

ORIGINAL ARTICLE

Plural marking in the second language: Atomicity, definiteness, and transfer

Sea Hee Choi*  and Tania Ionin

University of Illinois at Urbana–Champaign

*Corresponding author. Sea Hee Choi, 707 S. Mathews Ave., M/C 168, University of Illinois at Urbana–Champaign, Urbana, IL 61801. E-mail: schoi76@illinois.edu.

(Received 27 August 2019; revised 16 August 2020; accepted 24 August 2020; first published online 20 January 2021)

Abstract

This paper examines whether second language (L2)-English learners whose native languages (L1; Korean and Mandarin) lack obligatory plural marking transfer the properties of plural marking from their L1s, and whether transfer is manifested both offline (in a grammaticality judgment task) and online (in a self-paced reading task). The online task tests the predictions of the morphological congruency hypothesis (Jiang 2007), according to which L2 learners have particular difficulty automatically activating the meaning of L2 morphemes that are incongruent with their L1. Experiment 1 tests L2 learners' sensitivity to errors of *-s* oversuppliance with mass nouns, while Experiment 2 tests their sensitivity to errors of *-s* omission with count nouns. The findings show that (a) L2 learners detect errors with nonatomic mass nouns (*sunlights*) but not atomic ones (*furnitures*), both offline and online; and (b) L1-Korean L2-English learners are more successful than L1-Mandarin L2-English learners in detecting missing *-s* with definite plurals (*these boat*), while the two groups behave similarly with indefinite plurals (*many boat*). Given that definite plurals require plural marking in Korean but not in Mandarin, the second finding is consistent with L1-transfer. Overall, the findings show that learners are able to overcome morphological incongruency and acquire novel uses of L2 morphemes.

Keywords: atomicity; count/mass distinction; definiteness; first language transfer; generalized classifier language

We can count *chairs* but not *water*. The count/mass distinction is fully grammaticized in plural-marking languages such as English, but not in generalized classifier (GC) languages such as Korean and Mandarin Chinese. At the same time, Korean and Chinese do allow and/or require plural marking in certain contexts. This poses a testable question for second language (L2) acquisition of English by speakers of GC languages: do these learners transfer the properties of plural marking from their native language (L1) to their L2? The first goal of this paper is to investigate this question. The second goal is to examine whether L2 learners exhibit knowledge of plural marking, and/or L1-transfer with plural marking, across both offline

and online tasks. It is well established that learners do not perform uniformly across different types of tasks (e.g., Ellis, 2005). While untimed offline tasks such as grammaticality judgments have been argued (e.g., Ellis, 2005) to tap into explicit knowledge, online tasks such as self-paced reading have been argued to tap into more implicit or more automatized knowledge (e.g., Jiang, 2007). At the same time, many L2-studies have used offline tasks, such as grammaticality judgment and truth-value judgment tasks, to tap into learners' underlying, implicit knowledge (see Slabakova, 2008, 2016, for overviews). Our study uses both offline and online tasks with the same groups of learners in order to examine what learners know about English plural marking. The findings of our tasks largely converge (though there are some specific differences), and show that there is facilitative L1-transfer of plural marking from GC L1s, but only in contexts where plural marking is obligatory in the L1 (specifically, plural definite contexts in Korean). Our findings also show that learners exhibit sensitivity to errors with plural marking online as well as offline, even in contexts where no facilitative L1-transfer is expected; we argue that these findings pose a challenge to the morphological transfer hypothesis (Jiang 2007).

Plural marking cross-linguistically and in L2 acquisition

The count/mass distinction, atomicity, and definiteness in English, Korean, and Mandarin Chinese

The count/mass distinction is a morphosyntactic distinction that is encoded in the grammar of languages such as English, that is, languages that have obligatory plural marking (Chierchia, 1998). In English, plural *-s* is obligatory with count nouns when multiple entities are referred to, but ungrammatical with mass nouns. Only mass nouns can be used in bare (determiner-less) singular form ([1]), whereas count nouns cannot be used in bare singular form and always require a (definite or indefinite) determiner and/or plural marking ([2]).¹ Furthermore, count nouns can be directly combined with numerals, while mass nouns must combine with a measure noun phrase first ([3]).

- (1) I bought oil.
- (2) I bought *table/ a table/ tables.
- (3) three tables/ *three oils/ three bottles of oil.

The count/mass distinction is related to the cognitive object/substance distinction. In many languages, objects (e.g., *table*) are denoted by count nouns while substances (e.g., *oil*) are denoted by mass nouns. However, the object/substance distinction does not always map directly to the count/mass morphosyntax. In English, there are so-called object-mass nouns, which denote collections of objects, yet have mass morphosyntax: for example, *furniture*, *jewelry*, *footwear*, and so on (see Barner & Snedeker, 2005).²

Another concept that is tightly related to the count/mass distinction is *atomicity*, which is also called *individuality* or *boundedness* (see Bunt, 1985; Chierchia, 2010; Jackendoff, 1991; Landman, 1989, 1991; Langacker, 1999, 2008). According to Chierchia (2010, 2015), count nouns have stable (context-independent) atoms,

while mass nouns have unstable (context-dependent) atoms. For instance, while *table* has a uniform and recognizable minimal unit (a table), the minimal unit of *water* is vague and changes according to context (e.g., a cup of water vs. a drop of water). In this paper, we follow Chierchia (2010) in treating both count nouns (*table*) and object-mass nouns (*furniture*) as atomic (i.e., having stable atoms), but substance-mass nouns (*oil*) as nonatomic.

In GC languages such as Chinese and Korean any noun, object-denoting or substance-denoting, can be used in bare form, without determiners or plural marking, and numerals cannot combine with nouns directly, but require a classifier (see Li & Thompson, 1981), as shown in (4) and (5) for count nouns (there are some exceptions, namely, [+human] nouns in Korean, which can combine with the numeral directly). Furthermore, bare nouns in both languages are compatible with singular as well as plural meanings, as shown in (6) and (7).

- (4) han-kay-uy thakca
 one-Cl-Gen table
 “one table” [Korean]
- (5) yi zhang zhuozi
 one-Cl table
 “one table” [Mandarin]
- (6) Na-nun chayksang-ul sa-ss-ta.
 I-Top desk-Acc buy-Pst-Decl.
 “I bought a desk/desks.” [Korean]
- (7) Wo maile shuzhuo.
 I bought desk
 “I bought a desk/desks.” [Mandarin]

While, as shown above, singular nouns in GC languages can have plural meanings, such languages do nevertheless have plural marking, whose properties differ from those of plural marking in English. In Mandarin Chinese, the plural marker *-men* can be optionally combined with [+human] nouns ([8]) and is obligatory with pronouns (e.g., Iljic, 1994; Li, 1999). Its use with [-human] nouns is ungrammatical ([9]).

- (8) Wo yu daole pengyou(-men).
 I met friend(-Pl)
 “I met (the) friends.”
- (9) Wo maile wanju(*-men)
 I bought toy(*-Pl)
 “I bought toys.”

According to Li (1999), *-men* is only possible when the meaning of the (human) nouns is definite: for example, on Li’s analysis, *pengyou-umen* in (8) must mean “the friends” and not “friends.” However, while there is general agreement in the

literature about the restriction of *-men* to [+human] nouns, Li's claim that *-men*-marked nouns are definite is more controversial; for example, Jiang (2017) argues that *-men* is not inherently definite, and shows that it also has nondefinite, in particular, generic interpretations. For our purposes, what matters is that *-men* is restricted to [+human] nouns (regardless of whether it is further restricted only to definite [+human] nouns): all nouns tested in our study are inanimate and not compatible with *-men* in Mandarin.

In Korean, the plural marker *-tul* is used in a greater variety of contexts than the Mandarin plural marker. In particular, *-tul* can be combined with both animate and inanimate nouns. For our purposes, two properties of *-tul* are particularly important. First, while *-tul* is optional in indefinite contexts, as in (10), it is obligatory in definite contexts when multiple entities are being referred to (see Baek, 2002; Kim, 2005; Kwon & Zribi-Hertz, 2004): both in the presence of a demonstrative, as in (11), and in anaphoric (second mention) contexts; in the absence of *-tul*, (11) is a statement about only one book.³

(10) Yenghuy-nun ecey chayk-(tul)-ul sa-ss-ta.
 Yenghuy-Top yesterday book-Pl-Acc buy-Pst-Decl.
 "Younghee bought books yesterday."

(11) Yenghuy-nun ecey ku/i/ce chayk-*(tul)-ul sa-ss-ta.
 Yenghuy-Top yesterday that/this/that book-Pl-Acc buy-Pst-Decl.
 "Younghee bought these/those books yesterday."

The second property of *-tul* that is relevant for our purposes is its relationship with atomicity. Kim (2005) argues that *-tul* directly encodes the distinction between object-denoting (atomic) nouns and substance-denoting (nonatomic) ones. This proposal has received experimental support from Choi, Ionin, and Zhu (2018), who examined the use and judgments of *-tul* by Korean native speakers. They found that *-tul* was optionally used, and optionally allowed, with all object-denoting nouns: both those that are count in English (e.g., *chair*) and those that are mass in English (e.g., *furniture*). In contrast, *-tul* was both used and accepted to a much lower degree with substance-denoting nouns such as *oil*. Following Kim (2005), Choi et al. (2018) conclude Korean *-tul* directly encodes atomicity. In contrast Mandarin *-men*, which is restricted to [+human] nouns, does not encode atomicity; rather, atomicity has been argued to be reflected in the Mandarin classifier system (Cheng & Sybesma, 1998, 1999).

In this paper, we focus only on [-human], inanimate nouns, and on the differences between atomic and nonatomic, definite and indefinite interpretations. The relevant facts for these noun types in the three languages under discussion are summarized in Table 1.

Prior offline studies of plural marking in L2-English

There are many studies on the L2 acquisition of English plural marking, including Hiki (1990, 1991), Hua and Lee (2005), Snape (2008), and Choi et al. (2018). These

Table 1. Properties of plural marking in English, Korean, and Mandarin

	English	Korean	Mandarin
Indefinite inanimate nouns with plural interpretation	-s is obligatory	- <i>tul</i> is optional	- <i>men</i> is ungrammatical
Definite inanimate nouns with plural interpretation	-s is obligatory	- <i>tul</i> is obligatory	- <i>men</i> is ungrammatical
Relationship between plural marking and atomicity	Indirect: atomic mass nouns (e.g., <i>furniture</i>) exist	Direct: all atomic nouns can combine with - <i>tul</i>	Not applicable: atomicity is reflected in the classifier system

were offline studies, which focused on accuracy rather than reaction time (please see the next section for an overview of online studies in this domain). Hiki (1990, 1991) investigated how L1-Japanese L2-English learners judge the count/mass distinction in English by asking participants to edit a text that contained grammatical errors, and found that these learners have difficulty with abstract and substance-denoting nouns. Hua and Lee (2005) used two grammaticality judgment tasks and a forced-choice task, and found that L1-Chinese L2-English learners have more difficulty with abstract nouns than concrete nouns and more trouble with concrete mass nouns than with concrete count nouns. Snape (2008) obtained similar results with L1-Japanese L2-English learners and L1-Spanish L2-English learners of intermediate proficiency, who had more difficulty with mass nouns, incorrectly accepting mass plural nouns.

The only studies that specifically investigated the role of atomicity in L2 acquisition of the count/mass distinction are Inagaki (2014), MacDonald and Carroll (2018), and Choi et al. (2018). In Inagaki (2014) and MacDonald and Carroll (2018), participants completed a quantity judgment task (from Barner & Snedeker, 2005; see also Inagaki & Barner, 2009), in which they had to say which was “more”: two big entities (e.g., two big blobs of mustard or two big pieces of furniture—judgment by volume) or six small entities (e.g., six small blobs of mustard or six tiny pieces of furniture—judgment by number). Inagaki (2014) found that native speakers of Japanese, native speakers of English, and L1-Japanese L2-English learners all behaved the same on this task, judging substance-mass nouns such as *mustard* by volume, but judging object-denoting, atomic nouns by number, regardless of whether they are count (*chairs*) or mass (*furniture*) in English. MacDonald and Carroll (2018) obtained similar findings when they used this task with L1-Korean L2-English learners, compared to native speakers of English and Korean.

Choi et al. (2018) examined how L1-Korean and L1-Chinese L2-English learners acquire the count/mass distinction in English, using a grammar task in which participants had to supply the correct form of the noun in a sentence; use of the quantifiers *a lot of* and *more* ensured that all nouns had to be either mass (*more water/sunlight*) or plural (*more chairs*). Choi et al. found identical performance in the two L2-groups: the learners correctly supplied the plural form of count nouns (*chairs*) and the singular form of substance-mass nouns (*sunlight*) but incorrectly oversupplied the

plural form of object-mass nouns (*furnitures*). This was the case even though Korean allows *-tul* with atomic mass nouns such as *furniture* but Mandarin Chinese does not allow *-men* with inanimate nouns at all, as discussed above. Thus, the oversuppliance of *-s* with *furniture* in L2-English could not be attributed to L1-transfer from Korean, as L1-transfer would not predict such oversuppliance for L1-Mandarin L2-English learners. Choi et al. conclude that atomicity is a semantic universal, and that L2-English learners from GC L1s rely on atomicity when determining whether a noun is count or mass in English: as an initial hypothesis, such learners treat all atomic nouns as count. L1-transfer was argued not to play a role in this domain. A follow-up study by Tang, Fiorentino, and Gabriele (2019) found that while L1-Chinese L2-English learners associated English *-s* with atomicity (as in Choi et al., 2018), L1-French L2-English learners, coming from an L1 with obligatory plural marking, were subject to L1-transfer at the lexical level, allowing plural marking with those nouns that take plural marking in French, their L1.

Choi et al. (2018) considered only indefinite contexts, in which plural marking is optional in Korean; it is possible that L1-transfer would take place in definite contexts, in which Korean requires plural marking, as discussed above. Another limitation of Choi et al. (2018) was the use of highly explicit methodology (grammar correction), which may have probed learners' explicit, metalinguistic knowledge; it is not clear whether the results would be replicated with a task testing more implicit and/or automatized knowledge (see the next section).

Morphological congruency and prior online studies of plural marking in L2 acquisition

There are many theoretical proposals concerning the influence of the L1 on the L2 (e.g., Lardiere, 2008, 2009; Schwartz & Sprouse 1994, 1996); there are also many accounts of why inflectional morphology is particularly difficult for L2 learners (see, e.g., Franceschina, 2001, 2005; Hawkins & Liszka, 2003; Lardiere, 2008, 2009; Prévost & White, 2000). In this paper, we focus on a particular proposal, the morphological congruency hypothesis (MCH), which was proposed and tested in a series of studies by Jiang and colleagues (Jiang, 2004, 2007; Jiang, Hu, Chrabaszcz, & Ye, 2017; Jiang, Novokshanova, Masuda, & Wang, 2011). The reason we focus on the MCH is that it makes very specific, testable predictions for the role of L1-transfer with inflectional morphology, and for how learners should behave in online, time-pressured tasks.

According to the MCH (in its most recent instantiation in Jiang et al., 2017), a given L2 morpheme is *congruent* for an L2-learner if the meaning that it marks in the learner's L2 is also marked by a corresponding morpheme in the learner's L1. Otherwise, the morpheme is *incongruent*. Jiang et al. (2007) argue that incongruent morphemes are particularly difficult for L2 learners to acquire. The MCH is based on the notion of automatic activation: if a learner has nativelike command over a given morpheme, such as the plural *-s* in English, she will automatically activate the plural meaning of *-s* upon encountering it. This is why the MCH can only be tested with time-pressured tasks, which require the learner to activate the morpheme's meaning online. For the MCH, nativelike performance with a given morpheme in an untimed, offline task does not provide evidence that the learner has fully acquired a morpheme.⁴

Jiang and colleagues have tested the MCH in the realm of English plural marking, with a focus on L2-English learners from GC L1s, Chinese and Japanese. Jiang (2007) used a self-paced reading task (SPRT) with L1-Chinese L2-English learners, presenting them with both grammatical and ungrammatical sentences. The basic logic of the SPRT (see Jegerski, 2014) is that when native speakers encounter an error in the sentence that they are reading, they slow down in the region of the error (or the next, spillover, region), relative to a minimally different control condition with a grammatical sentence. In the relevant conditions in Jiang (2007), a plural noun phrase (NP) was presented with *-s* in the grammatical variant, and without *-s* in the ungrammatical variant; plurality was established by means of a partitive construction, for example, *some of the rabbit(s); a large number of the book(s)*. Jiang (2007) found that unlike native English controls, L1-Chinese L2-English learners did not slow down for missing *-s* in the ungrammatical conditions; in contrast, the learners did slow down for other types of errors (verb subcategorization errors). Jiang et al. (2011) replicated these findings with L1-Japanese L2-English learners, and furthermore showed that L1-Russian L2-English learners did slow down for missing *-s*; as Russian, like English, has obligatory plural marking, these findings provided support for the MCH: L2 learners could only recognize errors with plural marking online when they had a congruent plural morpheme in their L1. Jiang et al. (2017) provided further support for the MCH by means of a different methodology, a sentence-picture matching task. Once again, L1-Chinese L2-English learners did not exhibit sensitivity to presence versus absence of plural marking, whereas L1-Russian L2-English learners did.

Thus, there is much evidence from the studies by Jiang and colleagues that L1-Chinese L2-English learners do not have automatized knowledge of English plural marking. As discussed above, plural marking is disallowed with [-human] nouns in Chinese; as far as we can tell from the appendix in Jiang (2007), most of the stimuli contained [-human] nouns. An interesting question is what we would find in a study with L1-Korean L2-English learners, given that plural marking is much less restricted in Korean than in Mandarin. Another interesting question is how learners would perform on missing versus on incorrectly supplied *-s*. The studies by Jiang and colleagues all consider only errors of *-s* omission, and find a complete lack of online sensitivity in L1-Chinese L2-English learners. In contrast, Choi et al. (2018) found that L1-Chinese as well as L1-Korean L2-English learners were quite accurate at supplying *-s* with count nouns, and that their primary error was oversuppliance of *-s* with atomic mass nouns. However, this was an offline study with highly explicit methodology. Conversely, the studies by Jiang and colleagues used only online tasks, so we do not know whether the learners in these studies would exhibit knowledge of plural marking in an offline, more explicit task.

In our study, we attempt to bridge the gap between the studies of Jiang and colleagues, and Choi et al. (2018), by using both online and offline methodologies, and by examining both errors of plural marker misuse (Experiment 1) and those of plural marker omission (Experiment 2).

Research questions and study goals

This study examines what L1-Korean and L1-Mandarin L2-English know about English plural marking. We consider both atomic and nonatomic mass nouns

(indefinite only), and both definite and indefinite count nouns; only inanimate nouns are tested. We ask the following research questions:

RQ1: Do L1-Korean and/or L1-Chinese L2-English learners exhibit knowledge of English plural marking both offline and online?

RQ2: Do L1-Korean and/or L1-Chinese L2-English learners exhibit sensitivity both to errors of plural marker misuse (Experiment 1) and plural marker omission (Experiment 2), and does their sensitivity to errors vary based on the context or noun type?

RQ3: Is there evidence for L1-transfer from a GC L1 in this domain, and if so, is it manifested both offline and online?

Let us consider what the MCH predicts for the performance of these learner populations in these environments. As discussed above, Chinese is completely incongruent with English with regard to plural marking, at least on [-human] nouns. Therefore, L1-Chinese L2-English learners are expected to show no online sensitivity to errors of either *-s* oversuppliance or *-s* omission, for any noun type tested. As for Korean speakers, there are two logical possibilities.

If we consider the behavior of *-s* and *-tul* across all contexts, they are clearly incongruent: *-tul* is optional where *-s* is obligatory (see also Hwang & Lardiere, 2013). If *-tul* and *-s* are always incongruent, then we expect to see no L1-transfer, and L1-Korean L2-English learners would behave just like L1-Chinese L2-English learners, exhibiting no sensitivity to errors with *-s* online. Alternatively, it is possible that *-s* is congruent with *-tul* in those contexts where the presence of *-tul* is obligatory in Korean, that is, in definite contexts. Then, we would expect L1-Korean L2-English learners to be sensitive to missing *-tul* in those contexts. When these learners encounter a plural definite NP with missing *-s*, they should automatically react to the missing *-s*, as *-tul* would be required in a corresponding context in Korean; in contrast, when they encounter a plural indefinite NP with missing *-s*, they should not detect the missing *-s*, as in the corresponding context in Korean, *-tul* is only optionally used.

It is somewhat less clear what the MCH predicts for errors of *-s* oversuppliance on the part of L1-Korean L2-English learners. When these learners encounter an atomic mass noun marked with plural morphology, such as *furnitures*, they should be insensitive to the presence of *-s*, as the corresponding noun in Korean can optionally combine with *-tul*. In contrast, if they encounter a plural-marked nonatomic mass noun, such as *sunlights*, it is possible that the learners would detect the error of oversupplied *-s*, as *-tul* is ungrammatical in the corresponding contexts in Korean. We phrase this prediction somewhat tentatively, as none of the papers published on the MCH have anything to say about oversuppliance of inflectional morphemes, only about omission. However, we believe that this prediction follows naturally if *-tul* and *-s* are congruent. The above predictions are summarized in Table 2.

These predictions are for online performance. The MCH makes no predictions for offline performance, as it is concerned exclusively with learners' automatized, online performance. None of the studies by Jiang and colleagues cited above included offline tasks. However, we believe that it is very important to include

Table 2. Predictions for online performance with plural marking, under the MCH

Error type	Predictions for . . .	
	L1-Chinese L2-English learners (and L1-Korean L2-English learners if <i>-s</i> and <i>-tul</i> are not congruent)	L1-Korean L2-English learners, if <i>-s</i> and <i>-tul</i> are congruent in some contexts
<i>-s</i> oversuppliance with non-atomic mass nouns (<i>sunlights</i>)	NO	YES
<i>-s</i> oversuppliance with atomic mass nouns (<i>furnitures</i>)	NO	NO
<i>-s</i> omission with plural indefinites (<i>many book</i>)	NO	NO
<i>-s</i> omission with plural definites (<i>these book</i>)	NO	YES

YES, the learners are predicted to detect the error online, per the MCH. NO, the learners are predicted to *not* detect the error online, per the MCH

offline tasks in the design, in order to provide a baseline. If learners fail to detect an error online, we need to know whether this is because they lack the relevant knowledge at any level, or whether they do have the knowledge at an explicit level, but it has not (yet) become automatized. Including offline tasks in the design also allows for comparison with prior studies that used offline tasks.

Overall methodology and data analysis

This study tested L1-Korean and L1-Mandarin L2-English learners' sensitivity to English plural marking in both an offline and an online task. Both tasks incorporated two separate experiments: Experiment 1 on the role of atomicity with mass nouns, and Experiment 2 on the role of (in)definiteness with count nouns. This section describes the overall methodology, including participants, procedure, and the format of both experimental tasks, as well as the overall data analysis. Below we report on the specifics of the two experiments that were tested by those tasks.

Overall procedure and proficiency tasks

The participants were tested individually in a lab in a single session. After filling out the consent form, each participant completed a SPRT, a grammaticality judgment task (GJT), a vocabulary-check task, a proficiency cloze test, and a language background questionnaire. The vocabulary-check task and the cloze test were used only with L2 learners. The SPRT was administered prior to the GJT to reduce any possible influence from the GJT, which is more explicit than the SPRT. All the tasks were presented on a computer screen in a testing booth. The whole experiment took 30–40 min for English native speakers (NSs) and 40–60 min for L2 learners.

The proficiency test was taken from Ionin and Montrul (2010); it was a 40-item cloze test, adapted from a text passage from O'Neill, Cornelius, and Washburn (1991), in which every seventh word was removed and replaced by three multiple

choice options; while all three options made sense in the context, only one of them was grammatical. The NSs were not tested on this task, since Ionin and Montrul (2010) found that native English speakers tested in their study performed at ceiling. In the vocabulary-check task, the participants were given a list of English words, and asked to choose the correct translation of each word into their L1, choosing among three options. The words and the options were pseudorandomized. The vocabulary-check task contained all the target nouns used in the SPRT and the GJT for Experiment 1, in case any of these words were unfamiliar to the learners. The target nouns from Experiment 2 were not included in the vocabulary-check task, as they were all common, high-frequency count nouns. Both the proficiency test and the vocabulary-check test were administered via Google Forms.

Participants

The participants were 35 NSs of English, 31 L1-Korean L2-English learners, and 35 L1-Mandarin Chinese L2-English learners.⁵ All participants were recruited at a large Midwestern US university. The L2 learners recruited in the experiment were all born in their native countries (South Korea, China, and Taiwan) and had moved to the United States at age 16 or older. The participants' background information, as well as their scores on the proficiency test (as % correct out of 40) are reported in Table 3. The proficiency levels of the participants were from intermediate to advanced. An independent samples *t* test showed that the proficiency levels between the two L2-groups were not significantly different, $t(62) = -1.29, p = .198$.

SPRT and GJT format

Identical test materials were prepared for the SPRT and the GJT. Each task contained 128 sentences in total, with 32 sentences for Experiment 1, 32 sentences for Experiment 2, and 64 fillers, which were not related to either experiment. Because the target sentences for Experiments 1 and 2 were set up quite differently, they also served as fillers for each other. The 64 additional fillers consisted of a variety of unrelated constructions such as verb inflection, tense and the use of auxiliary verbs. Among fillers, half were grammatical while half were ungrammatical.⁶ Four experimental lists were constructed; the same fillers were used in all four lists. The items were pseudorandomized for order of presentation.

The SPRT was administered using E-prime, a software program developed by Psychology Software Tools (Pittsburgh, PA, USA). The sentences were presented word-by-word, and the participants were asked to press the space bar after each word to advance to the next word. Each item was followed by a yes/no comprehension question. The comprehension questions never targeted the critical nouns, but asked about other parts of the sentences. The SPRT was preceded by six practice items. Halfway through the task, there was a 5-min break.

The GJT was administered via Google Forms. Each sentence was presented as a whole, and the participants were instructed to read it and judge its grammaticality by selecting Yes or No. The GJT was untimed. In order to allow for a direct comparison between the results of the SPRT and the GJT, we administered the same experimental list for both tasks to each participant.

Table 3. Participants' background information

Group	N (male/ female)	Mean age at time of testing (range)	Mean length of stay in the United States in years (range)	Mean age of acquisition in formal setting (range)	Mean age of first exposure to English in the United States (range)	Mean cloze test scores as % correct (range)
NSs	35 (10/25)	21 (18–31)	n/a	n/a	n/a	n/a
Mandarin L2ers	35 (11/24)	20 (18–25)	5 (0.5–9)	11.5 (6–15.5)	18 (16–24)	80% (70%–95%)
Korean L2ers	31 (16/15)	27 (18–36)	6 (0.5–12)	10.3 (6–17)	24 (16–31)	78% (63%–90%)

Overall data analysis

All analyses were implemented in R (R Team, 2014). The GJT results in each experiment were analyzed using a mixed effects logistic regression; the *glmer()* function was used since the dependent variable was binary (YES vs. NO, coded 1 vs. 0; Jaeger, 2008).⁷ The SPRT reading time (RT) results in each experiment were analyzed using the *lmer* function in R (Bates, Mächler, Bolker, & Walker, 2014). Separate models were conducted for each region of interest. Prior to the actual data analysis, we trimmed the data to eliminate outliers. Following common practice in analyzing self-paced reading data (e.g., Wei et al., 2018), we adopted two-step procedure. First, we removed raw RTs that are below 100 ms as it is almost impossible to parse a sentence in less than 100 ms. Second, we replaced with the cutoff values any RT that was more than 3 *SD* from each participant's grand mean. This procedure resulted in 2.1% of the data being trimmed in Experiment 1, and 1.86% in Experiment 2.

We also analyzed comprehension question accuracy in the SPRT; following Foote (2011), we included in the analysis all participants who scored at least 70% correct on the comprehension questions. All participants met this criterion in both experiments; average comprehension accuracy was above 80% in both experiments for all three groups.

In both experiments, we analyzed RTs in the critical region (the target noun) as well as the postcritical or spillover region. In Experiment 1, the ungrammatical noun forms (e.g., *furnitures*) were always one or two letters longer than the corresponding grammatical forms. As it is well known that word length as well as (un)grammaticality increase RTs (Lago, Shalom, Sigman, Lau, & Phillips, 2015), the effects of word length were estimated from the entire data set in Experiment 1, and regressed from the raw RTs using a linear model (cf. Hofmeister, 2010). Thus, for data analysis for the critical regions and the spillover region in Experiment 1, we used length-regressed RTs, following Lago et al. (2015). In Experiment 2, the situation was the opposite: the target noun was always one or two letters *longer* in the grammatical than in the ungrammatical variant, as the plural form was grammatical in Experiment 2. The shorter word length in the ungrammatical conditions was expected to make it harder to find grammaticality effects, and/or to reduce the effect sizes. In order to check for any potential confounding effects of word length, the effect of word length was estimated from the entire data set in Experiment 2 using *lmer()* function in R. The results showed no significant effect of word length, so we did not use residual RTs in Experiment 2; to satisfy the assumption of normality, we log transformed raw RTs for the statistical modeling. For plotting, we used raw RTs in both experiments in order to allow for visual comparisons between the two data sets.

Within each analysis in each experiment, we initially used a maximal model including random intercepts and slopes for all fixed effects and their interactions, following the common practice in the psycholinguistics literature (Barr, Levy, Scheepers, & Tily, 2013), but the maximal model did not converge in any of the analyses. Thus, we simplified the model following the back-off procedure suggested by Bates et al. (2014). For each analysis, we report the final model that converged; it included by-subject and by-item random intercepts.

A total of six models are reported in this paper: two models analyzing the GJT data (one per experiment), and four models analyzing the SPRT data (one for the

critical region and one for the postcritical region in each experiment). In Experiment 1, the fixed effects introduced into each model were group (3 levels), atomicity (2 levels), and grammaticality (2 levels); in Experiment 2, the fixed effects were group (3 levels), definiteness (2 levels), and grammaticality (2 levels). Helmert coding was used for the variable of group: first, the English NS group was compared to the full L2-group, and subsequently the two L2-groups were compared to each other. Two-way and three-way interaction terms between the fixed factors were also included in each model. Significant interactions were followed by pairwise comparisons using the *emmeans()* function (Lenth, 2018; the Bonferroni correction for multiple comparisons is automatically implemented in R), as well as by a visual examination of the interaction plots, which were plotted using *plot_model()* function (Gelman, 2008).⁸ The results sections in the paper report the significant results; the full outputs of all models are available in the online-only supplementary materials.

Experiment 1: The role of atomicity in the L2-acquisition of English plural marking with mass nouns

In Experiment 1, we investigated whether atomicity affects L2 learners' sensitivity to plural marking both offline and online. Experiment 1 was intended as a partial replication of the offline study in Choi et al. (2018).⁹

Experiment 1: Test materials for the SPRT and GJT

In Experiment 1, only mass nouns were tested: atomic mass nouns such as *furniture* were contrasted with nonatomic ones such as *sunlight*. They were all used in bare form, with indefinite interpretation. This means that for native English speakers, the singular form of the noun is always grammatical, and the plural form is always ungrammatical. The target sentences for Experiment 1 were constructed by crossing the factor of *atomicity* (atomic vs. nonatomic) with the factor of *grammaticality* (grammatical singular form vs. ungrammatical plural form), giving rise to four conditions, with eight tokens per condition per list (32 total). A sample token set is provided in Table 4.

Sixteen atomic and 16 nonatomic mass nouns were used in the item construction. The sentence frames used for atomic and nonatomic mass nouns were different, as it is impossible to construct identical sentence frames that would work equally well for nouns with very different meanings. The overall frequency of the target nouns in the atomic and nonatomic conditions was matched using the Corpus of Contemporary American English; it was 8533.94 for atomic nouns, and 9159.94 for nonatomic nouns (out of 560 million words of text). A *t* test yielded no significant difference in frequency between the two noun types ($p = .83$).

Two test lists were created, using a Latin-square design so that each list contained eight of the atomic nouns and eight of the nonatomic nouns in the grammatical conditions, and the other eight atomic and eight nonatomic nouns in the ungrammatical conditions. As Experiment 2 necessitated four lists, this means that each list for Experiment 1 was repeated twice.

Table 4. Experiment 1 sample items: Breakdown of target sentences by region

	R1	R2	R3	R4	R5	R6 [CR]	R7 [CR+1]	R8	R9
Atomic, grammatical	At	the	store,	Mary	bought	furniture	for	her	family.
Atomic, ungrammatical	At	the	store,	Mary	bought	furnitures	for	her	family.
Nonatomic, grammatical	At	the	park,	Jill	enjoyed	sunlight	for	a	while.
Nonatomic, ungrammatical	At	the	park,	Jill	enjoyed	sunlights	for	a	while.

Note: CR, critical region. CR+1, postcritical region.

Table 4 shows how the target sentence in each condition breaks down by region in SPRT; the target noun was always in region 6 (r6), while the postcritical region was region 7 (r7).

Experiment 1: Predictions

Native English speakers are predicted to behave no differently in the atomic and the nonatomic conditions: they should reject sentences with *furnitures* and *sunlights* in the GJT, and slow down for ungrammaticality with both noun types in the SPRT, relative to the grammatical variant. The slowdown may take place in the critical and/or the postcritical region.

With regard to the learner groups, the predictions under the MCH are laid out in the top half of Table 2. L1-Chinese L2-English learners are predicted not to show sensitivity to *-s* oversuppliance with either noun type in the SPRT; L1-Korean L2-English learners are predicted to either behave just like their L1-Chinese counterparts, or alternatively, if *-s* and *-tul* are congruent, to recognize oversuppliance of *-s* with nonatomic mass nouns only.

The above predictions, made on the basis of the MCH, are markedly different from what was found by Choi et al. (2018) in their offline study, where both groups of learners overused *-s* with atomic mass nouns but not with nonatomic ones, indicating sensitivity to the semantic universal of atomicity. If their results are replicated in our study, then we expect both groups of learners to reject sentences with *-s* on nonatomic mass nouns, but not with atomic mass nouns, in the GJT, and to slow down for the oversuppliance of *-s* with nonatomic nouns, but not with atomic ones, in the SPRT. Such an SPRT result would go against the predictions of the MCH: there is no congruence between *-s* and the Chinese *-men*, hence no reason for L1-Chinese L2-English learners to detect oversupplied *-s* on any noun type in the SPRT.¹⁰

Experiment 1: Results

GJT results

Figure 1 shows the results of GJT, as the percentage of items that elicited a Yes response. English NSs showed almost at-ceiling performance in all conditions. Both L2-groups performed much like NSs in the grammatical conditions, but over-accepted sentences in the ungrammatical conditions. The overacceptance was much

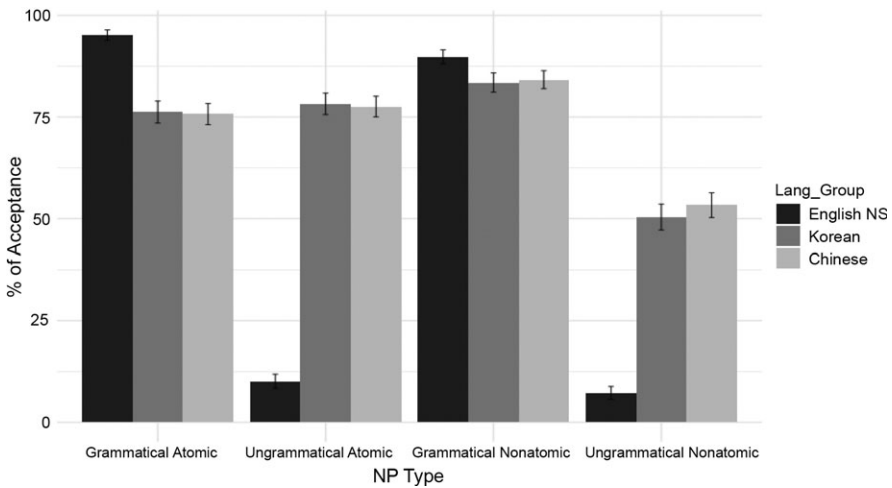


Figure 1. %Acceptance (YES response) to target items in the GJT in Experiment 1 (error bars indicate +/- standard deviation).

higher with atomic than with nonatomic nouns: in the former case, both L2-groups accepted grammatical and ungrammatical sentences to about the same extent.

In the analysis, we found a main effect of group (NS vs. L2: $z = 5.40$, $p < .0001$), due to greater acceptance rates for the L2 learners than the NSs; the two L2 groups did not differ from each other. There was also a main effect of grammaticality ($z = 20.84$, $p < .0001$), with higher acceptance rates for grammatical than ungrammatical sentences, as expected. We also found a main effect of atomicity ($z = 2.95$, $p = .0003$), due to higher acceptance rates with atomic than nonatomic nouns. There were also significant two-way and three-way interactions among group (NS vs. L2), grammaticality, and atomicity. The two L2-groups did not differ from each other in any condition. The source of the three-way interaction ($z = -4.47$, $p < .0001$) is that L2 learners accepted ungrammatical sentences more than the NSs, and furthermore that the L2 learners accepted ungrammatical sentences in the atomic condition much more than those in the nonatomic condition.

SPRT results

Figures 2, 3, and 4 show the average raw RTs for each group. The rectangular frames show the critical region and the spillover region on which we conducted statistical analyses.

In the analysis for the critical region (r6), there is a main effect of group (NS vs. L2; $t = 7.07$, $p < .0001$), which shows that the English NSs read significantly faster than L2 learners. There was a significant effect of grammaticality ($t = -4.75$, $p < .0001$; the target noun was read faster in the grammatical sentences than the ungrammatical sentences), and a significant effect of atomicity ($t = -2.41$, $p = .02$; the target noun was read faster in atomic than nonatomic conditions). We also found a significant interaction between group (NS vs. L2) and grammaticality

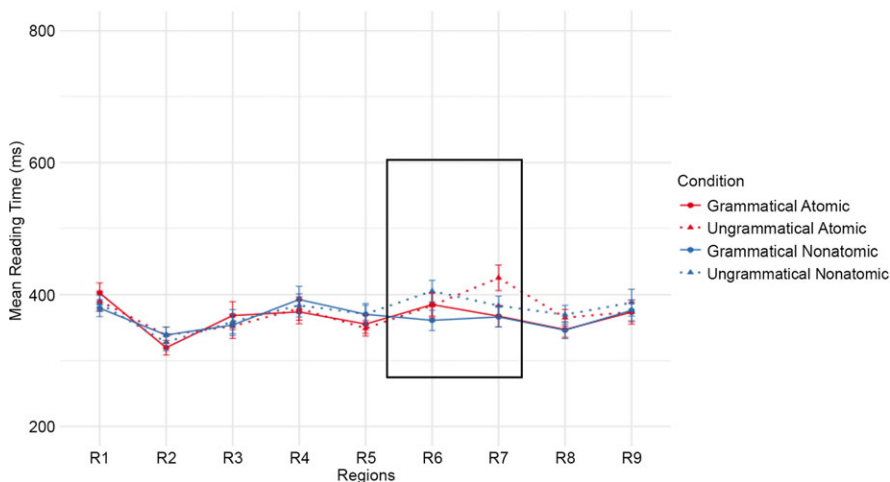


Figure 2. SPRT results, English NSs: average raw RTs, Experiment 1 (error bars show standard error).

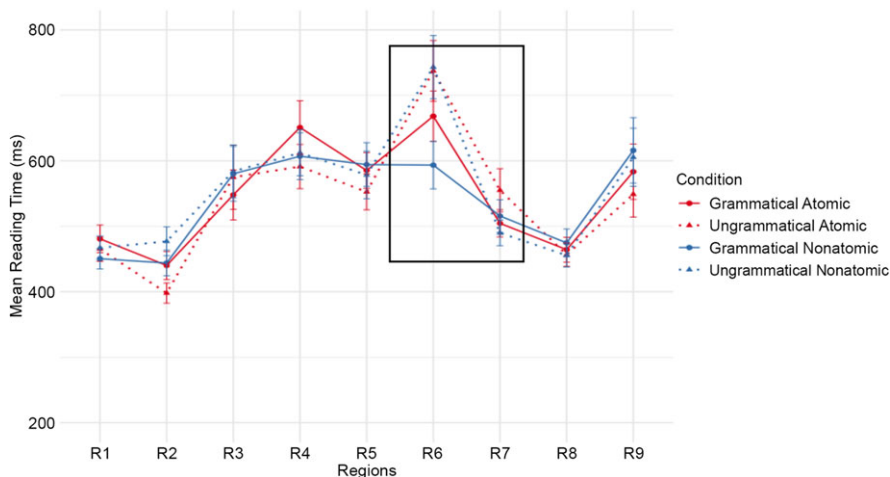


Figure 3. SPRT results, Korean L2ers: average raw RTs, Experiment 1 (error bars show standard error).

($t = -4.15, p < .0001$), which was due to L2 learners showing much bigger slow-downs than English NSs. There was also a significant interaction between group (Korean vs. Chinese) and atomicity ($t = -2.5, p = .01$), due to the Korean group showing a smaller RT difference between the atomic and the nonatomic conditions than the Mandarin group, even though both groups read faster in the atomic than the nonatomic condition. Finally, there was a significant interaction between grammaticality and atomicity ($t = 3.24, p = .001$), because all three groups, even NSs, slowed down for ungrammaticality in the nonatomic condition but not in the atomic condition.

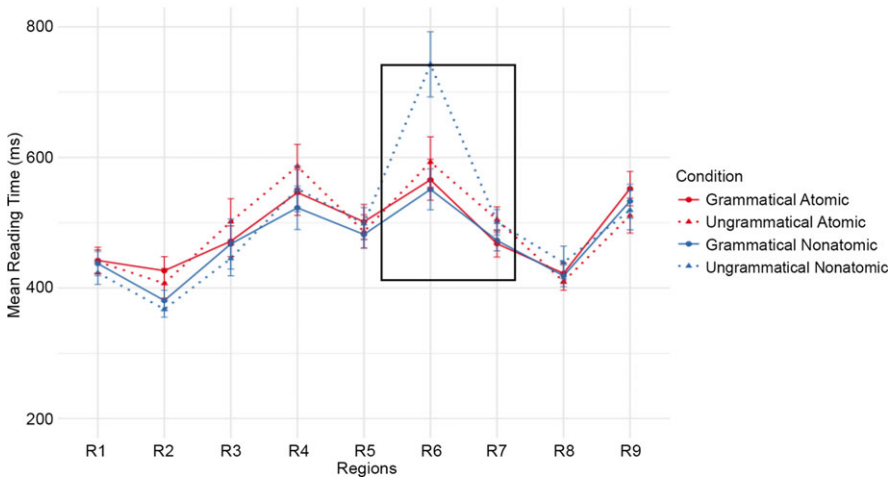


Figure 4. SPRT results, Mandarin L2ers: average raw RTs, Experiment 1 (error bars show standard error).

In the analysis for the spillover region (r7), there is a significant main effect of group (NS vs. L2; $t = 5.62$, $p < .0001$), which indicates that the English NSs read significantly faster than L2 learners; there was no difference between the two L2 groups. We also found a main effect of grammaticality ($t = -3.70$, $p < .0001$), with faster RTs in the grammatical than the ungrammatical conditions. There was also a marginal main effect of atomicity ($t = 1.84$, $p = .075$), with the spillover region read more slowly in the atomic than the nonatomic conditions. There was a significant interaction between grammaticality and atomicity ($t = -2.61$, $p = .009$), due to participants slowing down for ungrammaticality in the atomic condition more than in the nonatomic condition. This result was largely due to the NSs slowing down for ungrammaticality in the atomic condition, even though the interaction with group did not reach significance. We also found a marginal interaction between group (Korean vs. Chinese) and atomicity, as well as a marginal three-way interaction with both atomicity and grammaticality ($t = 1.93$, $p = .05$). This was because the Korean group read grammatical sentences faster than ungrammatical sentences in the atomic condition but the opposite in the nonatomic condition, while the Chinese group read grammatical sentences faster than ungrammatical sentences in both conditions; however, none of these r7 differences between grammatical and ungrammatical sentence RTs in the L2-groups reached significance in the pairwise comparisons.

To sum up, we see that NSs of English slow down for ungrammaticality in both conditions: the slowdown occurs in the critical region in the nonatomic conditions, but in the spillover region in the atomic conditions. In contrast, both L2-groups slow down for ungrammaticality in the nonatomic condition, in the critical region, but show no significant slowdown for ungrammaticality in the atomic condition. The behavior of the L2-groups is very similar; while there is a marginal three-way interaction between L2-group, grammaticality and atomicity, this interaction occurs in the spillover region, where the L2 learners make numerically very small distinctions among conditions. The slowdowns for the L2-groups occur in the

critical region, and here, we see a clear effect of atomicity, and the same patterns exhibited by both groups.

Discussion: Experiment 1

Both the GJT and the SPRT results show that L1-Korean and L1-Mandarin L2-English learners exhibit highly similar patterns. Both L2-groups were able to detect ungrammatical plural marking on nonatomic mass nouns such as *sunlight*, but not on atomic mass nouns such as *furniture*, online as well as offline. In contrast, English NSs were able to detect ungrammaticality in both atomic and nonatomic conditions, in both tasks. Interestingly, in the SPRT, the English NSs detected ungrammaticality faster with nonatomic than with atomic nouns. This may suggest that even English NSs, for whom the count/mass distinction is fully grammaticized, are sensitive to the semantic universal of atomicity on some level, so that detecting incorrect plural marking on atomic mass nouns takes more time and effort.

Returning to the L2-data, we see that performance of the two groups is very similar to each other, very similar on the offline versus online tasks (despite some minor differences), and also very similar to the results obtained by Choi et al. (2018), who used a different type of offline task (grammar correction). This relative convergence across studies and across tasks provides convincing evidence that atomicity is not subject to transfer. Even though the Korean plural marker *-tul* is optionally used with atomic nouns while the Chinese plural marker *-men* cannot be used with any inanimate nouns, the two groups were both influenced by atomicity. The behavior of the Chinese group in particular was not predicted under the MCH (see Table 2): even though *-s* is not congruent with *-men*, the L1-Chinese L2-English learners successfully detected oversupplied *-s* in the SPRT; furthermore, they detected this oversuppliance on nonatomic mass nouns but not on atomic ones, which is not traceable to transfer from Chinese.

Thus, we believe that the behavior of learners on *-s* oversuppliance goes against MCH predictions. However, all prior studies on the MCH were done with omission rather than oversuppliance of *-s*, making comparisons between Experiment 1 and prior studies impossible. We now turn to Experiment 2, which looked at omission rather than oversuppliance.

Experiment 2: The role of definiteness in the L2 acquisition of English plural marking with count nouns

Experiment 2 examines how L1-Korean and L1-Mandarin Chinese L2-English learners treat the English plural marker *-s* with definite versus indefinite count nouns, in order to investigate whether there is L1-transfer of plural marking with definites from Korean to English, and to further test the predictions of the MCH.

Experiment 2: Test materials for the SPRT and GJT

While in Experiment 1, the error was incorrect presence of plural *-s* with mass nouns, in Experiment 2, the error was omission of *-s* with plural count nouns. Experiment 2 crossed the factor of definiteness (definite vs. indefinite plural noun)

Table 5. Experiment 2: Breakdown of target sentences by region, sample token set

	R1	R2	R3	R4	R5	R6 [CR]	R7 [CR+1]	R8	R9	R10
Indefinite, grammatical	The	child	was	watching	many	boats	floating	on	the	sea
Indefinite, ungrammatical	The	child	was	watching	many	boat	floating	on	the	sea
Definite, grammatical	The	child	was	watching	those	boats	floating	on	the	sea
Definite, ungrammatical	The	child	was	watching	those	boat	floating	on	the	sea

NOTE: CR, Critical region. CR+1, Critical region +1.

with the factor of grammaticality (grammatical plural form vs. ungrammatical singular form). Thirty-two target token sets were constructed and distributed across four lists using a Latin-square design, with eight tokens per condition per list.

The target count noun always occurred after an unambiguously plural quantifier, to ensure plural rather than singular interpretation. For indefinites, half the items contained *many* and half contained *several*, while for definites, half the items contained *these* and half contained *those*. Table 5 shows how target sentences break down by region in the SPRT. The critical region was r6 and the postcritical (spillover) region was r7. All target nouns were inanimate.

Experiment 2: Predictions

English NSs are expected to succeed to detect missing plural *-s* in both the SPRT and the GJT. In the GJT, they are expected to respond “Yes” to the grammatical sentences and “No” to the ungrammatical sentences. In the SPRT, they are expected to slow down in the ungrammatical conditions compared to the corresponding grammatical conditions. The slowdown may take place in the target region (r6) and/or on the spillover region (r7). In contrast, L2 learners may overaccept ungrammatical sentences in the GJT as well as fail to slow down for ungrammaticality in the SPRT. Under the predictions of the MCH in Table 2, and given the prior findings of Jiang (2004, 2007), it is expected that L1-Chinese L2-English learners will fail to detect errors of missing *-s* online, with either definites or indefinites. L1-Korean L2-English learners are predicted to either pattern with L1-Chinese L2-English learners, or else to successfully detect errors of missing *-s* with definites only; the latter finding would provide evidence that *-s* is congruent with *-tul* in definite contexts.

The MCH makes no predictions for the GJT. Both learner groups may succeed in the GJT, as the obligatoriness of plural marking with count nouns is subject to explicit classroom instruction; it is also possible that we would see L1-transfer from Korean here, manifested in greater accuracy with plural definites than with plural indefinites on the part of the Korean group.

Experiment 2: Results

GJT results

As shown in Figure 5, English NSs showed at-ceiling performance in all conditions. In contrast, the L2-groups were close to ceiling in the grammatical conditions, but

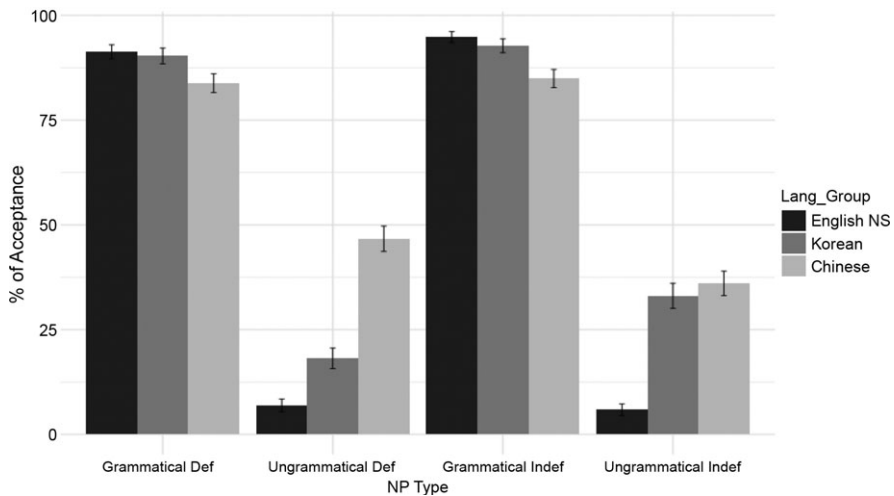


Figure 5. %Acceptance (YES response) to target items in the GJT in Experiment 2 (error bars indicate +/- standard deviation).

overaccepted sentences in the ungrammatical conditions. The Korean group exhibited much greater accuracy with definites than with indefinites, which was not the case for the Mandarin group.

In the analysis, we found a main effect of group (NS vs. L2; $z = 2.96$, $p = .003$), with higher acceptance rates by the L2 learners than by the NSs; the two L2-groups did not differ from each other. There was a significant effect of grammaticality ($z = 27.07$, $p < .0001$), with grammatical sentences eliciting significantly more acceptances than the ungrammatical sentences; there was also a marginal effect of definiteness ($z = -1.73$, $p = .084$), with slightly more acceptance in the indefinite than the definite conditions. There was a group (NS vs. L2) by grammaticality interaction ($z = -8.79$, $p < .0001$), due to overacceptance of ungrammatical sentences on the part of the L2 learners. For the Korean versus Chinese comparison, there were significant two-way and three-way interactions among group, definiteness, and grammaticality (for the three-way interaction, $z = -2.85$, $p = .0044$). Follow-up comparisons indicate that the acceptance rate of ungrammatical sentences in the indefinite ungrammatical condition was similar in two L2-groups, whereas in the definite condition, L1-Korean L2-English learners accepted ungrammatical sentences significantly less than L1 Mandarin L2-English learners. To sum up, while both L2-groups were less accurate than NSs, definiteness made a difference only for the Korean and not for the Mandarin group.

SPRT results

Figures 6, 7, and 8 show the average RTs for each group, respectively. The rectangular frames show the critical region and the spillover region on which we conducted statistical analyses.

In the analysis for the critical region, there was a main effect of group (NS vs. L2: $t = 6.28$, $p < .0001$), which indicates that the English NS group read significantly

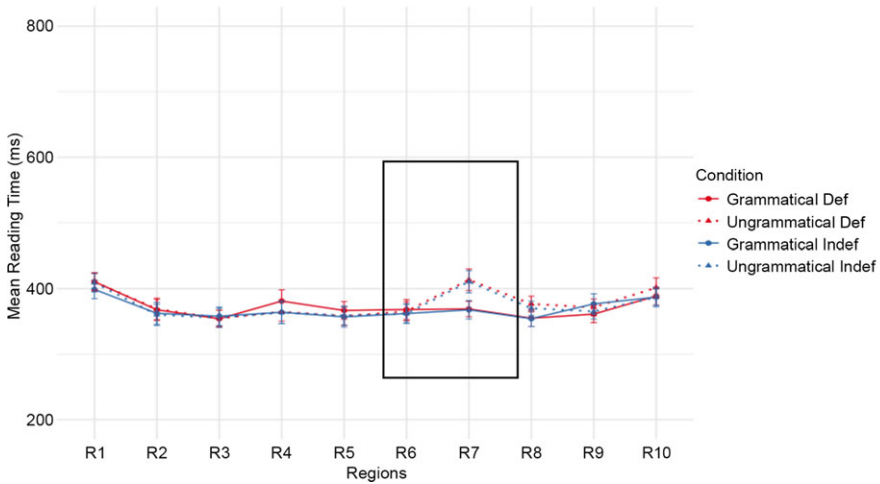


Figure 6. SPRT results, English NSs: average raw RTs, Experiment 2 (error bars show standard error).

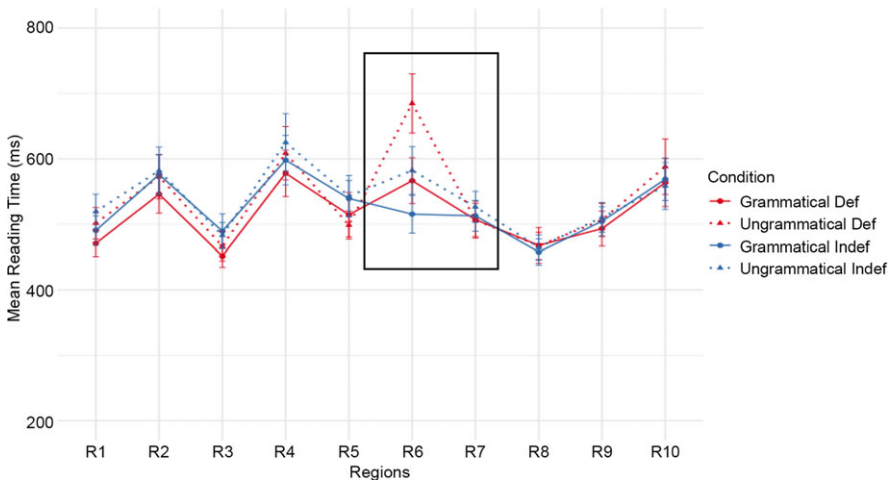


Figure 7. SPRT results, Korean L2ers: average raw RTs, Experiment 2 (error bars show standard error).

faster than L2 learners. There is also a significant effect of group (Korean vs. Chinese: $t = -1.99, p = .04$), due to the fact that the Mandarin group read faster than the Korean group. There was a significant effect of grammaticality ($t = -5.95, p < .0001$), with faster RTs in the grammatical conditions (because the target noun was shorter in the ungrammatical sentences). There was a significant effect of definiteness ($t = 2.91, p = .004$), due to the target noun in the definite sentences being read slower than in the indefinite sentences. We also found a significant interaction between group (NS vs. L2) and grammaticality ($t = -4.88, p < .0001$), due to

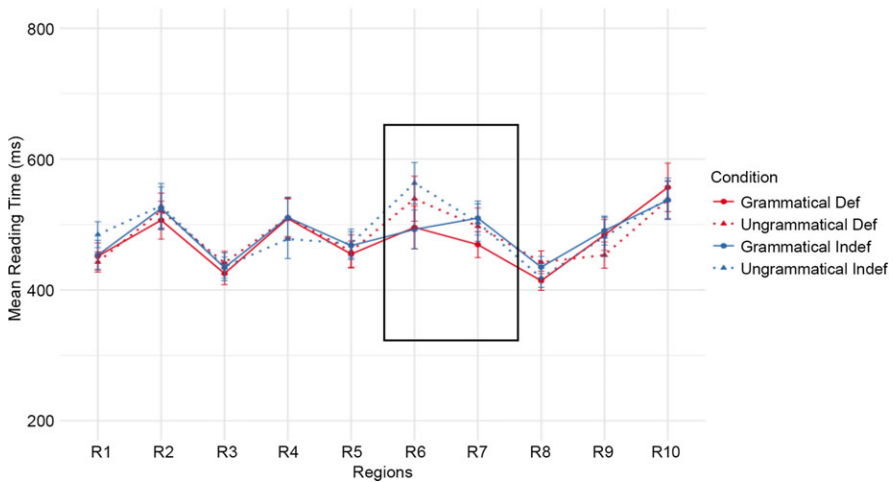


Figure 8. SPRT results, Mandarin L2ers: average raw RTs, Experiment 2 (error bars show standard error).

learners but not NSs slowing down for ungrammaticality in the critical region. Finally, we found a significant interaction between group (Korean vs. Chinese) and definiteness ($t = -4.59, p < .0001$), as well as a marginal three-way interaction among group (Korean vs. Chinese), definiteness, and grammaticality ($t = 1.82, p = .06$). Follow-up comparisons indicate that the Chinese group slowed down for ungrammaticality significantly in the indefinite conditions only, whereas the Korean group slowed down in both definite and indefinite conditions. From the interaction plots, we see that ungrammaticality affected both conditions in the Korean group, but had a greater effect with definites than with indefinites, whereas in the Chinese group, ungrammaticality affected indefinites more than definites.

In the spillover region (r7), there was a significant main effect of group (NS vs. L2; $t = 4.88, p < .0001$), which indicates that the English NSs read significantly faster than L2 learners. There was a main effect of grammaticality ($t = -3.36, p = .0008$), with faster RTs in the grammatical conditions. There was a significant interaction between group (NS vs. L2) and grammaticality ($t = 3.56, p = .0004$), due to the English NSs but not the L2 groups slowing down for ungrammaticality in the spillover region. There was also a marginal interaction between group (NS vs. L2) and definiteness ($t = -1.78, p = .07$), because English NSs read faster in the indefinite than the definite conditions, while the opposite was the case for the L2-groups. There were no interactions between group (Korean vs. Chinese) and any other factor in the spillover region.

In sum, we can see that English NSs showed no slowdown in the critical region, but slowed down for ungrammaticality in the spillover region in both definite and indefinite conditions. In contrast, the L2-groups showed significant slowdowns in the critical region rather than the spillover region. We speculate that the earlier slowdown of the L2-groups relative to the native speaker group is an indirect result of the overall longer RTs on the part of the learners (see Ionin, Choi, and Liu, 2019,

for similar findings in the online reading of articles). What concerns us more is the difference between the two L2-groups. The Chinese group slowed down for ungrammaticality in the indefinite condition only, whereas the Korean group did so for definites as well as indefinites.

Discussion: Experiment 2

Unlike Experiment 1, Experiment 2 found a difference between the two L2-groups in both the GJT and the SPRT. In the GJT, both L2 groups overaccepted sentences with missing plural marking to about the same extent with indefinites, whereas L1-Korean L2-English learners showed much better performance than the L1-Mandarin L2-English learners with definites. In the SPRT, the L1-Chinese L2-English learners slowed down for ungrammaticality only with indefinites, whereas the L1-Korean L2-English learners slowed down so with both definites and indefinites. As expected, English NSs detected ungrammaticality equally well with definites as with indefinites.

The group differences regarding definites are fully consistent with L1-transfer from the Korean *-tul*, which is obligatory with plural definites. We see transfer effects manifested both in the GJT, where the Korean learner group exhibited greater accuracy on definite plurals than the Chinese learner group, and in the SPRT, where a slowdown for ungrammaticality with definites was exhibited only by the Korean learner group. The behavior on definites in the SPRT is fully consistent with the MCH: due to the congruency between *-s* and *-tul* in definite contexts, the L1-Korean L2-English learners reacted to missing *-s* in plural definites; the L1-Chinese L2-English learners, in contrast, exhibited no slowdown for missing *-s* with plural definites.

The unexpected finding is that both learner groups slowed down significantly for missing *-s* on plural indefinites. This is not predicted by the MCH, and not consistent with L1-transfer, as neither Korean nor Chinese requires plural marking on indefinites: *-tul* is optional with Korean indefinites, whereas *-men* is disallowed on Chinese indefinites. We note that the unexpected results show up in the SPRT only: in the GJT, the Chinese group behaved the same with indefinites as with definites, whereas the Korean group had an advantage on definites, consistent with L1-transfer.

A possible reason for why both learner groups succeeded in slowing down for missing *-s* with indefinites in the SPRT was that the presence of a plural quantifier, *many* or *several*, provided a very strong cue to plurality; as noted by an anonymous reviewer, these words may have triggered the meaning of “more than one” in the learners, and caused the learners to attend more closely to plurality and to notice the missing *-s*. We note that the determiner was plural in the definite condition as well, *these/those*; however, it is possible that the plural demonstrative does not trigger the “more than one” meaning the way that *many/several* does.

We now consider how our findings compare to those of prior research. As discussed earlier, the studies by Jiang and colleagues found a complete lack of sensitivity to missing plural marking among L1-Chinese L2-English learners, whereas we found such sensitivity with plural indefinites for this population (though not with plural definites). The study that is most closely comparable to ours is Jiang (2007),

since this study also used an SPRT in which *-s* was missing on plural indefinite NPs. The L1-Chinese L2-English participants in Jiang (2007) were all studying at a US university at the time of testing, as were the learners in our study. It is not possible to compare the learner groups from the two studies directly with regard to proficiency, since Jiang used TOEFL scores to estimate proficiency, whereas we used an independent proficiency test.

The most likely reason for the different study results is the difference in stimuli construction. The stimuli in Jiang (2007), as in ours, placed the target plural NP in object position, and the stimuli sentences in the two studies were of comparable length. The main difference is the cue to plurality: whereas all of our indefinite plural items had the target noun preceded by *many* or *several*, Jiang's stimuli used the partitive construction of the form *X of the NP*, for example, *some of the rabbit(s)*. A great variety of quantifiers (and even a few nonquantificational expressions) were used in place of X: out of the 32 target stimuli used in Jiang (2007), listed in Appendix A, we counted 20 different types of partitives (each used once or twice). These were: *some/ several /a few /few /a large number /many /both /one /two /three /none /neither /all /any /a couple /every one /each /the last of /a lot of /more of the NP*. Note that while some of the lexical items (e.g., *many/several/two/three*) would be expected to trigger the "more than one" reading, others (e.g., *any/some/the last*) would not do so; and *one of* as well as *every one of* might have the opposite effect, triggering the singular reading for learners. It is possible that learners performed differently depending on the type of quantifier; it is also possible that partitive expressions are generally not as strong a cue to plurality as plural determiners. All of the above is of course speculation; the only way to know for certain whether the stimuli construction was responsible for the different study outcomes would be to run a single study that included both types of stimuli with the same group of participants.

General Discussion

We now revisit our three research questions from above. We focus on whether learners were able to make a distinction between grammatical and ungrammatical conditions, even while recognizing that they did not behave in a fully nativelike manner; the results are summarized in Table 6.

With regard to RQ1, learners make distinctions for all categories except atomic mass nouns (in both the GJT and the SPRT), and, in the case of the Chinese group, plural definites (in the SPRT). Thus, the learners do exhibit quite a bit of knowledge about English plural marking, not only offline but also online. With regard to RQ2, we find that learners overall do better in detecting omission of plural marking (Experiment 2) than in detecting oversuppliance (Experiment 1): the error of allowing plural *-s* with atomic mass nouns is the only one not detected even offline, in the GJT. It is not possible to compare directly between the experiments, but the results do suggest that detecting *-s* omission is somewhat easier for learners than detecting *-s* oversuppliance; this could be a result of L1-transfer (since both Korean and Chinese use plural marking much less than English does), or of count nouns being easier than mass nouns, consistent with the prior research (e.g., Hiki, 1990, 1991, Snape, 2008).

Table 6. Summary of findings, across experiments

Error type	L1-Chinese L2-English learners		L1-Korean L2-English learners	
	GJT	SPRT	GJT	SPRT
–s oversuppliance with nonatomic mass nouns (<i>sunlights</i>)	YES	YES	YES	YES
–s oversuppliance with atomic mass nouns (<i>furnitures</i>)	NO	NO	NO	NO
–s omission with plural indefinites (<i>many book</i>)	YES	YES	YES	YES
–s omission with plural definites (<i>these book</i>)	YES	NO	YES	YES

NOTE: YES, the learners made a distinction between grammatical and ungrammatical conditions in terms of acceptance rates (GJT) or RTs (SPRT). NO, the learners did not make a distinction between grammatical and ungrammatical conditions.

Finally, we consider RQ3, and the role of L1-transfer. The results of Experiment 1 show that even though Korean *-tul* is used (optionally) with inanimate atomic nouns, while Chinese *-men* is not, there is no evidence for L1-transfer; instead, both groups allow *-s* with atomic mass nouns in English, but not nonatomic ones. This provides convergent evidence for the proposal of Choi et al. (2018) concerning the role of atomicity as a semantic universal. In Experiment 2, in contrast, we found that the requirement for the Korean plural marker *-tul* in plural definite contexts facilitates L2 acquisition of plural marking in definite contexts in English; the advantage with definites for the Korean group relative to the Chinese group was manifested both offline and online.

As previously discussed, the results are only partially compatible with the predictions of the MCH (see Table 2). Focusing on the SPRT only, the MCH can successfully account for the finding that neither group detected oversupplied *-s* with atomic mass nouns in Experiment 1, as well as the finding that only the Korean group detected missing *-s* with plural definites in Experiment 2. The latter result suggests that *-s* and *-tul* are congruent for L1-Korean L2-English learners with regard to their obligatory use in definite contexts. At the same time, the MCH cannot explain why both groups successfully detected oversupplied *-s* with nonatomic mass nouns in Experiment 1, as well as missing *-s* with indefinites in Experiment 2. These findings indicate that L2-English learners can overcome morphological incongruence and activate the meaning of an L2 morpheme online despite L1/L2 differences.

We propose specific explanations for why learners can overcome incongruence for some uses of the English plural marker but not others. In the case of mass nouns, we follow the proposal of Choi et al.: once learners recognize that English has plural marking, they assume that the plural marker can be used with all atomic nouns; as a result, they are successful at detecting errors of *-s* oversuppliance with nonatomic mass nouns but not atomic ones. Our findings show that this success is manifested not only offline (as Choi et al. found) but online as well: when encountering a noun such as *sunlights*, learners automatically detect the conflict between the meaning of the noun (a nonatomic substance) and the presence of the plural marker. In contrast, they detect no conflict between the meaning of *furniture* and the presence of the plural marker, which is not surprising given that atomic nouns such as *furniture*

can be count cross-linguistically: there is nothing inherent to the meaning of the noun that prevents its combination with *-s*. Learners need to learn which atomic nouns disallow *-s* in English on a case-by-case basis: for example, *footwear* is mass but *shoes* is count, despite similar meanings.

The findings of Experiment 1 show that learners recognize that the English plural marker *can* be used with atomic nouns (but not nonatomic ones). Experiment 2 examined whether learners further recognize that the English plural marker *must* be used with atomic count nouns in order to express plural meanings. The findings here are twofold: with regard to plural definites, Korean but not Chinese learners of English detect errors of missing *-s*, consistent with L1-transfer; with plural indefinites, both groups are successful, and we have provisionally suggested that the plural quantifiers *many/several* provide the relevant cue.

In sum, we would like to make three points. First, a comparison between the Korean and Mandarin groups shows that speakers of different GC languages do not acquire English plural marking in the same way, but are subject to cross-linguistic influence for those contexts in which plural marking is obligatory in the L1. Second, morphological incongruency is not all-determining, and learners are able to acquire novel properties of the plural marker in their L2, based on such factors as semantic universals (atomicity) and cues in the input (plural quantifiers). Finally, third, while there are some differences in learners' performance on the online and offline tasks, the overall patterns of performance are remarkably similar: atomic nouns are more difficult than nonatomic across tasks, and, for the Chinese group, definites are more difficult than indefinites across tasks. This suggests that, when it comes to plurality, learners acquire the relevant knowledge at both more explicit and more implicit or automatized levels.

Conclusion and suggestions for further research

Our study findings suggest that L1-transfer affects only those contexts in which plural marking is obligatory in the L1 (definites in Korean), but not those where it is optional. The findings are only partially consistent with the MCH, and show that L2 learners can overcome the effects of morphological incongruence with regard to at least some aspects of English plural marking. The results also provide convergent evidence on the role of atomicity as a semantic universal in L2 acquisition.

A number of directions are open for future research. It would be fruitful to test speakers of a plural-marking language such as Russian or French on the tasks in this study. Jiang et al. (2011, 2017) found that L1-Russian L2-English learners performed like native English speakers with regard to their ability to detect errors of missing *-s* online; it would be fruitful to examine how this learner population would behave with regard to *-s* oversuppliance as well as to *-s* omission. In addition, as discussed above, future research should test the same group of L1-Chinese L2-English learners on different types of plural stimuli, the ones from our study as well as the ones from Jiang (2007). This would allow us to address the discrepancy in the findings between the two studies, and would also allow us to further test the role of cues to plurality.

Finally, as noted above, we found that learners are more accurate at detecting errors of missing plural marking with count nouns than at detecting errors of incorrect plural marking with mass nouns. There are two possible explanations for this:

that missing plural marking is easier to detect than incorrect plural marking, or that learners are better with count nouns than with mass nouns. Our study design does not allow us to tease apart these two explanations. In order to do so, one would need to keep the noun type (count vs. mass) constant, and focus on the error type (omission vs. misuse): for example, one could compare omission with plural count nouns (*many book*) to incorrect plural marking with singular count nouns (*a books*). If count nouns are generally relatively easy for L2 learners, we expect both error types to be detected equally well (and better than with *furnitures*). If the question is about omission versus oversuppliance, we might expect more success with *many book* than with *a books*.

To sum up, the present study lays the groundwork for continued investigation into L2 learners' ability to detect morphological errors of different types in their L2; while much prior work has looked at learners' online sensitivity to errors of omission, to the best of our knowledge this was the first study to consider errors of oversuppliance alongside errors of omission using a measure of real-time processing.

Notes

1. While there are some exceptions—count nouns used in bare form in certain contexts (Stvan, 2009)—they are not relevant for the purpose of our study.
2. There are other types of flexible nouns, which show variability in being count versus mass across languages (e.g., *spinach* is mass in English but count in French; *peas* is count in English but mass in Russian; see Inagaki, 2014; Inagaki & Barner, 2009). We do not discuss them in this paper.
3. Lardiere (2009), discussing the properties of Korean *-tul*, describes it as having a “specific” reading (p. 204); this seems to be based on the observations in Kwon and Zribi-Hertz (2004) that *-tul* is disallowed with readings that are clearly nonspecific, such as generic or narrow-scope. What matters for our purposes is that while *-tul* is allowed with definites and indefinites (in the latter case, restricted to specific indefinites only), it is always optional with indefinites. Lardiere (2009, pp. 200–201) cites examples from both Kim (2005) and Kwon and Zribi-Hertz (2004), which show that *-tul* is obligatory in plural definite contexts, and otherwise optional.
4. There is much debate about which tasks are more likely to tap into explicit versus into implicit knowledge, and in particular, about whether placing learners under time pressure results in learners accessing more implicit knowledge, as opposed to more automatized explicit knowledge (see, e.g., DeKeyser, 2003; Ellis, 2005; Godfroid et al., 2015; Vafaei, Suzuki, & Kachisnke, 2017). Here, we abstract away from this debate, but assume, following the work of Jiang and colleagues, that learners' ability to activate a morpheme online, quickly and automatically, is a mark of nativelike performance; whether this performance reflects implicit knowledge or automatized explicit knowledge is not directly relevant.
5. Four additional L1-Chinese L2-English learners were excluded from all of the data analyses, as follows. Two were excluded due to having very low comprehension question accuracy in the SPRT (less than 50%, while all other participants had accuracy above 75%), which suggests that they were not paying attention and/or did not have sufficient proficiency to complete the tasks. Two more were excluded due to scoring less than 60% correct on the proficiency test. Since all L1-Korean L2-English learners scored at least 60% on the proficiency test, this measure allowed us to have more closely proficiency-matched L2 groups.
6. The purpose of having fillers is to prevent participants from noticing the target structures. For more discussion, and for the rationale of balancing grammatical and ungrammatical sentences in the SPRT, please see Jegerski (2014).
7. We chose to code the GJT responses as YES versus NO, rather than as correct versus incorrect (recall that YES is correct for grammatical sentences but incorrect for ungrammatical sentences). This allows for a more direct comparison with the SPRT: in both tasks, we are looking at whether participants are able to distinguish between grammatical and ungrammatical sentences. In the GJT, distinguishing between the two means saying YES less often to the ungrammatical sentences; in the SPRT, distinguishing between the two means reading the target regions more slowly in the ungrammatical sentences.

8. We conducted additional analyses on the L2-groups only with proficiency score as a covariate. In the GJT, we found that there is no proficiency effect in experiment 1 ($p = .247$) or Experiment 2 ($p = .355$). In the SPRT, we found a significant effect of proficiency ($p = .0007$ in Experiment 1, $p = .015$ in Experiment 2) in the critical region but not in the postcritical region. The effects in the critical region were due to learners exhibiting shorter RT (i.e., reading faster) with more advanced proficiency. Additional models that included interaction terms between proficiency and other fixed effects did not find any significant interactions.
9. Preliminary results of experiment are also reported in Choi and Ionin (2017), a conference proceedings paper; the present paper compares performance on Experiment 1 to that in Experiment 2, whose results have not previously been published.
10. Of course, it is also possible that one or both groups would be successful with both noun types. This is very unlikely, given that the learners in Choi et al. (2018) were nontarget with atomic nouns in a highly explicit offline task.

References

- Baek, M. (2002). A study on Korean plural senses. *Discourse and Cognition*, 9, 59–78.
- Barner, D., & Snedeker, J. (2005). Quantity judgments and individuation: Evidence that mass nouns count. *Cognition*, 97, 41–66.
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68, 255–278.
- Bates, D. M., Mächler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. *Journal of Statistical Software*, 67, 47.
- Bunt, H. C. (1985). *Mass terms and model-theoretic semantics*. Cambridge: Cambridge University Press.
- Cheng, L. L. S., & Sybesma, R. (1998). Yi-wan tang, yi-ge tang: Classifiers and massifiers. *Tsing Hua Journal of Chinese Studies*, 28, 385–412.
- Cheng, L. L. S., & Sybesma, R. (1999). Bare and not-so-bare nouns and the structure of NP. *Linguistic inquiry*, 30, 509–542.
- Chierchia, G. (1998). Plurality of mass nouns and the notion of “semantic parameter.” In S. Rothstein (Ed.), *Events and grammar* (pp. 53–103). Dordrecht: Kluwer.
- Chierchia, G. (2010). Mass nouns, vagueness and semantic variation. *Synthese*, 174, 99–149.
- Chierchia, G. (2015). How universal is the mass/count distinction? Three grammars of counting. In A. Li, A. Simpson, & D. Tsai (Eds.), *Chinese syntax in a cross-linguistic perspective* (pp.147–175). Oxford: Oxford University Press.
- Choi, S., & Ionin, T. (2017). Acquisition and processing of mass nouns in L2-English by L2-learners from generalized classifier languages: Evidence for the role of atomicity. In J. Scott & D. Waughtal (Eds.), *Proceedings of the 41th annual Boston University Conference on Language Development*. Somerville, MA: Cascadilla Press.
- Choi, S. H., Ionin, T., & Zhu, Y. (2018). L1 Korean and L1 Mandarin L2 English learners’ acquisition of the count/mass distinction in English. *Second Language Research*, 34, 147–177.
- DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 313–348). Oxford: Blackwell.
- Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, 27, 141–172. doi: [10.1017/S0272263105050096](https://doi.org/10.1017/S0272263105050096)
- Foot, R. (2011). Integrated knowledge of agreement in early and late English–Spanish bilinguals. *Applied Psycholinguistics*, 32, 187–220. doi: [10.1017/S0142716410000342](https://doi.org/10.1017/S0142716410000342)
- Franceschina, F. (2001). Morphological or syntactic deficits in near-native speakers? An assessment of some current proposals. *Second Language Research*, 17, 213–247.
- Franceschina, F. (2005). *Fossilized second language grammars: The acquisition of grammatical gender* (Vol. 38). London: Benjamins.
- Gelman, A. (2008). Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine*, 27, 2865–2873.
- Godfroid, A., Loewen, S., Jung, S., Park, J. H., Gass, S., & Ellis, R. (2015). Timed and untimed grammaticality judgments measure distinct types of knowledge: Evidence from eye-movement patterns. *Studies in Second Language Acquisition*, 37, 269–297.

- Hawkins, R., & Liszka, S. (2003). Locating the source of defective past tense marking in advanced L2 English speakers. *Language Acquisition and Language Disorders*, *30*, 21–44.
- Hiki, M. (1990). The judgment of noun countability by Japanese college students: Where is the difficulty? *JACET Bulletin*, *21*, 19–38.
- Hiki, M. (1991). *A study of learners' judgment on noun countability*. Unpublished PhD thesis, Indiana University.
- Hofmeister, P. (2010). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes*, *26*, 376–405.
- Hua, D., & Lee, T. H. T. (2005). Chinese ESL learners' understanding of the English count-mass distinction. In L. Dekydtspotte, R. A. Sprouse, & A. Liljestrang (Eds.), *Proceedings of the 7th Generative Approaches to Second Language Acquisition Conference* (pp. 138–149). Somerville, MA: Cascadilla Proceedings Project.
- Hwang, S., & Lardiere, D. (2013). Plural-marking in L2 Korean: A feature-based approach. *Second Language Research*, *29*, 57–86.
- Iljic, R. (1994). Quantification in Mandarin Chinese: Two markers of plurality. *Linguistics*, *32*, 91–116.
- Inagaki, S. (2014). Syntax-semantics mappings as a source of difficulty in Japanese speakers' acquisition of the mass-count distinction in English. *Bilingualism: Language and Cognition*, *17*, 464–477.
- Inagaki, S., & Barner, D. (2009). **Countability in absence of count syntax: Evidence from Japanese quantity judgments.** In S. Inagaki et al. (Eds.), *The 8th annual conference of the Japanese Society for Language Sciences* (pp. 111–125). Tokyo: Kurozio.
- Inonin, T., Choi, S. H., & Liu, Q. (2019). Knowledge of indefinite articles in L2-English: Online vs. offline performance. *Second Language Research*. Advance online publication.
- Inonin, T., & Montrul, S. (2010). The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning*, *60*, 877–925.
- Jackendoff, R. (1991). Parts and boundaries. *Cognition*, *41*, 9–45.
- Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.
- Jegerski, J. (2014). Self-paced reading. In J. Jegerski & B. Van Pattern (Eds.), *Research methods in second language psycholinguistics* (pp. 20–49). New York: Routledge.
- Jiang, L. J. (2017). Mandarin associative plural-men and NPs with-men. *International Journal of Chinese Linguistics*, *4*, 191–256.
- Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634.
- Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, *57*, 1–33.
- Jiang, N., Hu, G., Chrabaszcz, A., & Ye, L. (2017). The activation of grammaticalized meaning in L2 processing: Toward an explanation of the morphological congruency effect. *International Journal of Bilingualism*, *21*, 81–98.
- Jiang, N., Novokshanova, E., Masuda, K., & Wang, X. (2011). Morphological congruency and the acquisition of L2 morphemes. *Language Learning*, *61*, 940–967.
- Kim, C. (2005). *The Korean plural marker tul and its implications*. Unpublished PhD dissertation, University of Delaware.
- Kwon, S.-N., & Zribi-Hertz, A. (2004). Number from a syntactic perspective: Why plural-marking looks “truer” in French than in Korean. *Empirical Issues in Formal Syntax and Semantics*, *5*, 133–58.
- Lago, S., Shalom, D. E., Sigman, M., Lau, E. F., & Phillips, C. (2015). Agreement attraction in Spanish comprehension. *Journal of Memory and Language*, *82*, 133–149.
- Landman, F. (1989). Groups. I. *Linguistics and Philosophy*, *12*, 559–605.
- Landman, F. (1991). *Structures for Semantics, SLAP* (Vol. 45). Kluwer: Dordrecht.
- Langacker, R. W. (1999). *Grammar and conceptualization*. Berlin: de Gruyter.
- Langacker, R. W. (2008). *Cognitive grammar: A basic introduction*. Oxford Scholarship Online. New York: Oxford University Press.
- Lardiere, D. (2008). Feature assembly in second language acquisition. In J. Liceras, H. Zobl, & Helen Goodluck (Eds.), *The role of formal features in second language acquisition* (pp. 106–140). Mahwah NJ: Erlbaum.
- Lardiere, D. (2009). Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research*, *25*, 173–227.

- Lenth, R.** (2018). Emmeans: Estimated marginal means, aka least-squares means. *R Package, Version 1(2)*. <http://github.com/rvleenth/emmeans>
- Li, C. N., & Thompson, S. A.** (1981). *Mandarin Chinese: A functional reference grammar*. Berkeley, CA: University of California Press.
- Li, Y.** (1999). Plurality in a classifier language. *Journal of East Asian Linguistics*, **8**, 75–99.
- MacDonald, D., & Carroll, S. E.** (2018). Second-language processing of English mass-count nouns by native-speakers of Korean. *Glossa: A Journal of General Linguistics*, **3**, 46.
- O’Neill, R., Cornelius, E., & Washburn, G.** (1991). *American kernal lessons: Advanced students’ book*. New York: Longman.
- Prévost, P., & White, L.** (2000). Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second Language Research*, **16**, 103–133.
- R Core Team.** (2014). R: A language and environment for statistical computing Retrieved from <http://www.r-project.org/>
- Schwartz, B. D., & Sprouse, R. A.** (1994). Word order and nominative case in non-native language acquisition: A longitudinal study of (L1 Turkish) German interlanguage. In T. Hoekstra & B. D. Schwartz (Eds.), *Language acquisition studies in generative grammar: Papers in honor of Kenneth Wexler from the 1991 GLOW workshops* (pp. 317–368). Amsterdam: Benjamins.
- Schwartz, B. D., & Sprouse, R. A.** (1996). L2 cognitive states and the full transfer/full access model. *Second Language Research*, **12**, 40–72.
- Slabakova, R.** (2008). *Meaning in the second language*. Berlin: de Gruyter.
- Slabakova, R.** (2016). *Second language acquisition*. Oxford: Oxford University Press.
- Snape, N.** (2008). Resetting the nominal mapping parameter in L2 English: Definite article use and the count–mass distinction. *Bilingualism: Language and Cognition*, **11**, 63–79.
- Stvan, L. S.** (2009). Semantic incorporation as an account for some bare singular count noun uses in English. *Lingua*, **119**, 314–333.
- Tang, W., Fiorentino, R., & Gabriele, A.** (2019). *Examining transfer in the acquisition of the count/mass distinction in L2 English*. Paper presented at the 15th meeting of the Generative Approach to Second Language Acquisition Conference, Reno, NV.
- Vafae, P., Suzuki, Y., & Kachisnke, I.** (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, **39**, 59–95.
- Wei, H., Boland, J. E., Brennan, J., Yuan, F., Wang, M., & Zhang, C.** (2018). Lexicalized structural priming in second language online sentence comprehension. *Second Language Research*, **34**, 395–416.