Semantic Universals and Variation
in L2 Article Choice

Heejeong Ko[1], Alex Perovic[2], Tania Ionin[3], and Ken Wexler[2]
[1]Seoul National University, [2]MIT, [3]University of Illinois at Urbana Champaign

1. Introduction

It has long been widely known that L2-learners make errors in article choice. In particular, L2-English learners sometimes misuse the definite article *the* with an indefinite DP, and sometimes misuse the indefinite article *a* with a definite DP. They also omit articles in obligatory contexts (see Huebner 1983, Master 1987, Parrish 1987, Thomas 1989, Kaneko 1996, Leung 2001, Ionin 2003, Ionin, Ko, and Wexler 2004, Ko, Ionin, and Wexler 2006, *inter alia*). Importantly, previous studies have shown that article misuse in L2-acquisition is not random, but tied to *semantic universals* which contribute to semantic interpretation of the target DP. In particular, it has been observed that the semantic features of *definiteness, specificity,* and *partitivity* play a significant role in determining L2-article choice (Ionin 2003, Ionin et al. 2004, Ko et al. 2006).

In this paper, we investigate how these semantic universals affect L2-English article choice with L2-English learners who come from two typologically different, article-less L1s: L1-Serbo-Croatian and L1-Korean learners of L2-English. While the effects of definiteness and specificity have been investigated with different L1-groups (Ionin et al. 2004), our previous investigations of partitivity (Ko et al. 2006) were limited to L1-Korean learners of English. One of the goals of the present paper, therefore, is to check whether partitivity effects hold across learners' L1s (as long as the L1 lacks articles). We also examine whether different semantic factors contribute to L2-article choice equally. Based on our experimental data, we argue that semantic universals are accessible to L2-learners regardless of their L1, but that some semantic factors may be more persistent than others in causing L2-article errors. The effects of each semantic factor may vary depending on learners' overall L2-proficiency. Specifically, it is argued that partitivity effects are more persistent than specificity effects; consequently, more advanced L2-learners are more sensitive to partitivity effects than specificity effects. We speculate that this difference results from the entailment relationships between different semantic factors.

2. Background

In this section, we present the theoretical background necessary for understanding the relevant semantic universals, as well as review previous studies on how these semantic universals affect L2-article choice.

2.1 Definiteness

Definiteness is a semantic feature which makes reference to the knowledge state of both the speaker and the hearer concerning a unique discourse referent, as stated informally in (1) (see Heim 1991 for formal definitions). English encodes this semantic notion in its article system. The English

article *the* marks [+definite] DPs, while the article *a* marks [-definite] DPs.[1]

(1) If a Determiner Phrase (DP) of the form [D NP] is [+definite], then the speaker assumes that the hearer shares the presupposition of the existence of a unique individual in the set denoted by NP.

The definite article *the* can be used in various contexts which satisfy the uniqueness presupposition described in (1). Most commonly, it can be used to refer back to a previously mentioned DP. Thus, in (2), indefinite *a* is used with a first-mentioned referent in the introductory sentence. When the same referent *black pen* is mentioned again in the following discourse, the definite article *the* is used. This is because the previous discourse ensures that the uniqueness presupposition is satisfied: the speaker can reasonably assume that the hearer shares the speaker's presupposition that there is a unique black pen under discussion.

(2) There is a black pen [...] I am going to use the black pen.

However, previous mention is not always necessary for felicitous use of *the*. When the uniqueness presupposition is fulfilled by mutual world knowledge of the speaker and the hearer, *the* can be used without any previous mention. Some relevant examples are given in (3) (Bloom 1970: 345; see also Hawkins 1978). For instance, *the moon* is unique for all speakers given our world knowledge; in a room containing just one desk, *the desk* is uniquely salient for the speaker and the hearer; items like *the engine* are uniquely salient as a result of our knowledge that there is usually only one engine in a car; ordinal expressions such as *the last X* satisfy the uniqueness presupposition by virtue of their meanings (i.e. there is only one last element in a set).

(3) a. I saw the moon today. It was really beautiful.
     b. Look at the desk in my room. It is really clean!
     c. I fixed the engine yesterday. The car runs really smoothly now.
     d. I like the last sentence of your novel very much.

If L2-learners used *the* only when the uniqueness presupposition is satisfied, as in (2) and (3), they would make no errors in L2-English article usage. They do, however, make errors and incorrectly use *the* and *a*. We argue that the semantic features of *specificity* and *partitivity* are relevant for understanding such errors.

2.2 Specificity

Specificity is a semantic feature that makes reference to the knowledge state of the speaker concerning a uniquely salient discourse referent, as described informally in (4), based on the proposal of Ionin (2003, 2006). Specificity crucially differs from definiteness in that specificity concerns the speaker's intent to refer *regardless of the hearer's knowledge status*.

(4) If a Determiner Phrase (DP) of the form [D NP] is [+specific], then the speaker intends to refer to a unique individual in the set denoted by the NP, and considers this individual to possess some noteworthy property (based on Fodor and Sag 1982; for formal definition, see Ionin 2003, 2006).

As discussed above, the English article system marks definiteness, not specificity. Hence, [+specific] DPs can be used with *the* or *a*, and [-specific] DPs can also be compatible with *the* or *a*. This is illustrated in (5) (based on test items from Ionin et al. 2004). In (5a), *the author* is [+definite] since the unique existence of a salient author is presupposed by the speaker and the hearer (see (3c)). It is also [+specific] since the speaker has concrete knowledge of the referent and intends to refer to that individual, satisfying the condition in (4). In contrast, in (5b), *the author* is [+definite] but [-specific].

---

[1] We limit our discussion to singular forms of English DPs. For plurals, *the* is used for definites and *some* or no article for indefinites; maximality effects hold for use of *the* with plurals (see Heim 1991 for details).

The author in (5b) satisfies the uniqueness presupposition, but the speaker lacks specific knowledge of the referent and has no intent to refer to a particular individual. In (5c), *a friend* is [-definite] since there is no uniquely salient referent shared by the speaker and the hearer (either as a result of previous discourse or by mutual world knowledge). It is, however, [+specific] since the speaker has a particular individual in mind (i.e. *Sam Brown*), and intends to refer to that individual. In (5d), *a friend* is [-definite] and [-specific]; the referent's existence is not uniquely salient to the speaker and the hearer, and the speaker does not have a particular individual in mind.

(5) a.　　[+specific, +definite]: target the
　　　　　I would like to meet <u>the author of that book</u> some day – I saw an interview with her on TV, and I really liked her!
　　b.　　[-specific, +definite]: target *the*
　　　　　I would like to meet <u>the author of that painting</u> – unfortunately, I have no idea who it is, since the painting is not signed!
　　c.　　[+specific, -definite]: target *a*
　　　　　I am here for a week. I am visiting <u>a friend from college</u> – his name is Sam Brown, and he lives in Cambridge now.
　　d.　　[-specific, -definite]: target *a*
　　　　　He is staying with <u>a friend</u> – but he didn't tell me who that is.

Though the English article system is not sensitive to the specificity distinction, it has been observed that L2-learners' article choice is in fact affected by the specificity feature. In particular, previous studies have shown that L2-learners overuse *the* with specific indefinites (e.g. (5c)) and overuse *a* with nonspecific definites (e.g. (5b)) (see Ionin 2003, Ionin et al 2004). This indicates that L2-learners erroneously associate *the* with the [+specific] feature (instead of [+definite]), and *a* with the [-specific] feature (instead of [-definite]). As our previous studies extensively discuss, this error pattern suggests that L2-learners have access to the specificity distinction, which is not overtly marked either by their L1 (Korean, Russian) or by the target language L2 (English). In Ionin et al. (2004), we proposed that such access to specificity is a result of UG-access to semantic universals.

2.3 Partitivity: a sub-type of presuppositionality

Presuppositionality is a semantic feature which makes reference to the presupposition of the speaker and the hearer concerning existence of a referent, as described informally in (6). Presuppositionality differs from definiteness in that it marks merely the presupposition of *existence of a discourse referent*, whereas definiteness marks the presupposition of *existence of a unique referent*.

(6) If DP of the form [D NP] is [+presuppositional], then the speaker assumes that the hearer shares the presupposition of the existence of a unique individual in the set denoted by NP. (based on Enç 1991, Diesing 1992; for formal definition, see Wexler 2003).[2]

As emphasized above, the English article system encodes only the definiteness distinction. Hence, presuppositionality cannot be morphologically marked by an article, but can only be inferred from the discourse. One can establish presuppositionality for indefinites in two ways: (i) by introducing in the previous discourse a set that the referent of the target DP belongs to (this is similar to previous-mentioned definites in (2), but without uniqueness), or (ii) by mutual world knowledge (this is similar to definites in (3), but without uniqueness). In this paper, we focus on the first sub-type of presuppositionality, calling it *partitivity* (cf. Enç 1991, Diesing 1992; regarding the second type of presuppositionality, refer to the discussion of D-linking in Pesetsky 2000).We consider the larger issue

---

[2] Enç (1991) suggests that there are two types of specificity: specificity encoded by partitive DPs, which are related to a previously mentioned set, and specificity encoded by elements such as *a certain* in English, which are non-partitive. Enç called both types 'specific'. We call these two types of DPs *partitive* and *specific*, respectively, and treat the two concepts as independent semantic features.

of presuppositional indefinites in the concluding section. See also Ko, Ionin, and Wexler (2007) for further discussion on this issue.

As exemplified in (7a), an indefinite DP such as *a puppy* is [+partitive] when the set *puppies* has been introduced in the previous discourse. In (7a), the speaker can safely assume that the hearer shares the presupposition of the existence of *a puppy* due to the mention of the set *puppies* in the previous discourse. In contrast, in (7b), set membership is not established, since *a comic book* does not belong to a set introduced by the previous discourse. The speaker cannot assume that the hearer shares the presupposition of the existence of *a comic book* – hence the DP is [-partitive].

(7)   a.   [+partitive, -definite]
           This pet shop had *five puppies and seven kittens*. Finally, John chose <u>a puppy.</u>
      b.   [-partitive, -definite]
           Kevin had to memorize two stories and three poems from his textbook. But, he spent the whole evening reading <u>a comic book</u>.

The semantics of partitivity is not overtly marked on the DP in English, but it has been observed that L2-learners's article choice is tied to the partitivity feature. Notably, L2-learners overuse *the* with [+partitive] indefinites (see Kaneko 1996 for L1-Japanese L2-English speakers; Ionin 2003 and Ko et al. 2006 for L1-Korean L2-English speakers; cf. section 3 for some differences among L1-groups). Partitivity has also been argued to play a role in child L1-acquisition of English (see Wexler 2003 on the findings of Maratsos 1976; see Ko et al. 2006 for an L1-L2 comparison). These results again indicate that learners' article errors reflect their access to semantic universals.

3.   Hypothesis and Predictions
3.1 Research Questions

From the previous studies discussed above, we conclude that semantic universals affect L2-article choice. Especially, L2-learners may associate *the* with [+specific] or [+partitive] features instead of the [+definite] feature, which leads to incorrect use of *the* with indefinites. The previous findings, however, have left many interesting questions unanswered.

First, given that multiple semantic factors contribute to L2-article choice, it is natural to ask whether each semantic factor "equally" contributes to article choice. In our previous study (Ko et al. 2006), we have shown that the effects of partitivity and specificity are independent; however, it has remained an open question whether there are any differences in how much these two factors influence L2-article choice, and if so, what these differences are due to. Second, we did not previously consider how L2-proficiency affects article choice with reference to semantic universals. All things being equal, we expect that advanced L2-learners would show more correct article use (i.e. adopt the definiteness distinction), but it remains open whether errors tied to specificity or partitivity would persist even in advanced L2-groups.

Third, our previous studies have argued that semantic universals are accessible by L2-learners whose L1 lacks articles (e.g. Ionin 2003, Ionin et al 2004, Ko et al. 2006; cf. Ionin, Zubizarreta and Bautista Maldonado (in press) for L1-transfer effects in L2-article choice). This leads us to ask how cross-linguistic variation interacts with access to semantic universals. The answer for this question has remained rather inconclusive. Ionin et al. (2004) show that both L1-Russian and L1-Korean speakers are affected by specificity; however, another part of the same study (reported in Ionin 2003, Appendix 4) showed that the effect of partitivity was strongly observed for L1-Korean speakers, but not for L1-Russian speakers. Kaneko (1996) showed that L1-Japanese learners of L2-English are affected by partitivity, but the effect of partitivity was relatively weak for L1-Spanish speakers of L2-English (consistent with the view that Spanish speakers transfer the semantics of definiteness from Spanish to English – see Ionin et al. (in press)). Another question that requires further study is whether the observed cross-linguistic variation in L2-article choice comes from L1-differences or from other factors, such as differences in L2-proficiency between the L1-groups.

3.2 Hypothesis

In this paper, we address the research questions introduced above by hypothesizing that semantic universals are accessible to L2-learners and that the variation in their effects is associated with L2-proficiency, not with L1-transfer, as long as only article-less L1s are considered. Our hypotheses are summarized in (8):

(8) Hypotheses: Semantic Universals in L2-Article Choice
A. Semantic universals are independent from each other.
B. Some semantic factors may be more persistent than others in causing L2 article errors.
C. There will be no cross-linguistic variation in the implementation of semantic universals when L1-transfer effects are controlled for.

The hypothesis (8A) states the basic premise that we adopt. We argue that semantic universals are independent features and we expect that the effect of each semantic universal cannot be attributed to the effect of the others (a view also taken in Ko et al. 2006). (8B) states our new hypothesis concerning L2-proficiency and the relative influence of different semantic factors on L2-article errors. We argue that although semantic universals are independent from each other, they may contribute to L2-article choice "unequally": some features can be more persistent than others in causing L2-article errors. Given that advanced L2-learners are less susceptible to errors in general, we expect that advanced learners may be less influenced by the semantic factors not instantiated by the English article system (specificity and partitivity). Less advanced learners, in contrast, are expected to be influenced by all the semantic features. Finally, it is possible that some error types persist to a greater degree than others among advanced learners (see section 5 for more discussion of why this would be). Finally, (8C) concerns our hypothesis that L2-learners have access to semantic universals regardless of their L1, as long as we control for L1-transfer by considering only article-less L1s.

In this paper, we test our hypotheses by comparing the performance of typologically different L1-groups on the same stimuli: L1-Serbo-Croatian (SC)[3] learners of L2-English and L1-Korean learners of L2-English. The SC data come from our new experiment, while the Korean data are taken from Ko et al. (2006). Since both SC and Korean lack articles (and have no other way of directly encoding either specificity or partitivity), L1-transfer effects are controlled for. Given our hypotheses in (8), we predict that both groups will be affected by semantic universals in L2-English article choice. If variation exists, it will be tied to differences in L2-proficiency, but not to L1-differences.

4. Experiment
4.1 Method

We have tested the effects of three semantic factors, *definiteness, specificity*, and *partitivity* with 30 L1-SC-learners of L1-English and 20 L1-Korean learners of L2-English (Korean data from Ko et al. 2006). The learners' L2-proficiency was measured by the Michigan test. All the 30 SC-speaking participants were advanced learners of English. Among the 20 Korean subjects, 16 learners were advanced, and 4 were intermediate. The test was piloted with 6 native speakers of English, who performed as expected: they used *the* in definite contexts, and *a* in indefinite contexts.

The task was a forced choice elicitation test. The subjects were presented with English dialogues and asked to choose an article for the target sentence in each dialogue on the basis of the context. The choices of *the, a,* and – (no article) were provided. The stimuli consisted of 80 English dialogues, falling into 20 context types, 4 items per context type. 10 context types targeted *a,* and 10 context types targeted *the.* The items were arranged into two pseudo-randomized test orders, each of which was given to half of the participants per group (i.e. 10 subjects for Koreans, 15 subjects for SC speakers). In this paper, we report the data from 4 indefinite context types, focusing on the effects of partitivity and

---

[3] The language of the former Yugoslavia, Serbo-Croatian, is now officially referred to as Serbian, Croatian or Bosnian; the linguistic phenomenon investigated here does not differ across the three dialects.

specificity.[4]

4.2 Stimuli and Predictions

Four indefinite context types were designed to test the effects of partitivity and specificity in article choice. The feature values of [±partitive] and [±specific] were crossed in the four contexts.

Two of the context types, exemplified in (9) and (10), contained partitive indefinites, where the referent of the target indefinite DP belongs to a set that has been mentioned in the previous discourse. All the partitives in these context types were implicit partitives: the relationship between the target referent and the set is implicit (*a team – a player*), not explicit (*five puppies – a puppy*) (see Ko et al. 2006 for more discussion). The two contexts differ with respect to whether the speaker intends to refer to a particular individual. The context type in (9) contained a [+partitive, +specific] indefinite, and the context type in (10) contained a [+partitive, -specific] indefinite.

(9) **[-definite, +partitive, +specific]**
　　Molly: So what did your guest Mr. Svenson do over the weekend?
　　Jamie:　**Well, he went to see our local softball team play.** He had a good time. **Afterwards, he met (a, the, --) player – she was very nice and friendly**. And she played really well!

(10)**[-definite, +partitive, -specific]**
　　Ben: I just saw Tom, and he looked really excited. Do you know why?
　　Melissa: **Yes – he was able to see the Boston Red Sox team while they were practicing.** And he is a huge fan! **He even got a signature from (a, the, --) player** – I have no idea which one. Tom was really excited!

The other two context types, exemplified in (11) and (12), contained non-partitive indefinites – the referent of the target indefinite DP does not belong to a set that has been mentioned in the previous discourse. The two contexts differ with respect to the value of specificity. The context in (11) contained a [-partitive, +specific] indefinite, and the context in (12) contained a [-partitive, -specific] indefinite.

(11)**[-definite, -partitive, +specific]**
　　Jennifer:　Hello, Helen? This is Jennifer!
　　Helen:　　Hi, Jennifer! It's wonderful to hear from you. I suppose you want to talk to my sister?
　　Jennifer:　Yes, I haven't spoken to her in years! I'd like to talk to her now if possible.
　　Helen:　　I'm very sorry, but she doesn't have time to talk right now. **She is meeting with (a, the, --) very important client from Seattle.** He is quite rich, and she really wants to get his business for our company! She'll call you back later.

(12)**[-definite, -partitive, -specific]**
　　Wife: Where is Peter? I haven't seen him all evening.
　　Husband: He is on the phone – he has been on it for hours.
　　Wife: That's not like Peter at all – he almost never uses the phone.
　　Husband: **But this time, he is talking to (a, the, --) girl – I have no idea who it is, but it's an important conversation to Peter.**

Under (8), we predict the following patterns in L2-English. First, given the hypothesis (8A) that semantic universals are independent from each other, we predict that L2-learners would overuse *the* when at least one of partitivity or specificity is present. In other words, we expect that when the context is [+partitive] or [+specific] ((9),(10),(11)), there will be an overuse of *the* with indefinites. Crucially, we expect that there will be no random overuse of *the* in [-partitive, -specific] contexts (12) where there is no semantic trigger for overusing *the*. Second, given that partitivity and specificity are independent

---

[4] In the other six indefinite contexts, we investigated how scope interacts with partitivity in L2-article choice. See Ko et al. (2007) for the results.

factors, we expect *additive effects*. We predict maximal overuse of *the* with indefinites in the [+partitive, +specific] context. The predictions concerning the effects of specificity and partitivity for each context are summarized in (13).

(13) Predictions for each context

For indefinites ([-definite])	[+partitive]	[-partitive]
[+specific]	Highest overuse of *the*	Overuse of *the*
[-specific]	Overuse of *the*	Correct use of *a*

Given that semantic factors may be effective unequally, we also expect that the more advanced learners may be less influenced by some semantic factors than others. On this view, [+partitive, -specific] and [-partitive, +specific] contexts would both trigger overuse of *the* in general, but there may be some differences between them depending on the strength of the effects of partitivity and specificity. If partitivity persists, we expect more overuse of *the* in [+partitive, -specific] contexts than [-partitive, +specific] contexts. If specificity persists, more overuse of *the* in [-partitive, +specific] contexts than [+partitive, -specific] contexts is expected. Lastly, we expect no substantial differences between the Korean and SC groups, as long as proficiency is controlled for.

4.3 Results: groups and individuals

The results concerning effects of specificity and partitivity are summarized in (14).

(14) Overuse of *the* (error %) per context

| For indefinites	[+partitive]	[-partitive]
[+specific]	Context (9) SC: 17.5% Korean: 38.8%	Context (11) SC: 7.5% Korean: 30%
[-specific]	Context (10) SC: 15% Korean: 16.3%	Context (12) SC: 7.5% Korean: 3.8%

Overall, the SC-group generally showed less overuse of *the* (12%) than the Korean group (22%). Both SC and Korean L2-learners overused *the* in [+partitive, +specific] contexts most often, as expected. To investigate the statistical significance of main effects, we conducted Repeated Measures ANOVAs on *the* use, using partitivity and specificity as within-subjects factors, and language and L2-proficiency as between-subjects factors. The ANOVAs show that there was a significant main effect of partitivity [$F(1,47)=11.73$, p=.001] and specificity [$F(1,47) =12.67$, p=.001], and there was a significant interaction between specificity and language [$F(1,47)=7.32$, p=.009]. There were no other main effects or interactions (e.g. no main effect of proficiency [$F(1,47) =1.87$, p=.18]; no main effects of language [$F(1,47) =2.13$, p=.151]; no interaction between specificity and partitivity [$F(1,47) =.17$, p=.68]).

To examine the main effects for each language group, we also conducted Repeated Measures ANOVAs on *the* use for each group separately, using partitivity and specificity as within-subjects factors, and L2-proficiency and test order as between-subjects factors. The ANOVAs on the SC-data show that there was a main effect of partitivity [$F (1, 28)=9.1$, p=.005], but there was no main effect of specificity [$F(1,28)=.22$, p=.64]. The ANOVAs on the Korean data show that there was a main effect of partitivity [$F(1,16) = 10.5$, p=.005] as well as specificity [$F(1,16)=12.72$, p=.003]. The ANOVA results suggest that the SC-group overused *the* significantly more often with [+partitive] contexts than with

[-partitive] contexts, just like the Korean group. Interestingly, however, specificity effects were observed (and quite strong) for the Korean group, but not for the SC-group. There was no significant interaction between specificity and partitivity in either group.

To understand the effect of partitivity and specificity for each individual learner, we categorized the individual patterns into five types: definiteness, specificity, partitivity, mixed, and unexpected.[5] The result of the individual analysis is given in (15). Only one subject falls in the unexpected pattern. All of the other four patterns were attested in both language groups, as shown in (15). More learners showed the definiteness (native-like) pattern and partitivity patterns in the SC-group than in the Korean group. The specificity pattern was observed more in the Korean group than in the SC-group.

(15) Individual results
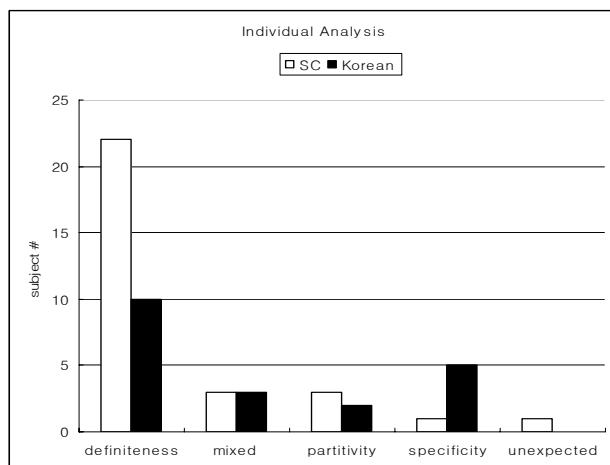
5. Discussion

Our data show that semantic universals play a significant role in L2-article choice. In particular, our experimental results show that both SC and Korean speakers are affected by the feature 'partitivity', and that they associate the [+partitive] feature with the definite article *the* in L2-English. This supports our hypothesis that partitivity is one of the semantic universals available for L2-learners regardless of the L1. In both SC and Korean data, there was no interaction between partitivity and specificity, which suggests that the two features are independent factors at work in L2-English article choice. Our results also show that both SC and Korean L2-learners overused *the* in [+partitive, +specific] contexts most often – which also indicates that partitivity and specificity contribute to overuse of *the* independently.[6]

---

[5] The individual analysis was conducted as follows: first, we selected subjects who made errors less than 25% (0-2 errors in total) for the relevant contexts. They belong to the definiteness pattern. For the remaining subjects, we computed the effects of specificity and partitivity. If the speaker made more errors in [+specific] than in [-specific] contexts (difference of at least 2 errors), and if there is no or little (1 error) difference between [+partitive] and [-partitive] contexts, s/he belongs in the specificity pattern. Similarly, if the speaker made more errors in [+partitive] contexts than in [-partitive] contexts (difference of at least 2 errors), and if there is no or little (1 error) difference between [+specific] and [-specific] contexts, s/he belongs in the partitivity pattern. If the speaker made at least 2 errors both in [+specific] and [+partitive] contexts, we categorized them into the mixed pattern. All remaining subjects belong in the unexpected pattern, including those subjects who made more than 2 errors in [-partitive, -specific] contexts, where no errors were expected. We acknowledge that there is no rule-of-thumb for categorizing the individual patterns, and the cut-off points that we chose are necessarily somewhat arbitrary. But our analysis suffices to show that the effect of specificity and partitivity holds on the individual as well as group levels.

[6] A cautionary note on the additive effects is in order, however. The difference between [+partitive, +specific] and [+partitive, -specific] contexts, and the difference between [+partitive, +specific] and [-partitive, -specific] contexts were statistically significant (p=.018 and p<.0001, respectively). The difference between [+partitive, +specific] and

We have also seen that there was almost no random overuse of *the* in [-partitive, -specific] contexts (e.g. (12)). This again supports the hypothesis that L2-article errors are not random, but tied to access to semantic universals.

Our results suggest that L2-learners face two independent problems in acquiring the semantics of English articles. One problem concerns learning that *the* requires common ground between speaker and hearer. The other concerns learning that *the* requires the uniqueness presupposition. If the learners have problems with the first issue, they show the specificity effect. They may overuse *the* in the absence of hearer knowledge/common ground. If the learners have problems with the second issue, the partitivity effect arises. The learners may erroneously associate *the* with the existence presupposition, which leads to incorrect overuse of *the* in partitive contexts.

One very puzzling finding is that there was a significant interaction between language and specificity. The Korean speakers were affected by both partitivity and specificity, as shown by the group results and individual analysis. Interestingly, however, the effect of specificity is rather weak (individual analysis) or non-existent (group analysis) for the SC-speakers. The SC-speakers seem to be sensitive only to partitivity effects.

One seemingly obvious interpretation of the data would be that there were L1-transfer effects in accessing specificity: namely, that the specificity feature is accessible to Koreans, but not to SC speakers. This argument, however, runs into serious problems. First of all, both Korean and Serbo-Croatian lack articles, and neither language encodes either specificity or partitivity. Therefore, it is difficult to claim that L1-transfer effects exist for SC-speakers, but not for Korean speakers. Furthermore, previous studies have found strong specificity effects for L1-Russian speakers of L2-English (Ionin 2003, Ionin et al. 2004). Serbo-Croatian is highly similar to Russian in the relevant respects, and there is no obvious reason why Russian speakers should be sensitive to specificity effects when SC-speakers are not.

We suggest that the observed differences between the SC-group and the Korean group are in fact due to the differences in overall L2-proficiency between the two groups. We argue that semantic universals are not equally persistent. Specifically, we suggest that the problem with the uniqueness presupposition (partitivity effects) is overcome later than the problem with the common ground (specificity effects) in the course of L2-acquisition. Thus, advanced learners may overcome specificity effects more easily than partitivity effects. Our SC-speakers were extremely advanced speakers of English, and their average Michigan test score was 28.5 points out of 30 points. In contrast, our Korean speakers were less advanced learners than SC-speakers, and their average Michigan test score was 24.3 points out of 30 points. Since the SC-speakers were more advanced than the Korean speakers in L2-English proficiency, the SC-speakers overcome specificity effects more easily than partitivity effects.

Our conjecture that L2-proficiency, instead of L1-differences, matters is supported by individual analysis as well. As shown in (16), the average L2-proficiency of the learners in the partitivity pattern (25.4) is much higher than the one in the specificity pattern (22.5). The average L2-proficiency of the learners in the definiteness (native-like) patterns (27.7) is comparable to the one in the mixed pattern (27.5). This indicates that advanced speakers may show the mixed pattern, but that when learners choose to adhere to only one semantic universal, more advanced learners fall into the partitivity pattern, while less advanced learners fall into the specificity pattern. Since the SC-group consisted of extremely advanced learners, they are less sensitive to the specificity effect, in comparison to the less advanced Korean group.

[-partitive, +specific] contexts were not statistically significant (p=.131) (data from both L1-groups). Hence, we cannot be absolutely certain that partitivity and specificity contribute to overuse of *the* additively.

(16) Individual analysis and L2-Proficiency
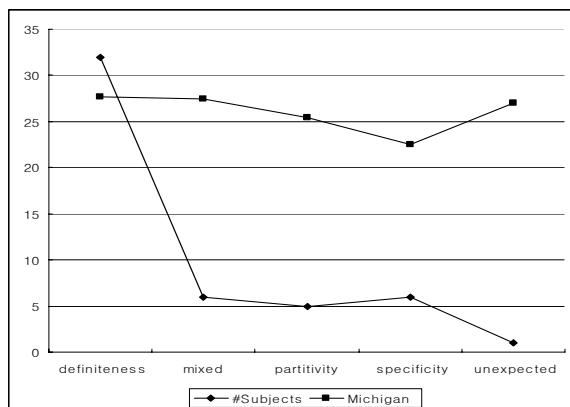
Our discussion raises many interesting questions for future research. We have argued that variation in implementing semantic universals is related to L2-proficiency, not to L1-differences, when L1-transfer effects are controlled for. To solidify this argument, further study is needed of beginner or intermediate SC-speaking L2-English learners and extremely advanced Korean L2-English learners. If our suggestion is on the right track, we expect that extremely advanced Korean learners would be less sensitive to specificity effects, and that beginner/intermediate SC-speaking learners would be sensitive to both specificity and partitivity effects.[7]

We have argued that partitivity effects are more persistent than specificity effects, which naturally raises the question of why that should be. We suggest that this difference is related to the semantic entailment relationship among the semantic universals.

Definiteness entails presuppositionality. Whenever a context satisfies the uniqueness presupposition (i.e., is definite), it also satisfies the presupposition of existence. Thus, it may be difficult for learners to decide which of these presuppositions *the* marks in English. If the learners associate *the* with presuppositionality, they will also correctly use *the* in definite contexts simply because all definites are [+presuppositional]. Hence, it would be hard for learners to abandon the incorrect assumption that *the* is associated with presuppositionality, because there is not very much positive evidence to the contrary. Since partitivity is a sub-type of presuppositionality, the learners' association of *the* with presuppositionality leads to *the* overuse with partitive indefinites.[8] On the other hand, definiteness does not entail specificity: there are non-specific definites (cf. Ionin et al. 2004). The learner may need ample positive evidence to understand that *the* does not mark specificity, since many contexts are both definite and specific, but it should be possible to tease these two semantic factors apart once learners understand that the common ground should be established to use *the*.

Another question that also remains open is whether we can obtain the effects of semantic universals in both L2A and L1A. We have argued that partitivity effects are more persistent than specificity effects in L2A. If our suggestion is on the right track, we would expect that a similar pattern would be observed in L1A as well. The existing data on L1A indicate that overuse of *the* in [+partitive] contexts is attested as late as age four (Maratsos 1976), but that overuse of *the* in potentially [+specific] contexts is practically gone by age three (Schaeffer and Matthewson 2005). To draw any strong

---

[7] In our currently ongoing study, we have obtained both partitivity and specificity effects with L1-Russian learners of L2-English. This further strengthens our claim that semantic universals are accessible to L2-learners regardless of the L1, when learners' L1 lacks articles.

[8] We should note that we have tested only presuppositional DPs, that are partitive. Partitivity by itself does not entail definiteness: while anaphoric definites (e.g. (2)) are [+partitive], non-anaphoric definites, like (3), are not. To solidify our claim that learners associate *the* with presuppositionality (which entails definiteness) rather than just its subtype partitivity, we need to test the effects of presuppositionality with indefinites in the absence of a previously mentioned set. Our prediction is that such non-partitive presuppositional indefinites will also trigger *the* overuse. We leave this as a future research question.

conclusion on the developmental patterns of partitivity and specificity in article errors, we need further studies on parallels between L1 and L2 article choice, with reference to the semantic universals, ideally in similar experimental settings.

6.   Conclusion

In this paper, we have investigated how semantic universals affect L2-article choice with two typologically different L1-groups: L1-Serbo-Croatian and L1-Korean learners of L2-English. Our results support the hypothesis that L2-article choice is not random, but reflects the learner's access to semantic universals. When there is no semantic trigger for using *the*, L2-learners rarely overuse *the* with indefinites. We have also argued that variation in L2-article choice is related to L2-proficiency. Partitivity effects are more persistent than specificity effects, and thus extremely advanced L2-learners (SC-speaking group) were more sensitive to partitivity effects than specificity effects. Our arguments suggest that previously observed variation in L2-article choice among different L1-groups should be more closely reexamined with reference to L2-proficiency and semantic universals. To this end, further studies on cross-linguistic as well as developmental data concerning article usage in L2A and L1A are required.

References

Bloom, Lois. 1970. *Form and Function in Learning to Talk*, Cambridge, MA: MIT Press.

Diesing, Molly. 1992. *Indefinites.* Cambridge, MA: MIT Press.

Enç, Mürvet. 1991. The semantics of specificity. *Linguistic Inquiry,* 22, 1-25.

Fodor, Janet and Ivan Sag. 1982. Referential and quantificational indefinites. *Linguistics and Philosophy,* 5, 355–398.

Hawkins, J.A. 1978. *Definiteness and indefiniteness.* London: Croom Helm.

Heim, Irene. 1991. Articles and definiteness. Published in German as "Artikel und Definitheit," in A. v. Stechow and D. Wunderlich (eds.), *Semantics: An international handbook of contemporary research.* Berlin: de Gruyter.

Huebner, Thomas. 1983. *A longitudinal analysis of the acquisition of English.* Ann Arbor: Karoma.

Ionin, Tania. 2003. *Article Semantics in Second Language Acquisition*, Doctoral dissertation, MIT, Cambridge, Mass.

Ionin, Tania, Heejeong Ko and Ken Wexler. 2004. Article semantics in L2-acquisition: the role of specificity. *Language Acquisition* 12: 3-69.

Ionin, Tania. 2006. *This* is definitely specific: specificity and definiteness in article systems. *Natural Language Semantics*, 14, 175-234.

Ionin, Tania, Maria Luisa Zubizarreta and Salvador Bautista Maldonado. in press. Sources of linguistic knowledge in the second language acquisition of English articles. To appear in a special issue of *Lingua.*

Kaneko, Yumina. 1996. Knowledge of the English article system in second language learning: To "the" or not to "the". Undergraduate thesis, Smith College.

Ko, Heejeong, Tania Ionin, and Ken Wexler. 2006. Adult L2-learners lack *the* maximality presupposition, too. In K. U. Deen et al (eds), *The Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition–North America, Honolulu, HI.* University of Connecticut Occasional Papers in Linguistics 4, 171-182.

Ko, Heejeong, Tania Ionin, and Ken Wexler. 2007. Parallels between L2 and L1 acquisition of article semantics. Ms. Seoul National University, University of Illinois at Urbana Champaign, and MIT.

Leung, Yan-kit Ingrid. 2001. The Initial State of L3A: Full Transfer and Failed Features? In X. Bonch-Bruevich et al. (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.

Maratsos, Michael P. 1976. *The Use of Definite and Indefinite Reference in Young Children.* Cambridge, UK: Cambridge University Press.

Master, Peter. 1987. *A Cross-Linguistic Interlanguage Analysis of the Acquisition of the English Article System*, Doctoral dissertation, University of California, LA.

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

Pesetsky, David. 2001. *Phrasal Movement and Its Kin,* Cambridge, MA: MIT Press.

Schaeffer, Jeannette and Lisa. Matthewson. 2005. Grammar and pragmatics in the acquisition of article system. *Natural Language and Linguistic Theory.*

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

Wexler, Ken. 2003. Maximal trouble. Paper presented at the CUNY Human Sentence Processing Conference. MIT. [In press as "Maximal trouble: cues don't explain learning", in E. Gibson and N. Pearlmutter (eds.), *The Processing and Acquisition of Reference*, Cambridge, MA: MIT Press].

Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)

edited by Roumyana Slabakova, Jason Rothman, Paula Kempchinsky, and Elena Gavruseva

Cascadilla Proceedings Project   Somerville, MA   2008

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Ko, Heejeong, Alex Perovic, Tania Ionin, and Ken Wexler. 2008. Semantic Universals and Variation in L2 Article Choice. In *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)*, ed. Roumyana Slabakova et al., 118-129. Somerville, MA: Cascadilla Proceedings Project.

or:

Ko, Heejeong, Alex Perovic, Tania Ionin, and Ken Wexler. 2008. Semantic Universals and Variation in L2 Article Choice. In *Proceedings of the 9th Generative Approaches to Second Language Acquisition Conference (GASLA 2007)*, ed. Roumyana Slabakova et al., 118-129. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1631.