library

A: I'd like to get something to read, but I don't know what myself.
B: Well, what are some of your interests? We have books on any subject.
A: Well, I like all sorts of mechanic things. Cars, trains, airplanes... Yes! I've decided.

A: I'd like to get ____ book about airplanes.

2.

A: I was late for work again today. My car wouldn't start!
B: Is it ok now?
A: No, it's at the mechanic's. I'll have to take the bus to work this week – again!

A: I need to buy ____ new car.

3.

In a pet shop
A: How may I help you? As you can see, we have a wide choice of animals here –
kittens, puppies, goldfish, and so on.
B: I don't know if you'll be able to help me.
A: Why not?
B: I have very specific demands.

B: I am looking for ____ white kitten with blue eyes.

4.

In a clothing store.
A: May I help you?
B: Yes, please! I've rummaged through every stall, without any success.

B: I am looking for ____ red hat.

Non-specific indefinites, narrow scope, RC-modification

5.

A: Roger was in a car accident last week - he hit somebody else's car. The police claim it was his fault, but Roger says he was innocent. He doesn't want to pay the fine.
B: So what is he going to do?

B: He is going to find _____ lawyer who can give him good advice.

6.

A: Tom has just been promoted. He has a new office, and a lot of responsibilities. I wonder how he'll handle it all.

B: He is planning to hire _____ secretary who will help him organize the work.

7.

A: I saw Anne in the library yesterday. Do you know what she was doing there?
B: She was studying for a history exam, and she couldn't remember the year that the American revolution took place.

B: So she was looking for _____ book about American history that could give her the information she needed.

8.

A: Sarah has always dreamed of being rich. She wants to have a lot of money, but she doesn't want to work.
B: So where will she get the money from?

A: She wants to marry _____ millionaire who will give her a lot of money.

Non-specific indefinites, no scope interactions, no RC-modification

9.

A: I am looking for our English teacher. Do you know where he is?

B: He is talking to _____ student.

10.

A: It was Julie's birthday yesterday.
B: Did her uncle Ted give her anything?

A: Yes, he brought her _____ necklace.

11.

A: John looked very happy today. Do you know why?

B: He got ____ dog for his birthday yesterday.

12.

A: Nick just went into the living room. Do you know what he is doing there?

B: He is watching ____ cartoon on television.

Specific indefinites, no scope interactions, RC-modification

13.

A: Alice just came to visit her nephew Andy.
B: Did she bring him anything?

A: Yes, she brought him ____ picture which shows some children playing games in their garden.

14.

A: I am looking for Professor Angela Smith. Do you know where she is?

B: She is meeting with ____ student who is taking her advanced physics class this semester.

15.

A: I thought that Timothy was in a bad mood today. Do you know if anything is wrong?

B: He quarreled with ____ friend that he really likes.

16.

A: I just saw Laura in the garden, sitting under a tree. Do you know what she is doing there?
B: She's developed a new hobby – bird-watching.

B: Right now she is watching____ bird which has pretty green wings.

Specific indefinites, wide scope over an operator, no RC-modification

17.

In a restaurant
A: Are you ready to order, sir? Or are you waiting for someone?
B: Can you please come back in about twenty minutes?

B: I am waiting for _____ woman from Brookline.

18.

In an airport, in a crowd of people who are meeting arriving passengers
A: Excuse me, do you work here?
B: Yes.
A: In that case, perhaps you could help me.

A: I am trying to find _____ girl with long red hair.

19.

In a «Lost and Found»
A: Can I help you? What are you looking for?

B: I am looking for _____ gold necklace.

20.

In a used bookstore
A: Do you need help? Are you looking for some book?
B: Not quite. I heard that this store also buys books.
A: Yes, that's correct.

B: I would like to sell you _____ book on French history.

Specific indefinites, wide scope over an operator, RC-modification

21.

A: I saw Roger in a bookstore yesterday. Do you know what he was doing there?

B: He was looking for _____ book about birds that was published three years ago.

22.

A: John is planning to get married next month. I don't approve of his choice of bride.
B: Why not?

A: He wants to marry _____ woman who has been divorced seven times and has five children.

23.

A: I heard that Mary was sick. Has she found medical help?
B: Yes, she has.

B: She is going to see _____ doctor who went to medical school at Harvard and now lives in Brookline.

24.

A: Yesterday, I saw Susan in the apartment building where elderly immigrants live. I wonder what she was doing there.
B: Susan goes there to give English lessons. She visits many of the tenants regularly.
A: Is she going there today?

B: Yes, she is planning to visit _____ woman who lives on the fifth floor.

Specific indefinites with *certain*, RC-modification

25.

A: Tom seemed very nervous to me. I think he is having problems in class. Do you know why he is so nervous right now?

B: He is going to meet with _____ certain professor who gave him a bad grade.

26.

A: Nancy went to the museum yesterday. She didn't look around at all, but went directly to the room with 18th century Dutch art.
B: Why?

A: Because she wanted to see _____ certain painting that she had heard a lot about.

27.

A: Billie looked very excited this morning. Do you know why?

B: He is planning to see _____ certain film that he has wanted to see for a long time.

28.

A: I just saw Lucy at a newsstand. She was there for a really long time - I wonder why?

B: She was looking for _____ certain magazine which has some interesting articles about Russia.

Previous-mention definites, no RC-modification

29.

A: Richard took a vacation in Italy last year. One time, he saw a beautiful sculpture on the street. Across the street there was a beautiful church. And he had only one frame left on the film in his camera!
B: What did he do?

A: He photographed _____ sculpture.

30.

A: Clara bought a kitten and a puppy yesterday. She played with them all day long.
B: Is she playing with them now as well?

A: No. She is giving some milk to_____ kitten.

31.

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 _____ magazine first.

32.

A: Last night, Paul decided to eat some dinner. He had just two things in his refrigerator – a pot of soup and a cheese sandwich.
B: So what did he eat?

A: He ate _____ sandwich.

Previous-mention definites, RC-modification

33.

A: Dorothy decided to buy a house last year. She looked at a lot of houses, and she really liked two: a small blue house and a big pink house.
B: So which house did she buy?

A: She bought _____ house which was small and blue.

34.

A: Samantha went to the library yesterday. She likes history a lot, so she got a book on ancient Greek history and another book on modern Chinese history.
B: Yes, I noticed that she is reading one of these books right now. I wonder which one?

B: She is reading _____ book which talks about modern Chinese history.

35.

A: Miranda bought two birds in the pet shop yesterday; one was healthy and one was a little bit sick.

B: What did she do when she brought the birds home?

A: She gave some seeds to _____ bird that was sick.

36.

A: Last night, after dinner, Andy decided to eat a banana. He found two bananas in his refrigerator. One was ripe and yellow and the other one was a little rotten.

B: So what did he do?

A: He ate _____ banana which was ripe and yellow.

Bare plurals

37.

A: Do you know what Rose had for supper yesterday evening?

B: She ate _____ potatoes.

38.

A: Angelica had a lot of free time this summer. Do you know what she did with it?

B: She read _____ magazines.

39.

A: I heard that George went to Italy last year. Did he bring any gifts for his sister?

B: He brought his sister _____ earrings.

40.

A: I wonder what George was doing in the toy store yesterday?

B: He was buying _____ dolls for his daughter.

Indefinite generics (subject position)

41.

A: I'm sorry I'm late. I saw an amazing thing! I was walking down a sidewalk and suddenly saw a farmer leading a cow down the middle of the street! This was in downtown Boston! So naturally I stayed to look.
B: I don't blame you.

B: _____ cow in Boston is not something you see every day.

42.

A: My boss has invited me to his house for dinner. What do you think I should bring?
B: Well, if it were me, I'd bring some wine.
A: You are right, that can hardly fail.

B: _____ bottle of wine is always a good gift.

43.

A: I'm so mad at my puppy, Bobby! While I was at work yesterday, he tore the curtains and broke my favorite vase!
B: Well what did you expect? You are at work all day, and he gets lonely – so he starts acting up.
A: You are right. I probably shouldn't have gotten a puppy, when I don't have time for him.

A: _____ young puppy needs a lot of time and attention.

44.

A: I visited my brother yesterday. His apartment was so dull! Mainly because he had a very dim overhead light. So I couldn't stand it anymore, and bought him a bright lamp. Now his place is much cozier.
B: That's not surprising.

B: _____ bright lamp makes any room more pleasant.

Definite generics (subject position)

45.

A: I went to the forest yesterday to pick some mushrooms. And I saw a bear there! That really surprised me.
B: Why?
A: Somehow I didn't think that there are any bears in Massachusetts.
B: Well you were wrong. After all, Massachusetts is in North America.

B: And _____ brown bear is very common in North America.

46.

A: I went to California recently. I saw pelicans there, can you believe it? They were sitting on some stones, not far away from the beach.
B: That's great! It doesn't surprise me. I was just reading a book about birds recently.
A: And what did you find out?

B: _____ pelican lives all along the coast of California.

47.

A: Is it true that there are a lot of tulips in Holland?
B: Yes. I read a book about Holland recently. Here's what it said.

B: _____ tulip is very popular in many Dutch gardens.

48.

In the zoo
A: What beautiful peacocks! I wonder where peacocks came from originally?
B: I think we can get that information in the book we bought in the gift shop... Let me look it up... Aha! I found it.

B: _____ peacock is native to South Asia.

Control for indefinite generics: indefinites in subject position

49.

A: Tell me, have you bought any exotic pets lately?
B: No. Why?
A: Look out the window.
B: What for?

A: _____ large tiger is standing in your garden.

50.

A: I visited a friend of mine last night. When it was time to leave, I couldn't find my hat. Suddenly I saw it crawling along the floor!
B: Crawling?
A: Yes! But it turned out to have a simple explanation right away.

A: ____ kitten climbed out from under my hat.

Control for definite generics: definites in subject position

51.

A: Mary had a visit from her neighbor yesterday. The neighbor brought her dog with her, which was a big mistake.
B: Why?
A: Because this dog saw Mary's cat right away.
B: And what happened?

A: ____ dog started barking and woke up Mary's baby.

52.

A: I went to see a girl I know yesterday and I couldn't find her house for the longest time.
B: Why? Did you have the wrong address?
A: No, I had the right one. It turned out, I passed by her house several times without noticing it.
B: How did that happen?

A: ____ house was standing behind several very tall trees.

Appendix 2B: Items from Elicitation Study 2

This appendix contains all of the test items from the second elicitation study, reported in Chapter 5. In this study, all of the contexts, as well as instructions, were given to the L2-learners in their L1, Russian or Korean. Only the last sentence and any sentence(s) following it were in English. The English versions of instructions and contexts are given here. The English versions of all of the items are given here. The sentences that were in English on the original test are highlighted in bold.

Instructions given to participants (English version)

This test consists of 56 short dialogues. Most of the dialogue is in Korean/Russian, except for the last sentence or sentence, which is/are in English. One of the English sentences is missing an article. A choice of possible articles is given in parentheses: it looks like either **(a, the, --)** or **(some, the, --)**. The dash **(--)** means no article is needed. Your job is to decide which of the three choices in the parentheses is most appropriate in the given context. The choice of article depends on the context, so please carefully read the entire dialogue – all of the Russian/Korean text, as well as all of the English text. You may sometimes feel that more than one of the choices provided is appropriate in the given context; in that case, please choose the variant that sounds best in the given context. It is important that you provide only one answer for each item. Please circle your answer.

The test consists of 56 items. There is a time limit of one hour for this test. Please complete the items in the order given. Please do not go back to or change your earlier answers. Please read each dialogue carefully, and then circle the answer that you feel is appropriate for that item; please do not spend too long on any given item. If you encounter unfamiliar English words, please refer to the translation sheet provided for you. Please do not use dictionaries or any other study aids.

Examples

1) Laura: Tell me, what does your daughter Erica want to be when she grows up?
   Ben: She wants to be an explorer. **She wants to go to (a, the, --) North Pole.**

2) Clara: My roommate got a new pet last week – a snake!
   Todd: Really? Aren't you worried?
   Clara: Why should I be?
   Todd: **(Some, The, --) snakes are poisonous! You might get bitten!**

Practice Items

1) Alex: Do you remember how you introduced me to your neighbor Robert yesterday? What does he do for a living?
   Charles: **Robert is (a, the, --) musician. He plays in our town orchestra.**

2) In a supermarket
   Client: Excuse me, can you help me?
   Salesperson: What can I do for you?
   Client: I have a complaint about your products. **(Some, The, --) milk is spoiled!**

Item types – arranged by category

Singular specific indefinites: scope over an intensional/modal operator

1) In an airport, in a crowd of people who are meeting arriving passengers
    Man: Excuse me, do you work here?
    Security guard: Yes.
    Man: In that case, perhaps you could help me. **I am trying to find (a, the, --) red-haired girl. She flew in from New York on Flight 239.**

2) In a restaurant
    Waiter: Are you ready to order, sir? Or are you waiting for someone?
    Client: Can you please come back in about twenty minutes? I am not ready to order yet. **I am planning to eat with (a, the, --) colleague from Princeton. She will be here soon.**

3) In a "Lost and Found"
    Clerk: Can I help you? Are you looking for something you lost?
    Customer: Yes… I realize you have a lot of things here, but maybe you have what I need. **You see, I am looking for (a, the, --) green scarf. I think that I lost it here last week.**

4) In a used bookstore
    Salesperson: Do you need help? Are you looking for some books?
    Client: Not quite. I heard that this store also buys books.
    Salesperson: Yes, that's correct.
    Client: Well, I might have something for you. **I would like to sell you (a, the, --) book about French history. It is very interesting.**

Singular specific wide-scope indefinites with *certain*

5) In a school
    Becky: Tom seemed very nervous to me. I think he is having problems in class. Do you know why he is so nervous right now?
    Ben: **He is going to meet with (a, the, --) certain professor – someone that Tom is really afraid of.**

6) Roger: I just saw Billie. He looked really excited! Do you know why?
    Anne: **Because he is planning to see (a, the, --) certain girl tonight – he really likes her!**

7) Julie: My friend Nancy went to the museum yesterday. She didn't look around at all, but went directly to the room with 18th century Dutch art.
    Rose: Why?
    Julie: **Because she wanted to see (a, the, --) certain painting – she had heard a lot about it.**

8) Timothy: I just saw Lucy at a newsstand. She was there for a really long time - I wonder why? Do you have any idea what she was doing there?
Gabrielle: **She was looking for (a, the, --) certain magazine - it contains interesting articles about Africa.**

Singular specific indefinites – explicit speaker knowledge, no scope interactions

9) At a university
Rob: Hi Katie - can you help me? I need to talk to Professor Christina Jones, but I haven't been able to find her. Do you know if she is here this week?
Katie: Well, I know she was here yesterday. **She met with (a, the, --) student - he is in my physics class.**

10) Roberta: I'm worried about my brother George. He is having a lot of personal problems.
William: I'm sorry to hear that. I hope he can talk to someone about his problems.
Roberta: He can. **This morning he went to talk to (a, the, --) friend - she's really smart and I hope she'll help him!**

11) Elise: I was at Angela's birthday party yesterday.
Karen: Oh, I didn't even know she was having a party! Did she get anything interesting?
Elise: **Well, her friend Lucy gave her (a, the, --) necklace. It's very beautiful – I know because I helped Lucy pick it out.**

12) Alex: I was looking for you and your roommate Mark yesterday. Where were you?
Jeff: We went to a department store. Mark needed to get some clothes for the winter.
Alex: So did Mark get anything?
Jeff: **Yes, he got (a, the, --) scarf - it is green with big purple stripes (I think it's rather ugly).**

Singular non-specific narrow scope indefinites

13) Sam: I'm having some difficulties with my citizenship application.
Julie: What are you going to do?
Sam: Well, I need some advice. **I am trying to find (an, the, --) experienced lawyer. I think that's the right thing to do.**

14) In a school
Student: I am new in this school. This is my first day.
Teacher: Welcome! Are you going to be at the school party tonight?
Student: Yes. I'd like to get to know my classmates. **I am planning to find (a, the, --) friend in this class! I don't like being all alone.**

15) In a clothing store.

Clerk: May I help you?

Customer: Yes, please! I've rummaged through every stall, without any success. **I am looking for (a, the, --) warm hat. It's getting rather cold outside.**

16) In a children's library

Child: I'd like to get something to read, but I don't know what myself.

Librarian: Well, what are some of your interests? We have books on any subject.

Child: Well, I like all sorts of things that move – cars, trains... I know! **I would like to get (a, the, --) book about airplanes! I like to read about flying!**

Singular non-specific indefinites – no scope interactions

17) At a university

Visitor: Excuse me - can you help me? I'm looking for Professor James Smith.

Secretary: I'm afraid he's not here right now.

Visitor: Is he out today?

Secretary: No, he was here this morning. **He met with (a, the, --) student... but I don't know where Professor Smith is right now.**

18) Tom: Hi, Susan. I'm looking for your roommate Lucy. Is she home tonight?

Susan: No, she's not home. **She went to visit (a, the, --) friend; please come back tomorrow afternoon.**

19) Samantha: How is Judy doing? I haven't seen her in months.

Pat: Well, just last week she had a really fun birthday party.

Samantha: I'm so sorry I missed that! Did she get nice things?

Pat: **Yes - for example, she got (a, the, --) bracelet; Judy really likes jewelry.**

20) Jessica: I tried calling Laura last night, but I couldn't get a hold of her. Do you know where she was?

Peter: Yes, she went shopping - summer is coming, and she needed some new things.

Jessica: Did she get anything?

Peter: **Yes, she bought (a, the, --) hat; she wanted to cover her head from the sun.**

Singular non-specific indefinites – express denial of speaker knowledge

21) At a university.

Professor Clark: I'm looking for Professor Anne Peterson.

Secretary: I'm afraid she is out right now.

Professor Clark: Do you know if she is meeting somebody?

Secretary: I am not sure. **This afternoon, she met with (a, the, --) student – but I don't know which one.**

22) Chris: I need to find Jonathan right away.
    Clara: I'm sorry, but he is not here.
    Chris: Do you have any idea where he is?
    Clara: **He went to visit (a, the, --) friend, but I don't know which one.**

23) Gertrude: My cousin Claudia just bought a new house.
    Richard: Does she need new furniture for it?
    Gertrude: Yes - just yesterday she went to a furniture store. This store has a
    lot of very nice things.
    Richard: Did she buy anything?
    Gertrude: **She bought (a, the, --) sofa, but I don't know which one.**

24) Gary: Last week, my nephew Ben went to the library for the first time in his life
    - he is six years old, and he just learned how to read.
    Melissa: Did he like the library? Did he find anything to read?
    Gary: Yes, he liked it. **He took out (a, the, --) book - but I don't remember
    which one.**

Singular partitive/presuppositional indefinites

25) Sam: My daughter Marian came to school very early yesterday. But she wasn't
    the first – she saw five students and two teachers there.
    Ed: So what did she do until classes started?
    Sam: Well, she didn't have much to do. **So she talked to (a, the, --) student.**

26) Robert: My little son Johnny went to the park yesterday. There were three little
    girls and four little boys already playing there.
    Lisa: Did Johnny play with them?
    Robert: Well, Johnny is usually very shy. **But finally, he said hello to (a, the, --)
    little girl.**

27) Clara: My friend George went to a jewelry store yesterday. The store didn't have
    a lot of things. All it had were three necklaces and two bracelets. But all of them
    were very beautiful.
    Beth: So did he get anything?
    Clara: **Yes, he bought (a, the, --) necklace for his wife.**

28) Charlene: Yesterday, my little niece Angela had a party - she's three years
    old. Six of her little friends came. Angela's parents provided gifts for all of the
    children: four teddy bears and three dolls. Angela got to choose what she wanted
    first.
    Bob: So what did Angela decide to get?
    Charlene: Well, she had to think about it for a while. **But finally, she chose
    (a, the, --) doll.**

Singular indefinites in subject position

29) Clovis: Excuse me, Ms. Williams - have you by any chance purchased an exotic domestic animal lately?
Ms. Williams: No, I have not. Why?
Clovis: **Because (a, the, --) large tiger is standing in your garden.**

30) Alexandra: I heard that you went on a trip. How was it?
Beatrice: Good, mostly. Except that one time I almost got into an accident!
Alexandra: Oh no! What happened?
Beatrice: Well, I was driving through a little town I didn't know. Everything was very quiet. **Then, suddenly, (a, the, --) small kitten ran out into the road in front of me!**

31) Rachel: I just went into Sarah's room to get something, and I found something rather surprising! I think Sarah has a secret admirer!
Angela: Why do you say that?
Rachel: **Because (a, the, --) beautiful flower is lying on Sarah's bed!**

32) Louise: Hi, Polly. How have you been?
Polly: Well, not very good. I had a strange adventure last summer. I was just walking down the street, and I heard people arguing in one of the houses.
Louise: So what happened then?
Polly: A very strange thing happened. **Suddenly, (a, the, --) big teapot fell on my head!**

Singular indefinites in the *there*-construction

33) Terry: I think I am going to go take a bath.
Anne: That might not be such a good idea. I was just in the bathroom. I don't think you should take a bath right now.
Terry: Why not?
Anne: **Because there is (a, the, --) big frog in our bathtub.**

34) Vicky: Something strange happened to me yesterday. I went to a grocery store, and I left my car in the parking. I left the windows open. Guess what I found when I came back?
Chester: What did you find?
Vicky: **There was (a, the, --) little puppy inside my car!**

35) Mike: I have really great friends! Just last week, they did something very nice for me.
Charlene: What did they do?
Mike: Well, it was my birthday. I came home from work, and all my friends were there! **And there was (a, the, --) huge cake on my kitchen table!**

36) Carla: I just visited your cousin Jodi's house. Jodi has excellent taste.
Ned: How do you know?
Carla: She has beautiful things all over her house. **For example, there is (a, the, --) beautiful picture on her living room wall.**

Plural specific indefinites

37) Charlene: Hi, Betsy. What are you doing out here? You look like you lost something!
Betsy: Sort of. **I am trying to find (some, the, --) old friends. They promised to meet me here!**

38) At a restaurant
Waitress: May I take your order, miss?
Client: Not yet. **I am planning to have lunch with (some, the, --) colleagues. They will be here in just a few minutes.**

39) In a "Lost and Found" at a theatre
Customer: I really need help.
Clerk: What's the problem? Have you lost something?
Customer: Yes... I am sure you have many things here, but perhaps you can look through them... **You see, I'm looking for (some, the, --) gold rings. I think I lost them here yesterday.**

40) Phone conversation
Jeweler: Hello, this is Robertson's Jewelry. What can I do for you, ma'am? Are you looking for a piece of jewelry? Or are you interested in selling?
Client: Yes, selling is right. **I would like to sell you (some, the, --) beautiful necklaces. They are very valuable.**

Plural non-specific indefinites

41) Brenda: I'm having a big dinner next week, for all of my friends.
Lawrence: That sounds like a lot of work!
Brenda: Well, I'm not going to cook myself. I don't even like to cook.
Lawrence: So what are you going to do instead?
Brenda: **Well, I am trying to find (some, the, --) good cooks. That way I won't have to do anything myself!**

42) Erica: You know, I really like animals.
Jennifer: Do you have a lot of animals?
Erica: No - all I have is a parrot! **But I am planning to get (some, the, --) new pets. I'm going to my local pet shop tomorrow.**

43) In a department store.
     Customer: Excuse me - I need some help.
     Salesperson: What can I do for you?
     Customer: **I am looking for (some, the, --) winter boots. But I haven't had any luck.**

44) Phone conversation
     Salesperson: Hello, Erik's Grocery Deliveries. What can I do for you?
     Customer: Well, I have a rather exotic order.
     Salesperson: We may be able to help you.
     Customer: **I would like to buy (some, the, --) green tomatoes. I'm making a special Mexican sauce.**

Singular definites in object position

45) Richard: I visited my friend Kelly yesterday. Kelly really likes animals - she has two cats and one dog. Kelly was busy last night - she was studying for an exam. So I helped her out with her animals.
     Maryanne: What did you do?
     Richard: **I took (a, the, --) dog for a walk.**

46) Gabrielle: My four-year-old son Eric has three friends – two little boys and one little girl. He visits them on the week-ends, and they play together.
     Sally: Did Eric visit any of his friends last Sunday?
     Gabrielle: **Yes, he visited (a, the, --) little girl last Sunday.**

47) Tom: My friend Brian and I went shopping yesterday. Brian bought one scarf and two sweaters.
     Cassie: What did he do with these things? Did he give any of them to anyone?
     Tom: Yes. **He gave (a, the, --) scarf to his sister.**

48) Sam: I took my daughter Ruth to the library yesterday. She got one book and two magazines there.
     Debra: Did she start reading as soon as she got home?
     Sam: Yes, she did. **And she read (a, the, --) book first.**

Plural definites

49) Angelique: Last week, I went to visit my friend Alice. She has three children – two little girls and one little boy. I really like children.
     Robert: Did you talk with any of Alice's children?
     Angelique: **Yes, I talked with (some, the, --) little girls.**

50) Alfred: My nephew Peter went to the circus yesterday. He saw one elephant and two monkeys there.
     Winifred: Which of these animals did he like best?
     Alfred: Well, he liked all of the animals. **But he liked (some, the, --) monkeys more.**

51) Arthur: Joan was in Chicago last week. She got lots of nice things here. For example, she got two sweaters and one new purse.
Jim: Did she keep these things for herself, or were they gifts?
Arthur: Gifts, mainly. **She gave (some, the, --) sweaters to her sister.**

52) Rosalyn: My cousin started school yesterday. He took one notebook and two new books with him to school, and he was very excited. He was so proud of having his own school things! But he came home really sad.
Jane: What made him so sad? Did he lose any of his things?
Rosalyn: Yes! **He lost (some, the, --) books!**

Singular definites in subject position

53) Robert: How was your walk in the park? Did you see anyone?
Jennifer: Well, the park was pretty empty. But I did see one little boy and one little girl.
Robert: What were they doing? Playing?
Jennifer: Not really. **(A, The, --) little girl was eating ice cream.**

54) Gina: I was at a circus yesterday. All of the acts were really great. I especially liked three clowns and one elephant.
Ruth: Tell me about what they did.
Gina: **Well, (an, the, --) elephant stood on his head!**

55) Matthew: I went shopping for winter clothes yesterday. I bought one hat and a pair of mittens.
Molly: How do you like your new clothes?
Matthew: They are quite warm. **But (a, the, --) hat is too small for me.**

56) Andrew: I took my daughter Louise to a bookstore yesterday. She got two magazines and one children's book.
Bill: How does she like the things that she got?
Andrew: She loves them! **In particular, (a, the, --) book is really interesting.**

Appendix 2C: Items from Elicitation Study 3

This appendix contains all of the test items from the third elicitation study, reported in Chapter 6. In this study, all of the contexts were given in English. The instructions were given to the L2-learners in their L1, Russian or Korean.

Instructions given to participants (English version)

This test consists of 76 short English dialogues. One of the last sentences in each dialogue is missing an article. A choice of possible articles is given in parentheses: it always looks like **(a, the, --)**. The dash **(--)** means no article is needed. Your job is to decide which of the three choices in the parentheses is most appropriate in the given context. The choice of article depends on the context, so please carefully read the entire dialogue –the text both before and after the missing article. You may sometimes feel that more than one of the choices provided is appropriate in the given context; in that case, please choose the variant that sounds best in the given context. It is important that you provide only one answer for each item. Please circle your answer.

If you come across unfamiliar words, please refer to the list of English words and corresponding translations on the enclosed vocabulary sheet. Please do not use your own dictionary.

There is a time limit of one and a half hours for this test. Please complete the items in the order given. Please do not go back to or change your earlier answers. Please read each dialogue carefully, and then circle the answer that you feel is appropriate for that item.

Practice Items
1) Alex: Your neighbor Robert is very nice. What does he do?
   Charles: Robert is (a, the, --) musician. He plays in our town orchestra.

2) Sam: What do you like to eat?
   Andy: I like lots of things. I especially like (a, the, --) ice cream!

3) Teacher: Who was Alexander Graham Bell?
   Student: He invented (a, the, --) telephone! Thanks to him, I can talk to my friends in other countries.

Item types – arranged by category

IA [+definite, +specific]
definite, wide scope, speaker knowledge

1. Conversation between two police officers
   Police officer Clark: I haven't seen you in a long time. You must be very busy.
   Police officer Smith: Yes. Did you hear about Miss Sarah Andrews, a famous lawyer who was murdered several weeks ago? We are trying to find (a, the, --) murderer of Miss Andrews – his name is Roger Williams, and he is a well-known criminal.

2. At a bookstore
   Chris: Well, I've bought everything that I wanted. Are you ready to go?
   Mike: Almost. Can you please wait a few minutes? I want to talk to (a, the, --) owner of this bookstore – she is my old friend.

3. At the end of a chess tournament
   Laura: Are you ready to leave?
   Betsy: No, not yet. First, I need to talk to (a, the, --) winner of this tournament – she is my good friend, and I want to congratulate her!

4. Eric: I really liked that book you gave for my birthday. It was very interesting!
   Laura: Thanks! I like it too. I would like to meet (a, the, --) author of that book some day – I saw an interview with her on TV, and I really liked her!

IB [+definite, -specific]
definite, narrow scope, no speaker knowledge

5. Conversation between a police officer and a reporter:
   Reporter: Several days ago, Mr. James Peterson, a famous politician, was murdered! Are you investigating his murder?
   Police officer: Yes. We are trying to find (a, the, --) murderer of Mr. Peterson – but we still don't know who he is.

6. At a supermarket
   Sales clerk: May I help you, sir?
   Customer: Yes! I'm very angry. I bought some meat from this store, but it is completely spoiled! I want to talk to (a, the, --) owner of this store – I don't know who he is, but I want to see him right now!

7. After a women's running race
   Reporter: Excuse me! Can you please let me in?
   Guard: What do you need?
   Reporter: I am a reporter. I need to talk to (a, the, --) winner of this race – I don't know who she is, so can you please help me?

8. At a gallery
Sarah: Do you see that beautiful landscape painting?
Mary: Yes, it's wonderful.
Sarah: I would like to meet (a, the, --) author of that painting – unfortunately, I have no idea who it is, since the painting is not signed.

IIA [+definite, +specific]
definite, no scope interactions, speaker knowledge

9. Paul: Do you have time for lunch?
Sheila: No, I'm very busy. I am meeting with (a, the, --) president of our university – Dr. McKinley; it's an important meeting.

10. *Meeting in a park*
Andrew: Hi, Nora. What are you doing here in Chicago? Are you here for work?
Nora: No, for family reasons. I am visiting (a, the, --) father of my fiancé – he is really nice, and he is paying for our wedding!

11. Reporter 1: Guess what? I finally got an important assignment!
Reporter 2: Great! What is it?
Reporter 2: This week, I am interviewing (a, the, --) governor of Massachusetts – Mitt Romney. I'm very excited!

12.
Kathy: My daughter Jeannie loves that new comic strip about Super Mouse.
Elise: Well, she is in luck! Tomorrow, I'm having lunch with (a, the, --) creator of this comic strip – he is an old friend of mine. So I can get his autograph for Jeannie!

IIB [+definite, -specific]
definite, no scope interactions, no speaker knowledge

13.
Bill: I'm looking for Erik. Is he home?
Rick: Yes, but he's on the phone. It's an important business matter. He is talking to (a, the, --) owner of his company! I don't know who that person is – but I know that this conversation is important to Erik.

14. *Phone conversation*
Mathilda: Hi, Sam. Is your roommate Laurie there?
Sam: No, she went to San Francisco for this week-end.
Mathilda: I see. I really need to talk to her – how can I reach her in San Francisco?
Sam: I don't know. She is staying with (a, the, --) mother of her best friend – I'm afraid I don't know who she is, and I don't have her phone number.

15. Mike: Guess what? You remember my friend Jessie, who is a reporter?
     Angela: Yes, what about her?
     Mike: She has a really important job right now, with a big newspaper.
Today, she is interviewing (a, the, --) governor of Arizona! I don't remember who
that is... but this is a really important assignment for Jessie!

16. Rose: Let's go out to dinner with your brother Samuel tonight.
     Alex: No, he is busy. He is having dinner with (a, the, --) manager of his office –
I don't know who that is, but I'm sure that Samuel can't cancel this dinner.

IIIA[-definite, +specific]
Indefinite, wide scope, speaker knowledge

17. In an airport, in a crowd of people who are meeting arriving passengers
     Man: Excuse me, do you work here?
     Security guard: Yes.
     Man: In that case, perhaps you could help me. I am trying to find
(a, the, --) red-haired girl; I think that she flew in on Flight 239.

18. In a restaurant
     Waiter: Are you ready to order, sir? Or are you waiting for someone?
     Client: Can you please come back in about twenty minutes? You see, I am
waiting. I am planning to eat with (a, the, --) colleague from work. She will be
here soon.

19. In a "Lost and Found"
     Clerk: Can I help you? Are you looking for something you lost?
     Customer: Yes... I realize you have a lot of things here, but maybe you have
what I need. You see, I am looking for (a, the, --) green scarf. I think that I lost it
here last week.

20. *phone conversation*
     Jeweler: Hello, this is Robertson's Jewelry. What can I do for you, ma'am?
Are you looking for some new jewelry?
     Client: Not quite – I heard that you also buy back people's old jewelry.
     Jeweler: That is correct.
     Client: In that case, I would like to sell you (a, the, --) beautiful silver
necklace. It is very valuable – it has been in my family for 100 years!

IIIB [-definite, +specific]
indefinite, wide scope, embedded context

21. At an airport
     Security guard1:   I saw that you just talked to an old man who looked very
nervous. What did he want?
     Security guard2:   He said that he is trying to find (a, the, --) little girl from
American Airlines flight 142. He said it's his granddaughter. I couldn't help him,
unfortunately – flight 142 is not here yet.

22. *phone conversation*

    Sam's mother: Hi, Sam. How are you doing?
    Sam: Hi, mom. I'm good. I have a new roommate – his name is George.
    Sam's mother: Do you like him? Do you see him a lot?
    Sam: He is nice. I don't see him very much. I know that I will not see him tonight. He said that he is planning to have dinner with (a, the, --) girl from work tonight; I don't know who she is, but George was very excited about seeing her!

23. In a "Lost and Found"

    Clerk1: That lady you were talking with looked very upset. What was the matter?
    Clerk2: She was upset because I couldn't help her. She said that she is looking for (a, the, --) gold necklace. She said that she lost it here last night, and that it's really valuable; unfortunately, I couldn't find it.

24. Phone conversation

    Art dealer: Hello? How may I help you?
    Agent: Hello. I am calling on behalf of my client, Ms. Kathy Rogers. Ms. Rogers said that she would like to sell you (a, the, --) famous 19th century painting; she said that she just bought it last week in France. She didn't tell me what it is, but she praised it highly.

IIIC [-definite, -specific]
indefinite, narrow scope, no speaker knowledge

25. In a children's library

    Child: I'd like to get something to read, but I don't know what myself.
    Librarian: Well, what are some of your interests? We have books on any subject.
    Child: Well, I like all sorts of things that move – cars, trains... I know! I would like to get (a, the, --) book about airplanes! I like to read about flying!

26. In a school

    Student: I am new in this school. This is my first day.
    Teacher: Welcome! Are you going to be at the school party tonight?
    Student: Yes. I'd like to get to know my classmates. I am planning to find (a, the, --) new good friend! I don't like being all alone.

27. In a clothing store.

    Clerk: May I help you?
    Customer: Yes, please! I've rummaged through every stall, without any success. I am looking for (a, the, --) warm hat. It's getting rather cold outside.

28. Sam: I'm having some difficulties with my citizenship application.
    Julie: What are you going to do?
    Sam: Well, I need some advice. I am trying to find (an, the, --) lawyer with lots of experience. I think that's the right thing to do.

IVA [-definite, +specific]
certain, **wide scope, speaker knowledge**

29. Rick: Did you have a good week-end?
    Bonnie: Yes, thanks! For example, last night I went to the cinema. I wanted
    to see (a, the, --) certain movie; it's a British movie that I've read a lot about.

30.
    Louise: You seem nervous about something.
    Dorothy: I am very nervous! Tomorrow morning, I am going to see (a, the, -
    -) certain lawyer... He always gives me bad news!

31. Robert: Hi, Cathy. Do you have time to talk?
    Cathy: Sorry, not right now – I am about to leave. I am planning to have
    coffee with (a, the, --) certain colleague; she is very punctual, so I should be on
    time!

32.
    In a library
    Librarian: May I help you, miss?
    Client: Yes, please. I am looking for (a, the, --) certain book; it's by John
    Wyndham, and is called "The chrysalids".

IVB [-definite, +specific]
certain, **wide scope, embedded context**

33. Julie: My friend Nancy went to the museum yesterday.
    Rose: Did she go to see anything in particular?
    Julie: Yes – she went to the room with 18th-century Dutch art. She said that she
    wanted to see (a, the, --) certain painting there; I have no idea what it is, but
    Nancy said it's really wonderful!

34. In a school
    Becky: Tom seemed very nervous to me. I think he is having problems in class.
    He looked really nervous just now!
    Ben: I am not surprised. He said that he is going to meet with (a, the, --) certain
    professor; I don't know who it is, but Tom is really afraid of this person!

35. Roger: I just saw Billie. He looked really excited!
    Anne: Of course! He said that he is planning to see (a, the, --) certain girl tonight;
    I don't know who she is, but I know that Billie really likes her!

36. Timothy: I just saw Lucy at a newsstand. She was there for a really long time
    - I wonder why? Do you have any idea what she was doing there?
    Gabrielle: She said that she was looking for (a, the, --) certain magazine; I
    don't know what it is, but Lucy said it had some interesting articles.

VA [-definite, +specific]
indefinite, no scope interactions, speaker knowledge

37. *Phone conversation*
Christina: Hello, you've reached Christina Jones's office.
Rob: Hi, Christina. This is Rob. Do you have time to talk?
Christina: Not right now. I'm sorry, but I'm busy. I am meeting with (a, the, --) student from my English class – he needs help with his homework, and it's important.

38. Meeting on a street
Roberta: Hi, William! It's nice to see you again. I didn't know that you were in Boston.
William: I am here for a week. I am visiting (a, the, --) friend from college – his name is Sam Bolton, and he lives in Cambridge now.

39. Reporter 1: Hi! I haven't seen you in weeks. Do you have time for lunch?
Reporter 2: Sorry, no. I'm busy with a story about local medicine. Today, I am interviewing (a, the, --) doctor from Bright Star Children's Hospital – he is a very famous pediatrician, and he doesn't have much time for interviews. So I should run!

40. Gary: I heard that you just started college. How do you like it?
Melissa: It's great! My classes are very interesting.
Gary: That's wonderful. And do you have fun outside of class?
Melissa: Yes. In fact, today I'm having dinner with (a, the, --) girl from my class – her name is Angela, and she is really nice!

VB [-definite, -specific]
Indefinite, no scope interactions, no speaker knowledge

41. *At a university*
Professor Clark: I'm looking for Professor Anne Peterson.
Secretary: I'm afraid she is busy. She has office hours right now.
Professor Clark: What is she doing?
Secretary: She is meeting with (a, the, --) student, but I don't know who it is.

42. Chris: I need to find your roommate Jonathan right away.
Clara: He is not here – he went to New York.
Chris: Really? In what part of New York is he staying?
Clara: I don't really know. He is staying with (a, the, --) friend – but he didn't tell me who that is. He didn't leave me any phone number or address.

43. Gertrude: Guess what? My cousin Claudia is in Washington, D.C. this week.
    Richard: That's great. What's she doing there?
    Gertrude: She is doing some interviews for her newspaper. She is
    interviewing (a, the, --) politician; I'm afraid I don't know who, exactly. I'll find
    out when I read her article!

44. Karen: Where's Beth? Is she coming home for dinner?
    Anne: No. She is eating dinner with (a, the, --) colleague; she didn't tell me who
    it is.

VIA partitive indefinite

45. Gabrielle: My son Ralph didn't have anything to read last week-end. So, he
    went to the library.
    Charles: Did he find something to read?
    Gabrielle: Yes – he took out three books and four children's magazines. And
    as soon as he came home, he read (a, the, --) book.

46. Marian: I came to school very early yesterday.
    Jim: So were you the first person there?
    Marian: No. I saw five other students and two teachers at the school. I
    didn't have anything to do. So I talked to (a, the, --) student.

47. Rudolph: My niece Janet likes animals a lot. Last week, she decided to get a
    pet and went to a pet shop.
    Lisa: Did she find any pets that she liked?
    Rudolph: Yes – she saw three beautiful puppies and six lovely kittens. She
    couldn't decide! Finally, she bought (a, the, --) kitten.

48. Sophie: I spent last week-end in my summer cottage. I wanted to walk
    around and to swim.
    Elise: So did you have a good time?
    Sophie: No! When she got to my cottage, the weather was terrible! I
    couldn't go swimming or walking. And I didn't have anything to do inside my
    cottage – nothing to read, nothing to watch. So, finally, I went to the library, and
    got out two books and three videos. After I came home, I watched (a, the, --)
    video.

VIB simple indefinite [-definite, -specific]

49. Judy: Last Saturday, I didn't have anywhere to go, and it was raining.
    Samantha: So what did you do?
    Judy: First, I cleaned my apartment. Then I ate lunch. And then I read (a,
    the, --) book.

50. Eric: My friend Tom was in his office at the university, but he really didn't want to work.

Bill: So what did he do?

Erik: Well, he walked around my department. He had some coffee and checked his e-mail. And he talked to (a, the, --) student.

51. Mary: I heard that it was your son Roger's birthday last week. Did he have a good celebration?

Roger: Yes! It was great. He got lots of gifts – books, toys. And best of all – he got (a, the, --) puppy!

52. Tom: How was your trip to New York?

Susan: Great! I went to many museums, and ate in lots of wonderful restaurants. I also visited many friends. And I saw (a, the, --) play.

VIC simple definite [+definite, +specific]

53. Vicky: Where were you yesterday? I tried to call you, but you weren't home.

Rachel: I went to a bookstore yesterday.

Vicky: Oh, what did you get?

Rachel: I got lots of things – several magazines, two red pens, and an interesting new book. After I came home, I read (a, the, --) book.

54. Sarah: Yesterday, I took my granddaughter Becky for a walk in the park.

Claudia: How did she like it?

Sarah: She had a good time. She saw one little girl and two little boys in the park. Claudia is a little shy. But finally, she talked to (a, the, --) girl.

55. Molly: How is your grandpa Sam's farm doing?

Tom: All right, thanks. Last summer, grandpa needed some new animals, so he went to an animal market.

Molly: Did he find any?

Sam: Yes – he found a big cow and a small, friendly horse. But he didn't have enough money for both. In the end, he bought (a, the, --) horse.

56. Alice: What did you do last night?

Robin: I watched TV.

Alice: What did you watch?

Robin: Well, on one channel, I found an interesting German film. On another channel, I found an exciting news program. Finally, I watched (a, the, --) film.

Fillers

VIIA. Universal definites

57. Laura: I'd like to go for a walk. Is it nice outside?
Jenny: I think so – I can see (a, the, --) sun!

58. Andrea: I went for a walk last night – I really enjoyed it.
Jodi: Were you scared walking when it was so dark?
Andrea: It wasn't dark! I saw (a, the, --) moon.

59. Child: Can you please give me a blue pencil?
Mother: Here you go. What are you drawing?
Child: I am drawing (a, the, --) sky.

60. Debra: What are you planning to do after you graduate from college? Are you going to get a job?
Alex: Not yet. My parents gave me a wonderful graduation gift: a year-long trip. I am going to travel around (a, the, --) world!

VIIB. Geographic names

61. Rose: Did you have a good trip to California?
Bill: Yes, it was wonderful. I saw lots of interesting things. And I swam in (a, the, --) Pacific Ocean. It was quite warm.

62. Tom: You know my uncle Ed? He is a doctor, and once, he went on an expedition!
Louis: Where did he go?
Tom: He went to (a, the, --) South Pole! He spent a year down there!

63. Rick: I haven't seen your sister Clara in a long time.
Marilyn: That's because she is away. She is doing research in South America. She is living near (a, the, --) Amazon River. She studies birds that live in that area.

64. Teacher: Tell me about London.
Student: London is in (a, the, --) United Kingdom. It's a very big city.

VIIC. Proper names

65. Louise: I just saw a movie about a ship that was hit by an iceberg, a long time ago. But I can't remember what this ship was called!
Betsy: It was called (a, the, --) Titanic. It was very famous!

66. Roberta: What did you on your last vacation?
    Fred: I went to Egypt!
    Roberta: Wow, that's really exciting. What did you see there?
    Fred: I saw (a, the, --) Great Pyramids. They are really huge!

67. Janet: I just came back from Paris. It's so beautiful!
    Peter: What did you see in Paris?
    Janet: I saw lots of beautiful buildings. And I went to lots of museums. I went to
    the Louvre. It's such a wonderful museum!

68. Leo: My grandfather is a veteran.
    Chris: In what war he did he fight?
    Leo: He fought in (a, the, --) Second World War. He fought for four years!

VIID. Definite generics

69. Peter: Is Sally home? I need to talk to her right away.
    Kim: You'll have to wait a few minutes. She is talking on (a, the, --) telephone.
    I'll tell her you are here.

70. Angela: How was your dinner with your aunt and uncle last night?
    Charles: Quite boring. They are very nice people, but we don't have much
    to talk about. So we talked about (a, the, --) weather. And about my uncle's
    health.

71. Pauline: There is so much happening in our world today.
    Rob: I know! It's hard to follow everything.
    Pauline: Well, I watch (a, the, --) news every day. That way, I know what's
    happening.

72. Sam: Hi, Ben. I didn't know that you were in Boston! How did you get here?
    Ben: I drove here from my home in Virginia.
    Sam: That's a long way! Were you bored?
    Ben: A little. I listened to (a, the, --) radio while I drove. That made my trip
    more exciting.

VIII. Possessive fillers

73. Julie: What did you do last night?
    Peter: Not much. I just worked on (a, the, --) my physics homework.

74. Louise: I tried to call you yesterday, but the line was busy.
    Angela: My husband was talking to (a, the, --) his mother.

75. Ron: Where is your little daughter?
    Janine: She is playing with (a, the, --) her dolls.

76. Sam: What's wrong?
    Ed: I'm so sorry. I broke (a, the, --) your favorite teacup.

Superlatives and ordinals

IXA [+definite, +specific]
Superlatives with speaker knowledge
77. *In a college*
    Leon: What classes are you taking next semester?
    Laura: Japanese Literature, Calculus, Physics – and French History; for French
    history, I hope to get (a, the, --) best professor in this college – Dr. Leary.
78.
    Betsy: What are you going to study when you go to college?
    Kendra: I will study Italian and Spanish films. I especially want to study (a,
    the, --) most wonderful director in Italy – Federico Fellini.

    Jim: What are you going to do this summer?
    Rose: I'm taking a trip to Paris! I am going to go to the Louvre. I want to see (a,
    the, --) most famous painting in Europe – the Mona Lisa! I studied it in my art
    history class.

79. Eric: What are you doing for Thanksgiving week-end?
    Ed: I am going out west – to Arizona and New Mexico.
    Eric: That sounds like fun! What are you going to do there?
    Ed: Lots of things – hike, climb mountains… And I also plan to visit (a, the, --
) oldest house in the U.S. – it's in Santa Fe, New Mexico; I heard that it's a very
    interesting place.

IXB [+definite, -specific]
Superlatives without speaker knowledge

80. Sarah: I want to know what the meaning of life is.
    Julie: How can you find out?
    Sarah: I want to talk to (a, the, --) most intelligent person on Earth!
    Unfortunately, I don't know who that is…

81. George: My mother needs to have an operation.
    Anne: Are you worried?
    George: A little. But I'm doing something about that! I am trying to find (a,
    the, --) best doctor in Boston – I don't know who that is, but I will find out!

82. *In a library*
    Client: Excuse me, can you help me?
    Librarian: Certainly. What can I do for you?
    Client: I would like to read (a, the, --) most interesting book in this library
    – but I don't know what that is. Can you tell me?

83. Jane: Someday, I'd like to travel all over North America! And see everything –
    rivers, deserts, mountains, lakes, forest…
    Ellen: That sounds wonderful!
    Jane: And I hope to climb (a, the, --) tallest tree in North America; I don't know
    what it is or where it is, so I'll have to find out!

XA [+definite, +specific]
Ordinals with speaker knowledge

84. Jerry: What are you going to do this week-end?
    Lucy: I am going to read! I plan to read (a, the, --) third book about Harry Potter
    – "Harry Potter and the Prisoner of Azkaban".

85. *At a college*
    Andy: Guess what? I might be moving from my old office.
    Sandy: Really? Where?
    Andy: I want to get (a, the, --) fifth office in this hall, #205: it's really big, and it
    has two windows!

86. *At a supermarket*
    Lesley:It's almost our turn in line.
    Sarah: Not yet. I'll be right back.
    Lesley:Where are you going?
    Sarah: I am going to talk to (a, the, --) fourth person in this line – it's my friend
    Peter!

87. *At a college party*
    Roger: Are you enjoying this party?
    Chester: Yes, it's a lot of fun! Lots of students!
    Roger: Yeah, I know. I feel sorry for (a, the, --) only professor in this
    room – Dr. Richardson. He looks bored!

XB [+definite, -specific]
Ordinals without speaker knowledge

88. Julie: How did you spend your summer vacation?
Ruth: I read a lot! I started reading a really interesting trilogy. Now I am trying to find (a, the, --) second book of this trilogy – I don't know what its name is, or what it's about, unfortunately... But I really want to know what happens next, so I have to find it!

89. *In an apartment building*
Manager: How can I help you?
Client: I'd like to rent an apartment in this building. I want to live on (a, the, --) fourth floor of this building – I don't know what it looks like, but I like living up high!

90. *In a line for movie tickets*
George: That clerk just made an announcement. Did you hear what he said?
Ella: He said that there are only five tickets left! We are lucky, since we are next! And there are so many people behind us.
George: Yeah, we are lucky. You know, I feel sorry for (a, the, --) sixth person in this line – even though I have no idea who that is.

91. Sarah: Our company just formed a committee to look into women's role in the office.
Clara: That's wonderful!
Sarah: But almost all the people on this committee are men! I'd like to talk to (a, the, --) only woman on this committee – I don't know who she is, but I want to find out what she thinks about this situation.

Appendix 3: Translation task

This appendix discusses the translation task used in the second study of L2-English article choice (Chapter 5). The characteristics of the L2-learners who took this task, including their proficiency levels, are discussed in Chapter 5.

The translation task was originally designed to test L2-learners' implicit knowledge of English articles: unlike the elicitation tasks, the translation task did not reveal to the learners that they were being tested on articles. The predictions for the translation task were the same as for the elicitation task discussed in Chapter 5: L2-learners were predicted to correctly use *the* with (specific) definites; to correctly use *a* with non-specific indefinites; and to overuse *the* with specific indefinites. (This task did not test non-specific definites).

However, as discussed below, the translation task encountered several methodological problems. While the results were in the predicted direction, they should be interpreted with caution because of the methodological problems. This task is therefore not discussed in the body of the thesis. A follow-up study used a different (and more successful) methodology for testing L2-learners' implicit knowledge of English articles: collection of written production data, discussed in Chapter 7.

I now briefly discuss the design of the translation task and the corresponding results.

1. Translation task: design

In the translation task, the learners were asked to read a story in their L1 and to translate some of the sentences into English. The 36 sentences to be translated were numbered and underlined. Unbeknownst to the learners, the sentences were designed to elicit articles in different context types. The translation task was always administered before the elicitation task, so that the L2-learners did not know they were being tested on articles. The translation task was accompanied by a translation sheet which translated potentially unfamiliar words from the learners' L1 into English.

1.1. Story format

The story that the learners read was about a young immigrant woman's week-end trip to New York. The versions of the story given to the Russian and Korean participants differed only in the characters' names and backgrounds: the characters in the Russian version of the story were immigrants from Russia, with Russian names, and the characters in the Korean version were immigrants from Korea, with Korean names. Whenever example items are given here, the Russian character names are used. The full story, with Russian character names, is given in Section 3.2. The version given here is entirely in English, with the target translations of the underlined sentences.

The target sentences in the story were divided into nine contexts of four items each. Five of the contexts are summarized below. The remaining four contexts addressed use of generics, and will not be discussed here. All examples given here are the target English translations of the items. The target DP, which was always in a postverbal position, is underlined.

Two of the contexts contained wide-scope indefinites. In one of them, the DP was followed by additional information about the referent, as in (1a); in the other, the

continuation provided no information about the referent, as in (1b). Both contexts could in principle be specific and could thus receive *the* in L2-English; we tested whether including additional information about the referent (as in (1a)) contributed to overuse of *the*. (Recall that a continuation discussing the referent of the target DP contributed to overuse of *the* in the elicitation tasks reported in Chapters 5 and 6).

Another context involved a narrow-scope, clearly non-specific indefinite (1c), and was predicted to receive only *a* in L2-English.

1. a) She is going to marry <u>a famous doctor</u> – he lives in our neighborhood.
   b) Lena bought <u>a pretty toy</u> for her cousin – she loves her cousin very much.
   c) I hope to buy <u>a new car</u> next year.

Two more contexts were designed to elicit *the* with definites. One (2a) contained *previous mention* definites, and the other (2b) contained *associative use* definites: definites which satisfy the uniqueness presupposition through association. For instance, in (2b), *the door* is associated with *her relatives' apartment.* Since an apartment typically has only one (front) door, the uniqueness presupposition is satisfied. Both types of definites are specific, since the speaker clearly has knowledge of the referent's identity, and this identity is noteworthy: it is important that Lena ate the sandwich *that her aunt gave her,* rather than some other sandwich; it is similarly important that Lena knocked on the door *of her relatives' apartment,* rather than on just any door.

2. a) [previous mention of *a sandwich*] Lena thanked her aunt and ate <u>the sandwich</u>.
   b) She knocked on <u>the door of her relatives' apartment</u>.

All of the definite and indefinite target items in the translation task were singular. This was done because Korean does not have a singular/plural distinction. In the Korean version of the task, all of the target DPs were bare, unmodified by classifiers or numerals, and were therefore understood as singular.

1.2. Pilot data

The translation task was piloted with seven L1-English controls. These were the same L1-English speakers who took the elicitation task discussed in Chapter 5. Since the L1-English speakers could not be asked to translate the story from some other language to English, the pilot proceeded as follows. The L1-English speakers were given the English version of the story, including the 36 underlined target sentences. The target sentences were ungrammatical: they were missing articles, verbal inflection, and auxiliaries. The L1-English speakers were told that these sentences needed to be corrected for grammar (rather than content). The control speakers were instructed to add articles, verbal inflection, and auxiliaries in order to make the sentences grammatical.

The first version of the story administered to the L1-English speakers failed to elicit the target article in all contexts. The version was changed three times before it succeeded in eliciting the target article from the L1-English speakers. The final version of the story was administered to the L2-learners. This is the version given in Section 3.

2. Translation task: results and discussion

I now move on to a brief discussion of the translation task. Before discussing the results, I consider some methodological problems with this task.

The translation task proved rather difficult for the participants: when giving feedback after the test, many L2-learners reported that the translation task was hard for them and that they had difficulty with much of the vocabulary. The task was also very time-consuming – the L2-learners were given 60 minutes to complete it, and nearly all used the entire 60 minutes; in contrast, most learners finished the elicitation task in less time than the allotted 60 minutes. Given the difficulty and time-pressure of the translation task, it is likely that many participants focused their attention on the underlined target sentences, ignoring the surrounding context; since an evaluation of a DP as definite or specific is crucially dependent on the context, this is a serious problem. Thus, the results reported in this section should be taken only as an approximation of L2-learners' understanding of English articles.

2.1. Results of the intermediate/advanced L2-learners

Of the 37 intermediate/advanced L1-Russian participants, two participants failed to complete large portions of the translation task, simply leaving lines on the translation sheets blank; their results had to be excluded from the analysis. Of the 37 intermediate/advanced L1-Korean participants, one participant did not complete the translation task at all, saying that she was found it too difficult. All results reported below, therefore, are for 35 L1-Russian speakers and 36 L1-Korean speakers.

The results of the translation task are given in (3) and (4) for the L1-Russian and L1-Korean speakers, respectively. The percentages of *the* use, *a* use, and article omission are reported. The category *other* contains a variety of translations that do not fit neatly under the previous three categories. For instance, if an L2-learner used a plural instead of a singular form of the DP in the translation, this was counted as *other*. The only instances of plural form use occurred in the narrow scope indefinite category, and all were bare plurals. In the case of definites by association, L2-learners frequently used a possessive: for instance, instead of writing *the door of her relatives' apartment,* L2-learners often wrote *the apartment's door* or *her relatives' apartment door.* In cases such as these, it is impossible to tell whether *the* marks definiteness on the head noun (*door*) or the possessor (*apartment*). Therefore, all cases of possessives were counted under *other.*

3. Translation task results: L1-Russian intermediate/advanced L2-learners (N=35)

Category (example number)	%the	%a	%omission	%other
definites – previous mention (2a)	71%	19%	8%	2%
definites – association (2b)	57%	9%	12%	22% (incl. 19% use of possessives)
wide scope indefinites, with continuation (1a)	18%	57%	25%	0%
wide scope indefinites, no continuation (1b)	16%	59%	24%	1%
narrow scope indefinites (1c)	5%	70%	23%	2% (all bare plurals)

4. Translation task results: L1-Korean intermediate/advanced L2-learners (N=36)

Category (example number)	%the	%a	%omission	%other
definites – previous mention (2a)	63%	22%	13%	2%
definites by association (2b)	71%	8%	13%	9% (incl. 6% use of possessives)
wide scope indefinites, with continuation (1a)	10%	70%	19%	1%
wide scope indefinites, no continuation (1b)	8%	72%	19%	1%
narrow scope indefinites (1c)	8%	67%	19%	5% (incl. 4% use of bare plurals)

2.1.1. Article use with definites

As the above tables show, L2-learners omitted articles across all contexts, to a greater extent than they did with singular contexts in the forced choice elicitation task; this is not surprising, since the learners were paying less attention to article use in the translation task and may have omitted articles as a result of performance pressure. More surprisingly, there was fairly high use of *a* with previous-mention definites in the translation task – an effect not found in the elicitation tests. Notably, use of *a* was much lower with definites in association contexts.

This difference most likely stems from the nature of the task. In order to determine definiteness on DPs in association contexts, the learners had to read only the underlined target sentence: upon reading *door of her relatives' apartment,* they could deduce (from their world knowledge) that an apartment should have only one door, and mark uniqueness via *the.* In contrast, previous-mention definites get their definiteness from the preceding context rather than from world knowledge: an L2-learner will mark *sandwich* as definite only if she realizes that a sandwich has been previously mentioned. It is quite probable that L2-learners often read only the target sentence, disregarding the preceding context (which they did not have to translate): in that case, they would not know that a word like *sandwich* was a previous-mention definite, and would treat it instead as a first mention indefinite – hence the use of *a.*

2.1.2. Article use with indefinites

Turning to indefinites, we see that there was more overuse of *the* with wide-scope than with narrow-scope indefinites, but that this difference was not very great, especially in the case of the L1-Korean speakers. Nor did a continuation with additional information about the referent contribute to overuse of *the* for wide-scope indefinites. In order to determine what contributed to overuse of *the,* I now look at performance on individual item types. (5) through (7) list individual items, and (8) gives the L2-learners' performance on each item. Only percentages of use of *the* and *a* are reported, excluding instances of omission.

5. *wide scope indefinite with continuation:*
   a) You must congratulate your second cousin Katya! – her aunt exclaimed – she is getting married! She is going to marry <u>a famous doctor</u> – he lives in our neighborhood!

b) Lena and Masha left the museum, and took the subway downtown. They went to a tall white building – Lena's uncle's office was in it.

c) "I've also long wanted a pet," – she said – "I've wanted one ever since I was little. Once, I wanted to get a beautiful white cat – I saw it at our local pet shop!"

d) Lena and Anna turned around. They saw a small gray mouse – it was running across Broadway!

6. *wide scope indefinite, no continuation:*

a) Lena went to a toy store to get Masha a gift. She chose a pretty doll – Lena loves her little cousin very much.

b) Then Masha switched to a new topic. – Lena, guess what? I got a dragon for my birthday – and lots of other gifts, too!

c) They stopped by a jewelry boutique. Lena bought a necklace – she likes jewelry.

d) Lena's aunt was still awake. She was reading a magazine – she usually goes to bed late.

7. *narrow scope indefinite:*

a) She had heard a lot about the Broadway musicals, the many art museums, and the various other attractions of this city. Lena wanted to see a musical.

b) When I grow up – little Masha said suddenly - I want to marry an actor.

c) Anna looked at the lady's dogs with envy. "It's such a pity that my dorm won't let one keep pets. I would like to own a dog very much."

d) Anna asked Lena if she had a car. "No" – Lena answered – "But I hope to buy a car next year."

8. Performance on individual item types

	L1-Russian speakers		L1-Korean speakers	
	%the	%a	%the	%a
wide scope, continuation				
(5a) – famous doctor	11%	43%	14%	64%
(5b) – tall white building	43%	37%	19%	58%
(5c) – beautiful white cat	11%	71%	3%	78%
(5d) – small gray mouse	6%	83%	6%	81%
wide scope, no continuation				
(6a) – pretty doll	9%	57%	0%	89%
(6b) – dragon	11%	69%	17%	67%
(6c) – necklace	20%	57%	6%	81%
(6d) – magazine	29%	57%	8%	78%
narrow scope				
(7a) – musical	9%	46%	11%	44%
(7b) – actor	6%	74%	3%	78%
(7c) – dog	3%	80%	14%	67%
(7d) – car	3%	80%	6%	78%

From (8), we see that, in the case of wide-scope indefinites with a continuation, two items – the *famous doctor* and *tall white building* items – were more likely to elicit *the*

from both groups of L2-learners than the other two items. These two items are in fact more likely to contain specific indefinites: the speaker has clear knowledge of the doctor's identity in (5a) and the building's identity in (5b); both the doctor and the building are important for the discourse. On the other hand, the identity of the cat in (5c) is not particularly important for the discourse, and the identity of the mouse in (5d) is completely irrelevant.

In the case of wide-scope indefinites with no continuation, no particular pattern is observed. The same holds for narrow-scope indefinites.

2.2. Results of the beginner learners

Of the 13 L1-Russian beginner learners, two failed to complete large portions of the translation task. The results reported in (9) are therefore for 11 L1-Russian beginners plus the single L1-Korean beginner.

9. Translation task results: all beginners (N=12)

Category (example number)	%the	%a	%omission	%other
definites – previous mention (2a)	27%	27%	35%	10%
definites by association (2b)	25%	13%	33%	29% (incl. 10% use of possessives)
wide scope indefinites, with continuation (1a)	15%	23%	63%	0%
wide scope indefinites, no continuation (1b)	17%	31%	46%	6%
narrow scope indefinites (1c)	4%	35%	56%	0%

Not surprisingly, the beginners' results include more article omission than the results of the intermediate/advanced learners, and there is more noise in the data. The beginners do, however, make a distinction between wide scope and narrow scope indefinites, as expected.

2.3. Summary

To summarize, the results of the translation task are rather noisy. There is a general indication that L2-learners use *the* more with definites than with indefinites, and overuse *the* more with specific than with non-specific indefinites. However, the results of this task need to evaluated with caution, since the difficult nature of the task probably made the participants disregard much of the context crucial for article evaluation. The high overuse of *a* with previous-mention definites (compared to definites by association) in particular suggests that the learners often disregarded the context. Thus, the results of the translation task can be taken as only an approximation of article choice in L2-English.

3. The story

This section gives the English version of the story administered to the L2-English learners.

3.1. Codes

The underlined sentences in the story below fall into nine types, which are given in (10). The types corresponding to the codes 1a, 1b, 1c, 2a, and 2b were discussed in Section 1.1. The other four types are related to genericity. Type 3a consists of kind-denoting DPs, which in English can be expressed either through bare plurals (*Penguins come from Antarctica*) or singular definites (*The penguin comes from Antarctica*). In the Russian version of the story, these DPs were expressed by bare plurals, since Russian does not have singular kind-denoting DPs of this type. In the Korean version, these DPs were expressed by bare NPs with no demonstratives or classifiers; Korean does not have plural morphology.

Type 3b consists of indefinite singular DPs in characterizing sentences (e.g., *A policeman always helps people*). Type 3c consists of generic definite singular DPs in object position, such as *the dollar* and *the telephone*. Such DPs are expressed by bare singulars in Russian and by bare NPs in Korean.

Finally, type 4 consists of number phrases (e.g., *three friends*); these items were included in order to check whether the L2-learners used plurality marking in obligatory contexts.

10. 1a: wide-scope indefinite, no additional information about the referent
　　1b: wide-scope indefinite, additional information about the referent
　　1c: narrow-scope indefinite
　　2a: definite by previous mention
　　2b: definite by association
　　3a: generic, subject position: either singular definite or bare plural
　　3b: generic, subject position, singular indefinite
　　3c: generic, object position, singular definite
　　4: number control

The story given in the next section contains the codes next to the target DPs. The story administered to the L2-learners did not contain these codes.

3.2. The story

<div align="center">A Trip to New York</div>

Lena is a young woman who came to America from Moscow three years ago. She lives in Boston. Lena is twenty-seven years old; she is a graduate student. Lena has three close friends(4). One of these friends, Anna, lives in New York.

Last summer, Lena went to New York for the first time in her life. She had heard a lot about the Broadway musicals, the many art museums, and the various other attractions of this city. Lena wanted to see a musical (1c). She also wanted to visit some of New York's famous museums.

Lena decided to stay with her aunt, who lives in Brooklyn. Her aunt has a young daughter named Masha. Lena went to a toy store to get Masha a gift. She chose a pretty doll(1a) -- Lena loves her little cousin very much. The night before she left, Lena packed her suitcase. She also made two sandwiches (4) for her trip.

Lena went to New York by bus. The bus was very full and she had to sit in the very back. Next to Lena sat an old man. This man was traveling with two little boys (4). They were playing cards.

Lena looked out of the window for a while, but soon she got bored. She took a book and a magazine from her bag, and tried to decide which one looked more interesting. The magazine looked boring. So Lena opened the book(2a). Lena hardly noticed the time passing by as she read. "A good book (3b) always makes travel more pleasant" – she thought.

Eventually, Lena finished reading her book and got into a conversation with her elderly neighbor. Lena told him that she was planning to stay with her relatives, who she hadn't seen in a long time. "I'm not sure what I should bring them" – she told the old man. "Well, if I were you, I would bring them something sweet, like chocolates or a cake – everyone loves dessert," – said the old man. "You are right," – Lena exclaimed – "A box of chocolates (3b) is a good gift for anyone!"

Just then, a police car went past the bus and attracted the attention of the old man's grandsons. "When I grow up" – the older boy exclaimed – "I am going to be in the police!" "Why?" - Lena asked him. "Because I like helping people. And a good policeman (3b) always helps people!" - the boy answered.

"And what do you want to be?" – Lena asked the younger boy. "I won't be anything" – he announced – "I'll just have a lot of animals. I like all the animals that you find in a circus, like giraffes and monkeys; I'll keep them all at home, in Boston!" "But, - Lena objected - Giraffes (3a) come from Africa. Boston would be too cold! And monkeys (3a) live in groups. You probably won't have enough room in your apartment!"

So, talking back and forth, they finally found themselves in New York, at the bus terminal. Lena bid her neighbors good-bye and got on the subway. While Lena was waiting for the train, she noticed a cat on the platform. She stroked the cat(2a).

Then the train finally arrived, and Lena went to Brooklyn. It took the train a long time to get there, and by the time Lena got to her aunt's apartment, it was nearly seven in the evening. She knocked on the door (2b) of her relatives' apartment. Lena's little cousin Masha opened the door for her. Lena's relatives were very happy to see her.

While Lena ate her supper, her aunt brought her up to date on all the family news.

– You must congratulate your second cousin Katya! – her aunt exclaimed – she is getting married! She is going to marry a famous doctor (1b) – he lives in our neighborhood!

– When I grow up – little Masha said suddenly - I want to marry an actor (1c). Then Masha switched to a new topic. – Lena, guess what? I got a dragon (1a) for my birthday – and lots of other gifts, too!

– What do you mean – dragon? – Lena wondered. - (Dragons (3a)exist only in fairy tales!

– Actually, it's a lizard – Masha admitted – her name is Camilla.

- That's wonderful – said Lena – when I was your age, I didn't have any pets. But I really wanted one – I wanted a penguin!

- That's funny! – said Masha - Penguins (3a) live in Antarctica! Not in Russia!

- Well, I didn't know that when I was little – said Lena.

Soon everyone went to bed. The next morning, Lena went to Manhattan, and took little Masha with her. First, they took a walk in Central Park. They saw a really tall tree

there. <u>They measured the shadow (2b)of this tree</u>. Masha really enjoyed this project. Then Masha played near the tree for a while, and Lena read a book.

- What are you reading? - Masha asked her cousin.

- I am reading a book about Spain - Lena said - it has a lot of interesting information in it. I just found out a really interesting fact.

- What is it? - Masha asked.

- <u>Spain is home to the dollar(3c)</u>! - Lena exclaimed.

Then, the two cousins went to a museum of a history. They looked at an exhibit about the development of communications technology. Masha read a plaque inside the exhibit window. "Alexander Graham Bell – she read – what a funny name!" "What is he famous for?" – she asked her cousin. "Don't you know?" – Lena was surprised - "<u>He invented the telephone(3c)</u>! We should be very grateful to him!"

Then Lena and Masha went to a room which contained various prehistoric artifacts. A plaque on the wall told them a little bit about life in prehistoric times. "Wow, - Masha said – people had a really hard life back then, didn't they?" "Yes, - Lena agreed – they didn't have a lot of the things that we take for granted." "Like what?" – Masha asked. "Well, for instance, they didn't have any kind of vehicles" – Lena told her – "<u>This was before people invented the wheel(3c)</u>!"

Then, Lena and Masha went to the museum gift shop. Masha found it boring, but Lena saw a couple of interesting things there – a cup with the museum's name on it, and a book about automobiles. <u>Lena bought the cup (2a)</u>. She found the book too expensive.

Lena and Masha left the museum, and took the subway downtown. <u>They went to a tall white building (1b)</u> – Lena's uncle's office was in it. The uncle was finishing work and was going to take Masha home to dinner. As for Lena, she went to see her friend old Anna. They hadn't seen each other since Lena had left Moscow.

Anna was very happy to see her friend. The two friends decided to go for a walk together. They stopped by a jewelry boutique. <u>Lena bought a necklace (1a) – she likes jewelry.</u>

Then Lena and Anna walked to the discount theater ticket booth, and got in line. In front of them stood a lady. <u>This lady had three beautiful dogs (4).</u> The girls petted them. Anna looked at the lady's dogs with envy. "It's such a pity that my dorm won't let one keep pets. <u>I would like to own a dog (1c)very much.</u>" Lena felt sorry for her friend. "I've also long wanted a pet," – she said – "I've wanted one ever since I was little. <u>Once, I wanted to get a beautiful white cat (1b) – I saw it at our local pet shop</u>! But my parents said no. <u>A pet (3b)needs a lot of attention.</u>"

Suddenly, the dogs started barking! Lena and Anna turned around. <u>They saw a small gray mouse (1b) – it was running across Broadway</u>! Lena took out her camera, but it was already too late.

Then their turn came. Lena and Anna were quite lucky – they managed to buy tickets for "Les Miserables." The two friends had a quick dinner and went to the theatre. Other people were driving up to the theater in cabs or in their own cars. Anna asked Lena if she had a car. "No" – Lena answered – "<u>But I hope to buy a car (1c) next year</u>. It's funny, isn't it – most Americans can't imagine life without a car. Where were cars first made in the U.S., do you know?" "As far as I know," – Anna answered – "<u>The home of the automobile (3c) is Michigan.</u>"

The girls had a great time at "Les Miserables." When it was over, they hugged each other good-bye, and Anna promised to visit her friend in Boston. Lena went back to her relatives'. She rang the bell (2b) on her relatives' door.

Lena's aunt was still awake. She was reading a magazine(1a) – she usually goes to bed late. Then, Lena's aunt decided to feed Lena. She opened the refrigerator (2b) in her kitchen. She took out some bread and cheese, and made a sandwich for Lena. Lena thanked her aunt and ate the sandwich(2a).

She offered Lena some supper, and asked Lena about her day. "I saw my friend Anna, I went to the museum, and I saw "Les Miserables", a famous musical"– said Lena. Then it was time to go to sleep. Lena had had a very exciting first day in New York.

Appendix 4: Presuppositionality and word order in L2-English

In this appendix, I report some properties of L2-English article choice which are not related to either definiteness or specificity. These are *presuppositionality* and *word order*.

1. Presuppositionality and L2-English article choice

In the last two studies of L2-English article choice, reported in Chapters 5 and 6, we investigated the effects of *partitivity / presuppositionality* on L2-English article choice. An indefinite is presuppositional if the set denoted by its restrictor NP is presupposed to exist in the discourse (and to be non-empty). It is additionally partitive if the set has been explicitly mentioned in the discourse.

This is illustrated in (1): on the most likely reading of (1a), *a cat* denotes a member of the previously established set of cats, so *a cat* is partitive (and presuppositional). In (1b), on the other hand, *a cat* is not partitive.

1. a. Several cats were sitting on my doorstep. I stroked <u>a cat</u> (=one of the cats)
   b. Several cats were sitting on my doorstep. I later also saw <u>a (different) cat</u> in the park.

Partitivity in indefinites may receive morphological expression in some languages. As discussed in Chapter 2, Enç (1991) argues that when an indefinite in Turkish bears accusative-case, it necessarily gets a presuppositional reading (but see Kelepir 2001 for counter-examples). Membership in a previously established set has also been found to contribute to overuse of *the* in early L1-English (Maratsos 1976, Schafer and de Villiers 2000; see the overview and discussion in Chapter 3).

Additionally, Chapter 3 reported on the study of Kaneko (1996), which found that L1-Japanese speakers overused *the* to a very high extent in contexts like (2), compared to other indefinite contexts. L1-Spanish speakers did not show the same effect.

2. Once there was a boy. He wanted to write a letter. He went to his mother. She showed him some pencils. So he took ___ pencil. And he wrote his letter.

This background led us to test the effects of presuppositionality / partitivity on article choice among L1-Russian and L1-Korean learners of English.

1.1. Partitivity in Elicitation Study 2

The learners who participated in Study 2 (reported in Chapter 5), were tested on the context given in (3), in which the indefinite is presuppositional: the existence of a set of girls is presupposed. The indefinite here is also a covert partitive.

3. *presuppositional/partitive indefinite*
   Robert: My little son Johnny went to the park yesterday. There were three little girls and four little boys already playing there.
   Lisa: Did Johnny play with them?
   Robert: Well, Johnny is usually very shy. **But finally, he said hello to (a, the, - -) little girl.**

Additionally, one of the non-specific context types tested in Chapter 5 also contained indefinites which may be interpreted as presuppositional. This was the "denial of speaker knowledge" context, repeated in (4): the D-linked phrase *which one* in this context implies the existence of a salient set of students. If presuppositionality is found to affect L2-English article choice in contexts such as (3), it may also affect article choice in contexts like (4), albeit to a smaller degree. While the existence of a salient set is explicitly stated in (3) (in which we learn of the existence of three little girls), it is only implied in (4); the indefinite in (4) is not partitive.

4. *non-specific indefinite, explicit denial of speaker knowledge, possibly presuppositional*
   Professor Clark: I'm looking for Professor Anne Peterson.
   Secretary: I'm afraid she is out right now.
   Professor Clark: Do you know if she is meeting somebody?
   Secretary: I am not sure. **This afternoon, she met with (a̱, the, --) student – but I don't know which one.**

We compared performance on the clearly partitive context in (3), the potentially partitive context in (4), and the simple (non-specific, non-partitive) indefinite context, repeated in (5).

5. *non-presuppositional indefinite*
   Visitor: Excuse me - can you help me? I'm looking for Professor James Smith.
   Secretary: I'm afraid he's not here right now.
   Visitor: Is he out today?
   Secretary: No, he was here this morning. **He met with (a̱, the, --) student... but I don't know where Professor Smith is right now.**

The results are reported in (6). This table reports percentages of *the* and *a* use in each context; the remaining percentage points represent article omission.

6. Effects of presuppositionality: intermediate/advanced L2-learners

	L1-Russian speakers		L1-Korean speakers	
context	*%the*	*%a*	*%the*	*%a*
Presuppositionality established (3)	19%	75%	30%	70%
Presuppositionality implied (4)	10%	89%	13%	83%
No presuppositionality (5)	14%	85%	2%	95%

We see the following pattern. For L1-Russian speakers, presuppositionality has little noticeable effect: overuse of *the* with presuppositional indefinites (3) is very low, and is not significantly different from overuse of *the* with simple indefinites (5)[1] (p = .22).

In contrast, overuse of *the* with presuppositional indefinites is quite high (30%) for L1-Korean speakers. Overuse of *the* with non-presuppositional indefinites (5) is significantly lower than overuse of *the* with either true partitive indefinites (3) (p < .0001) or indefinites which may be presuppositional (4) (p < .0001).

---

[1] There is a marginally significant difference in overuse of *the* with true partitive indefinites (3) and indefinites which have denial of speaker knowledge (4) (p = .06).

1.2. Presuppositionality in Elicitation Study 3

We tested the effects of presuppositionality again in the third study of L2-English article choice (reported in Chapter 6). In Study 3, the stimulus items were all in English, while in Study 2, the items had been in the learners' L1, with only the target sentence in English. We wanted to be certain that the presuppositionality effect observed in Study 2 was not linked to something in the original Korean contexts, or to code-switching.

The elicitation test in the third study involved four items in the context of partitive indefinites, an example of which is given in (7). Performance on these items was compared to performance on simple non-partitive indefinites, an example of which is given in (8).

7. *partitive indefinite*
   Rudolph: My niece Janet likes animals a lot. Last week, she decided to get a pet and went to a pet shop.
   Lisa: Did she find any pets that she liked?
   Rudolph: Yes – she saw three beautiful puppies and six lovely kittens. She couldn't decide! **Finally, she bought (a̲, the, --) kitten.**

8. *non-partitive indefinite*
   Mary: I heard that it was your son Roger's birthday last week. Did he have a good celebration?
   Roger: Yes! It was great. He got lots of gifts – books, toys. And best of all – **he got (a̲, the, --) puppy!**

The results are reported in (9). As in the previous study, overuse of *the* in partitive indefinite contexts is greater for L1-Korean than for L1-Russian speakers.

9. Effects of presuppositionality: intermediate/advanced L2-learners

	L1-Russian speakers		L1-Korean speakers	
context	*%the*	*%a*	*%the*	*%a*
Partitive indefinite (7)	30%	60%	39%	59%
Non-partitive indefinite (8)	15%	69%	6%	91%

The difference in article use between partitive and non-partitive contexts was non-significant for L1-Russian speakers, whether use of *the* or use of *a* was measured. The difference was highly significant for the L1-Korean speakers, both for use of *the* ($p < .001$) and for use of *a* ($p < .001$).

1.3. Presuppositionality/partitivity in written production

The above findings led me to look at partitive contexts in the production data of L1-Korean speakers (see Chapter 7 for details concerning the data collection). I looked for all of the contexts in which L2-learners mentioned a set and then referred to a member of this set. There were eight such contexts in the data. In five of them, L2-learners appropriately used forms like *one of the X* or simply *one*. These are given in (10).

10. a) There are *3 closets* in my room. Q̲n̲e̲ is big enough to walk in. The others have shelves, so I put some clothes for my family.

b) There are *three desks*, a book shelf and a queen size bed. <u>One of the desks</u> has the shape of "L" and the others have the shape of the square.

c) The fabric contained *two different pictures*. <u>One</u> is a little brown bear and the other is a couple kissing each other.

d) There are *two rooms* in my house. <u>One</u> is for sleeping, the other is for studying.

e) There are *two desks*, three chairs and two tall book shelves. On <u>one of desks</u>, I put my computer that I use everyday to search the internet or check my e-mail and on the others, I arranged some books.

In two more contexts, given in (11), L2-learners put *a,* which is also quite appropriate. Interestingly, L1-English coders often did not consider *a* to be the appropriate article for these contexts. In the case of (11a), two of the four coders put *a* before *room* and two put *the.* In the case of (11b), all four coders put *the,* completely disregarding the fact that two closets had been mentioned.

11. a) My house has *two rooms* and a dining room. <u>A room</u> for my son, another for I and my husband.

b) I have a queen size bed, a vanity mirror and chair, a drawer, a TV set on a cabinet, a bookshelf, and *two closets*. My bed is white and fluffy. My wife and I bought it together two years ago. We are quite satisfied with it. The vanity mirror and chair was bought as a gift fro my wife. She likes it very much but it is in the way to <u>a closet</u>.

Finally, there was a single partitive context in which an L2-learner put *the.* This is given in (12). Use of *the* here should be inappropriate given the immediate previous mention of two desks. However, two of the four L1-English coders also put *the* in this context (the two others put *a*).

12. *Two desks* are lined up toward the wall. One 17''monior on <u>the desk</u>.

Thus, it appears that L2-learners do not use *the* as a marker of partitivity in production: *the* occurs in only one of the eight partitive contexts, and L1-English speakers in fact allowed *the* in this context.

L2-learners also appropriately used overt partitives like *one of X* or *some of X* in ten other contexts where the relevant set is not previously mentioned. This suggests that overt partitive forms are quite productive in the learners' grammar.

A few examples are given in (13). Curiously, there was one instance (13d) in which the learner put *the* before *one.* This use of *the,* however, appears to reflect specificity rather than partitivity: the learner has used an overt partitive to mark the fact that the news was not *the saddest,* but rather *one of the saddest.* The learner may be using *the* to signal intent to refer to a very particular instance of sad news (the sad news concerning the stolen watch).

13. a) On my last birthday, I receive a special gift from <u>one of my best friends</u>.

b) Atlanta airport is <u>one of the biggest airports in U.S.</u>

c) <u>One of the professors of UF</u> was waiting for my family and helped us to set up afterwards.

d) One year later, the watch was stolen. it was <u>the one of the saddest news in my life</u>.

To sum up, there is no evidence that L2-learners use *the* to encode partitivity in production. Additionally, none of the cases of *the* overuse with indefinites reported in Chapter 7 for the Korean speakers are presuppositional in nature: in using *the* with first-mention referents like *frame, baseball player,* and *cruise* (see Chapter 7 for examples), the learners cannot be presupposing that their reader shares knowledge of a salient set of frames, baseball players, or cruises.

1.4. Discussion

The above summary of results shows that, on the one hand, Korean speakers (but not Russian speakers) overuse *the* to a large extent in elicitation, and also show slight *the* overuse in non-partitive presuppositional environments. However, the same speakers appropriately use overt partitives in written production, and show no evidence of associating *the* with partitivity or presuppositionality.

1.4.1. A possible explanation: effects of topic-marking

The production data indicate that *the* is not associated with partitivity. However, the partitive contexts tested in the elicitation task are not only partitive: they are also contrastive. They always specify that something happened to a member of one set but not another: e.g., in (3), Johnny talked to *one of the little girls,* and not to *one of the little boys.* In contrast, most of the uses of overt partitives in production are not contrastive in this way; instead, the learners often use overt partitives to contrast one member of a set to another member of the set (as in (10)).

It may be relevant that overuse of *the* in elicitation is tied to *contrast between sets,* since this type of contrast is expressed by topic-marking in Korean.

An example of Korean topic-marking is given in (14), where the topic marker gives *book* a contrastive interpretation[2].

14. *John-i   chayk-un   ilk-ass-ta*
    John-NOM   book-TOP read-PST-DEC
    'John read a book (but not a magazine).'

The indefinites in the Korean variants of the partitive contexts in (3) and (7) are compatible with the topic marker, as shown in (15), the Korean variant of (3). The presence of topic-marking allows for a contrastive reading: Johnny talked to a girl rather than to a boy. (Topic marking is not obligatory here, however).

15. Korean translation of the last line in (3)
    kul-sey,   Johnny-nun   maywu   swucwupum-ul   manh-i   thaci.
    Well,   Johnny-TOP   very   shyness-ACC   much   bear
    haciman,   kyelkwuk   (han)   sonye-eykey-nun   insa-lul   hayss-e
    however,   finally   (one)   little.girl-to-TOP   greeting-ACC did
    "Well, Johnny is usually very shy. But finally, he said hello to a little girl."

It is less clear to what extent the "denial of speaker knowledge" items such as (4) are compatible with topic-marking in Korean. In these cases, no contrasting sets are present. According to Heejeong Ko (p.c.), it is still possible to use topic marking in the Korean equivalents of such contrasts if there is understood to be a contrast concerning the event.

---

[2] All of the Korean examples in this section are due to Heejeong Ko (p.c.)

Thus, in (16), the topic marker indicates that Professor Clark did at least one thing (talked to a student), as opposed to doing nothing.

16. Korean translation of the last line in (4)

Cal molu-keyssney-yo.   Onul ohwuey,   kunyeka   (han) haksayng-un
Well. don't.know-POL. This afternoon, she-NOM (one) student-TOP
manna-ess-ciyo —   haciman,   nwukwuinci-nun   molla-yo.
meet-PST-POL   however,   who it is-TOP   don't.know-POL.
"I am not sure. This afternoon, she met with a student – but I don't know which one."

On the other hand, the other indefinite contexts used in our studies, both specific and non-specific, are not compatible with topic-marking in Korean, since they cannot be understood as contrasting one event to another, or one set to another (see also the discussion in Chapter 3 concerning topic-marking in Korean). The reason that the "denial of speaker knowledge" contexts can be construed as contrastive is that they all involved questions of the form *Did she do anything / Do you know if she did anything / Do you know if she talked to anyone,* which set up a potential contrast between events. This was not the case for the other indefinite (or definite) contexts in the studies.

It is thus possible that topic-marking in Korean affects use of *the* in Korean speakers' L2-English. It is questionable whether this is in fact the case for the "denial of speaker knowledge" in (4), since the contrastive interpretation is not readily available – for instance, (4) could simply be expressing the fact that the professor had talked to somebody, without contrasting this event to any other. Overuse of *the* in these contexts was in fact fairly low (13%), and is not, on its own, sufficient to motivate a proposal linking topic-marking and *the* overuse.

On the other hand, in contexts that are clearly partitive, overuse of *the* is very high for the L1-Korean speakers, and use of topic-marking in the Korean equivalents is quite felicitous. It is thus possible that the learners are using *the* to mark contrastive topic. Since Russian does not have topic-marking, L1-Russian speakers do not show this effect.

According to Heejeong Ko (p.c.), the Korean equivalents of some of the partitive indefinite contexts in our tasks could also receive topic-making for a different reason: to indicate that their referents are *discourse-old* (by virtue of belonging to a previously mentioned set). Such use of topic-marking would be more felicitous if the indefinite undergoes scrambling. This form of topic-marking is inapplicable in the "denial of speaker knowledge" contexts. (See also Buring 1999 on different kinds of topics).

The account linking topic-marking to *the* overuse may be applicable to Kaneko's (1996) findings of *the* overuse in partitive contexts among Japanese speakers, since Japanese also has topic-marking. Kaneko's contexts are not contrastive, (see (2)), and it would be necessary to investigate whether topic-marking can be used in the relevant non-contrastive partitive contexts in Japanese.

More testing is necessary before a definitive conclusion concerning *the* overuse in partitive contexts can be reached. First, it is necessary to see whether *the* overuse would occur in partitive contexts which are not contrastive, to distinguish the potential effects of *contrastive topic marking* from those of *old information topic marking.* Second, it would be necessary to test learners on contexts that are clearly contrasting one event to another but that are not partitive, in order to distinguish the effects of contrastive topic-marking from the effects of partitivity. Third, it would be important to see whether the results of

the elicitation study can be replicated with a different methodology, which would give learners a greater freedom in the choice of responses. We saw that in the production task, where learners were free to use any morphological form that they chose, they typically used overt partitives in partitive contexts. However, they were not given this option in the elicitation task, where they had to choose between *the* and *a* (and the null article). It is important to see whether learners would overuse *the* in elicitation even when they have the freedom to use an overt partitive instead. If that is the case, then the overuse of *the* may have resulted from a confound in the elicitation task. I discuss some possible confounds in the next sections.

1.4.2. Possible methodological problems with partitive contexts

There is suggestive evidence that *a* is not actually considered a fully felicitous response in partitive contexts by L1-English speakers. While the English speakers in our studies consistently used *a* in partitive contexts (except for one control speaker who chose both *the* and *a*), recall that the study of Schafer and de Villiers (2000) found that L1-English adults overused *the* with partitives as much as 30% of the time (see Chapter 3), in contexts like (17)[3]. These adults used *a* only 25% of the time and *one of the* 45% of the time[4]. Thus, *a* was clearly a dispreferred response.

> 17. Three ducks and two dogs were walking across a bridge. One of the animals fell off the bridge and said "Quack". Guess which?/What was it?

Similarly, in the three partitive contexts in our production data in which L2-learners put *a* or *the* (see ex. 11 and 12), L1-English speakers also allowed both articles, with *the* being used more than *a,* on average. Recall that Maratsos (1976) also found adult L1-English speakers overusing *the* with partitives, but only when the speaker's attention was focused on a particular member of the set.

Thus, there may be a general problem with partitive contexts – even adult L1-English speakers may for some mysterious reason allow *the* in partitive contexts. In an area where L1-English speakers sometimes make mistakes[5], L2-English learners may well make more mistakes.

However, this cannot explain why L1-Korean speakers make more mistakes than L1-Russian speakers. While L1-Russian speakers did overuse *the* to a greater extent in partitive than in simple indefinite contexts, this difference was not significant; L1-Russian speakers had lower rates of *the* overuse with partitives than with specific indefinites. In contrast, overuse of *the* with partitives was as high or higher than overuse of *the* in any other indefinite context for the L1-Korean speakers. The partitive context was also the only context in which L1-Korean speakers performed noticeably worse than L1-Russian speakers, in either study.

---

[3] I obtained anecdotal support for Schafer and de Villiers' findings when discussing the example in (17) with the audience of the UCLA Psychobabble Seminar. At least two L1-English members of the audience said that they would use *the* in this context.

[4] These numbers are averages: Schafer and de Villiers report that in one of the two types of partitive contexts, *a* use was 20% and *one of the* use was 50%; in the other type of partitive context, the numbers were 30% and 40%, respectively.

[5] It is questionable whether the word "mistakes" is appropriate when native English speakers' performance is discussed, since native speakers' intuitions should reflect the underlying grammar. I use "mistakes" in this instance to mean "responses not expected under our current understanding of article semantics."

It is possible that there is an additional methodological confound contributing to overuse of *the* with partitives on the part of the L1-Korean speakers but not on the part of the L1-Russian speakers. This confound is traceable to the difference in plurality marking between the learners' L1: Russian has plurality marking on nouns (as well as adjectives), while Korean does not. There is evidence that speakers of languages with no plurality marking do not attend to plurality marking in English as much as speakers of languages which do have plurality marking: Wakabayashi (1997) found that L1-Japanese speakers had more difficulty than L1-Spanish speakers detecting the ungrammaticality of English sentences in which plurality marking was missing, such as (18)[6].

18. Jack went to the market yesterday. He bought *five apple* and one big pineapple to
make a dessert.   (Wakabayashi 1997:166)

This may have implications for performance in partitive contexts. Suppose that Korean speakers, unlike Russian speakers, do not consistently attend to presence vs. absence of plurality marking. In most contexts in our elicitation studies, this would not affect their performance. Even if they mistakenly treat a singular DP as plural, or vice-versa, this has no effect on whether or not the context is definite vs. indefinite, specific vs. non-specific. Moreover, the singular contexts usually make it fairly clear that a single individual is being discussed, and there is nothing to bias the learners into treating the DP as plural.

On the other hand, the partitive contexts *may* bias the learners into interpreting them as plural. Take a partitive context like (19) (from Study 3). Suppose that the learner ignores the fact that there is no plurality marking on *book,* and interprets the last sentence as meaning *...And as soon as he came home, he read the books.* This is a fully felicitous sentence in this context, and use of *the* is fully licensed with the plural. The context may be biasing the learners into treating *book* as plural because the plural reading is more felicitous than the singular reading: on the plural reading, *the book[s]* is a perfectly normal previous-mention definite. In contrast, *a book* is not fully felicitous: the partitive reading would be better conveyed by *one of the books.*

The same explanation can apply to Kaneko's (1996) findings with Japanese learners of English, since Japanese also does not have plurality marking, and since Kaneko's contexts (e.g., (2)) are also compatible with a plural reading of the target DP.

19. Gabrielle: My son Ralph didn't have anything to read last week-end. So, he
went to the library.
Charles: Did he find something to read?
Gabrielle: Yes – he took out three books and four children's magazines. **And
as soon as he came home, he read (a, the, --) book.**

Thus, it is possible that overuse of *the* with partitives among L1-Korean (and possibly L1-Japanese) speakers is due to a combination of two factors: (1) the general slight infelicity for use of *a* in partitive contexts, suggested by performance of L1-English speakers in previous studies; and (2) a possible tendency on the part of the L1-Korean speakers to ignore plurality marking in English. This is a speculative explanation, and it

---

[6] The Japanese speakers in this study were also less accurate than the Spanish speakers in detecting missing indefinite articles. This difference is irrelevant for my purposes, since both Korean and Russian lack articles; as seen in Chapter 8, both groups allow article omission in production.

is not unproblematic. Two of the eight partitive contexts tested in Studies 2 and 3 (ex. (7), and a similar example from the Study 2) make it fairly clear that only a single individual is being talked about (e.g., in (7), Janet intends to buy only one pet), yet *the* was still overused in these contexts.

While there was some omission of plurality marking in production (11 of the 40 L1-Korean speakers omitted –*s* in one obligatory context each), there was no relationship between omission of –*s* and overuse of *the* in partitive contexts in the elicitation task. On the other hand, only three of the 30 L1-Russian speakers allowed omission of plurality marking in production, consistent with the view that Korean speakers have more difficulty with plurality marking in English than do Russian speakers. One of these three L1-Russian speakers showed a very high rate of plurality omission (omitting –*s* in 5 instances, as opposed to one instance each for the other two L1-Russian speakers); interestingly, this learner showed a very high (75%) overuse of *the* in partitive contexts in elicitation. This is consistent with the view that learners who disregard plurality marking overuse *the* with partitives – but data from one learner are insufficient to prove this argument.

Knowledge of plurality was also tested in the second study, in the Translation task reported in Appendix 3. Learners had to translate numeral phrases such as *three friends* from their L1 to their L2. The intermediate/advanced L1-Korean speakers omitted plurality marking in 11% of such contexts, compared to 3% for L1-Russian speakers. While this is consistent with the view that L1-Korean speakers have difficulty with English plural morphology, 11% is not very high.

In order to find out whether disregard of plurality affected the L1-Korean speakers' performance in partitive contexts, it is necessary to do two things. First, partitive contexts should be set up that make it unambiguous that only one member of the set is under discussion, and focus the learners' attention on this fact. Second, as discussed earlier, it is necessary to give the learners freedom to use overt partitives, to ensure that overuse of *the* isn't tied to the infelicity of *a*.

If, after all of the methodological controls, it is found that L1-Korean speakers still overuse *the* in partitive contexts (while L1-English controls do not), then a proposal linking topic-marking in Korean to overuse of *the* may need to be developed. This proposal would be independent of the specificity discussion, since topic-marking and partitivity are not linked to specificity. It would be necessary to explain how the Article Choice Parameter on the one hand and transfer of topic-marking on the other interact in L2-English. It would also be necessary to explain why partitivity plays a role in child L1-English (see Maratsos 1976, as explained by Wexler 2003), even though topic-marking is not involved, and to relate overuse of *the* in partitive contexts in L1- and L2-acquisition. For the time being, I leave the issue open.

2. Word order

Another effect investigated in the second study (reported in Chapter 5) was the role of word order in article choice. As discussed briefly in Chapter 2, Russian word order interacts with information status, so that preverbal elements tend to be old information (cf. Bailyn 1995, Ch.3, and the references cited therein). A bare unmodified preverbal DP is typically interpreted as definite, as shown in (20). In (20a), the preverbal subject has to be old information, and would be translated into English as *the boy*. In order to receive a

new information / indefinite reading, the subject would have to be placed post-verbally, as in (20b).

20. a.　*Mal'čik*　　*vošël*　　*v*　*komnatu.*
　　　Boy-NOM　　enter-PST　in　room
　　　"The boy entered the room."
　　b.　*V*　*komnatu*　*vošël*　　*mal'čik.*
　　　In　room　　enter-PST　boy-NOM
　　　"A boy entered the room."

In Korean, on the other hand, preverbal DPs can be either definite or indefinite, as shown in (21).

21. *Sonyen-i*　*pang-ey*　*dule-wa-ss-ta.*
　　Boy-NOM room-in　enter-come-PST-DEC
　　"The/a boy entered the room."　　　　　　　　　　　　　(Heejeong Ko, p.c.)

In order to avoid word order effects for L1-Russian speakers in our investigations of definiteness and specificity, we always put the target item in object position, which is not obviously associated with either definiteness or indefiniteness in Russian SVO order. However, we also included an item type with an indefinite in subject position, illustrated in (22). We predicted that if there is any transfer of word order effects, L1-Russian speakers but not L1-Korean speakers should overuse *the* in this item type. Note that the indefinite in (22) is a non-specific indefinite, since the identity of the referent (in this case, the teapot) is unimportant for the purposes of the discourse.

22. *non-specific indefinite in subject position*
　　Louise:　Hi, Polly. How have you been?
　　Polly:　Well, not very good. I had a strange adventure last summer. I was just walking down the street, and I heard people arguing in one of the houses.
　　Louise:　So what happened then?
　　Polly:　A very strange thing happened. Suddenly, (a, the, --) big teapot fell on my head!

We included two control context types. The first involved *there*-constructions, illustrated in (23). The *there*-construction can be seen as a rough equivalent (in meaning, not structure) to the postverbal subject construction in Russian. If transfer takes place, it should cause overuse of *the* with subject indefinites but not with indefinites in the *there*-construction.

23. *non-specific indefinite in the* there-*construction*
　　Terry:　I think I am going to go take a bath.
　　Anne:　That might not be such a good idea. I was just in the bathroom. I don't think you should take a bath right now.
　　Terry:　Why not?
　　Anne:　Because there is (a, the, --) big frog in our bathtub.

The second control item type had a previous-mention definite in subject position, as shown in (24), to ensure that the L2-learners correctly put *the* with definite subjects.

The results are reported in (25). Only percentages of *the* and *a* use are reported, with the remaining percentage points referring to article omission. As shown in (25), overuse

of *the* with indefinite subjects (compared to the *there*-construction) was indeed fairly high for L1-Russian speakers, while this was not the case for L1-Korean speakers. The difference in use of *the* in subject position with indefinites vs. in the *there*-construction was significant for L1-Russian speakers (p < .001) but not for L1-Korean speakers (p = .21).

24. *definite subject*

> Gina: I was at a circus yesterday. All of the acts were really great. I especially liked three clowns and one elephant.
> Ruth: Tell me about what they did.
> Gina: Well, (an, <u>the</u>, --) elephant stood on his head!

25. Subject position vs. *there*-construction: intermediate/advanced L2-learners

	L1-Russian speakers		L1-Korean speakers	
	the	*a*	*the*	*a*
Definite, subject position (24)	84%	14%	74%	24%
Indefinite, subject position (22)	33%	61%	9%	84%
Indefinite, *there*-construction (23)	15%	80%	5%	93%

I conclude that the relationship between word order and information structure in a learner's L1 affects article choice in the learner's L2. There are two possibilities for the nature of this transfer effect. One is that L1-Russian speakers assume that English, like Russian, uses word order to encode information structure; disregarding the context, they interpret any preverbal element as old information, and therefore definite.

Another possibility is that, while L2-learners do not consider indefinites in the preverbal position to be definite, they do consider them to be obligatorily specific. As discussed in Chapter 2, indefinites are able to appear preverbally in Russian if they have lexical modifiers such as *kakoj-to* 'some' or the specificity marker *odin,* 'one'. This is shown in (26). The fact that specific indefinites headed by *odin* are compatible with the preverbal position, while bare indefinites typically aren't, may lead L1-Russian speakers who are learning English to also treat the preverbal position as specific. They would then overuse *the* with preverbal indefinites, as with all specific indefinites.

26. a) *Odin   mal'čik   zvonil.*
       one   boy-NOM   called
       "A boy called."

   b) *Kakoj-to   mal'čik   zvonil.*
       some   boy-NOM   called
       "Some boy called."

More investigation into Russian word order is needed: since definites, specific indefinites, and *some*-indefinites (as well as overt partitives and strong quantifiers) can all appear preverbally in Russian, it is far from clear what the exact requirement on the preverbal position is. For now, there is some indication that L1-Russian speakers transfer the properties of the preverbal position in Russian onto the preverbal position in English.

References

Abusch, Dorit. 1994. The scope of indefinites. *Natural Language Semantics*, 2, 83–135.

Abusch, Dorit and Mats Rooth. 1997. Epistemic NP modifiers. *Proceedings of SALT 7*.

Azar, Betty Schrampfer. 1993. *Chartbook: a reference grammar. Understanding and using English grammar,* 2nd edition. Englewood Cliffs, NJ: Prentice Hall Regents.

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

Beck, Maria-Luise. 1998. L2 acquisition and obligatory head movement: English-speaking learners of German and the Local Impairment Hypothesis. *Studies in Second Language Acquisition* 20, 311-348.

Becker, Misha. 1999. The *some* indefinites. In Gianluca Storto (ed.), *UCLA Working Papers in Linguistics*, 3, Syntax at sunset 2.

Belletti, Adriana. 1990. *Generalized Verb Movement.* Torino, Italy: Rosenbers & Sellier.

Bickerton, Derek. 1981. *Roots of Language.* Ann Arbor: Karoma.

Birdsong, David. 1992. Ultimate attainment in second language acquisition. *Language,* 68, 706-755.

Birdsong, D. and M. Molis. 1998. Age and maturation in L2A: A replication of Johnson & Newport (1989). Unpublished ms., University of Texas.

Bley-Vroman, R. 1989. What is the logical problem of foreign language learning? In S. Gass and J. Schachter (eds.), *Linguistic perspectives on second language acquisition.* Cambridge, UK: Cambridge University Press.

Borer, Hagit. 2003. Things that count: Null D. Chapter 5 In Hagit Borer, *Structuring Sense: An Exo-Skeletal Trilogy. Book I: Determining Structures.*

Borthen, Kaja. 1998. Linguistic evidence prompting the linguistic necessity of Donnellan's referential-attributive distinction? The use of bare superlatives in Norwegian. Oral presentation at the national summer school in semantics and pragmatics in Trondheim, June 1998. (Also available as a manuscript.)

Borthen, Kaja. 2003. Some semantic properties of Norwegian bare singulars. Manuscript.

Broselow, E. and D. Finer. 1991. Parameter setting in second language phonology and syntax. *Second Language Research,* 7, 35-59.

Brown, Roger. 1973. *A first language.* Cambridge, MA: Harvard University Press.

Buring, Daniel. 1999. Topic. In P. Bosch and R. van der Sandt (eds.) *Focus -- Linguistic, Cognitive, and Computational Perspectives.* Cambridge, UK: Cambridge University Press, 142-165

Choe, H-S. 1995. Focus and Topic Movement in Korean and Licensing. In K. Kiss (ed.), *Discourse configurational languages.* Oxford: Oxford University Press, 269-334.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1981. *Lectures on Government and Binding: The Pisa Lectures,* Dordrecht: Foris.

Chomsky, Noam. 1993: A minimalist program for linguistic theory. In K. Hale, and S.J. Keyser, editors. *The view from Building 20: Essays in linguistics in honor of Sylvain Bromberger.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program.* Cambridge, MA: MIT Press.

Chierchia, Gennaro. 1998. Reference to kinds across languages. *Natural language semantics,* 6, 339-405.

Chierchia, Gennaro. 2001. A puzzle about indefinites. In Carlo Cecchetto, Gennaro Chierchia, and Maria Teresa Guasti (eds.), *Semantic Interfaces: Reference, Anaphora, and Aspect*. Stanford: CSLI.

Chung, Sandra and William A. Ladusaw. 2003. *Restriction and Saturation*. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and functional heads: a cross-linguistic perspective.* New York: Oxford University Press.

Clahsen, H. and U. Hong. 1995. Agreement and null subjects in German L2 development: new evidence from reaction-time experiments. *Second Language Research*, 5, 1-29.

Clahsen, H. and P. Muysken. 1986. The availability of Universal Grammar to adult and child learners: a study of the acquisition of German word order. *Second Language Research*, 2, 93-119.

Cooper, R.., E. Olshtain, R. Tucker, and M. Waterbury. 1979. The acquisition of complex English structures by adult native speakers of Arabic and Hebrew. *Language Learning*, 29, 255-275.

Cziko, Gary A. 1986. Testing the Language Bioprogram Hypothesis: A review of children's acquisition of articles. *Language*, 62, 878-898.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

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.

du Plessis, J., D. Solin, L. Travis and L. White. 1987. UG or not UG, that is the question: a reply to Clahsen and Muysken. *Second Language Research*, 3, 56-75.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry*, 22, 1-25.

Epstein, Samuel D., Suzanne Flynn and Gita Martohardjono. 1996. Second language acquisition: Theoretical and experimental issues in contemporary research. *Behavioral and Brain Sciences*, 19, 677-758.

Epstein, Samuel D., Suzanne Flynn and Gita Martohardjono. 1998. The strong continuity hypothesis in adult L2 acquisition of functional categories. In S. Flynn, G. Martohardjono and W. O'Neil (eds.), *The Generative study of second language acquisition*. Hillsdale, NJ: Erlbaum, 61-77.

Eubank, Lynn. 1993/1994. On the transfer of parametric values in L2 development. *Language Acquisition* 3, 183-208.

Eubank, Lynn, Janine Bischof, April Huffstutler, Patricia Leek, and Clint West. 1997. "Tom eats slowly cooked eggs": Thematic-verb raising in L2 knowledge. *Language Acquisition* 6, 171-199.

Evans, Gareth. 1980. Pronouns. *Linguistic Inquiry*, 11, 337-362.

Finer, Daniel. 1991. Binding parameters in second language acquisition. In Lynn Eubank (ed.), *Point counterpoint: Universal Grammar in second language acquisition*, 351-373.

Finer, Daniel and Ellen Broselow. 1986. Second language acquisition of reflexive-binding. In *Proceedings of NELS 16*. Amherst: University of Massachusetts, Graduate Linguistics Student Association, 154-168.

Flege, J. and S. Liu. 2001. The effect of experience on adults' acquisition of a second language. *Studies in Second Language Acquisition*, 23.

Flynn, Suzanne. 1983. A study of the effects of principal branching direction in second language acquisition: The generalization of a parameter of universal grammar from first to second language acquisition. Ph.D. dissertation, Cornell University.

Flynn, Suzanne. 1984. A universal in L2 acquisition based on a PBD typology. In F. Eckman, I. Bell and D. Nelson (eds.), *Universals of second language acquisition.* Newbury House.

Flynn, Suzanne. 1987. *A parameter-setting model of L2 acquisition.* Reidel.

Flynn, Suzanne. 1996. A parameter-setting approach to second language acquisition. In W. Ritchie and T. Bhatia (eds.), *Handbook of second language acquisition.* San Diego: Academic Press, 121-158.

Flynn, Suzanne, Claire Foley and Donna Lardiere. 1991. The Minimality principle in adult second language acquisition. Paper presented at the Second Language Research Forum, USC, Los Angeles.

Flynn, Suzanne, Claire Foley and Barbara Lust. 2000. Evidence for grammatical mapping from second language acquisition. Paper presented at the LSA, Chicago, IL.

Flynn, Suzanne and Barbara Lust. 1980. Acquisition of relative clauses in English. *Cornell Working Papers in Linguistics,* 1, 33-46. Ithaca, NY.

Flynn, Suzanne and Gita Martohardjono. 1994. Mapping from the initial state to the final state: The separation of universal principles and language specific properties. In B. Lust, J. Whitman, and M. Suner (eds) *Syntactic theory and first language acquisition: Crosslinguistic perspectives.* Hillsdale, NJ: Lawrence Erlbaum, 319-335.

Flynn, Suzanne, Inna Vinnitskaya and Claire Foley. 2001. Role of prior language knowledge in the acquisition of a third language. Paper presented at the second International Conference on Third Language Acquisition and Trilingualism, Fryske Akademy.

Fodor, Janet and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy,* 5, 355–398.

Givón, Talmy. 1981. On the development of the numeral 'one' as an indefinite marker. *Folia Linguistica Historica,* 2.1.

Givón, Talmy. 2001. *Syntax: Volume 1.* Amsterdam: John Benjamins.

Haspelmath, Martin. 1997. *Indefinite pronouns.* Oxford: Claredon Press.

Hawkins, J.A. 1978. *Definiteness and indefiniteness.* London: Croom Helm.

Hawkins, Roger. 2001. *Second language syntax.* Blackwell.

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.

Heim, Irene. 2003. Semantics and pragmatics of article choice: A brief introduction. Paper presented at the CUNY Sentence Processing Conference, MIT, March 2003.

Heim, Irene and Ken Wexler. 2000. Class notes for "Topics in semantics and acquisition", Fall 2000, MIT.

Hirakawa, M. 1990. A study of the L2 acquisition of English reflexives. *Second Language Research,* 6, 60-85.

Huebner, T. 1983. *A longitudinal analysis of the acquisition of English.* Ann Arbor: Karoma.

Huebner, T. 1985. System and variability in interlanguage syntax. *Language Learning* 35, 141-163.

Hulk, Aafke. 1991. Parameter setting and the acquisition of word order in L2 French. *Second Language Research,* 7, 1-34.

Ionin, Tania and Ken Wexler. 2002. Optional verb raising in L2 acquisition: Evidence from L1-Russian learners of English. In J. Costa and M.J. Freitas (eds.), *Proceedings of the GALA' 2001 Conference on Language Acquisition.* Lisbon: Associação Portuguese de Linguistica, 134-141.

Johnson, Jacqueline and Elissa Newport. 1989. Critical period effects in second language learning: the influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology,* 21: 60-99.

Johnson, Jacqueline and Elissa Newport. 1991. Critical period effects on universal properties of languages: the status of subjacency in the acquisition of a second language. *Cognition,* 39, 215-258

Kadmon, Nirit. 1990. Uniqueness. *Linguistics and Philosophy,* 13, 273-324.

Kaneko, Yumina. 1996. Knowledge of the English article system in second language learning: To "the" or not to "the". Undergraduate thesis, Smith College.

Kaplan, D. 1978. Dthat. In P. Cole (ed.), *Syntax and semantics: vol. 9, Pragmatics.* New York: Academic Press.

Karmiloff-Smith, Annette. 1979. *A functional approach to child language.* Cambridge, UK: Cambridge University Press.

Kelepir, Meltem. 2001. Topics in Turkish syntax: Clausal structure and scope. Ph.D. dissertation, MIT. Distributed by MITWPL.

Kim, Young-Eun, 2000, Focus and old information: Polarity focus, contrastive focus, and contrastive topic. Ph.D. dissertation, University of Texas.

Kratzer, Angelika. 1998. Scope or pseudo-scope? Are there wide-scope indefinites? In S. Rothstein (ed.), *Events in Grammar,* 163-196. Dordrecht: Kluwer.

Krifka, Manfred, Francis Jeffry Pelletier, Gregory N. Carlson, Alice ter Meulen, Godehard Link, and Gennaro Chierchia. 1995. Genericity: an introduction. In Gregory N. Carlson and Francis Jeffry Pelletier (eds.), *The generic book,* Chicago, University of Chicago Press, 1-124.

Kripke, Saul. 1977. Speaker reference and semantic reference. In P. French, T. Uehling, and H. Wettstein (eds.), *Contemporary perspectives in the philosophy of language,* Minneapolis: University of Minnesota Press, 6-27.

Lardiere, D. 1998. Dissociating syntax from morphology in a divergent end-state grammar. *Second Language Research,* 14, 359-375.

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-Bruevich, 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.

Leventhal, Vitaly. 2000. *Zanimatel'nyj anglijskij: Learning English creatively for Russian-speaking students.* New York: Edulink.

Lidz, Jeffrey. 1999. The morphosemantics of object case in Kannada. *Proceedings of WCCFL 18,* 325-336.

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

Ludlow, Peter and Gabriel Segal. 2002. On a unitary semantical analysis for definite and indefinite descriptions. In A. Bezuidenhout & M. Reimer (eds.), *Descriptions: Semantic and pragmatic perspectives*. Oxford: Oxford University Press.

Lust, Barbara. 1994. Functional projection of CP and phrase structure parameterization: An argument for the strong continuity hypothesis. In B. Lust, J. Whitman, and M. Suner (eds.), *Cross linguistic perspectives. Volume I: Heads, projections, and learnability*. Hillsdale, NJ: Erlbaum, 85-118.

Lust, Barbara. 1999. Universal Grammar: The strong continuity hypothesis in first language acquisition. In W.C. Ritchie and T.K. Bhatia (eds.), *Handbook of child language acquisition*. New York: Academic Press, 111-155.

Lust, Barbara. In preparation. *Universal Grammar and the initial state: Cross-linguistic studies of directionality*. Bradford Books/MIT Press.

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.

Maclin, Alice. 1987. *Reference guide to English*. New York: Holt, Rinehart and Winston.

Maratsos, Michael P. 1976. *The use of definite and indefinite reference in young children*. Cambridge, UK: Cambridge University Press.

Martohardjono, Gita. 1992. Wh-movement in the acquisition of a second language: Workshop on the Acquisition of Wh-Movement, University of Massachusetts, Amherst, MA.

Martohardjono, Gita. 1993. Wh-movement in the acquisition of a second language: A cross-linguistic study of three languages with and without movement. 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.

Matthewson, Lisa. 1998. *Determiner systems and quantificational strategies: evidence from Salish*. The Hague: Holland Academic Graphics.

Matthewson, Lisa. 1999. On the interpretation of wide-scope indefinites. *Natural Language Semantics*, 7, 79–134.

Matthewson, Lisa and Jeannette Schaeffer. 2000. Grammar and pragmatics in the acquisition of article systems. in J. Gilkerson, M. Becker, and N. Hyams (eds.) *UCLA Working Papers in Linguistics: Language Development and Breakdown 1*, University of California, Los Angeles,1-39.

Matthewson, Lisa, Tim Bryant and Tom Roeper. 2001. A Salish Stage in the Acquisition of English Determiners: Unfamiliar 'Definites'. *The Proceedings of SULA: The semantics of under-represented languages in the Americas*. University of Massachusetts Occasional Papers in Linguistics, 25.

Moore, Harumi. 2003. Japanese *ka*-indefinites and the semantic domain of *unknown to the speaker*. *Journal of pragmatics*, 35, 589–614.

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.

Otsu, Y. and K. Naoi. 1986. Structure-dependence in L2-acquisition. In Japanese Association of College English Teachers, Tokyo: Keio University.

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

Parodi, T., B.D. Schwartz and H. Clahsen. 1997. On the L2 acquisition of the morphosyntax of German nominals. *Essex Research Reports in Linguistics,* 15, 1-43.

Pesetsky, David. 1989. Language-particular processes and the earliness principle. Unpublished manuscript, MIT.

Peterson, C.L. 1974. *Communicative and narrative behavior of preschool-age children.* Unpublished Ph.D. dissertation, University of Minnesota.

Petrova, Milena. 2003. Bulgarian 'some' indefinites. To appear in the Proceedings of the High Desert Linguistics Conference-V, University of New Mexico.

Poesio, Massimo and Renata Vieira. 1998. A Corpus-Based Investigation of Definite Description Use. *Computational Linguistics,* 24, 183-216.

Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry,* 20(3), 365-424.

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.

Raimes, Ann. 1990. *How English works.* New York: St. Martin's Press.

Raimes, Ann. 1992. *Grammar troublespots: an editing guide for students,* 2nd edition. New York: St. Martin's Press.

Reinhart, Tanya. 1997. Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy,* 20, 335–397.

Roberts, Craige. 2000. Uniqueness in definite noun phrases. Ms. Ohio State University. Available at http://semanticsarchive.net/Archive/WMyMDYxN/uniqueness.pdf

Robertson, Daniel. 2000. Variability in the use of the English article system by Chinese learners of English. *Second Language Research,* 16, 135-172.

Robertson, Daniel and Antonella Sorace. 1999. Losing the V2 constraint. In E. Klein and G. Martohardjono (eds.), *The development of second language grammars: a generative approach.* Amsterdam: John Benjamins, 317–361.

Ruys, E.G. 1992. *The Scope of Indefinites.* Ph.D. dissertation, Utrecht University.

Schachter, Jacquelyn. 1989. Testing a proposed universal. In S. Gass and J. Schachter (eds.), *Linguistic perspectives on second language acquisition.* Cambridge: Cambridge University Press, 73-88.

Schachter, Jacquelyn. 1990. On the issue of completeness in second language acquisition. *Second Language Research,* 6, 93-124.

Schachter, Jacquelyn. 1996. Maturation and the issue of Universal Grammar in second language acquisition. In W. Ritchie and T. Bhatia (eds.), *Handbook of second language acquisition.* San Diego: Academic Press, 159-193.

Schaeffer, Jeannette. 1999. Articles in English Child Language. Paper presented at the LSA, Los Angeles.

Schafer, Robin and Jill de Villiers. 2000. Imagining articles: What *a* and *the* can tell us about the emergence of DP. S. C. Howell et al. (eds.), *BUCLD 24 Proceedings.* Somerville, MA: Cascadilla Press.

Schlenker, Philippe. 2002. Indexicality, logophoricity, and plural pronouns. Ms., UCLA.

Schlenker, Philippe. 2003. A Plea for Monsters. *Linguistics & Philosophy,* 26, 29-120

Schwartz, Bonnie D. 1998. The second language instinct. 1998. *Lingua,* 106, 133-160.

Schwartz, Bonnie D. and Rex A. Sprouse. 1994. Word order and nominative case in non-native language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In Teun Hoekstra and Bonnie D. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar: Papers in Honor of Kenneth Wexler from the 1991 GLOW Workshops*. Amsterdam: John Benjamins, 317-368.

Schwartz, Bonnie D. and Rex A. Sprouse. 1996. L2 cognitive states and the Full Transfer/Full Access model. *Second Language Research*, 12, 40-72.

Schwartz, B.D. and A. Tomaselli. 1990. Some implications from an analysis of German word order. In W. Abraham, W. Kosmeijer and E. Reuland (eds.), *Issues in Germanic Syntax*. Berlin: Walter de Gruyter, 251-274.

Schwarz, Bernhard. 2001. Two kinds of long-distance indefinites. Manuscript.

Sharvy, Richard. 1980. A more general theory of definite descriptions. *The Philosophical Review*, 89, 607-624.

Sherman, J.C. and B. Lust. 1993. Children are in control. *Cognition*, 46, 1-51.

Singleton, D. and Z. Lengyel (eds.). 1995. *The age factor in second language acquisition: a critical look at the Critical Period Hypothesis*. Clevedon & Philadelphia, Multilingual Matters.

Snyder, William. 2002. Parameters: The view from child language. In Y. Otsu (ed.), *Proceedings of the Third Tokyo Conference on Psycholinguistics*. Tokyo: Hituzi Shobo, 27-44.

Sorace, Antonella. 2000. Syntactic optionality in non-native grammars. *Second Language Research*, 16, 93–102.

Sprouse, Rex A. and Bonnie D. Schwartz. 1998. In defense of Full Transfer in German-English and French-English interlanguage: Comparative L2 acquisition research. *Proceedings of BUCLD 22*. Somerville, MA: Cascadilla Press, 726-736.

Sridhar, S.N. 1990. *Kannada*. New York: Routledge.

Stalnaker, Robert. 1972. Pragmatics. In Harman and Davidson (eds.), *Semantics of natural language*. Dordrecht: D. Reidel, 380-397.

Strawson, P. 1952. *Introduction to logical theory*. London: Methuen.

Tarone, Elaine and Betsy Parrish. 1988. Task-related variation in interlanguage: The case of articles. *Language Learning*, 38, 21-45.

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

Thomas, Margaret. 1989b. The interpretation of English reflexive pronouns by non-native speakers. *Studies in Second Language Acquisition*, 11, 281-303.

Thomas, Margaret. 1991. Universal Grammar and the interpretation of reflexives in a second language. *Language*, 67, 211-239.

Thornton, Rosalind and Ken Wexler. 1999. *VP ellipsis, Binding Theory and interpretation in child grammar*. Cambridge, MA: MIT Press.

Tsimpli, I. and A. Roussou. 1991. Parameter resetting in L2? *UCL Working Papers in Linguistics*, 3, 149-169.

Vainikka, Anne and Martha Young-Scholten. 1994. Direct access to X'-Theory: Evidence from Korean and Turkish adults learning German. In Teun Hoekstra and Bonnie D. Schwartz (eds.), *Language Acquisition Studies in Generative Grammar: Papers in Honor of Kenneth Wexler from the 1991 GLOW Workshops*. Amsterdam: John Benjamins, 265-316.

Vainikka, Anne and Martha Young-Scholten. 1996. Gradual development of L2 phrase structure. *Second Language Research*, 12, 7-39.

Wakabayashi, S. 1997. The acquisition of functional categories by learners of English. Unpublished doctoral dissertation, University of Cambridge.

Wexler, Ken. 1998. Very early parameter setting and the unique checking constraint: A new explanation of the optional infinitive stage. *Lingua*, 106, 23-79.

Wexler, Ken. 2003. Maximal trouble. Paper presented at the CUNY Sentence Processing Conference, MIT, March 2003.

Wexler, Ken and Rita Manzini. 1987. Parameters and learnability in binding theory. In T. Roeper and E. Williams (eds.), *Parameter setting*. Dordrecht: Reidel, 41-76.

White, Lydia. 1990/91. The verb-movement parameter in second language acquisition. *Language Acquisition*, 1, 337-360.

White, Lydia. 1992. Long and short movement in second language acquisition. *Canadian Journal of Linguistics/Revue Canadienne de Linguistique*, 37, 273-286.

White, Lydia. 1996. Universal Grammar and second language acquisition: Current trends and new directions. In W. Ritchie and T. Bhatia (eds.), *Handbook of second language acquisition*. San Diego: Academic Press, 85-120.

White, Lydia. 2000. Second language acquisition: From initial state to final state. In J. Archibald (ed.), *Second language acquisition and linguistic theory*. Blackwell, 130-155.

White, Lydia and Fred Genesee. 1996. How native is near-native? The issue of ultimate attainment in second language acquisition. *Second Language Research*, 12, 238-265.

White, Lydia and Alan Juffs. 1998. Constraints on wh-movement in two different contexts of non-native language acquisition: competence and processing. In S. Flynn, G. Martohardjono and W. O'Neil (eds.), *The generative study of second language acquisition*. Hillsdale, NJ: Earlbaum, 111-129.

Winter, Yoad. 1997. Choice functions and the scopal semantics of indefinites. *Linguistics and Philosophy*, 20, 399-467.

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.