Applied Psycholinguistics (2019), **40**, 1377–1404
doi:10.1017/S0142716419000304

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Incidental and explicit learning of L2 derivational morphology and the nature of acquired knowledge

Keiichi Ishikawa*

Kwansei Gakuin University
*Corresponding author. Email: ishikawa.k@kwansei.ac.jp

(Received 20 May 2018; revised 16 June 2019; accepted 21 June 2019)

Abstract

The aim of the current study was to compare both incidental and explicit auditory learning of second language derivational morphology by measuring the accuracy and the reaction time of a grammaticality judgment task. Furthermore, the study was set up to examine the nature of acquired knowledge using subjective measures of awareness during the testing phases and postexperimental verbal reports. The delayed effects of learning were investigated by testing participants immediately after the learning and 1 week later also. The results showed a significant learning effect for the incidental and explicit learners immediately after exposure, but only the explicit learners maintained the learning effects a week later. Both types of learners showed no significant difference in the reaction time. Incidental learners primarily developed implicit knowledge, while explicit learners relied on explicit knowledge to a large extent, part of which became unconscious later. The differences of learning under incidental and explicit learning conditions are discussed in terms of the maintenance of knowledge.

Keywords: delayed effects; explicit learning; implicit knowledge; incidental learning; second language derivational morphology

The main issue the present study attempts to address is how people acquire and generalize new morphological systems. This is a critical aspect of language competence: the ability to produce an unlimited number of expressions by combining a limited number of linguistic units (Diessel, 2004; Pinker & Jackendoff, 2005). This creativity is reflected in the process of generalization across individual exemplars to which people are exposed that evidently emerges in morphological learning (Tamminen, Davis, Merkx, & Rastle, 2012; Tamminen, Davis, & Rastle, 2015). For example, the vast majority of new English words (e.g., *unreasonable*) are created by combining a limited set of stems (e.g., *reason*) with affixes (e.g., *un–*, *–able*) based on certain rules (Algeo, 1991). The hallmark of this productivity lies in the capability to generalize newly learned morphemes (e.g., *un–*, *–able*) to create novel words

(e.g., *ungoogleable*) in appropriate grammatical categories (e.g., "*This is an ungoogleable question*").

It is generally accepted that children acquire their first language unconsciously. It has long been a subject of debate, however, whether adults learn a second language (L2) without awareness or intention. Krashen (1981, 1985) claimed that language acquisition takes place through unconscious, implicit processing of input and that explicit learning plays a small role in the development of implicit knowledge. Schmidt (1990, 1995, 2001) proposed that noticing is sufficient, and in some cases necessary, for language learning to take place, though incidental learning is possible when attention is focused on what is to be learned. Schmidt also proposed two levels of awareness regarding conscious and unconscious language learning: the level of noticing, where learners register some aspects of input, and the level of understanding, where learners recognize rules or regularities underlying the structure of the input. Explicit knowledge has been argued to foster and facilitate the development of implicit knowledge (Ellis, 2005, 2015; Suzuki & DeKeyser, 2017), and conscious learning may lead to unconscious knowledge through proceduralization and automatization of explicitly learned knowledge (DeKeyser, 1997, 2015). Automatized explicit knowledge and implicit knowledge can be distinguished by the awareness criterion (Suzuki, 2017; Suzuki & DeKeyser, 2017). Although both types of knowledge can be accessed rapidly, automatized explicit knowledge involves awareness about linguistic forms, while implicit knowledge does not.

The definitions of terms used in this article in regard to the process and product of learning are stated here. Implicit learning is a process whereby people acquire knowledge without the intention to learn and without awareness of the knowledge they have acquired, while explicit learning is a process in which people are aware that they are acquiring knowledge. In this context, incidental learning refers to learning without any intention to learn, but it does not necessarily exclude the awareness of what has been learned. Implicit knowledge refers to the product of learning (knowledge) people are unaware and unconscious of, whereas explicit knowledge signifies the conscious knowledge people are aware of and often able to verbalize (Grey, Williams, & Rebuschat, 2014; Hulstijn, 2005; Leow, 2018; Leow & Zamora, 2017; Rebuschat, Hamrick, Riestenberg, Sachs, & Ziegler, 2015; Rebuschat & Williams, 2012; Williams, 2009).

Recent studies have shifted their interest toward measuring the knowledge acquired under different learning conditions (i.e., the product of learning), rather than the process of learning as the stage of encoding information (Hama & Leow, 2010; Rebuschat et al., 2015; Suzuki, 2017; Suzuki & DeKeyser, 2017). This is to determine whether the knowledge acquired during exposure or leaning phases is implicit or explicit in nature (Grey et al., 2014; Rebuschat, 2013; Rebuschat & Williams, 2012; Rogers, 2017; Rogers, Révész, & Rebuschat, 2016).

A growing body of research has investigated what aspects of a L2 can be learned incidentally, including form–meaning mappings (Hama & Leow, 2010; Kerz, Wiechmann, & Riedel, 2017; Leung & Williams, 2011, 2012, 2014; Rebuschat et al., 2015; Williams, 2005), word order (Grey et al., 2014; Rebuschat & Williams, 2012; Ruiz, Tagarelli, & Rebuschat, 2018; Tao & Williams, 2018; Williams & Kuribara, 2008), case marking (Brooks & Kempe, 2013; Grey et al., 2014; Robinson, 2005; Rogers, 2017, 2019; Rogers et al., 2016), verb meaning

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

(Paciorek & Williams, 2015), and phonological stress patterns (Chan & Leung, 2018; Graham & Williams, 2016). These studies have examined learning under incidental learning conditions and the conscious status of the resultant knowledge after incidental exposure (i.e., the product of learning).

To date, however, no research has been conducted on incidental learning of L2 derivational morphology and its nature of acquired knowledge. Unlike inflectional morphology, derivational morphology creates new words and adds novel meanings in different grammatical categories due to its combinatory system containing stems and suffixes (Marslen-Wilson, 2007; Tyler & Nagy, 1989). Derivational processes and derived words make a significant impact on processing and expanding vocabulary among children, adults, and L2 learners (Carlisle, 2000; Jarmulowicz, 2006; Silva & Clahsen, 2008). Previous studies, however, have not examined how derivational morphological system is learned under incidental or explicit learning conditions. Marsden, Williams, and Liu (2013) conducted priming research regarding recognition memory of suffixed nonwords; however, the meanings added to the suffixes they used were functional (e.g., singular or plural, present or past) rather than derivational (causing a change of word class). Marsden et al. exposed participants to suffixed nonwords (e.g., *gatot*), instead of sentences, to investigate the morphological representations under the priming paradigm. What is relevant to incidental learning was their finding that the generalization task using a picture-matching test (e.g., oral and visual presentation of *smafot* with three pictures) showed that participants learned the meanings of suffixes (e.g., *–ot* in *gatot*) even when their attention was oriented to the stems (e.g., *gat–* in *gatot*). The following two sections provide a brief introduction to some key concepts that are widely discussed in light of awareness assessment and learning conditions with a particular focus on delayed effects of learning.

Measurement of awareness

Researchers have used different procedures to operationalize and measure the construct of awareness. Studies concerned with the stage of encoding information (the process of learning) employ concurrent measures of awareness such as think-aloud protocols in which participants are requested to say aloud whatever they are thinking while performing experimental tasks (Bowles & Leow, 2005; Hama & Leow, 2010; Rebuschat et al., 2015). One limitation to think-aloud protocols is the issue of reactivity influencing learners' cognitive processes and accuracy of task performance. In contrast, in order to measure whether participants have acquired implicit or explicit knowledge (the product of learning), nonconcurrent retrospective verbal reports are conducted after testing (Rebuschat et al., 2015; Rebuschat & Williams, 2012; Williams, 2005). Retrospective verbal reports prompt participants to verbalize any rules or patterns they might have noticed during the experiment. Knowledge is considered unconscious or implicit if participants demonstrate the effect of learning despite being unable to verbalize the underlying rule in the exposure items. One problem with this measurement, however, is that participants may fail to recall something after a long period of exposure and testing. Another drawback is that participants may not report knowledge because they lack the confidence to do so

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

(Hamrick & Rebuschat, 2012; Rebuschat, 2013; Rebuschat et al., 2015; Rebuschat & Williams, 2012). It is noteworthy that both measures of using think-aloud protocols and verbal reports rely on verbalization, which does not identify the construct of awareness that is not verbalized (Rebuschat et al., 2015).

In order to refine the measurement of the conscious status of knowledge, recent studies have adopted two types of subjective measures of awareness during testing phases: confidence ratings and source attributions (Dienes, Altmann, Kwan, & Goode, 1995; Dienes & Scott, 2005; Hamrick & Rebuschat, 2012; Rebuschat, 2013; Rebuschat et al., 2015; Rogers et al., 2016). Confidence ratings ask participants to indicate their level of confidence (e.g., *no confidence, somewhat confident, very confident*, or *absolutely certain*) for each decision they make during the testing phrase. Source attributions require them to identify the source of their judgment (e.g., guess, intuition, memory, or rule knowledge). Knowledge can be considered unconscious if participants' confidence is unrelated to their accuracy (called the *zero-correlation criterion*), or if a negative correlation between confidence and accuracy is observed, which would provide evidence for better performance without conscious knowledge (Rebuschat, 2013). Knowledge can also be considered implicit if they believe themselves to be guessing when their classification performance is significantly above chance (called the *guessing criterion*). The intuition and guess attributions are both considered to correspond to unconscious knowledge (Dienes & Scott, 2005; Rebuschat, 2013).

For instance, Rogers et al. (2016) used the two types of subjective measures. Rogers et al. (2016) investigated whether novel morphological case markings can be learned by adults via incidental exposure and whether the resulting knowledge is implicit or explicit in nature. The study used a semiartificial language based on English sentences with Czech case marking. The testing phase included a grammaticality judgment test, along with subjective measures of awareness. At the end of the experiment, participants also provided retrospective verbal reports. The results of grammaticality judgment tests (GJT) showed that the experimental group demonstrated significantly better learning effects than the control group for accusative cases, but not for nominative cases. The verbal reports found that all participants noticed the changing morphological inflections during the training phase, although none of them were able to verbalize the underlying rule system. Nevertheless, the analysis of the subjective measures of awareness demonstrated that participants developed both implicit and explicit knowledge; confidence ratings suggested the conscious knowledge evidenced by above chance performance when participants chose the categories *somewhat confident* and *very confident*, while the analysis of source attributions revealed the unconscious knowledge evidenced by above chance performance only based on "intuition."

Learning conditions and delayed effects

A number of studies investigating differences between implicit and explicit instruction have demonstrated that learning under explicit-type conditions is more effective than learning under implicit-type conditions (Goo, Granena, Yilmaz, & Novella, 2015; Norris & Ortega, 2000; Spada & Tomita, 2010). For example,

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

Robinson (1996) compared the learning of English grammar rules by adult learners of English as a second language under four training conditions: implicit (explained to participants as a memory test), incidental (as an exercise in reading for meaning), rule-search (as an exercise in finding rules), or instructed (where participants were asked to read through the rules). The results of the GJT (e.g., "*Across the street raced Tom*" judged as being grammatical, and "*On Saturday night danced Charlie*" judged as being ungrammatical) showed that instructed learners outperformed the other groups. Similar results were obtained by Robinson (1997) in which the dative alternation in English (e.g., "*John donated the piano to the church*" as grammatical, and "*John donated the church the piano*" as ungrammatical) was learned under four different conditions: implicit, incidental, enhanced (target items were boxed for each sentence), or instructed. Instructed learners were superior to learners in other conditions, in particular in generalizing the knowledge developed through training to new items.

Andringa and Curcic (2015) investigated how a morphosyntactic rule in Spanish is explicitly or implicitly acquired by adult native speakers of Dutch. The rule is called differential object marking according to which direct objects are preceded by the preposition *a* ("to") if they are animate. In the 11-min instruction, both explicitly and implicitly instructed learners were presented with the audio of 52 target trials providing exposure to the rule. The difference between the two groups was that only the explicit group received the explanation of the target rule. In the explicit instruction, 2 target trials were replaced by screens explaining the rule to the participants in Dutch. The results showed that explicitly instructed learners were more accurate than implicitly instructed ones in the GJT. They conducted source attributions that confirmed that explicitly instructed learners based their decision on the rule 71% of the time, while implicitly instructed learners did so 12% of the time. In their study, however, participants in the implicit learning condition, as well as those in the explicit group, were informed in advance of the grammaticality test that the knowledge was tested after the instruction, which makes it hard to characterize their learning as implicit or even incidental.

Denhovska and Serratrice (2017) examined the acquisition of the gender agreement patterns in Russian by adult native speakers of English. Participants in the incidental learning condition were visually presented with Russian sentences along with English translations. They were asked to read the Russian sentences and translations without performing any other task. Participants in the explicit condition received a metalinguistic explanation of the gender agreement in Russian and were provided with example sentences. They were asked to memorize the agreement rule and were informed that they would be tested afterward. The training time for both conditions was the same (15 min). The results showed that participants in both conditions were similarly accurate in the GJT, although learners in the explicit condition performed significantly better than those in the incidental condition in the production task where they had to fill in the blanks with appropriate inflections.

Unlike the above four studies examining grammar acquisition, Hamrick and Rebuschat (2012) compared the learning effects between intentional and incidental learning conditions in the statistical word learning paradigm. Each condition had 30 native speakers of English, and participants in both conditions were exposed to 57 trials in which two images (e.g., *panda* and *glass*) were displayed on the screen and

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

two pseudowords (e.g., *houger* and *femod*) were auditorily presented. The location of the image and the order of the pseudoword were not related. The only clue to the right referent of the word was to keep track of the co-occurrences of the images and the words across trials. Participants in the intentional learning condition were instructed to learn the meanings of the words and were told that they would be tested afterward. In contrast, participants in the incidental condition were not informed about the purpose of the experiment or about being tested afterward, and were asked to indicate instead how many objects on each slide were animate. In the testing phase, participants in both groups were presented with four pictures on the screen and a spoken pseudoword, and were required to select the appropriate referent as quickly and accurately as possible. Performance on the picture-matching task was regarded as the measure of learning. Awareness was measured by means of confidence ratings and source attributions. The results showed that although there was a clear learning effect for both groups, the learning effect was larger under intentional learning conditions. The zero-correlation criterion was met in the case of the incidental group, but it was not met for the intentional group. Source attributions with three categories (*guess, intuition*, and *memory*) revealed that both groups performed significantly above chance when choosing any of the three categories, including when the intentional group members based their decisions on guessing. This suggests that some unconscious knowledge developed even under intentional learning conditions.

As a longitudinal study, Morgan-Short, Steinhauer, Sanz, and Ullman (2012) examined how explicit and implicit training affected the learning of an artificial language's grammar, which was auditorily presented to participants using the GJT and electrophysiological (event-related potential) measures. Learners in the explicit condition were provided with metalinguistic explanation of the artificial language's rules along with 33 example sentences, while those in the implicit condition received 127 example sentences only. The initial training time was 13.5 min for both groups. After the initial training, learners in both groups received additional practice in comprehension and production over 1 to 5 days until they reached low proficiency (accuracy significantly above chance) and then high proficiency (80% accuracy or more). Performance of the behavioral judgment task was not significantly different between the groups at either low or high proficiency, although event-related potential measures revealed that only the implicitly trained learners showed brain activity typical of native speakers at a high proficiency.

With respect to the delayed effects of learning, Tamminen et al. (2012) and Tamminen et al. (2015) discussed the role of consolidation in language learning and linguistic generalization (i.e., a process of integration or generalization of newly learned knowledge, which develops slowly over a period of time). In their studies, native speakers of English were taught new words consisting of a familiar stem and a novel affix (e.g., *sleepafe*) with a definition (e.g., "*sleepafe* is a participant in a study about the effects of sleep"). The training phase included a typing task and a recall task that lasted roughly 1 hr. In the testing phase, the sentence congruency task and the recognition memory task were given. Reading aloud latencies were measured in the sentence congruency task, and the accuracy as to whether the word was trained or untrained was obtained in the recognition memory task. The results showed that although the accuracy of recognition declined a week later, the semantic congruency

effect (i.e., longer reading latency for incongruent sentences compared to congruent ones) was observed a week after training. This suggests that generalization takes place following a delay between learning and testing. These studies, however, were not concerned with whether participants acquired implicit or explicit knowledge.

In a study on the effects of input- and output-based instruction on the learning of Spanish object pronouns, Morgan-Short and Bowden (2006) found that input-based instruction in the interpretation test made no significant change from immediate to delayed tests. Output-based instruction, however, showed a significant loss, and the control group revealed a significant gain although the two experimental groups demonstrated significant gains from the pretests.

Grey et al. (2014) investigated incidental learning of word order and case marking in a semiartificial language called *Japlish*, which followed the word order and case-marking rules of Japanese but uses English vocabulary. The verb was always at the end of the sentences and nouns were case marked for subject (*–ga*), object (*–o*), and indirect object (*–ni*; e.g., "*Stacey-ga that picture-o painted*," "*Those seeds-o Ann-ga parrot-ni fed*"). During the exposure phase (approximately 20 min), participants were presented audio of 128 sentences during which they decided if the sentences were semantically plausible or not. Performance was assessed with the acceptability judgment task (AJT) examining the learning of word order and the picture-matching tasks (PMT), which tested the learning of case marking that were administered immediately after exposure and 2 weeks later. On the immediate tests, there was a significant learning effect on the AJT, but not on the PMT. The delayed tests, however, showed maintenance of the learning effects for the AJT and significantly improved performance for the PMT. It was assumed that consolidation played a role when the establishment of the form–meaning connections (between *–ga*, *–o*, *–ni* and subject, object, indirect object, respectively) was required for the PMT. They took subjective measures of awareness only for the AJT. Confidence ratings showed their performance was above chance when they felt *very confident*. Source attributions revealed above chance accuracy based on intuition.

The results from the above studies comparing different learning conditions have generally shown that learning under explicit conditions is superior to incidental conditions. Very few studies, however, have compared how the incidental and explicit learning of grammatical knowledge develops through a relatively short period of exposure, together with the measurement of awareness using confidence ratings and source attributions (see Andringa & Curcic, 2015; Hamrick & Rebuschat, 2012). Only a small number of studies, as shown above, were conducted to discover the difference in learning between immediate and delayed testing. Further, no previous research has compared the delayed effects of the learning of grammatical rules between incidental and explicit learning conditions. The present study addresses this gap by examining the learning of the L2 derivational system and its durability under different learning conditions.

The present study

The present research extended the study by Rogers et al. (2016) in several ways: to focus on derivational morphology, include an explicit learning group, and examine

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

the delayed effects after 1 week. As noted earlier, derivational morphology is a key factor in developing the ability to generalize novel words and grammar (Marslen-Wilson, 2007; Tamminen et al., 2012, 2015; Tyler & Nagy, 1989). The comparison between incidental and explicit learning also has pedagogical implications (Andringa & Curcic, 2015; Ellis, 2015; Hulstijn, 2005; Leow, 2018), considering that in the current study the stimulus materials were based on real English suffixes and that participants were intermediate learners of English. Furthermore, it is worthwhile to examine the delayed effects of learning[1] or consolidation under incidental and explicit learning conditions (Grey et al., 2014; Morgan-Short & Bowden, 2006; Morgan-Short et al., 2012; Tamminen et al., 2012, 2015), especially when the research is concerned with productivity and generalization of newly learned grammatical patterns.

In Rogers, Révész, and Rebuschat (2015), participants were asked in the third experiment to repeat sentences and target foreign words aloud during exposure to promote awareness and induce a learning effect. This is because in the first and second experiments, participants did not demonstrate learning effects without repeating target foreign words (repeating sentences aloud was included in the second experiment). The current study excluded repeating sentences and target words for the incidental learning group due to the possibility of participants noticing the rule hidden in the exposure sentences. Considering that the present study's experimental sentences were real English sentences with familiar words and one nonword keeping the basic word order, adding the task that directs participants' attention to the target nonword would have likely made them more aware of the pattern in the target nonword.

Several studies regarding the learning of grammatical rules have measured reaction times (RTs). Although RT data are often less directly linked to the main research question, they serve as complementary measures for exploring the differences between participant groups or different structures (Jiang, 2012). RT data can also measure the speed and processing load to reach correct grammatical judgments, and can shed light upon the nature of linguistic knowledge that individuals possess (Gass & Mackey, 2007). Robinson (1996) measured RTs in addition to learners' accuracy of GJT and reported no significant difference between learners on implicit or instructed learning conditions. In Robinson (1997), however, incidental learners were significantly slower than instructed learners in making correct grammaticality judgments on untrained sentences, and it was assumed that incidental learners needed additional time to find rules. Andringa and Curcic (2015) also used RT measures besides accuracy data and concluded, based on no significant difference between the two groups, that explicit knowledge of the target grammatical structure does not provide explicit learners with any speed advantage in deciding on correct answers.

The results on RT measures so far have thus been mixed. It can be assumed that there might be a trade-off between the strategies (conscious or unconscious) employed by both types of learners. If participants in the incidental learning condition attempt to look for rules to decide on the grammaticality of testing sentences, they tend to be slow in responding. However, if they use the implicit knowledge at least partially developed during exposure, they may have faster access to knowledge (Suzuki, 2017; Suzuki & DeKeyser, 2017). Likewise, participants in the explicit

learning condition can be very quick in accessing explicit rules taught in learning phases, but at the same time, applying the conscious rules to grammatical judgment may take time. The present study also attempts to address this issue by measuring RTs for different learning groups, which was not included in the study by Rogers et al. (2016), though self-paced reading tasks with RTs were used in some studies investigating the learning of case marking such as Rogers (2019) and VanPatten and Smith (2018).

With a view to validate awareness assessment measures, the current experiment also included retrospective verbal reports at the end of the experiment, as well as two types of subjective measures of awareness during testing phases: confidence ratings and source attributions (Dienes et al., 1995; Dienes & Scott, 2005; Rebuschat, 2013; Rebuschat et al., 2015).

The aim of the current study is thus to compare incidental and explicit auditory learning of L2 derivational suffixes by measuring the accuracy and the RT of a GJT administered to Japanese learners of English at an intermediate level. A further aim was to explore whether the resultant knowledge was implicit or explicit in nature through the analysis of subjective measures of awareness during the testing phases, in addition to postexperimental verbal reports. The study also examined the delayed effects of learning by testing participants immediately after learning and 1 week later. In particular, the research aims to address the following four research questions:

1. To what extent are adult learners able to acquire L2 derivational morphology incidentally through auditory exposure?
2. To what extent is their performance comparable to that of explicit learners?
3. What is the nature of the acquired knowledge: implicit or explicit?
4. Are there any delayed effects of learning demonstrated by the different learning groups?

Method

Participants

For the current study, 60 Japanese university students were recruited as participants (43 females, 17 males) with an age range from 18 to 21 ($M = 19.80$, $SD = 1.09$). All participants were native speakers of Japanese and had studied English as a foreign language for more than 6 years.[2] Their scores of English proficiency tests administered in Japan were 400–470 at TOEFL ITP, 440–600 at TOEIC L&R, or Eiken Grade Pre-2 and Grade 2. Participants were approximately equivalent to A2/B1 levels in the Common European Framework of Reference for Languages based on the conversion table provided by the Ministry of Education, Culture, Sports, Science and Technology, Japan (2015).

Participants were paid for their participation. They filled out the consent forms on the first day (Day 1), at which time they were reminded that they had to participate again a week later (Day 8). The consent form also explained that the task was to perform an action after listening to a sentence (via headphones connected to the computer) and provided no information about the purpose of the study.

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

Table 1. Example sentences with three grammatical categories

Grammatical category	Examples
Noun	She shows some mezəsmənt at home. The teacher likes his dəfɪʒən at school.
Adjective	She is very bæpisfəl at home. The doctor is a very fəmɪsɪk person.
Verb	She tries to nebəsaɪz her friend at home. The doctor wants to vədɪsɪfaɪ his patient.

Participants drew lots to be randomly assigned to three groups, with 20 participants in each group: incidental learning group, explicit learning group, and the control group.

Stimuli

The stimuli were 144 sentences including nonwords attached with 12 suffixes based on real English suffixes forming three grammatical categories (Noun: –mənt, –nəs, –ʃən [ʒən], –əti; Adjective: –fəl, –ɪʃ, –əl, –ɪk; Verb: –eɪt, –ən, –aɪz, –ɪfaɪ). The suffixes were selected from 32 major English suffixes reported in Harwood and Wright (1956). The stems of the nonwords consisted of two syllables: consonant plus vowel plus consonant plus /ɪ/, /i/, /ɪs/, /ɪz/, /is/, /iz/, or /əs/. The consonant plus vowel plus consonant parts of the nonwords were constructed with reference to Noble (1961), ensuring a variation of vowels and consonants. The participants' first language (Japanese) has similar derivational suffixes (Shibatani, 1990; Tsujimura, 2007). For example, when the suffix *–sa* is attached to an adjective "*yasashii*" (*kind*), the derived form "*yasashisa*" (*kindness*) serves as a noun.

These nonwords were inserted into sentence frames designed for the three word classes. Noun frames were "She shows some __ at home," "He shows his __ in the library," "The teacher likes his __ at school," and "The doctor likes to show his __." Adjective frames were "She is very __ at home," "He is very __ in the library," "The teacher is very __ at school," and "The doctor is a very __ person." Verb frames were "She tries to __ her friend at home," "He wants to __ his book," "The teacher tries to __ at school," and "The doctor wants to __ his patient." The sentence frames mentioned above were thought to be appropriate for intermediate learners of English to understand while listening without written forms. All the stimuli are presented in Appendix A. Examples of sentences with three grammatical categories are shown in Table 1.

Forty-eight out of those 144 sentences were used to create an exposure list for the incidental group. Another 48 sentences were used as grammatical sentences for the testing phase. The remaining 48 sentences were made ungrammatical by inserting the nonwords into inappropriate categories in the sentences (e.g., "The doctor is a very bəfɪsnəs person"). Half of the testing sentences (24 sentences) were completely matched except for the suffixes between grammatical and ungrammatical sentences:

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

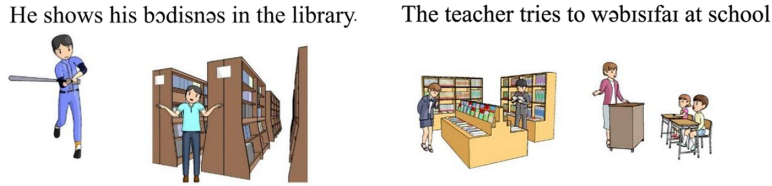

Figure 1. Examples of sentences and pictures used for incidental learners.

"The doctor is a very bəfisfəl person" versus "The doctor is a very bəfisnəs person." The other half (the remaining 24 sentences) had different stems and suffixes to prevent participants from paying too much attention to the suffix itself: "The teacher likes his səfɪʒən at school" versus "The teacher likes his hənısıfaı at school." Two versions of sentence lists (each including different 48 sentences) were prepared for the two testing phases (Day 1 and Day 8) and counterbalanced across participants.

These stimulus sentences were converted into audio files using text-to-speech software (Globalvoice English 3, HOYA). Half of them were created with an American male voice and the other half with a British female voice and were counterbalanced across grammatical categories for each learning and testing set. Postexperimental interviews revealed that no participant noticed the sound was synthetic speech. There are also several studies including Hirai and O'ki (2011) and Matsuda (2017) where the synthetic speech generated via the same software did not result in comprehension problems for L2 learners compared with natural human speech, and was found to be effective for language learning (particularly for nonproficient Japanese learners of English).

Procedure

The experiment consisted of two phases: the learning phase and the testing phase. In the learning phase, the incidental learning group listened to 48 grammatical sentences twice (96 sentences in total). After they listened to each sentence, participants were requested to choose one of the two pictures on the computer screen that they thought matched the content of the sentence. Whether their choice was correct or not was indicated by a different sound: a chime or a beep. The provision of this kind of feedback presumably prompted the semantic processing of the sentences while avoiding raising awareness as to what they were learning. Furthermore, their choice could be made just on the basis of the English words in the sentences such as "at home," "in the library," "at school," and "doctor," which was intended to prevent participants paying much attention to the novel words and the suffixes. The order of the presentation of the sentences was randomized for each participant, and the learning period lasted about 15 min. Example sentences and pictures are illustrated in Figure 1 (all the sentences and the pictures used in the learning phase for the incidental group are presented in Appendix B).

The explicit learning group received an explanation in Japanese (using PowerPoint slides) that English has three types of suffixes that form nouns,

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

adjectives, and verbs. Participants then listened to auditory examples of 12 suffixes and sentences including the suffixes (three grammatical categories × four suffixes) with explanations written on PowerPoint slides. No suffix or sentence was presented in the written mode. The explanation was repeated three times for about 15 min, almost the same amount of time for the incidental learning group. Participants were not allowed to go back to the previous slides. All the slides presented in the learning phase for the explicit condition are shown in Appendix C.

Meanwhile, the control group had no learning phase. The stimulus presentation was controlled by SuperLab 5.0 software with a response pad (RB-740) (Cedrus Corporation).

Testing phase

The three groups took the GJT twice: immediately after the learning phase (Day 1) and a week later (Day 8). Participants listened to one of the two versions of 48 sentences including new nonwords with the suffixes, and judged whether each sentence was grammatically correct or incorrect as quickly and accurately as possible by pressing designated buttons on the response pad. Participants had to complete this task without being given any information about which aspect of the sentence was grammatical or not. It was assumed, however, that the use of basic English word order and the highly familiar words used in the experimental sentences (except for the suffixed nonwords) prompted similar expectations among the three groups regarding a grammaticality judgment (Hamrick & Sacks, 2018). The sentences were presented in randomized order for each participant. Half of the sentences were grammatically correct (e.g., "The doctor likes to show his bəfɪzmənt"), and the other half were grammatically incorrect (e.g., "The teacher tries to səfɪsɪk at school"). There was no feedback in the testing phase. The correct response rates and RTs were obtained during the testing phase.

After each trial, participants were asked to indicate on a 4-point scale how confident they were in their decision (*no confidence, somewhat confident, very confident*, or *absolutely certain*) and what the basis of their judgment was (*guess, intuition, memory*, or *rule*). Explanations about what the English terms in the 4-point scale meant were provided in Japanese. No participant seemed to be confused about the differentiation between the terms. Even the distinction between *guessing* and *intuition* was relatively clear to participants. The phenomenology of *intuition* was labeled as "*chokkan*" in Japanese to mean knowing that a judgment is correct, but not knowing why, while guessing was labeled as "*atezuppou*" in Japanese to mean not knowing either (Dienes & Scott, 2005).

Different stimulus sentences were used on Day 1 and Day 8. After the testing phases on Day 8, participants filled out a debriefing questionnaire that asked whether they might have noticed any rules or patterns in the sentences they had heard, and if so, when they noticed them (during the training, the first testing phase, or the second testing phase). This was followed by an oral interview that elaborated on the questionnaire. Table 2 summarizes the overall design of the experiment.

Table 2. The design of the experiment

Participants	Learning phase	Testing phase	
		Day 1	Day 8
Incidental learning group	Exposure to sentences and picture matching (15 min)	Grammaticality judgment test 1. Correct response rate 2. Reaction time 3. Subjective measures of awareness (Retrospective verbal reports on Day 8)	
Explicit learning group	Explanation with example sentences (15 min)		
Control group	No learning phase		

Table 3. Results of a logit mixed-effects model for accuracy

	Fixed effects				Random effects	
					By subject	By item
Parameters	Estimate	*SE*	*z*	*p*	*SD*	*SD*
Intercept	0.66	0.24	2.80	.005	<.001	.470
Day	−0.25	0.14	−1.78	.076	—	—
Group	−0.21	0.10	−1.97	.049*	—	.058
Day × Group	0.10	0.07	1.56	.120		

Note: Formula: acc~Day*Group + (1|Subject) + (Group|Item); family = binomial.
* $p < .05$.

Results

GJT

Accuracy

The accuracies for Day 1 were 56.4% ($SD = 6.70$) for the incidental group, 56.8% ($SD = 4.88$) for the explicit group, and 51.5% ($SD = 5.48$) for the control group. The accuracies for Day 8 were 50.8% ($SD = 8.11$) for the incidental group, 56.4% ($SD = 7.81$) for the explicit group, and 52.4% ($SD = 5.44$) for the control group. The accuracies were analyzed by a logit mixed-effects model with crossed random effects for subject and item (Linck & Cunnings, 2015) using the lme4 package (version 1.1-18-1) of R (version 3.5.1; R Core Team, 2018). The dependent variable was a GJT performance specified by a binary outcome. The model included day (Day 1 and Day 8) and group (incidental, explicit, and control) as fixed effects. Because this maximal model failed to converge, the model was simplified by removing by-subject random slopes for day, which produced a stable model. The results are reported in Table 3. There were the significant fixed effect of group and the marginally nonsignificant fixed effect of day, but their interaction was not significant.

Learning was assessed by *A*-prime (*A'*) scores on the GJT.[3] The *A'* scores were calculated based on the proportions of hits (correct acceptance of a grammatical sentence) and false alarms (incorrect acceptance of an ungrammatical sentence).

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

Table 4. A' scores for incidental, explicit, and control groups on Days 1 and 8

	Day 1		Day 8	
Group	*M*	*SD*	*M*	*SD*
Incidental	0.609	0.107	0.537	0.115
Explicit	0.616	0.080	0.617	0.099
Control	0.527	0.102	0.539	0.080

Table 5. Results of a linear mixed-effects model for A'

	Fixed effects			Random effects
				By subject
Parameters	Estimate	*SE*	*t*	*SD*
Intercept	0.77	0.07	10.79	.046
Day	−0.10	0.04	−2.35*	—
Group	−0.08	0.03	−2.52*	—
Day × Group	0.04	0.02	2.07*	

Note: Formula: aprime~Day*Group + (1|Subject). *$|t| > 2.0$, indicating a significant effect (Gelman & Hill, 2007).

They provide a more sophisticated measure than simply reporting accuracy scores for the GJT according to signal detection theory (Grier, 1971; Linebarger, Schwartz, & Saffran, 1983; Macmillan & Creelman, 2005; Snodgrass, Levy-Berger, & Haydon, 1985). An *A'* score of .50 was taken to mean a chance performance. Table 4 shows *A'* scores for three groups on Days 1 and 8. The analysis of the scores was conducted using a linear mixed-effects model with only the subject as a random effect. The fixed effects were day and group. A fixed effect was considered significant if the absolute value of the *t* statistic was greater than or equal to 2.0 (Gelman & Hill, 2007; Linck & Cunnings, 2015). The results are summarized in Table 5.

The main effects of day and group and their interaction were significant. To interpret this interaction, Bonferroni pairwise comparisons were performed. The analysis produced the following results. The incidental group and the explicit group performed significantly better than the control group ($p = .028$, $d = .79$, 95% confidence interval; CI [.007, .158], and $p = .015$, $d = .97$, 95% CI [.014, .165], respectively) on Day 1. On Day 8 the explicit group performed significantly better than the incidental group ($p = .041$, $d = .75$, 95% CI [.002, .157]) and the control group ($p = .048$, $d = .87$, 95% CI [.001, .155]). Regarding the effect of time, the performance of only the incidental group significantly declined from Day 1 to Day 8 ($p = .016$, $d = .65$, 95% CI [.014, .130]), while the explicit and the control groups did not change their performance significantly after a week interval ($p = .974$, $d = .01$ and $p = .663$, $d = .13$, respectively).

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

Table 6. Reaction times (ms) for three groups on Days 1 and 8

	Day 1		Day 8	
Group	*M*	*SD*	*M*	*SD*
Incidental	2513	859	1532	556
Explicit	2907	1394	2284	1428
Control	2613	1411	1736	867

Table 7. Results of a linear mixed-effects model for reaction times (RT)

	Fixed effects			Random effects	
				By subject	By item
Parameters	Estimate	*SE*	*t*	*SD*	*SD*
Intercept	3571.80	666.40	5.36	1836.15	26.81
Day	−963.61	319.84	−3.01*	835.85	—
Group	−14.05	308.94	0.05	—	62.92
Day × Group	58.39	148.11	0.39		

Note: Formula: RT~Day*Group + (Day|Subject) + (Group|Item). *$|t| > 2.0$, indicating a significant effect (Gelman & Hill, 2007).

RT

The analysis was based on correct responses to the testing items only. Participants' RTs were measured from the end of the sentences until the participant pressed the button. Outlying RTs, at cutoff limits of ±2.5 *SD* from each participant's mean, were removed (2.01% of the data were removed). Table 6 presents RTs for three groups on Days 1 and 8. The analysis of RT was conducted using a linear mixed-effects model with crossed random effects for subject and item with day and group as the fixed effects. A fixed effect was considered significant if the *t* value was above 2.0 (Gelman & Hill, 2007; Linck & Cunnings, 2015). The results were summarized in Table 7. Only the effect of day was found to be significant, whereas the effects of group and the interaction between day and group were not found to be significant.

Retrospective verbal reports

The reports were composed of what participants wrote on the debriefing questionnaires and also their oral interviews in which participants were prompted to describe whatever they noticed during or after the experiments. No one in the control group could verbalize or describe anything related to the English derivational system. Participants made comments such as "*The same (or similar) sentences were heard*," "*The same words were used in the sentences*," "*The same subjects (he, she, doctor, teacher) were repeated*," "*The word order was the same*," and "*The voices were balanced between male and female*."

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

In the incidental group, 4 out of the 20 participants reported that they had decided the grammaticality based on the endings of words during the testing phases, although none of them could verbalize any correct rules or appropriate patterns included in the training and testing sentences. This means they might have possessed awareness at the level of noticing, but did not demonstrate awareness at the level of understanding (Schmidt, 1990, 1995, 2001).[4] Their accuracies on the GJT were 56.3%, 56.3%, 62.5%, and 64.6%. They were all better than the average of the control group (51.5%) and the same or above their group mean (56.4%). This indicates that awareness, even at the level of noticing, facilitates learning (Rogers, 2017; Rogers et al., 2016; Schmidt, 1990, 1995, 2001). These verbal reports should be viewed with caution, however, given the 1-week time difference between the learning phase and the administration of the verbal reports. Other comments included "*The same (or similar) sentences were heard*," "*The same words are used in the sentences*," "*The same subjects and places were used*," "*The voices were balanced between male and female*," "*I was focusing on familiar words*," "*I was focusing on verbs*," "*There were one or two unfamiliar words*," and "*Place names were often at the ends of the sentences*."

In contrast, most of the participants in the explicit group mentioned something regarding the derivational system or the word endings. Examples included "*I paid attention to word endings*" and "*I thought about word classes of unfamiliar words*." Other comments included "*The same (or similar) sentences were heard*," "*The word order was the same most of the time*," "*The voices were balanced between male and female*," "*I was thinking about word classes*," "*Familiar words were used*," "*Many adjectives were used*," and "*I was focusing on word order*."

Generally, participants did not report much in the debriefing session, arguably because they lacked the confidence in judging the sentence grammaticality correctly as evidenced by the very low confidence level in the ratings reported below. Meanwhile, no participant mentioned the similarity of the derivational suffixation between English and Japanese. The majority of the participants' comments seem to reflect the simple sentence structures containing familiar words used in the experimental materials.

Subjective measures of awareness

Whether the knowledge acquired during the learning phases is implicit or explicit was assessed using two types of subjective measure of awareness: confidence ratings and source attributions. Binominal tests were used to determine whether performance was above chance (Jackson, 2018; Marsden et al., 2013).

Confidence ratings. Table 8 reveals that the incidental group performed significantly above chance when they reported to be somewhat confident and to have no confidence only on Day 1. Yet the explicit group performed significantly above chance when they were somewhat or very confident on both Days 1 and 8. A logit mixed-effects model was used to compare the relationship between accuracy and confidence for the incidental group at Day 1 and the explicit group at Days 1 and 8, after confidence ratings were regrouped into two categories: less confidence (1 and 2) and more confidence (3 and 4; Rogers, 2017). Accuracy was specified as a binary outcome with confidence level as a fixed effect. The results show that in all three cases

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

Table 8. Accuracy (%) and number of responses across confidence ratings for three groups on Days 1 and 8

	Day 1			Day 8		
Group	Accuracy	Number	*p*	Accuracy	Number	*p*
Incidental						
No confidence	54.6	179	.043	53.8	213	.060
Somewhat confident	57.8	260	<.001	49.8	201	.520
Very confident	54.1	79	.141	52.9	64	.234
Absolutely certain	62.1	18	.068	36.4	8	.857
Explicit						
No confidence	53.7	239	.054	53.4	238	.071
Somewhat confident	59.8	219	<.001	56.7	200	.005
Very confident	62.1	59	.007	64.6	84	<.001
Absolutely certain	46.5	20	.620	47.6	10	.500
Control						
No confidence	52.5	224	.144	53.5	240	.066
Somewhat confident	49.9	196	.500	52.7	205	.132
Very confident	51.8	43	.330	45.1	32	.762
Absolutely certain	50.0	22	.440	53.1	17	.298

confidence and accuracy was significantly related ($p < .001$). This negative correlation (a higher accuracy with less confidence) did not satisfy the zero-correlation criterion (Dienes & Scott, 2005), but may suggest better performance with unconscious knowledge in the opposite way (Rebuschat, 2013).

Source attributions. As can be seen from Table 9, learners in the incidental condition performed significantly above chance when they based their decisions on guessing, intuition, and rule knowledge only on Day 1. An above chance performance when choosing "*Guess*" met the guessing criterion for implicit knowledge, indicating those learners developed some unconscious knowledge. On the other hand, learners in the explicit condition performed significantly above chance when based on memory and rule knowledge on Day 1, indicating their knowledge was largely conscious. However, they performed above chance when basing decisions on intuition, memory, and rule knowledge on Day 8. Their above-chance performance when choosing "*Intuition*" implies that they developed at least some unconscious knowledge (Dienes & Scott, 2005; Rebuschat, 2013).

Discussion

The present study investigated the effects of incidental and explicit learning conditions of the English derivational system by Japanese learners of English and the nature of the knowledge they acquired immediately after learning as well as a week

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

Table 9. Accuracy (%) and number of responses across source attributions for three groups on Days 1 and 8

	Day 1			Day 8		
Group	Accuracy	Number	*p*	Accuracy	Number	*p*
Incidental						
Guess	55.4	133	.041	52.3	170	.187
Intuition	54.3	195	.046	50.8	185	.357
Memory	55.6	104	.054	45.9	56	.792
Rule	61.5	107	.001	55.9	81	.067
Explicit						
Guess	54.4	147	.064	51.5	175	.275
Intuition	51.9	192	.218	55.3	188	.022
Memory	63.8	120	<.001	61.8	107	.001
Rule	62.9	83	.001	63.6	68	.002
Control						
Guess	49.9	187	.459	53.4	204	.084
Intuition	52.0	179	.209	51.5	206	.258
Memory	52.2	60	.288	53.3	65	.208
Rule	53.2	66	.209	54.0	27	.240

later. The findings of the present research with regard to the four research questions are discussed in this section.

Research Question 1: To what extent are adult learners able to acquire L2 derivational morphology incidentally through auditory exposure?

The incidental learners exhibited significant learning effects, at least immediately after the brief auditory exposure to the sentences in the learning phase. Because new nonwords were used in the test items, the participants presumably assumed that it was something about the nonword that was crucial to their decision. Therefore, unlike the training task, attention was drawn to the nonwords (possibly the suffixes) in this case, to the extent that they could be segmented without learners' awareness. In addition, the exposure sentences were composed of familiar words (except for the targeted nonwords) in the basic English word order, as many participants mentioned in their retrospective verbal reports. This may have provided salience to nonwords, including the target structure (Ellis, 2016; Endress, Scholl, & Mehler, 2005; Romberg & Saffran, 2013; Schmidt, 1990). Furthermore, as noted in the Stimuli section, the participants' first language (Japanese) also has a similar derivational system to English. Implicit learning of a certain system in a L2 is more likely to occur if that system is present in their first language (Leung & Williams, 2014).

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

With respect to the limited amount of learning demonstrated in the current research (i.e., 56.4% and 56.8% for incidental and explicit learners, respectively, immediately after learning, and 56.4% for explicit learners after a week), the current study does not differ much from similar studies reporting the accuracy of incidental learning. Grey et al. (2014) reported the average accuracy of 57.5% for the AJT and 56.3% for the delayed PMT. In Rogers et al. (2016) the average accuracy of the GJT was 55.4%. The similar amount of learning in the present study was notable considering that participants were found to be less confident and had comparatively higher degrees of unconscious knowledge compared to participants of the two previous studies. Thus, language acquisition may arguably take place like this at least in the initial stages of learning: a little above chance performance accumulates over time and thus develops a more fine-tuned learning system with a higher accuracy.

Research Question 2: To what extent is their performance comparable to that of explicit learners?

Considering the immediate measurements of learning, performance of the incidental learners was on par with that of the explicit learners. It was not the case, however, after a 1-week delay where incidental learners declined to a chance-level performance, while explicit learners still remained reliable in that they performed significantly better than the incidental and the control groups. The plausible reason for the difference regarding the delayed effects will be discussed in the response to the fourth research questions below.

One might wonder why the explicit learners were far from perfect (56.8% accuracy). There are three aspects of their learning conditions that seem relevant to this question. First, learners in the explicit condition did not receive any trials to which they had to respond intentionally. Andringa and Curcic (2015), for instance, provided explicitly instructed learners with 56 trials in which they performed PMT. They reached the accuracy of 83%. Explicit learners in the current study just saw the rule explanation on the computer screen and listened to the suffixes and sentences including suffixed nonwords. Second, the number of sentences they heard was small: 36 sentences compared to 96 sentences for the incidental condition despite the same amount of time spent on learning. Third, learning via the aural mode rather than written input could have been cognitively challenging throughout the course of learning for participants (Hama & Leow, 2010; Leow, 1995), in particular for Japanese learners of English who are arguably less proficient at understanding spoken English because they are, for the most part, taught English by reading and writing (Watanabe, 1988). The final possibility appears to be able to account for the generally low performance of all the participants in the present study.

Research Question 3: What is the nature of the acquired knowledge: Implicit or explicit?

Incidental learners were found to develop implicit knowledge given their lack of confidence and their judgment based on guess and intuition. The negative correlation between confidence and accuracy caused by a large number of choices of the "less confidence" category may be indicative of better performance with less

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

confidence (Rebuschat, 2013). At the same time, their performance at above chance levels based on rule knowledge suggests they also developed some explicit knowledge, although none of the participants were able to describe the derivational rule in the retrospective verbal reports. Considering that participants in the control group did not possess this kind of explicit knowledge, the explicit knowledge developed by incidental learners was arguably created through incidental exposure in the learning phase (Grey et al., 2014; Rebuschat et al., 2015; Rebuschat & Williams, 2012; Rogers et al., 2016). Furthermore, it was found that the nature and durability of the resultant explicit knowledge seems to differ substantially between the two types of learners, given that the explicit knowledge immediately developed by incidental learners diminished afterward, but the knowledge possessed by explicit learners persisted throughout the 1-week experiments. The reason for the learners' difference in the robustness of explicit knowledge is not clear, but it is worth noting that only the explicit learners seemed to demonstrate the consolidation effects or to engage in conscious hypothesis testing, as will be mentioned later in the discussion of Research Question 4.

Explicit learners were found to possess conscious knowledge both in the immediate and the delayed tests. Interestingly enough, explicit learners also developed unconscious knowledge 1 week later evidenced by their reliance on intuition (DeKeyser, 2015; Ellis, 2005). Hamrick and Rebuschat (2012) showed, too, that learners in an intentional learning condition developed some unconscious knowledge even immediately after exposure, evidenced by their above chance performance based on intuition.

It should be noted here that L2 learners' intuition is different in nature from native speakers' intuition, which develops to the level of automaticity after many years of exposure. A limited amount of exposure does not enable L2 learners to develop the intuition that native speakers possess (Leow & Hama, 2013; Rogers, 2017). In light of this, Rogers (2017) interpreted intuition as indicative of low levels of awareness (i.e., explicit knowledge). However, these measures were developed in the context of artificial grammar learning where there is also low levels of exposure. Therefore, in the present results, based on Dienes and Scott (2005) and Rebuschat (2013), the intuition attribution was considered to belong to unconscious knowledge.

The two types of learners were found to not be significantly different in the speed with which they accessed the knowledge they had acquired. The results of RT measures only showed that both types of learners responded faster a week later than the first day. As there was not a significant effect of the group or its interaction with time, it is not easy to explain why no significant differences in RT was detected between incidental and explicit conditions. Studies have suggested that incidental learners can be quick to access their unconscious knowledge when they possess implicit knowledge or can be slow if they attempt to find rules, whereas explicit learners may be quick to rely on their conscious knowledge explicitly explained beforehand, even though they do not possess implicit knowledge deployed to make rapid access possible (Robinson, 1997; Suzuki & DeKeyser, 2017). In the present study, given that only the explicit learners showed learning effects a week later, their judgment (which was found to be above chance) based on intuition, in addition to memory and rule knowledge on Day 8, might have contributed to their faster access

to their knowledge.[5] Decisions based on intuition reflecting implicit knowledge are likely to enable rapid access to unconscious knowledge (Suzuki & DeKeyser, 2017).

Another question is concerned with the relationship between the grammatical cues and suffixes the incidental group learned immediately after exposure. It seems that participants in the incidental group learned nonadjacent dependencies between the grammatical cues and suffixes over an intervening nonword (Gomez, 2002; Mintz, 2002, 2003; Romberg & Saffran, 2013; St. Clair, Monaghan, & Christiansen, 2010). For example, in the learning phase, they may have simply been learning the invariant relationships between "his" and "nəs," "very" and "fəl," and "to" and "aɪz." In the testing phase, they judged the sentence including the association, which broke this invariance such as the one between "his" and "aɪz" as ungrammatical above chance. Whether the participants in the present study did or did not learn the semantic functions of the suffixes cannot be known. Marsden et al. (2013) demonstrated, however, that after brief exposure to suffixed nonwords (e.g., *sifedec* presented orally and visually), surface forms of suffixes were learned without adding meanings to stems or suffixes. Furthermore, their subsequent experiments showed that when the meaning was given to either the stem or the suffix, the meaning of the other could be learned partly incidentally, which may indicate the learning of form–meaning connections without attention to target structures. Whether the nonadjacent dependency learning seen in the current study leads to the learning of grammatical categories is an interesting but difficult question to tackle (Mintz, 2002, 2003; St. Clair et al., 2010). Given that the participants in the present experiment were intermediate-level learners of English (as detailed in the Method section), it seems highly likely that they already possessed a sense of grammatical categories in English. They might have known that words with different natures (i.e., something to distinguish between nouns, adjectives, and verbs) follow the different grammatical units ("his" and "some" vs. "very" vs. "wants to" and "tries to").

Research Question 4: Are there any delayed effects of learning demonstrated by the different learning groups?

To reiterate some points mentioned above, incidental learners were found to not retain their learning for a week. This indicates that they may need additional exposure, for example, on the day after the immediate learning or at least another time within the week. Nevertheless, the learning gained by explicit learners was found to endure, and part of the resulting knowledge became unconscious later on. These results corroborate previous observations that explicit types of instruction are more efficient and the effects are more durable than implicit types (Andringa & Curcic, 2015; Ellis, 2015; Norris & Ortega, 2000; Spada & Tomita, 2010). Although these previous studies were not concerned with whether implicitly instructed participants were really unaware of what they had learned, the current research provides further evidence that explicit learning helps to develop more reliable knowledge.

The finding that the explicit knowledge gained by explicit learners became partly unconscious at the time of delayed testing is in line with the interface position, which postulates that explicit knowledge, with a certain degree of proceduralization and automatization, impacts and facilitates implicit knowledge (DeKeyser 2015;

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

Suzuki & DeKeyser, 2017). This finding may also be accounted for by consolidation effects (Grey et al., 2014, Tamminen et al., 2012, 2015). The question arises as to whether consolidation took place in the testing phase for learners in explicit learning conditions. In order to explore this issue, the test phase (48 sentences presented) on Day 1 was divided into two sessions: the first half of the test (24 sentences) and the second half of the test (24 sentences).[6] The accuracy of the incidental group was 56.5% for the first session and 56.3% for the second session, with no significant difference between the sessions, t (19) = 0.068, p = .946, d = .02, 95% CI [6.20%, 6.61%]. The accuracy of the explicit group was 53.5% for the first session and 60.0% for the second session, with no significant difference between the sessions, either, t (19) = 1.656, p = .114, d = .65, 95% CI [1.70%, 14.6%]. However, participants in the explicit condition revealed a medium effect size (d = .65; Plonsky & Oswald, 2014), in contrast with the very small effect size (d = .02) shown by participants in the incidental condition. There is some indication that participants in the explicit condition were improving their performance during the test phase; here they might have been engaging in conscious hypothesis testing in the testing phase (Grey et al., 2014; Tamminen et al., 2012, 2015). This might partially explain the reason for the explicit learners' learning durability. The medium size effect for the explicit group diminished on Day 8, t (19) = 0.066, p = .948, d = .02, 95% CI [6.35%, 6.77%], with their accuracy between the first half (56.5%) and the second half (56.3%). It seems that after the consolidation of their explicit knowledge on Day 1, learners in the explicit condition developed the stable knowledge of the derivational rule that was employed throughout the course of the test. Given the explicit learners' potential awareness raising during the test phase, one would expect the reliance on explicit knowledge to increase from Day 1 to Day 8.[7] To examine this issue, proportions of the correct responses for explicit learners across the four source categories (*guess, intuition, memory*, and *rule*) were compared between Day 1 and Day 8. None of the four categories revealed a significant difference, although the responses for the guess category approached that level (28.1% on Day 1 and 33.9% on Day 8); t (19) = 1.95, p = .066, d = .24, 95% CI [0.43%, 12.0%]. This indicates that participants in the explicit learning conditions did not change the extent of the reliance on explicit knowledge, even though they might have improved their performance in the testing phase on Day 1.

Morgan-short et al. (2012) demonstrated that in terms of neural measures, implicit training, as opposed to explicit training, leads to brain processing typical of native speakers, both on the first day and 1 to 5 days after instruction. Morgan-short et al. used almost the same amount of time in implicit and explicit training in the initial instruction of their experiment as the present research (13.5 min), and supplemented it by a considerable amount of practice. Performance of learners of both implicit training and explicit training were found to not differ in terms of behavioral measures at both times. This took place presumably because the participants in both groups in their study received extensive training over 1 to 5 days, and learners in the implicit condition were most likely aware of the grammar rules they were being trained on.

Grey et al. (2014) found the maintenance of learning effects of word order 2 weeks later, even without additional exposure to the stimulus sentences. The durability of the learning shown in their study might be associated with the observation

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

that 68% of their learners were able to state the word order rule for simple sentences on the debriefing questionnaires, and that only the learners who reported to be *very confident* in the confidence ratings performed the AJT significantly above chance immediately after exposure. Regarding case marking, 38% of Grey et al.'s participants correctly stated case-marking rule for *–ga* and *–0*, and 35% for *–ni*. The significant difference of case-marking learning between those conscious learners and those who did not report case-marking rules was not found with immediate testing, but did emerge at delayed testing. In contrast, no one in the incidental group in the current study stated the correct rule of the derivational suffixes in the debriefing session and the above chance level performance was limited to the learners who reported having *no confidence* or being *somewhat confident*. This comparison suggests that reportable knowledge and confidence may be key factors in maintaining or improving learning effects.

The finding in the current research that incidental learning resulted in similar learning effects to explicit learning, at least immediately after learning, may suggest the potential of implicit types of learning in acquiring L2 derivational systems. Because the test items were not included in the learning items, this made it possible to test learners' abilities to generalize the acquired knowledge to new items they encountered. Whereas explicit learners generalize newly learned rules persistently, presumably through a consolidation process, even incidental learners can develop their ability to generalize their acquired knowledge immediately after exposure to input. How to make it durable is a crucial issue to be explored in future work. From a totally different perspective, the present results may lead to another interpretation. Neither of the types of learning is particularly effective (no more than 60% of average accuracy) in the absence of meaningful input and/or practice opportunities. A long period of training with the large amount of practice is needed to make skills develop (DeKeyser, 1997, 2015).

There are several other issues that should be addressed in future research. First, the GJT used in the study did not tap into the production process of language. Although the implicit form of knowledge seems difficult to exploit in generation tasks (e.g., Jimenez, Mendez, & Cleeremans, 1996), it would be interesting to investigate whether the knowledge acquired under incidental or explicit conditions can transfer to production ability. Second, the present study did not consider individual differences, such as working memory, procedural memory, and personality (e.g., Denhovska & Serratrice, 2017; Denhovska, Serratrice, & Payne, 2016; Jackson, 2018; Tagarelli, Ruiz, Vega, & Rebuschat, 2016), which may influence different aspects of incidental and explicit learning. Finally, the current experiment did not examine the impact of prior knowledge about the target rule participants may have possessed. Instead, it compared the results of learning groups with those of the control group who had no learning phase. It is plausible to see the exposure materials based largely on a real L2 that participants are learning as a pedagogical advantage that secures ecologically valid forms of learning. An interesting avenue for future work would be to investigate how the derivational system of a completely new language can be learned under an incidental or explicit learning condition.[8]

https://doi.org/10.1017/S0142716419000304 Published online by Cambridge University Press

Conclusion

The current study has demonstrated that L2 derivational morphology can be auditorily learned, either incidentally or explicitly, immediately after exposure, and that learning can be maintained for a longer time only when learned explicitly. Incidental learners were found to develop primarily unconscious knowledge, whereas explicit learners were found to rely principally on conscious knowledge, part of which becomes unconscious later on.

New evidence presented in this study for incidental learning of derivational morphology and the consolidation effects gained through explicit knowledge adds a theoretical contribution to the growing literature on incidental and explicit learning of L2 grammar. The findings of this study utilizing a real L2 system as experimental stimuli may have direct pedagogical implications for L2 learning. Explicit instruction of grammatical rules may be more effective and durable than incidental learning, at least only with a one-time exercise. Incidental learning, albeit with immediate effects, seems to require extensive exposure over time to be effective.

Acknowledgments. I thank Associate Editor Dr. Annie Tremblay and three anonymous reviewers for their helpful comments on earlier versions of the manuscript. I am grateful to Dr. John Williams for his advice on the experimental procedure. This research was supported by a grant from the Japan Society for the Promotion of Science (No. 15K02768).

Supplementary Material. To view the supplementary material for this article, please visit https://doi.org/10.1017/S0142716419000304

Notes

1. A reviewer pointed out that one weakness of research in second language acquisition is that time of testing is manipulated within subjects rather than between subjects and that as a result of this, repeated testing becomes a potential confounding variable.

2. Participants also had studied one of the following languages as a third language for 6 months to 2 years: German, French, Spanish, Chinese, and Korean. The present study, however, did not ask each of the participants which third language they had studied.

3. I wish to thank one reviewer for suggesting the use of nonparametric A' instead of d' given the low overall performance of the experimental groups.

4. These participants might have made a link between the surface features and grammaticality, but they might have simply lacked the metalanguage or confidence to explain this during the debriefing session. Thanks to one reviewer for suggesting this possibility.

5. One reviewer commented that faster access on Day 8 was a testing effect given that there was not a significant main effect of group.

6. I would like to thank one reviewer for suggesting this analysis.

7. I thank a reviewer for posing this issue.

8. One reviewer commented that this study is still a semiartificial grammar experiment and lacks ecological validity. Thanks to the reviewer for this feedback.

References

Algeo, J. (1991). *Fifty years among the new words: A dictionary of neologisms*. Cambridge: Cambridge University Press.

Andringa, S., & Curcic, M. (2015). How explicit knowledge affects online L2 processing: Evidence from differential object marking acquisition. *Studies in Second Language Acquisition*, **37**, 237–268.

Bowles, M. A., & Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology: Expanding the scope of investigation. *Studies in Second Language Acquisition*, **27**, 415–440.

Brooks, P. J., & Kempe, V. (2013). Individual differences in adult foreign language learning: The mediating effect of metalinguistic awareness. *Memory and Cognition*, **41**, 281–296.

Carlisle, J. F. (2000). Awareness of the structure and meaning of morphologically complex words: Impact on reading. *Reading and Writing*, **12**, 169–190.

Chan, R. K. W., & Leung, J. H. (2018). Implicit knowledge of lexical stress rules: Evidence from the combined use of subjective and objective awareness measures. *Applied Psycholinguistics*, **39**, 37–66.

DeKeyser, R. M. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, **19**, 195–221.

DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94–112). New York: Routledge.

Denhovska, N., & Serratrice, L. (2017). Incidental learning of gender agreement in L2. *Journal of Psycholinguistic Research*, **46**, 1187–1211.

Denhovska, N., Serratrice, L., & Payne, J. (2016). Acquisition of second language grammar under incidental learning conditions: The role of frequency and working memory. *Language Learning*, **66**, 159–190.

Dienes, Z., Altmann, G., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **21**, 1322–1338.

Dienes, Z., & Scott, R. (2005). Measuring unconscious knowledge: Distinguishing structural knowledge and judgment knowledge. *Psychological Research*, **69**, 338–351.

Diessel, H. (2004). *The acquisition of complex sentences*. Cambridge: Cambridge University Press.

Ellis, N. C. (2005). At the interface: Dynamic interactions of explicit and implicit language knowledge. *Studies in Second Language Acquisition*, **27**, 305–352.

Ellis, N. C. (2015). Implicit AND explicit language learning: Their dynamic interface and complexity. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 3–23). Amsterdam: Benjamins.

Ellis, N. C. (2016). Salience, cognition, language complexity, and complex adaptive systems. *Studies in Second Language Acquisition*, **38**, 341–351.

Endress, A. D., Scholl, B. J., & Mehler, J. (2005). The role of salience in the extraction of algebraic rules. *Journal of Experimental Psychology: General*, **134**, 406–419.

Gass, S. M., & Mackey, A. (2007). *Data elicitation for second and foreign language research*. New York: Routledge.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, **13**, 431–436.

Goo, J., Granena, G., Yilmaz, Y., & Novella, M. (2015). Implicit and explicit instruction in L2 learning: Norris & Ortega (2000) revisited and updated. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 443–482). Amsterdam: Benjamins.

Graham, C. R., & Williams, J. N. (2016). Implicit learning of Latin stress regularities. *Studies in Second Language Acquisition*, **40**, 3–29.

Grey, S., Williams, J. N., & Rebuschat, P. (2014). Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition*, **36**, 1–34.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. *Psychological Bulletin*, **75**, 424–429.

Hama, M., & Leow, R. P. (2010). Learning without awareness revisited: Extending Williams (2005). *Studies in Second Language Acquisition*, **32**, 465–491.

Hamrick, P., & Rebuschat, P. (2012). How implicit is statistical learning? In P. Rebuschat & J. N. Williams (Eds.), *Statistical learning and language acquisition* (pp. 365–382). Berlin: de Gruyter.

Hamrick, P., & Sacks, R. (2018). Establishing evidence of learning in experiments employing artificial linguistic systems. *Studies in Second Language Acquisition*, **40**, 153–169.

Harwood, F. W., & Wright, A. M. (1956). Statistical study of English word formation. *Language*, **32**, 260–273.

Hirai, A., & O'ki, T. (2011). Comprehensibility and naturalness of text-to-speech synthetic materials for EFL learners. *JACET Journal*, **53**, 1–17.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning. *Studies in Second Language Acquisition*, **27**, 129–140.

Jackson, D. O. (2018). The potential relationship between openness and explicit versus implicit L2 knowledge. *Journal of Psycholinguistic Research*. Advance online publication. doi:10.1007/s10936-018-9603-6

Jarmulowicz, L. (2006). School-aged children's phonological production of derived English words. *Journal of Speech, Language, and Hearing Research*, **49**, 294–308.

Jiang, N. (2012). *Conducting reaction time research in second language studies*. New York: Routledge.

Jimenez, L., Mendez, C., & Cleeremans, A. (1996). Comparing direct and indirect measures of sequence learning. *Journal of Experimental Psychology: Leaning, Memory, and Cognition*, **22**, 948–969.

Kerz, E., Wiechmann, D., & Riedel, F. B. (2017). Implicit learning in the crowd: Investigating the role of awareness in the acquisition of L2 knowledge. *Studies in Second Language Acquisition*, **39**, 711–734.

Krashen, S. (1981). *Second language acquisition and second language learning*. Oxford: Pergamon.

Krashen, S. (1985). *The input hypothesis: Issues and Implications*. Lincolnwood, IL: Laredo Publishing.

Leow, R. P. (1995). Modality and intake in second language acquisition. *Studies in Second Language Acquisition*, **17**, 79–89.

Leow, R. P. (2018). ISLA: How implicit or how explicit should it be? Theoretical, empirical, and pedagogical/curricular issues. *Language Teaching Research*. Advance online publication. doi:10.1177/1362168818776674

Leow, R. P., & Hama, M. K. (2013). Implicit learning in SLA and the issue of internal validity: A response to Leung and Williams's "The implicit learning of mappings between forms and contextually derived meanings." *Studies in Second Language Acquisition*, **35**, 545–557.

Leow, R. P., & Zamora, C. C. (2017). Intentional and incidental learning. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 33–49). New York: Routledge.

Leung, J. H. C., & Williams, J. N. (2011). The implicit learning of mappings between forms and contextually derived meanings. *Studies in Second Language Acquisition*, **33**, 33–55.

Leung, J. H. C., & Williams, J. N. (2012). Constraints on implicit learning of grammatical form-meaning connections. *Language Learning*, **62**, 634–662.

Leung, J. H. C., & Williams, J. N. (2014). Crosslinguistic differences in implicit language learning. *Studies in Second Language Acquisition*, **36**, 733–755.

Linck, J. A., & Cunnings, I. (2015). The utility and application of mixed-effects models in second language research. *Language Learning*, **65**, 185–207.

Linebarger, M., Schwartz, M., & Saffran, E. M. (1983). Sensitivity to grammatical structure in so-called agrammatic aphasics. *Cognition*, **13**, 361–392.

Macmillan, N., & Creelman, C. (2005). *Detection theory: A user's guide* (2nd ed.). Mahwah, NJ: Erlbaum.

Marsden, E. J., Williams, J., & Liu, X. (2013). Learning novel morphology: The role of meaning and orientation of attention at initial exposure. *Studies in Second Language Acquisition*, **35**, 619–654.

Marslen-Wilson, W. D. (2007). Morphological processes in language comprehension. In G. Gaskel (Ed.), *The Oxford handbook of psycholinguistics* (pp. 175–193). Oxford: Oxford University Press.

Matsuda, N. (2017). Evidence of effects of text-to-speech synthetic speech to improve second language learning. *JACET Journal*, **61**, 149–164.

Ministry of Education, Culture, Sports, Science and Technology, Japan. (2015). Retrieved from http://www.mext.go.jp/b_menu/shingi/chousa/shotou/117/shiryo/__icsFiles/afieldfile/2015/11/04/1363335_2.pdf

Mintz, T. H. (2002). Category induction from distributional cues in an artificial language. *Memory & Cognition*, **30**, 678–686.

Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition*, **90**, 91–117.

Morgan-Short, K., & Bowden, H. W. (2006). Processing instruction and meaningful output-based instruction: Effects on second language development. *Studies in Second Language Acquisition*, **28**, 31–65.

Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman M. T. (2012). Explicit and implicit second language training differentially affect the achievement of native-like brain activation patterns. *Journal of Cognitive Neuroscience*, **24**, 933–947.

Noble, C. E. (1961). Measurements of association values (a), rated associations (a'), and scaled meaningfulness (m') for the 2100 CVC combinations of the English alphabet. *Psychological Reports*, **8**, 487–521.

Norris, J. M., & Ortega, L. (2000). Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, **50**, 417–528.

Paciorek, A., & Williams, J. (2015). Implicit learning of semantic preferences of verbs. *Studies in Second Language Acquisition*, **37**, 359–382.

Pinker, S., & Jackendoff, R. (2005). The faculty of language: What's special about it? *Cognition*, **95**, 201–236.

Plonsky, L., & Oswald, F. L. (2014). How big is "big"? Interpreting effect sizes in L2 research. *Language Learning*, **64**, 878–912.

R Core Team. (2018). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rebuschat, P. (2013). Measuring implicit and explicit knowledge in second language research. *Language Learning*, **63**, 595–626.

Rebuschat, P., Hamrick, P., Riestenberg, K., Sachs, R., & Ziegler, N. (2015). Triangulating measures of awareness: A contribution to the debate on learning without awareness. *Studies in Second Language Acquisition*, **37**, 299–334.

Rebuschat, P., & Williams, J. N. (2012). Implicit and explicit knowledge in second language acquisition. *Applied Psycholinguistics*, **33**, 829–856.

Robinson, P. (1996). Learning simple and complex second language rules under implicit, incidental, rule-search, and instructed conditions. *Studies in Second Language Acquisition*, **18**, 27–67.

Robinson, P. (1997). Generalizability and automaticity of second language learning under implicit, incidental, enhanced, and instructed conditions. *Studies in Second Language Acquisition*, **19**, 223–247.

Robinson, P. (2005). Cognitive abilities, chunk-strength, and frequency effects in implicit artificial grammar and incidental L2 learning: Replications of Reber, Walkenfeld, and Hernstadt (1991) and Knowlton and Squire (1996) and their relevance for SLA. *Studies in Second Language Acquisition*, **27**, 235–268.

Rogers, J. (2017). Awareness and learning under incidental learning conditions. *Language Awareness*, **26**, 113–133.

Rogers, J. (2019). Levels of awareness, depth of processing, and the learning of L2 case markings. In R. P. Leow (Ed.), *The Routledge handbook of second language research in classroom learning* (pp. 76–88). New York: Routledge.

Rogers, J., Révész, A., & Rebuschat, P. (2015). Challenges in implicit learning research: Validating a novel artificial language. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 275–299). Amsterdam: Benjamins.

Rogers, J., Révész, A., & Rebuschat, P. (2016). Implicit and explicit knowledge of inflectional morphology. *Applied Psycholinguistics*, **37**, 781–812.

Romberg, A. R., & Saffran, J. R. (2013). All together: Concurrent learning of multiple structures in an artificial language. *Cognitive Science*, **37**, 1290–1318.

Ruiz, S., Tagarelli, K. M., & Rebuschat, P. (2018). Simultaneous acquisition of words and syntax: Effects of exposure condition and declarative memory. *Frontiers in Psychology*, **9**, 1168. doi:10.3389/fpsyg.2018.01168

Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, **11**, 129–158.

Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning* (pp. 1–63). Honolulu, HI: University of Hawai'i Press.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge: Cambridge University Press.

Shibatani, M. (1990). *The languages of Japan*. Cambridge: Cambridge University Press.

Silva, R., & Clahsen, H. (2008). Morphologically complex words in L1 and L2 processing: Evidence from masked priming experiments in English. *Bilingualism: Language and Cognition*, **11**, 245–260.

Snodgrass, J. G., Levy-Berger, G., & Haydon, M. (1985). *Human experimental psychology*. New York: Oxford University Press.

Spada, N., & Tomita, Y. (2010). Interactions between type of instruction and type of language feature: A meta-analysis. *Language Learning*, **60**, 263–308.

St. Clair, M. C., Monaghan, P., & Christiansen, M. H. (2010). Leaning grammatical categories from distributional cues: Flexible frames for language acquisition. *Cognition*, **116**, 341–360.

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics*, **38**, 1229–1261.

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, **67**, 747–790.

Tagarelli, K. M., Ruiz, S., Vega, J. L. M., & Rebuschat, P. (2016). Variability in second language learning: The roles of individual differences, learning conditions, and linguistic complexity. *Studies in Second Language Acquisition*, **38**, 293–316.

Tamminen, J., Davis, M. H., Merkx, M., & Rastle, K. (2012). The role of memory consolidation in generalisation of new linguistic information. *Cognition*, **125**, 107–112.

Tamminen, J., Davis, M. H., & Rastle, K. (2015). From specific examples to general knowledge in language learning. *Cognitive Psychology*, **79**, 1–39.

Tao, Y., & Williams, J. N. (2018). Generalization of syntactic knowledge in semiartificial language learning. *Language Learning*, **68**, 1001–1031.

Tsujimura, N. (2007). *An introduction to Japanese linguistics* (2nd ed.). Oxford: Blackwell.

Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of Memory and Language*, **28**, 649–667.

VanPatten, B., & Smith, M. (2018). Word-order typology and the acquisition of case marking: A self-paced reading study in Latin as a second language. *Second Language Research*. Advance online publication. doi:10.1177/0267658318785652

Watanabe, K. (1988). Sentence stress perception by Japanese students. *Journal of Phonetics*, **16**, 181–186.

Williams, J. N. (2005). Learning without awareness. *Studies in Second Language Acquisition*, **27**, 269–304.

Williams, J. N. (2009). Implicit learning in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319–353). Bingley: Emerald Group Publishing.

Williams, J. N., & Kuribara, C. (2008). Comparing a nativist and emergentist approach to the initial stage of SLA: An investigation of Japanese scrambling. *Lingua*, **118**, 522–553.

Cite this article: Ishikawa, K. (2019). Incidental and explicit learning of L2 derivational morphology and the nature of acquired knowledge. *Applied Psycholinguistics* **40**, 1377–1404. https://doi.org/10.1017/S0142716419000304