L2 English Generics: Japanese Child Returnees’ Incomplete Acquisition or Attrition?

Neal Snape¹, Makiko Hirakawa², Yahiro Hirakawa³, Hironobu Hosoi¹, and John Matthews⁴
¹Gunma Prefectural Women’s University, ²Bunkyo University, ³Tokyo Institute of Technology, and ⁴Chuo University

1. Introduction

Recent proposals in second language (L2) acquisition have referred to interface phenomena; specific linguistic properties that are particularly vulnerable in adult L2 acquisition (e.g., Gabriele, 2009; Lardiere, 2009; Slabakova, 2008; Sorace, 2005).¹ The Interface Hypothesis, for example, posits a distinction between properties that are internal to the grammar (syntax, semantics, phonology) and properties that require the integration of grammatical knowledge with discourse constraints or world knowledge, domains considered external to the grammar (e.g., Belletti, Bennati, & Sorace, 2007; Sorace, 2005; Sorace & Filiaci, 2006; Tsimpli & Sorace, 2006). Sorace, Serratrice, Filiaci, and Baldo (2009) showed that bilingual children who were acquiring Spanish and Italian, two languages that are similar with respect to pro nominal forms, had difficulty with properties related to the syntax-discourse interface despite the overlap between the languages. Other studies have focused on the semantics-pragmatics interface (Slabakova, 2011), and the phonology-morphology interface (Goad & White, 2004, 2008). Slabakova (2011) investigated sensitivity to scalar implicatures in English by Korean L2 learners, who were instructed to complete a grammaticality judgment task. Expressions like some and all form an implicational scale (Horn, 1972), resulting in scalar implicatures, with the stronger term, all, implying the weaker term, some, but not vice versa. Although some could logically mean some and possibly all, it is not usually interpreted in that way. Rather, if the weaker term is used, the stronger one is assumed not to hold. For example, if a speaker utters a sentence like (1a), the hearer infers that (1b) was not intended; rather, (1c) is the intended meaning.

(1) a. Some elephants have trunks.
   b. All elephants have trunks.
   c. Some but not all elephants have trunks.

Slabakova (2011) found differences between Korean L2 learners’ performance in their first language (L1) (Korean) and their L2 (English), as well as differences between their L2 English performance and that of child and adult English speakers: the L2 learners were more likely to give responses indicating that the implicature had been computed than the native speaker groups were. In other words, they rejected sentences like (1a) about 60% of the time, in contrast to the other groups who accepted these about 65% of the time. The results suggest that L2 learners do not have problems computing implicatures as such, since their favored response was the one that showed that the

¹ See White (2011) and Rothman and Slabakova (2011) for research in L2 acquisition relating to interfaces.
implicature had been calculated. Slabakova argues that the difference found between native speakers and L2 learners reflects the fact that, due to diminished processing resources, L2 learners are less able to conjure up logical alternatives and so stick with the original implicature (some implies not all, so sentences like (1a) are false). Goad and White (2004) investigated the role of prosodic transfer in L2 acquisition. They found that a Turkish-speaking endstate L2 learner of English continued to omit functional morphology (past tense -ed, articles the and a) even after years of exposure to the target language; in other words, the L2 learner had reached a point where acquisition does not progress – a state of fossilization in the interlanguage grammar. Part of the explanation for a fossilized state in this Turkish-speaker’s English is the prosodic differences between Turkish and English. Goad and White claim that the interface between syntactic and prosodic representations leads to non-target outputs in production (resulting in omission or stressing of functional morphology).

The present study focuses on the syntax-semantics interface, in particular, the mapping of semantic features to appropriate forms in the L2. Studies that have investigated the syntax-semantics interface in L2 acquisition have focused on a range of properties. One example of this interface phenomenon is articles in English: L2 learners who come from article-less L1s are typically less successful in acquiring articles in the L2 compared to those L2 learners from L1s with an article system, as they are unable to map the semantic features ±definite to the morphological markers of definiteness, the and a. (Ionin, Ko, & Wexler, 2004; Snape, 2008). Another area of interest, in relation to the acquisition of articles is genericity (Ionin, Montrul, 2010; Ionin, Montrul, Kim, & Philippov, 2011; Kupisch, 2012; Kupisch & Pierantozzi, 2010; Montrul & Ionin, 2010; Serratrice, Sorace, Filiaci, & Baldo, 2009). One might expect that since L2 learners have difficulties at the syntax-semantics interface with definiteness, they will also experience problems with genericity. In the present study we examine the acquisition/attrition of English generic reference by Japanese child returnees.

The next section provides details about English and Japanese generic NPs. The differences between NP-level genericity and sentence-level genericity in English are outlined according to Krifka, Pelletier, Carlson, ter Meulen, Link, and Chierchia’s (1995) account.

2. Generics in English and Japanese

There are two types of generic NPs in English. Examples of each are provided in (2):

(2) a. The dinosaur is extinct.
   b. #A dinosaur is extinct.
   c. #The dinosaurs are extinct.
   d. Dinosaurs are extinct.

The two NP types in (2a) and (2d) can refer to natural kinds while indefinite singulars (see example 2b) cannot. A natural kind represents a single abstract individual such as an animal or plant species by name or by a singular definite description along with a kind predicate such as be extinct. The definite singular has a [+species] feature that defines a natural kind, according to Vergnaud and Zubizarreta (1992). The definite description for kinds is otherwise known as a generic at the NP-level (Krifka et al., 1995). The definite plural in (2c) can only refer to a set of dinosaurs within the species, not the species as a whole.

(3) a. #The dog is man’s best friend.
   b. A dog is man’s best friend.
   c. #The dogs are man’s best friend.
   d. Dogs are man’s best friend.

Indefinite singulars are compatible only with characterizing sentences: they can be used to make a general statement about members of the kind dog as in example (3b), unlike definite singulars in (2a). Example (3a) is a combination of a generic (the dog) in a characterizing sentence, but unlike (2a), there is no kind predicate that allows the dog to refer to the entire species. Thus, given the choice between (3a) and (3b), (3b) is generally preferred for providing a description about someone or something as the generic force of characterizing sentences comes from an operator at the sentence-level, not from the subject NP. The definite plural in (3c), like in (2c), can only refer to a set of dogs which are man’s best
friend, not dogs in general. Bare plurals are not restricted to NP-level or sentence-level genericity (see examples 2d and 3d). Bare plurals function much in the same way as the definite singular in (2a) and the indefinite singular in (3b) (Carlson, 1977; Chierchia, 1998).

There is no article system present in Japanese. According to Kuroda (1992) and Lyons (1999), definiteness and genericity are expressed through the use of the nominative marker ga and the topic marker wa. The standard way to refer to generics at the NP-level and generics at the sentence-level in Japanese is with a bare NP followed by topic marker wa, as illustrated in the examples (4) and (5).

**NP-level generics**

(4)  

a. #Sono kyoryu-wa zetsumetsushi-ta.  
   that dinosaur-TOP die out-PAST TENSE  
   ‘That dinosaur died out’.

b. #Aru kyoryu-wa zetsumetsushi-ta.  
   a certain type of dinosaur-TOP die out-PAST TENSE  
   ‘A certain type of dinosaur died out’.

c. Kyoryu-wa zetsumetsushi-ta.  
   dinosaur(s)-TOP die out-PAST TENSE  
   ‘The dinosaur died out. / Dinosaurs died out’.

**Sentence-level generics**

(5)  

a. #Sono jyagaimo bitamin C to amino-san-o fukunde i-ru.  
   that potato vitamin C and amino acids-ACC contain-ASP-NON-PAST  
   ‘The potato contains vitamin C and amino acids’.

b. #Aru jyagaimo bitamin C to amino-san-o fukunde i-ru.  
   a certain type of potato vitamin C and amino acids-ACC contain-ASP-NON-PAST  
   ‘A certain type of potato contains vitamin C and amino acids’.

c. Jyagaimo-wa bitamin C to amino-san-o fukunde i-ru.  
   potato(s)-TOP vitamin C and amino acids-ACC contain-ASP-NON-PAST  
   ‘The potato contains vitamin C and amino acids’. / ‘Potatoes contain vitamin C and amino acids’.

The bare noun and use of topic marker wa provides a generic interpretation to a sentence and the nominative marker ga can only be interpreted as referring to someone or something specific (Kuroda, 1992). Hence, the examples (4c) and (5c) appear with the topic marker wa. The deictic demonstrative sono (that) and aru (a certain type) appear in examples (4) and (5), but they can only be used to express a reading where the speaker refers to a certain set of dinosaurs or potatoes.

To summarize, both English and Japanese can express genericity via different ways. English has an article system and bare plurals; in contrast, Japanese is an article-less language which employs topic marker wa to provide a generic interpretation. For Japanese L2 learners of English, it means they must acquire the syntactic functional category Determiner Phrase (DP). Lardiere (2009) argues that even if a language does not have articles, like Japanese, there are demonstratives and other nominal expressions that are interpreted as definite or generic. In the case for L2 learners of English, the Japanese speakers must reassemble features made available to them via the L1 grammar under the Feature Reassembly Hypothesis (Lardiere, 2009). For example, the [+definite] feature must be mapped onto the appropriate morphology in the L2, i.e., the definite singular. In addition, the definite singular for NP-level generics has a [+species] feature which learners must map onto the definite singular. The Feature Reassembly Hypothesis is consistent with the Full Transfer / Full Access model (Schwartz & Sprouse, 1996). However, hypotheses such as the Representational Deficit Hypothesis (Hawkins, 2005) state that adult L2 learners who are post-critical period will transfer properties from the L1 but be unable to acquire any new syntactic category, such as DP. The returnees in our study started acquiring English within the critical period. We predict, following Lardiere’s (2009) Feature Reassembly Hypothesis, the returnees have successfully acquired DP and show evidence that the [+species] feature has been mapped onto the
definite singular. They will further demonstrate that they understand how the indefinite singular can only refer to descriptions, not kinds. The next section gives an overview of recent adult L2, bilingual, and L2 attrition studies.

3. Adult L2 Learners, Bilinguals, and Child L2 Acquisition

Ionin and Montrul (2010) examined the role of L1 transfer in the interpretation of definite plurals by Spanish L2 learners of English and Korean L2 learners of English. As English only allows bare plurals in subject position to have a generic interpretation, Ionin and Montrul predicted that the Spanish speakers would accept definite plurals more than the Korean speakers, since in Spanish, definite plurals are grammatical and bare plurals are ungrammatical. The results of a truth-value judgment task revealed that both Spanish and Korean learners who were both advanced and had received more immersion in the target language were highly accurate in their judgments of definite plurals as they were able to reject them and accept bare plurals. Montrul and Ionin (2010) investigated the role of transfer from the stronger language by focusing on the interpretation of definite articles in Spanish and English by Spanish heritage speakers (i.e., minority language-speaking adult bilinguals). They found that Spanish heritage speakers face problems using articles with generic subject noun phrases due to L1 transfer effects. Kupisch and Pierantozzi (2010) tested 6–10-year-old German–Italian bilingual children in both German and Italian. They designed 8 stories to investigate the interpretation of definite plurals. They found that the bilingual children gave more generic interpretations in Italian than in German and the authors argue that their findings can be interpreted in favour of language separation. Serratrice et al. (2009) focused on bilingual children’s acquisition of specificity and genericity. They used an acceptability judgment task to investigate plural NPs in specific and generic subject nominations. Their study tested whether simultaneous English–Italian bilinguals (aged 6;2-10;10) accept ungrammatical Italian sentences with a bare NP. They found that the bilingual children based in the UK were significantly more likely than monolingual children to accept ungrammatical bare nouns in a generic context. All English-Italian bilinguals performed worse in the generic condition than in the specific condition. To date, there have been relatively few studies that have investigated child L2 learners, with the exception of Zdorenko and Paradis (2008) and Kolb (2013). Zdorenko and Paradis (2008) examined the role of definiteness and specificity in child L2 learners’ article choices. The children in their longitudinal study came from a range of different L1 backgrounds, including Chinese, Korean, and Japanese (article-less languages) and Spanish, Arabic, and Romanian (article languages). They found that all children substituted the for a in indefinite specific contexts regardless of L1 background and all children were more accurate with the use of the in definite contexts than with a in indefinite contexts, regardless of L1 background. Generic uses of articles were not included, as the oral narrative task they used was not designed to elicit generics. Kolb (2013) administered a truth-value judgement task to L1 English speaking children acquiring L2 French. There was a short context provided to each participant that involved a puppet adding something extra to the story. For example, the puppet continued the story by saying it was funny and finally supplying the target sentence ‘Les requins sont dangereux’ (The sharks are dangerous). The participants accepted the definite plural as generic (recall English definite plurals are definite; only bare NPs allow a generic interpretation in English). Kolb found that cross-linguistic influence is less likely if there is an increase in the length of exposure for acquiring the generic reading. All the L2 learners with more than 5.8 years of exposure accepted definite plurals (in subject position) in French as generic.

4. L2 Attrition

Up until this point we have focused on the acquisition of articles. However, the opposite of acquisition is attrition or language loss, in this case, loss of the L2. Some of the first studies (Cole, 1929, for French) in L2 attrition investigated the question of loss of foreign language skills by high school and college students during summer vacations. More recent studies have focused on a range of phenomena from frequency of input, loss of morphological complexity, lexical and morphosyntactic influence from the dominant language, and a reduction in registers of use (see Bardovi-Harlig & Stringer, 2010 for an overview). Tomiyama (2009) conducted a study of two Japanese siblings, focusing on the attrition of their second language, English. The older sibling (Eugene) was 5;8 and the younger one (Lily) was 2;8. They were both born in Japan to Japanese native speaker parents: their L1
is Japanese and their L2 is English. They lived and attended a school in the U.S. for four years. Japanese was spoken at home during the period in the U.S. except for the time when they were in the company of monolingual English speakers. The issue of which language is the dominant language and whether or not there is L2 maintenance after the children have left the L2 environment are two important factors in avoiding L2 attrition. Upon the children’s return to Japan, Tomiyama (2009) examined the two children’s L2 grammatical complexity, grammatical accuracy, lexical complexity, and lexical productivity by administering a storytelling task. The task is in the form of a book which is part of Mayer and Mayer’s (1971) frog series. Attrition researchers commonly adopt it to elicit narratives. It has 28 detailed pictures but no written text. The participants were given a few minutes each to look over the pictures before they started to narrate a story based on the pictures. The task was not timed. Data was collected over a period of 31 months in total. Tomiyama found that overall both children showed similar attrition patterns. Grammatical complexity was well maintained as they both used main clauses with subordinate clauses, e.g., He seemed to get confused because all of a sudden the turtle is back alive. This is counted as one T-unit with two clauses (i.e., he seemed to get confused; because all of a sudden the turtle is back alive). However, Tomiyama discovered some attrition in the area of lexical complexity which is a ratio of the number of word types to the square root of two times the word tokens. Lexical productivity was well maintained by both children. The lexical productivity measure is obtained by dividing the total number of word tokens by the total number of clauses. For grammatical accuracy, Eugene managed to keep the same level throughout the study but Lily’s grammatical accuracy peaked in session 6 when her grammatical complexity was at its minimum. Tomiyama describes Lily’s performance in grammatical accuracy as a trade-off with grammatical complexity as Lily was only able to maintain one, grammatical complexity, and not the other, grammatical accuracy. Tomiyama’s findings show the importance of gaining a high proficiency level in the L2 before returning to the home country, as it seems that Eugene and Lily have passed the critical threshold level: “beyond that threshold, they become immune to interference or decay” (Neisser, 1984, p. 33).

To our knowledge, there has been little attention paid to specific areas of grammar in previous L2 attrition studies of Japanese returnees. Tomiyama (1999) looked at Japanese child returnees’ L2 attrition, but in a qualitative study. No focus was on grammar per se. Yoshitomi (1999) examined article use by four child returnees, but this was only in spoken production. In session one, the older returnee only produced 63.6% of definite articles and 25% of indefinite articles, which could mean that the returnee has incomplete acquisition (Montrul, 2008) of the article system. Yoshitomi (1999) found that two returnees who spent a longer time abroad were less accurate with articles compared with the two who had spent less time abroad. Yoshitomi (1999) looked at possible L2 attrition of grammar in all four participants and found no evidence of attrition. However, crucially, it is an open question as to whether the L2 grammar was fully acquired. It is further unclear what the role of L1 transfer effects is for returnees, as the returnees may have “a special kind of interlanguage, manifesting English usage which differs from that of both native speakers and foreign learners, as a special combination of native-like and of distinctly non-native patterns” (Berman & Olshtain, 1983, p. 223). The returnees may display knowledge of generics that is different from native speakers but also different from typical Japanese second language learners.

5. Our Study

This is the first longitudinal study we aware of that looks exclusively at child returnees knowledge of generics in L2 English. Returnees, like Eugene and Lily, are those children who leave their home country at a young age and return to home after a prolonged sojourn abroad. Japanese children typically go abroad because of a parent job transfer (Kanno, 2003). Little is known about specific areas of grammar, such as the acquisition of complex semantic functions of articles, genericity being one of the functions.
5.1. Research Questions

Our research question is as follows:

RQ 1: Are the child returnees able to make a distinction between two types of genericity (natural kinds and characterizing sentences) and recognize that this distinction is morphologically encoded by English articles?

5.2. Participants

4 child returnees and 35 adult native speaker controls took part in our study. Table 1 provides details about the participants.

Table 1
Details of the Four Child Returnees and Native Controls

<table>
<thead>
<tr>
<th>Age at testing (years)</th>
<th>English proficiency level</th>
<th>Age of first exposure to English (years)</th>
<th>Age at first testing</th>
<th>No. of years living in the U.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CS</td>
<td>TOEIC = 905</td>
<td>19</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>YS</td>
<td>mean = 15.00 (range = 11-19)</td>
<td>TOEIC = 785 ~ 990</td>
<td>15</td>
<td>8</td>
</tr>
<tr>
<td>KS</td>
<td>mean = 7.25 (range = 3-12)</td>
<td>TOEIC = 785 ~ 990</td>
<td>11</td>
<td>8</td>
</tr>
<tr>
<td>TA</td>
<td>TOEIC = 600</td>
<td>15</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Native English speakers (n=35)</td>
<td>mean = 21.70 (range = 19-24)</td>
<td>n/a</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The three sisters, CS, YS, and KS, are classified as sequential bilinguals as they started to acquire English after the age of 3. TA is classified as a simultaneous bilingual. The definition of a simultaneous bilingual is a child who acquires two languages at the same time, at least before 3 years of age (Montrul, 2008). A sequential bilingual child, on the other hand, is someone who begins acquiring the second language after the first language has been acquired, otherwise termed as a child L2 learner (Meisel, 2011). Three of the four child returnees in the current study started acquiring the second language after their first language, Japanese, had been acquired, which makes them sequential bilinguals. The fourth returnee started to acquire the second language at the age of 3 by attending an English-speaking kindergarten, making her a simultaneous bilingual. We first met with the four returnees two months after their return from the U.S. to Japan (Time 1). The second time we met with them was six months later (Time 2). The background questionnaires revealed that all four child returnees, CS, YS, KS, and TA, have high levels of proficiency in English, according to the Test of International Communication for English (TOEIC). The maximum one can achieve on TOEIC is a score of 990. At the time of testing KS was only 11 years old and had not taken the TOEIC test but she had taken EIKEN which is a test for practical English proficiency. For each session (Time 1 and Time 2) we administered C-tests; C-tests are a standard way of testing participants in L1 and L2 attrition studies. The results of the C-tests show that each participant has a slightly higher score in Time 1.

2 We acknowledge that children rather than adults would be more directly comparable.
3 YS and KS took a test equivalent to TOEIC called EIKEN, a test for practical English proficiency, and their scores were within the high TOEIC range.
4 The C-tests were created by M. Schmid and are freely available online at http://www.let.rug.nl/languageattrition/
compared with Time 2. TA scored the lowest on the C-tests for both testing sessions whilst CS scored the highest. YS and KS scores are in-between CS and TA. First exposure to English for CS began in Japan at senior high school from age 11. YS and TA were first exposed to English when they started attending high school in the U.S. and for KS it was when she entered kindergarten in the U.S. On average, the four returnees received a high amount of exposure to English (roughly 30 hours a week) outside of their homes as they frequently socialized with friends. The parents of all four returnees only spoke Japanese at home to the children during their stay in the U.S. The main reason for only speaking Japanese is because the parents are Japanese and that is their first language. The other reason for maintaining a ‘one language at home policy’ was because the parents were concerned that once they returned to Japan their children would fall behind in their literacy skills, especially their knowledge of the written systems in Japanese, hiragana and kanji characters. Another way to bolster their level of kanji was to send their children once a week (usually on weekends) to a Japanese school in the U.S.

5.3. Task

The written acceptability judgment task consisted for 40 items, 20 control items, and 20 test items. Each item featured a short story context and five possible continuation sentences. The task was administered as a paper-based task for the child returnees and as a web-based survey tool for the native control group. The task includes control items, which are singular and plural anaphoric contexts (non-generics), and the test categories, which include NP-level generics (well-defined kinds) and sentence-level generics (general descriptions). All the continuation sentences were designed to solely focus on generics in subject position. None of the continuation sentences had generics in object position. Example test items from our task, along with the correct and incorrect answers, are provided in (6) to (9): The participants were asked to rate each of the five sentences on a scale from 1 to 4 in relation to the preceding context: 1 means the continuation sentence is unacceptable and 4 means it is an acceptable continuation. The instructions clearly demonstrated that one or more sentences could receive the same rating, as they were not being asked to rank their responses. The control items were designed to ascertain whether learners had acquired certain functions of English article use and the anaphoric use of the definite article (second mention of a referent) is one such function that L2 learners typically acquire from early on. The control items were in singular and plural contexts as this could potentially show whether or not learners could distinguish between singular and plural NPs. An example of a singular anaphoric context is in (6). Only one sentence is an acceptable continuation of the preceding context as it is the t-shirt with a unique or unusual property of being made of plastic and the color red within the set of the three t-shirts. The expected target sentence (6a) for native speakers is shaded as 4 and the other four sentences are shaded as 1. This is the only acceptable continuation sentence of the preceding context as it is anaphoric singular. No shading was given in the actual test instrument apart from an example at the beginning of the task. The sentences from a – e in the task were all randomized for each short context to avoid any ordering effects and two versions of the task were produced: The first half of version 1 was used as the last half for version 2 and vice versa. All the sentences from (6a – d) are grammatical sentences except for the bare singular in (6e) as singular count nouns require an article.

(6) Control category: Anaphoric singular

Takako loves shopping. Yesterday she bought three t-shirts: One is red and two are yellow. However, they are different from regular t-shirts. For example ….

| a. | the red t-shirt is plastic. | 1 | 2 | 3 | 4 |
| b. | the red t-shirts are plastic. | 1 | 2 | 3 | 4 |
| c. | red t-shirts are plastic. | 1 | 2 | 3 | 4 |
| d. | a red t-shirt is plastic. | 1 | 2 | 3 | 4 |
| e. | red t-shirt is plastic | 1 | 2 | 3 | 4 |

An example of a plural anaphoric context is in (7). As in the singular anaphoric context above, there is a set of suits and out of the set of three, two are distinct as they are grey (the other one is black) and they are tailor made. The sentences from b – e are not acceptable as continuations of the preceding context.
(7) Control category: Anaphoric plural

Yahiro has three suits: one black suit and two grey suits. He takes great care of his
grey suits because they are unusual. You see ….  

a. the grey suits are tailor made. 1 2 3 4
b. the grey suit is tailor made. 1 2 3 4
c. grey suits are tailor made. 1 2 3 4
d. a grey suit is tailor made. 1 2 3 4
e. grey suit is tailor made. 1 2 3 4

The test categories (NP-level and sentence-level genericity) are set up in the same way as the
control categories. The only difference between the two categories is that there are two possible target
continuations for the test categories. In (8) the definite singular and the bare plural are acceptable for
NP-level genericity so they are both rated as 4. All of the NPs in the test category referred to a well-
deﬁned kind such as a species of animal. The contexts for the test categories are very different to the
control (anaphoric) contexts as there is no previous referent in the discourse and the definite singular
and bare plural can only be interpreted as referring to a well-deﬁned kind. The indefinite singular and
definite plural are unacceptable continuations of the preceding context and the bare noun in (8e) is
ungrammatical, therefore they are expected to receive a rating of 1:

(8) Test category: NP-level generic

I have been studying biology today and I found out that many species are no longer
alive. For example, I found out ….  

a. the dinosaur is extinct. 1 2 3 4
b. dinosaurs are extinct. 1 2 3 4
c. a dinosaur is extinct. 1 2 3 4
d. the dinosaurs are extinct. 1 2 3 4
e. dinosaur is extinct. 1 2 3 4

The NP in (9) is a sentence-level description that provides a generic interpretation rather than a
well-deﬁned kind with a kind predicate. The indeﬁnite singular and the bare plural are acceptable for
sentence-level genericity so they are given a rating of 4. The deﬁnite singular, deﬁnite plural, and bare
noun are unacceptable so they are scored as 1:

(9) Test category: Sentence-level generic

My 3 year old daughter’s birthday is coming soon. I don’t know what to buy her for a
present. However, my wife read in a recent survey of young children, …..  

a. a colourful toy is popular with young kids. 1 2 3 4
b. colourful toys are popular with young kids. 1 2 3 4
c. the colourful toy is popular with young kids. 1 2 3 4
d. colourful toy is popular with young kids. 1 2 3 4
e. the colourful toys are popular with young kids. 1 2 3 4

If the participants were unsure of the acceptability of a certain sentence they could choose 2 (less
acceptable) or 3 (more acceptable).

5.4. Procedure

Parents of all four of the child returnees were given a background questionnaire to complete before
testing. To date, the four returnees have attended three sessions. The detailed background questionnaire
was given out at the first session only. The C-tests were administered in each session in order to obtain
the child returnees’ proficiency level in English. Five short C-tests were administered and the total amount of time it took for the participants to complete them was 25 minutes. For each session, the returnees completed a written acceptability judgment task and after they had finished they were asked if they wanted to take a short break. The C-tests were given to the returnees after the short break. Since our participants have returned to Japan it is possible that they will undergo L2 attrition as English is likely to be spoken much less in Japan on a daily basis.  

5.5. Predictions

We have five predictions for the participants in our study.

Prediction #1: The English native speakers are predicted to be target-like on English articles and allow definite singular NPs in the context of NP-level genericity, allow indefinite singular NPs in the context of sentence-level genericity and bare plurals in the context of both NP-level genericity and sentence-level genericity.

Prediction #2: If the child returnees fail to map the [+species] feature to the definite singular, they will assign a low rating to the definite singulars in the context of NP-level genericity.

Prediction #3: The child returnees are predicted to assign a high rating to indefinite singular NPs in the context of sentence-level genericity.

Prediction #4: The child returnees are expected to be target-like with bare plurals in the context of both NP-level genericity and sentence-level genericity.

Prediction #5: The child returnees may show signs of L2 attrition between Time 1 and Time 2.

5.6. Results

The group results in Table 2 below show that for the native control group, there is a clear preference for the definite singular in the anaphoric singular category (3.9) and for the definite plural in the anaphoric plural category (3.9). The bare plural, bare singular, and indefinite singular are all rated below 2. The analysis of generics we adopted from Krifka et al. (1995) is supported as the native controls have a clear preference for the definite singular (the dinosaur = 3.3) and bare plural (dinosaurs = 3.9) in the NP-level test category and prefer the indefinite singular (a colourful toy = 3.8) and the bare plural (colourful toys = 3.8) in the sentence-level test category.

Table 2
Results for Native English Controls

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericity</th>
<th>sentence-level genericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>3.9</td>
<td>1.5</td>
<td>3.3</td>
<td>1.8</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>1.5</td>
<td>1.7</td>
<td>1.5</td>
<td>3.8</td>
</tr>
<tr>
<td>bare singular</td>
<td>1.2</td>
<td>1.1</td>
<td>1.4</td>
<td>1.2</td>
</tr>
<tr>
<td>definite plural</td>
<td>1.4</td>
<td>3.9</td>
<td>2.5</td>
<td>1.7</td>
</tr>
<tr>
<td>bare plural</td>
<td>1.4</td>
<td>1.5</td>
<td>3.9</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Note. (n=35): Mean ratings by context and sentence type.

The results of the four child returnees are below. Table 3 shows the eldest returnee, CS, at Time 1.

---

5 We explore the possibility of L2 attrition in these child returnees in the Discussion section of the paper.
Table 3

Results of CS at Time 1

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericity</th>
<th>sentence-level genericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>3.2</td>
<td>1.6</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>2.9</td>
<td>1.3</td>
<td>2.2</td>
<td>2.2</td>
</tr>
<tr>
<td>bare singular</td>
<td>2.8</td>
<td>1.5</td>
<td>3.0</td>
<td>3.1</td>
</tr>
<tr>
<td>definite plural</td>
<td>1.4</td>
<td>3.3</td>
<td>3.1</td>
<td>2.6</td>
</tr>
<tr>
<td>bare plural</td>
<td>1.8</td>
<td>2.9</td>
<td>3.3</td>
<td>3.0</td>
</tr>
</tbody>
</table>

Note: Mean ratings by context and sentence type.

The results in Table 3 show that for the control categories, CS clearly has acquired number as she rates the definite singular (the red shirt = 3.2) higher than the definite plural (the red shirts = 1.4) for the anaphoric singular category. But, there are no differences in her ratings in the same category of the indefinite singular (a red shirt = 2.9) which implies she has not fully mastered the article system. This is also evident from her ratings for the anaphoric plural category; there is no distinction made in her ratings between the definite plural (the grey suits = 3.3) and the bare plural (grey suits = 2.9), but she correctly rates the definite plural (3.3) higher than the definite singular (the grey suit = 1.6). For the two test categories, CS has no clear preference for any article: there is no clear preference for any article for the NP-level generics or the sentence-level generics. In other words, there are no differences between the ratings between the definite singular (the dinosaur = 2.9) and the indefinite singular (a dinosaur = 2.2) for the NP-level generics category and no differences between the indefinite singular (a colourful toy = 2.2) and the definite singular (the colourful toy = 2.8) for the sentence-level generics category. We interpret the findings thus far for CS that she has no preference for either the definite singular or the indefinite singular in either test category. Furthermore, there are no differences in the ratings of the definite plural and the bare plural for both of the test categories where we would expect higher ratings for the bare plural. Table 4 shows the results for CS at Time 2.

Table 4

Results of CS at Time 2

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericity</th>
<th>sentence-level genericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>4.0</td>
<td>1.0</td>
<td>2.2</td>
<td>2.5</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>2.4</td>
<td>1.0</td>
<td>2.0</td>
<td>3.2</td>
</tr>
<tr>
<td>bare singular</td>
<td>3.1</td>
<td>1.1</td>
<td>2.0</td>
<td>3.1</td>
</tr>
<tr>
<td>definite plural</td>
<td>1.1</td>
<td>3.4</td>
<td>3.1</td>
<td>2.0</td>
</tr>
<tr>
<td>bare plural</td>
<td>1.0</td>
<td>3.2</td>
<td>3.3</td>
<td>2.3</td>
</tr>
</tbody>
</table>

Note: Mean ratings by context and sentence type.

Table 4 reveals that there are no differences between Time 1 and Time 2. There are no signs of L2 attrition in the ratings of articles. CS still correctly identifies ‘number’ as the definite singular is rated higher than the definite plural in the anaphoric singular control category. The same is found in the anaphoric plural control category as CS rates the definite plural higher and gives a low rating to the definite singular. For the test categories, there is no clear preference for one particular article in the NP-level and sentence-level categories. Tables 5 and 6 are the results of YS.
Table 5
Results of YS at Time 1

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericity</th>
<th>sentence-level genericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>3.8</td>
<td>1.0</td>
<td>1.9</td>
<td>2.2</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>2.3</td>
<td>1.0</td>
<td>1.5</td>
<td>2.8</td>
</tr>
<tr>
<td>bare singular</td>
<td>2.2</td>
<td>1.0</td>
<td>1.5</td>
<td>2.2</td>
</tr>
<tr>
<td>definite plural</td>
<td>1.2</td>
<td>3.8</td>
<td>3.7</td>
<td>3.2</td>
</tr>
<tr>
<td>bare plural</td>
<td>1.2</td>
<td>3.2</td>
<td>3.3</td>
<td>3.8</td>
</tr>
</tbody>
</table>

Note. Mean ratings by context and sentence type.

Table 5 shows that YS is better than CS at rating the definite singular for the anaphoric singular control category. She rates the definite singular (3.8) higher than the indefinite singular (2.3). For the anaphoric plural category, YS makes no distinction in her ratings of the definite plural (3.8) and the bare plural (3.2). In the test categories there is no preference for one particular article for NP-level generics as the definite singular, indefinite singular, and bare singular are all given low ratings (below 2). For the sentence-level generics, the indefinite singular (2.8) is only marginally rated higher than the definite singular (2.2) and the bare singular (2.2). YS is no different to CS in her ratings of the bare plural for NP-level generics and sentence-level generics. She rates bare plurals and definite plurals highly for both test categories (above 3).

Table 6
Results of YS at Time 2

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericity</th>
<th>sentence-level genericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>3.9</td>
<td>1.0</td>
<td>2.9</td>
<td>2.8</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>3.0</td>
<td>1.0</td>
<td>2.3</td>
<td>3.3</td>
</tr>
<tr>
<td>bare singular</td>
<td>2.3</td>
<td>1.0</td>
<td>1.9</td>
<td>2.8</td>
</tr>
<tr>
<td>definite plural</td>
<td>1.0</td>
<td>4.0</td>
<td>3.6</td>
<td>3.3</td>
</tr>
<tr>
<td>bare plural</td>
<td>1.0</td>
<td>2.9</td>
<td>3.7</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Note. Mean ratings by context and sentence type.

The results in Table 6 for YS at Time 2 show a similar pattern to Time 1. One main difference however, is that YS correctly rates the definite singular (2.9) higher in the NP-level category at Time 2 than Time 1 (1.9). She rates the definite singular higher than the indefinite singular (2.3) and the bare singular (1.9). Tables 7 and 8 provide the results of KS.

Table 7
Results of KS at Time 1

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericity</th>
<th>sentence-level genericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>3.6</td>
<td>3.0</td>
<td>2.3</td>
<td>2.1</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>1.4</td>
<td>1.7</td>
<td>2.1</td>
<td>2.8</td>
</tr>
<tr>
<td>bare singular</td>
<td>1.2</td>
<td>1.6</td>
<td>1.6</td>
<td>1.5</td>
</tr>
<tr>
<td>definite plural</td>
<td>3.3</td>
<td>3.9</td>
<td>2.6</td>
<td>2.6</td>
</tr>
<tr>
<td>bare plural</td>
<td>1.2</td>
<td>1.1</td>
<td>3.8</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Note. Mean ratings by context and sentence type.
In the control categories, KS has more of a problem with number rather than definiteness. She rates the definite singular (3.6) and the definite plural (3.3) equally highly for the anaphoric singular category and for the anaphoric plural category she rates the definite plural (3.9) and the definite singular (3) highly. She correctly assigns a low rating to the indefinite singular (1.4) in the anaphoric singular category and correctly gives a low rating to the bare plural (1.1) in the anaphoric plural category. Like CS and YS, there is no high rating for the definite singular in the NP-level generic test category, though the indefinite singular (2.8) is rated a little higher than the definite singular (2.1) and bare singular (1.5) in the sentence-level generic test category.

Table 8
Results of KS at Time 2

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericity</th>
<th>sentence-level genericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>3.9</td>
<td>3.2</td>
<td>2.7</td>
<td>2.4</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>1.9</td>
<td>1.7</td>
<td>2.4</td>
<td>3.1</td>
</tr>
<tr>
<td>bare singular</td>
<td>1.1</td>
<td>1.5</td>
<td>1.7</td>
<td>1.1</td>
</tr>
<tr>
<td>definite plural</td>
<td>2.8</td>
<td>3.4</td>
<td>3.0</td>
<td>2.5</td>
</tr>
<tr>
<td>bare plural</td>
<td>1.0</td>
<td>1.4</td>
<td>4.0</td>
<td>3.6</td>
</tr>
</tbody>
</table>

Note. Mean ratings by context and sentence type.

In Table 8, the results at Time 2 show no change from Time 1 for KS. The ratings in the control categories and the test categories are very similar. She continues to have a problem with identifying ‘number’ for the anaphoric control categories as she incorrectly rates both the definite singular and definite plural highly in both anaphoric singular and anaphoric plural categories. The final set of results is of TA.

Table 9
Results of TA at Time 1

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericity</th>
<th>sentence-level genericity</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>2.7</td>
<td>1.5</td>
<td>1.2</td>
<td>2.2</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>2.2</td>
<td>1.0</td>
<td>1.6</td>
<td>1.7</td>
</tr>
<tr>
<td>bare singular</td>
<td>3.0</td>
<td>1.4</td>
<td>2.8</td>
<td>2.2</td>
</tr>
<tr>
<td>definite plural</td>
<td>1.2</td>
<td>2.5</td>
<td>2.5</td>
<td>1.9</td>
</tr>
<tr>
<td>bare plural</td>
<td>1.3</td>
<td>3.4</td>
<td>1.6</td>
<td>3.1</td>
</tr>
</tbody>
</table>

Note. Mean ratings by context and sentence type.

The results in Table 9 indicate that TA experiences difficulties in her ratings for the control and test categories. The bare singular (3), indefinite singular (2.2), and the definite singular (2.7) are all rated highly for the anaphoric singular category. However, she correctly assigns lower ratings to the bare plural (1.3) and the definite plural (1.2). For the anaphoric plural category, she rates the bare plural (3.4) higher than the definite plural (2.5), and correctly rates the singulars 1.5 or less. The control category results show that she has acquired ‘number’ but still has not mastered the article system. For the test categories, the bare singular is either rated higher than (2.8 for NP-level generics) or equal to (2.2 for sentence-level generics) the definite singular (1.2) and indefinite singular (1.7) articles.
Table 10

Results of TA at Time 2

<table>
<thead>
<tr>
<th></th>
<th>anaphoric singular</th>
<th>anaphoric plural</th>
<th>NP-level genericty</th>
<th>sentence-level genericty</th>
</tr>
</thead>
<tbody>
<tr>
<td>definite singular</td>
<td>2.0</td>
<td>2.4</td>
<td>3.1</td>
<td>3.0</td>
</tr>
<tr>
<td>indefinite singular</td>
<td>1.9</td>
<td>3.4</td>
<td>2.2</td>
<td>2.7</td>
</tr>
<tr>
<td>bare singular</td>
<td>1.8</td>
<td>2.3</td>
<td>2.9</td>
<td>2.3</td>
</tr>
<tr>
<td>definite plural</td>
<td>3.1</td>
<td>2.2</td>
<td>2.3</td>
<td>2.1</td>
</tr>
<tr>
<td>bare plural</td>
<td>2.8</td>
<td>2.1</td>
<td>1.1</td>
<td>1.7</td>
</tr>
</tbody>
</table>

*Note.* Mean ratings by context and sentence type.

The final set of results shown in Table 10 are of TA at Time 2. In comparison with Time 1, there are not many differences across the ratings in the control and test categories. TA seems to have difficulty identifying ‘number’ and ‘definiteness’ as the indefinite singular (3.4) is given the highest rating in the anaphoric plural category. In the test categories, TA rates all the possible choices highly apart from the bare plural (1.1 for NP-level generics and 1.7 for sentence level generics).

6. Discussion

Despite a prolonged period spent abroad in an English-speaking environment, the four child returnees experience a number of difficulties with the English article system. The results show that they have difficulties with rating the definite singular for NP-level generics, and to a lesser extent, have problems with rating the indefinite singular for sentence-level generics. Furthermore, the two control categories, anaphoric singular, and anaphoric plural, pose problems for the four returnees as options other than the definite singular and the definite plural are rated highly.

The performance by the four child returnees on the acceptability judgment task suggests that they do not have a basic mastery of the article system and the more complex uses of articles, like generics, is too difficult to acquire at this stage. We might expect child L2 learners who have received far less input (child EFL learners, for example) to perform poorly on such a task. But, it is somewhat surprising that the four child returnees in the current study still have such difficulties with anaphoric contexts. Yoshitomi (1999) argues that attrition starts soon after child or adult returnees leave the target language environment and come back to their home country. Our study, in contrast, shows that the child returnees’ knowledge of the English article system is not necessarily attrited because Time 2 is not in any way greatly different to Time 1. If L2 attrition had taken place between the two testing times we would expect to see a change between Time 1 and Time 2 as six months had elapsed. However, no observable changes were found which suggests that rather than attrition, the returnees have not yet fully acquired the article system.

The returnee named TA in our study seems to have an advantage since she started learning English from a very early age, but actually this may not be the case. It may be that Japanese is the stronger language despite having lived in an L2 environment. Other variables such as degree of social interaction with English native speakers etc. would have to be considered. Maybe L2 competence has a stronger influence over age of acquisition and L2 exposure. We also acknowledge that it is difficult to conclude since this may be an individual problem.

7. Conclusion

In conclusion, the findings from our written acceptability judgment task demonstrated that the child returnees have still not mastered the English article system – possibly incomplete acquisition rather than L2 attrition. Their choices differed across the control (anaphoric) and test (generic) categories to the choices made by the native control group. However, for the generics, there is evidence to suggest that the definite singular for NP-level generics remains more difficult than the indefinite singular for sentence-level generics for all the returnees as the definite singular was never rated very highly. Our findings suggest that the returnees fail to map the [+species] feature to the definite singular even after many years of exposure to L2 English. We believe that despite the non target-like
performance on definite singulars, the results are consistent with the Feature Reassembly Hypothesis as there is evidence to suggest that the returnees recognize that there is a difference between the definite singular and the indefinite singular in regards to definiteness for the anaphoric singular test items. The feature [+definite] is in the process of being mapped onto the definite singular in their interlanguage grammars as the definite singular is given a higher rating by all four returnees. Nevertheless, the [+species] feature is semantically complex and has yet to be successfully mapped onto the definite singular. The problem with generics is not likely due to any type of syntactic deficit (Hawkins, 2005) but rather where the syntax and semantics interface with each other.

One reason for the observed differences within the child returnees’ results is the role of Japanese and English in everyday life. Some children will actively use two languages while others will limit themselves to one. This second type of child may be called a passive bilingual, someone who seems to understand both languages but chooses to speak mainly one (Kasuya, 1998). The dominant language for all the returnees is Japanese, now that they have returned to Japan, but it was the dominant language used at home during their stay in the U.S. Therefore, though the returnees attended English-speaking schools and had English-speaking friends, the more active language may have been Japanese because of its use at home. Despite English being the weaker language there is no sign of L2 attrition perhaps due to their ongoing English education in Japanese high schools and university.

We acknowledge that one of the limitations of our study is that the results may only adequately describe the four returnees in this study since three out of the four belong to the same family. A larger scale study of returnees needs to be conducted in order to confirm or disconfirm the current findings. A further limitation is the possible influence of contexts. It is impossible to know exactly how often the returnees were exposed to the types of generics discussed in our study. There would have been plenty of exposure to articles as they are highly frequent in the input but NP-level generics are not (Biber, Johansson, Leech, Conrad, & Finegan, 1999). This may explain why the returnees in our study have a hard time choosing the definite singular for NP-level genericity; bare NPs are ungrammatical and the returnees seem to be unaware that the indefinite singular and definite plural are ungrammatical when referring to species.

References


Selected Proceedings of the 2012 Second Language Research Forum: Building Bridges between Disciplines

edited by
Ryan T. Miller, Katherine I. Martin, Chelsea M. Eddington, Ashlie Henery, Nausica Marcos Miguel, Alison M. Tseng, Alba Tuninetti, and Daniel Walter

Cascadilla Proceedings Project Somerville, MA 2014

Copyright information
Selected Proceedings of the 2012 Second Language Research Forum: Building Bridges between Disciplines
© 2014 Cascadilla Proceedings Project, Somerville, MA. All rights reserved
A copyright notice for each paper is located at the bottom of the first page of the paper. Reprints for course packs can be authorized by Cascadilla Proceedings Project.

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, 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: