CrossMark

Applied Psycholinguistics **37** (2016), 781–812
doi:10.1017/S0142716415000247

Implicit and explicit knowledge of inflectional morphology

JOHN ROGERS
Qatar University

ANDREA RÉVÉSZ
University College London Institute of Education

PATRICK REBUSCHAT
Lancaster University

Received: September 19, 2014 Accepted for publication: April 6, 2015

ADDRESS FOR CORRESPONDENCE
John Rogers, Foundation Program Department of English, Qatar University, PO Box 2713, Doha, Qatar. E-mail: rogers@qu.edu.qa

ABSTRACT
This study set out to test the degree to which second language inflectional morphology can be acquired as a result of incidental exposure and whether the resulting knowledge is implicit (unconscious) or explicit (conscious) in nature. Participants were exposed to an artificial language system based on Czech morphology under incidental learning conditions. In the testing phase, a grammaticality judgment test was utilized to assess learning. In addition, subjective measures of awareness and retrospective verbal reports were used to measure whether the acquired knowledge was conscious or not. The results of the experiment indicate that participants can rapidly develop knowledge of second language inflectional morphology under incidental learning conditions in the absence of verbalizable rule knowledge.

First language acquisition is frequently cited as a process that occurs implicitly (without awareness) and results in a tacit knowledge base (e.g., Ellis, 2002, 2008, 2011; Reber, 1967; Williams, 2009). However, whether implicit processes are involved in second language (L2) learning has been the focal point of much debate in recent years (see, e.g., the contributions in Andringa & Rebuschat, 2015; Hulstijn & Ellis, 2005; Rebuschat, 2015). Included in this debate is a renewed interest in the product of learning, specifically the degree that different exposure conditions lead to the development of implicit and/or explicit knowledge. This interest originates from Krashen's (1982) claims that L2 acquisition[1] is primarily an unconscious (i.e., implicit) process, comes about as a result of processing input for meaning, and explicit language instruction and metalinguistic awareness play little role in the acquisition process. Although some 30 years have passed since

© Cambridge University Press 2015 0142-7164/15

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Krashen first made these claims, researchers and theorists have yet to reach a consensus on the role of awareness in L2 acquisition. This lack of consensus is partly attributable to methodological issues in obtaining data that validly and reliably reflects the conscious status of the process of learning as well as learners' knowledge (R. Ellis, 2005).

It is important here to introduce several items of terminology that are central to the nature of this ongoing debate. With regard to the construct of *awareness*, the most influential model of attention and awareness in L2 research stems from Schmidt's noticing hypothesis (1990, 1993, 1994a, 1995, 2001, 2010). In this model, Schmidt differentiates between two levels of awareness that are argued to be facilitative to the learning process (Schmidt, 1994a). First, *awareness at the level of noticing* (noticing) refers to a lower level of awareness that involves the conscious registration of surface features of the input. For Schmidt, this level of awareness is largely isomorphic with attention (Schmidt, 1995). Second, reflecting a higher level of awareness, *awareness at the level of understanding* reflects awareness that includes rule and/or metalinguistic knowledge. Schmidt's ongoing position is that noticing is necessary for L2 acquisition to take place (e.g., 1990, 1993, 2001, 2010), whereas awareness at the level of understanding is facilitative, but perhaps not necessary.[2]

In addition to differentiating between levels of awareness that can occur during the learning process, researchers have used additional terminology to describe the process of learning (Schmidt, 1994a; cf. Leow, 2015). The term *incidental learning* refers to learning that occurs without the intent to learn. A narrower definition of this term was proposed by Schmidt to be the learning "of one thing (e.g., grammar) when the learner's primary objective is to do something else (e.g., communicate)" (1994a, p. 16). From a methodological perspective, incidental learning has been operationalized as an experimental condition in which participants are informed of neither the learning target nor that they will be tested later in the experiment (Hulstijn, 2003, 2013; Williams, 2009). Frequently, experiments operating in this area also include a "cover story" or task in their training conditions that is designed to orient participants' attention on the meaning of the input, rather than toward the grammatical features that will later be tested (e.g., see the plausibility judgment task in Rebuschat & Williams, 2012). In this view, incidental learning is best seen as an experimental condition designed to influence the manner in which participants approach the task(s), and consequently process the input, in the training phase of the experiment.[3]

The term *implicit learning* is often conflated with incidental learning (Hulstijn, 2007). This confusion can partly be attributed to implicit learning as learning that takes place incidentally, that is, without intention. However, implicit learning is further characterized by a lack of awareness throughout the learning process (Williams, 2009). In other words, when learning incidentally, participants may or may not become aware of the linguistic focus of the experiment. The key criterion when learning *implicitly* is that participants remain unaware of what is being learned throughout the experiment. *Explicit learning* is often equated with *intentional learning* in that it refers to learning that occurs with intent and conscious awareness of what is being learned (DeKeyser, 2003). An example of explicit learning would be scenarios in which people are deliberately trying to

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

learn, for example, students studying grammatical rules on which they will be later assessed at school.

In contrast to implicit and explicit learning, the terms implicit and explicit knowledge refer to the product of learning. Implicit knowledge is unconscious knowledge; in other words, implicit knowledge lies outside of awareness, is not verbalizable, and, as such, can only be inferred from behavior. Explicit knowledge is knowledge that lies within awareness and is often, though not always, verbalizable (e.g., DeKeyser, 2009; Dörnyei, 2009; Reber, 1993; Williams, 2009). Following DeKeyser (2009), the present article will treat the term implicit to be synonymous with unconscious, and explicit to be synonymous with conscious.

Although there has been a long-standing interest in L2 research in the development of implicit and explicit knowledge on theoretical grounds (cf. R. Ellis, 2005), the measurement of implicit and explicit knowledge has proven challenging methodologically for various reasons, which include, for example, differences in the definition and operationalization of these terms (cf. Dörnyei, 2009). Another complicating factor is that participants might employ both implicit and explicit knowledge to varying degrees in the testing phase of the experiment, depending on the nature of the task they are asked to perform (Bialystok, 1982; DeKeyser, 2009). To counter these difficulties, researchers have continued to develop an increasing range of instruments designed to probe learners' awareness in regard to both the process and the product of learning (for reviews, see Rebuschat, 2013; Robinson, Mackey, Gass, & Schmidt, 2012).

The current study has two objectives. The first is to examine the degree that L2 inflectional morphology can be acquired under incidental learning conditions. The second goal is to examine the conscious status of the product of learning. In other words, this study aims to examine whether the knowledge resulting from this incidental exposure is implicit or explicit in nature. Before reviewing the relevant literature on the acquisition of L2 morphology, the following section provides a brief overview of two instruments that are designed to tap into awareness of the product of learning: retrospective verbal reports and subjective measures of awareness (for a comprehensive review, see Rebuschat, 2013).

MEASUREMENT OF IMPLICIT AND EXPLICIT KNOWLEDGE

Several studies within the area of L2 research have attempted to tap participants' awareness of linguistic features that result from incidental exposure. Traditionally, in these studies, awareness has been operationalized as the ability to verbalize knowledge and has been measured through the use of retrospective verbal reports. As part of retrospective verbal reports, participants are asked to verbalize any rules or regularities they might have noticed during the experiment. In studies adopting retrospective verbal reports, knowledge is typically considered implicit if participants are able to demonstrate learning (e.g., through performance on a grammaticality judgment test) yet are unable to verbalize the underlying rule that is guiding their performance (see, e.g., Rebuschat & Williams, 2012). For example, within the artificial grammar learning (AGL) paradigm (e.g., Reber, 1967), participants are typically exposed to strings of letters under incidental learning conditions. Following the exposure phase, the participants are given a surprise

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

grammaticality judgment test (GJT) in which they are asked to judge new strings as being grammatical or ungrammatical based on the strings of letters they were exposed to during training. Following the grammaticality judgment task, participants complete the verbal reports, noting any rules or regularities they might have noticed during the experiment.

Although verbal reports do provide direct evidence of the degree to which participants are able to verbalize their knowledge, the veridicality of these reports has frequently been called into question (e.g., Leow, 2002; Schmidt, 1994b; Shanks & St. John, 1994). This lack of sensitivity has led to calls for researchers to include multiple measures of awareness within their experimental designs to increase validity (cf. Leow & Hama, 2013; Rebuschat, 2008, 2013; Robinson, 2008). More recent experiments within the fields of cognitive psychology and L2 research, which have set out to investigate the conscious status of the product of learning, have begun to triangulate data by employing subjective measures of awareness, specifically source attributions and confidence ratings (for overviews, see Dienes, 2008, 2010, 2013; Rebuschat, 2008, 2013) in addition to retrospective verbal reports.

Source attributions can be obtained alone (e.g., Dienes, 2004; Marsden, Williams, & Liu, 2013) or in addition to confidence ratings (e.g., Dienes & Scott, 2005; Rebuschat & Williams, 2012). Source attributions ask participants to identify the source of their judgment (e.g., guess, intuition, memory, *or* rule knowledge) for each decision they make during the testing phase (e.g., judging a sentence to be grammatical or ungrammatical). Knowledge can be considered implicit if performance is significantly above chance levels when participants base their decisions on intuition or when they believe themselves to be guessing (Dienes, Altmann, Kwan, & Goode, 1995). Above chance performance when participants attribute their decision to guessing is called the *guessing criterion*. 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 phase. Dienes et al. (1995; see also Dienes, 2008; Rebuschat, 2013, for discussions) hold that knowledge can be considered unconscious if there is no correlation between level of confidence and performance. This is called the *zero-correlation criterion*.

The results of subjective measures of awareness have been argued to reflect two different types of knowledge, with source attributions reflecting structural knowledge and confidence ratings reflecting judgment knowledge (see Dienes, 2004, 2008, 2010, 2013; Rebuschat, 2013; Rebuschat, Hamrick, Riestenberg, Sachs, & Ziegler, 2015; Rebuschat, Hamrick, Sachs, Riestenberg, & Ziegler, 2013, for discussions). According to Dienes (2004) structural knowledge represents the knowledge that participants acquire of the underlying structure of the stimuli during the experiment (e.g., knowledge of the sequence of letters in an AGL experiment). Later, when confronted with a new string (e.g., in the testing phase of the experiment), participants are able to develop a new type of knowledge, judgment knowledge, when they judge whether the structure of the new item shares the same structural characteristics as the items from the training phase (Dienes, 2013). Although it is assumed that conscious structural knowledge leads to conscious judgment knowledge (Dienes & Scott, 2005), it has been argued that unconscious structural knowledge can lead to judgment knowledge that is conscious

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

or unconscious (for a discussion, see Rebuschat et al., 2013). Thus, by logical extension, it is possible for source attributions to indicate that knowledge is fully conscious (conscious structural knowledge and conscious judgment knowledge), fully unconscious (unconscious structural knowledge and unconscious judgment knowledge), or partially unconscious (unconscious structural knowledge and conscious judgment knowledge).

A number of recent studies within L2 research have included subjective measures of awareness in their experimental design (e.g., Grey, Williams, & Rebuschat, 2014, 2015; Hamrick, 2013; Hamrick & Rebuschat, 2012, 2013; Rebuschat, 2008; Rebuschat & Williams, 2012; Serafini, 2013). The results from these studies have corroborated claims (e.g., Schmidt, 1994b) that awareness is not as straightforward of an issue as the results from retrospective verbal reports might suggest; in particular, that the development of implicit and explicit knowledge is not mutually exclusive. For example, recent research (e.g., Hamrick, 2013; Hamrick & Rebuschat, 2013; Rebuschat, 2008, experiment 6; Tagarelli, Borges Mota, & Rebuschat, 2011) has pointed toward the possibility of participants developing both implicit and explicit knowledge as a result of exposure, regardless of whether they were exposed to incidental or intentional conditions.

In short, the studies listed above, which have included subjective measures of awareness in their experimental design, are reflective of the renewed attention in L2 research toward the incidental learning of L2 grammar as well as the type of knowledge that is developed as a result of incidental exposure. However, most of the research carried out in this arena has so far addressed the incidental learning of L2 word order (e.g., Rebuschat & Williams, 2012; Williams & Kuribara, 2008) or noun–determiner systems (e.g., Hama & Leow, 2010; Leung & Williams, 2011, 2012; Rebuschat et al., 2013; Williams, 2005). To date, very little research has investigated the learning of L2 inflectional morphology resulting from incidental exposure with the aim of examining whether the resulting knowledge was implicit or explicit in nature (see Grey et al., 2014, for an exception). The following section provides a brief review of the research that has examined the acquisition of L2 morphology.

ACQUISITION OF L2 MORPHOLOGY

Within the area of L2 research, there has been a long-standing interest in the acquisition of L2 morphology. The sustained importance and relevance of the issue is well reflected in that two journals, *Language Learning* and *Language Teaching Research*, have devoted special issues to research on L2 morphology in recent years. This continued interest can partly be attributed to the large volume of published studies that have demonstrated that L2 learners process inflectional morphology differently from native speakers (e.g., Clahsen, Felser, Neubauer, Sato, & Silva, 2010; Jiang, 2004, 2007) and face considerable difficulties in acquiring both receptive and productive knowledge of L2 morphology (DeKeyser, 2005; Hopp, 2013; Larsen-Freeman, 2010; White, 2009).

It is important that research has also attested that, although often with substantial effort, L2 learners can process and acquire L2 inflections. However, the extent of difficulty that learners encounter in the processing and subsequent acquisition of

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

L2 morphemes is hinged on a variety of factors (Gor, 2010). These factors include type and token frequency of the target inflection in the input (e.g., Ellis, 2002; Ellis & Schmidt, 1997, 1998), first language (L1) background (Ellis & Sagarra, 2010a, 2010b, 2011; Rast, 2010), individual difference variables (e.g., Brooks, Kempe, & Sionov, 2006; Kempe & Brooks, 2008; Kempe, Brooks, & Kharkhurin, 2010), metalinguistic awareness of the target morphological constructions (Brooks & Kempe, 2013; Grey et al., 2014; Marsden et al., 2013; Robinson, 2005), and type of instructional condition (e.g., Alanen, 1995; DeKeyser, 1995).

Although a number of studies have investigated the learning of L2 case-marking systems (e.g., Brooks, Kempe, & Donachie, 2011; Brooks et al., 2006; Hinz, Krause, Rast, Shoemaker, & Watorek, 2013; Kempe & Brooks, 2008; Kempe et al., 2010; Kempe & MacWhinney, 1998; Rast, 2008, 2010), thus far, very little research has examined the degree to which L2 case marking can be acquired under incidental conditions, in particular after a minimal amount of exposure, and even fewer studies have set out to examine the type of knowledge that is acquired as a result of this exposure. To our knowledge, this gap in the literature has been addressed by only four studies to date (Brooks & Kempe, 2013; Grey et al., 2014; Robinson, 2002, 2005).

Grey et al. (2014) adapted the artificial language developed by Williams and Kuribara (2008) in order to examine the acquisition of both L2 word order and case marking. Like Williams and Kuribara, Grey et al. used "Japlish," a semiartificial language consisting of English words and Japanese syntax, in which the English words were inflected according to their function in the sentence (e.g., *Stacey-ga that picture-o painted*). In the training phase, participants ($N = 36$) in Grey et al. were exposed to 128 Japlish sentences auditorily, and were asked to judge the semantic plausibility of each of these sentences. The testing phase, in Williams and Kuribara's original (2008) study, consisted solely of a GJT that assessed the learning of L2 word order. Grey et al. modified the testing phase of the study to include a GJT, designed to test the learning of L2 word order, as well as a picture-matching task, which assessed the learning of the morphological case markers. In the picture-matching task, participants had to indicate if the sentence (e.g., *Man-o woman-ga necklace-ni gave*) matched the picture displayed on the monitor (in this case, a man giving a necklace to a woman). The testing phase of this study was further modified to include both immediate and delayed posttests. In addition, confidence ratings and source attributions were taken alongside the GJT, but not with the picture-matching task. This was followed by retrospective reports at the end of the testing phase.

The results from the GJT showed a significant learning effect for word order on both immediate and delayed posttests. An analysis of the confidence ratings and source attributions from the GJT indicated that participants were most accurate when basing their decisions on intuition, rather than on explicit knowledge of the L2 syntax. In other words, the participants developed some implicit knowledge of the L2 word order as a result of incidental exposure. In the case of the acquisition of case marking, the results from the picture-matching task showed a significant learning effect for case marking only on the delayed posttest. In addition, an analysis of the retrospective verbal reports indicated that the overall learning effect of L2 case markings was driven by participants who expressed awareness of the

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

morphological system. This suggests that the knowledge of L2 case markings acquired in this experiment was largely explicit in nature. However, because the subjective measures of awareness were only taken on the GJT (assessing knowledge of word order), and not on the picture-matching task (assessing knowledge of the case markers), it is not possible to say if participants were partly basing their decisions on intuition, rather than solely on explicit knowledge of the underlying morphological rule system.

In sum, the results of Grey et al. (2014) indicate that it is possible to acquire some knowledge of a novel morphological system following incidental exposure. However, a few limitations of this study need to be acknowledged. First, as mentioned above, although subjective measures of awareness were taken alongside the GJT, the picture-matching task was not accompanied by such measures. As a result, any conclusions regarding the nature of the knowledge acquired with regard to case markers can solely be inferred based on the results of the offline verbal reports, jeopardizing the construct validity of the findings. Second, the design of the picture-matching task did not allow for an analysis of performance on individual morphological markers. For example, in the testing phase, the participants had to indicate if the presented sentence matched the picture displayed on the screen. Because the stimulus sentences for this test (e.g., *Man-o woman-ga necklace-ni gave*) contained both nominative and objective case-markers, it was not possible to distinguish the quality of participants' performance on the two morphological markers. The learning effect on one of the case endings might have been masked by poor performance on the other case ending.

In contrast to Grey et al. (2014), an earlier study by Robinson (2002, 2005) did afford comparisons across isolated morphological case markers. In this study, a miniature Samoan system was used to examine the acquisition of L2 morphology, specifically locative, ergative, and noun incorporation case markers. Participants in this study, L1 speakers of Japanese, were randomly assigned to an experimental ($n = 38$) or a control ($n = 17$) group. Participants in the experimental group were trained to criterion on translation equivalents of the Samoan lexis that would be used in the next phase of the study. Following this training, a Samoan sentence was presented to the participants on the monitor for 10 s. Based on the content of this sentence, participants had to answer a yes/no comprehension question, and were given feedback after each answer. This training included 10 trials, each of which consisted of 45 sentences. Thus, participants were exposed to a total of 450 target structures. Finally, participants completed immediate, 1-week, and 6-month delayed written and aural GJTs, as well as a retrospective written awareness questionnaire. The results of this study indicated a significant learning effect only for the locative markers on the immediate and delayed posttests. However, like in Grey et al. (2014), the results of the retrospective questionnaire indicated that participants were largely able to verbalize the rule governing the use of the locative marker, suggesting that the overall learning effect was driven by explicit knowledge.

Finally, Brooks and Kempe (2013) set out to examine the role of cognitive predictors of individual differences and metalinguistic awareness in the learning of gender agreement and case marking. Like in Robinson (2002, 2005), Brooks and Kempe (2013) utilized a miniature natural language as the stimulus material

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

for this experiment (in this case, an artificial system based on Russian with two gender agreement and three case-marking rules). The exposure phase of this experiment consisted of six training sessions administered over a 2-week period. Each training session lasted on average around 45–60 min and consisted of various comprehension and production tasks designed to draw participants' attention to the meaning of Russian phrases containing the target linguistic structures. Following the final training phrase, participants ($N = 77$) completed an immediate posttest that required the participants to verbally produce new Russian phrases following a prompt, thus demonstrating their ability to generalize their knowledge of the target inflectional system to novel test items. These responses were scored on the accuracy of the participants in producing the correct gender and case markers. Following testing, participants completed retrospective verbal reports designed to elicit the degree that they noticed the inflected morphological forms during the exposure phase as well as their metalinguistic awareness of the underlying morphological rules of the artificial language.

An analysis of the participants' performance across the later training blocks indicated a near ceiling performance on tasks that targeted comprehension of Russian phrases, suggesting that the extended exposure phase resulted in high levels of recognition knowledge on the part of the participants. In regard to production, however, an analysis of the exposure and testing phases indicated a high degree of variance in the performance of the participants. Further analyses indicated that the performance of individuals was mediated by their ability to verbalize the rules governing the case-marking system. Taken together, the results of Brooks and Kempe (2013), Robinson (2002, 2005), and Grey et al. (2014) clearly show that explicit knowledge of inflectional morphology was linked to improved performance. Furthermore, these results suggest that awareness plays an important, if not vital, role in the learning of inflectional morphology, regardless of the intentionality of the learning condition.

SUMMARY AND RESEARCH QUESTIONS

In summary, despite the attention in recent years to both the incidental learning of L2 grammar and the acquisition of L2 morphology in particular, few studies (Grey et al., 2014; Robinson, 2002, 2005) have directly examined the acquisition of L2 case marking under incidental learning conditions after a minimal amount of exposure. Furthermore, in both of these studies, the verbal reports provided little evidence to substantiate the claim that L2 case markers can be acquired in the absence of conscious awareness.

Building and expanding on the work of Grey et al. (2014) and Robinson (2002, 2005), we designed an experiment to further investigate the extent to which incidental learning conditions can promote the acquisition of L2 case marking. In addition, we explored whether any knowledge acquired after incidental exposure is implicit or explicit in nature. It should be reiterated here that previous research by Grey et al. (2014) did include subjective measures of awareness on tasks measuring knowledge of the L2 word order system, but not on the picture-matching task, which was designed to measure knowledge of the case marking. Thus, a methodological improvement of our research was that, unlike Grey et al. and

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Robinson, we utilized subjective measures of awareness in addition to retrospective verbal reports (for a review, see Rebuschat, 2013). This allowed us to come to a fuller understanding of the type of knowledge that resulted from incidental exposure. Thus, this research not only contributes to the discussion surrounding the issue of incidental acquisition of L2 morphology but also informs the thorny debate of whether a dissociation between learning and awareness is possible. In particular, this study examines whether it is possible to develop knowledge of L2 grammar, namely, case marking, in the absence of conscious awareness of what has been acquired. In addition, on a methodological level, we sought to investigate whether subjective measures can provide additional insight into the conscious status of the acquired knowledge. Acquisition/learning was operationalized here as a statistically significant performance, when compared with a control group, on an immediate GJT.

We formed two research questions. To what extent can L2 case marking be acquired under incidental learning conditions? What type of knowledge is acquired as a result of this exposure, implicit or explicit?

METHOD

This experiment focused on the incidental learning of Czech morphology, specifically the nominative marker *–a* and the accusative marker *–u*, by means of an artificial language paradigm. This experiment consisted of an exposure phase, a testing phase, and a short debriefing session, during which subjects completed a short written questionnaire and an oral interview. Prior to the outset of this experiment, ethical approval was granted by the review board of the host institution. In addition, two pilot studies were carried out in order to validate both stimulus materials and experimental procedure.

Participants

Fifty-one undergraduate native speakers of English at a university in the United Kingdom volunteered to take part in this study. Nine of these volunteers were excluded from the present study on the basis of having previous experience with a language with a rich morphological system, such as Latin. The remaining 42 participants (26 female, 16 male) were randomly assigned to an experimental ($n = 21$) and a control ($n = 21$) group. None of the participants selected for this study majored in linguistics or foreign languages. The ages of the participants ranged from 18 to 25 years ($M = 20.03$, $SD = 1.86$).

Stimulus material

Training set. A semiartificial morphological system, based on Czech case marking, was used to generate the stimulus material for this experiment. As shown in Table 1, the system consisted of English phrases and a Czech noun, which was inflected according to one of two cases (nominative *–a* and accusative *–u*), depending on its function in the sentence (subject or object).

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Table 1. *Descriptions and examples of the two morphological categories*

Morphological Category	English Equivalent	Examples
Nominative	Subject	*The britva cut David's face at the sink last night.*
Accusative	Direct object	*Peter used a britvu in the bathroom today.*

A total of 24 Czech nouns, foreign words to the participants, were used in the training set. All of the Czech words were regular, feminine nouns that end with the inflection *–a* in their nominative form. All of the nouns followed the same pattern of declension. Only nouns with relatively "concrete" meanings were chosen in order to ensure that the nouns could be easily represented visually through images during training (see Appendix A).

For the exposure phase, 48 clip-art images were collected. Twenty-four of these images corresponded to the foreign Czech words used in the training set, and 24 images were distractor images that did not correspond to any of the foreign words. Each of the 48 images was used three times in the exposure phase of the study. The distribution of the distractor images was balanced throughout the exposure phase so that they did not occur more than once with any particular Czech noun.

Two stimulus sentences were created for each of the 24 Czech nouns in the training set. Of these 2 sentences, 1 sentence was designed so that the Czech noun occurred in the nominative case, and 1 sentence included the noun in the accusative case. That is, each Czech noun occurred twice in the training set, each time in a different sentence and each time with a different function (nominative or accusative) and case marking. The entire training set was repeated a total of three times during the exposure phase of the experiment. In other words, subjects were exposed to 144 sentences in total, 72 sentences for each of the two morphological categories.

In addition, to the inflected Czech noun, the word order in each sentence was scrambled according to two syntactic patterns (see Table 2 for syntactic patterns and example sentences). The word order was scrambled for several reasons. First, scrambled word order is more reflective of natural languages with rich morphologies. Like Czech, such languages rely on case endings rather than word order to mark functions within the sentence. Second, increasing the overall syntactic complexity of the stimulus material would increase the overall complexity of the artificial system, thus making it more likely for implicit learning to occur (e.g., Reber, 1993; Rebuschat, 2013). Third, the syntactic patterns allowed for controlling the position of the foreign or Czech word within the sentence. This ensured that the position of the foreign word could not serve as a reliable indicator of its function within the sentence.

All sentences in the training phase were designed so that an approximate meaning of the foreign word could be inferred by the participants from the rest of the sentence. In the construction of the training and testing sets, care was taken to control for the length of the sentences. Each sentence had exactly 12 syllables,

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Table 2. *Templates, sample sentences, and frequencies for the two syntactic patterns*

Pattern	Template	Frequency in Training Set
Pattern 1	$[[AP]_{TEMP} > [NP]_{OBJ} > [VP] > [NP]_{SUBJ} > [PP]]$	
Nominative	*Last summer the grass ate the koza in the field.*	(12)
Accusative	*Last month the kasu opened Patrick with the key.*	(12)
Pattern 2	$[[AP]_{TEMP} > [NP]_{SUBJ} > [VP] > [NP]_{OBJ} > [PP]]$	
Nominative	*Last year the studna supplied water to the town.*	(12)
Accusative	*Last night the sailor raised the kotvu with the chain.*	(12)

Note: AP, Adjective phrase; NP, noun phrase; VP, verb phrase; PP, prepositional phrase.

and an independent-samples *t* test confirmed that number of words per sentence was not a reliable indicator of morphological category during the exposure phase, $t\ (46) = 0.44$, $p = .66$. In addition, we controlled for frequency of the English words used in the stimulus materials. These were among the 2,000 most frequent English words as determined by Lextutor's vocabulary profile program (Cobb, 2012). The sentences were also comparable in terms of lexical diversity (types per tokens, $M = 0.04$), lexical density (content words per total number of words, $M = 0.57$), and average word length (number of syllables per word, $M = 1.28$, $SD = 0.48$).

Testing set. The testing set consisted of 48 new sentences. These consisted of novel Czech words inserted in English sentences that had not occurred in the training set. Our test items thus allowed us to test whether subjects could generalize the acquired knowledge to novel sequences. The sentences of the testing set were designed with the same considerations as the ones in the training set, controlling for the total number of syllables per sentence, as well as lexical and syntactic complexity. A frequency analysis of the testing set indicated that the difference in average stimulus length between grammatical ($M = 9.46$) and ungrammatical ($M = 9.34$) was nonsignificant, $t\ (46) = 0.87$, $p = .39$. There was also no significant difference between the sentence length of the stimulus materials used in the training ($M = 9.33$ words) and the testing ($M = 9.40$ words) phases, $t\ (94) = -0.56$, $p = .58$. This indicates that sentence length could not serve as a reliable predictor of grammaticality during the testing phase.

The ungrammatical items in the testing set were generated by replacing the correct case marking with the other case marker that had also been present in the training set. Out of 48 total items in the testing phase, 24 were nominative (*–a*), and 24 accusative (*–u*). Twelve of the nominative items were grammatical and 12 were ungrammatical. The 12 ungrammatical nominative items were created by replacing the nominative marker (*–a*) with the accusative marker (*–u*), for example, *in summer the lisku ate the eggs in the field*. The same procedure was followed in creating the accusative items, for example, *Last weekend he baked the*

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

kachna in the oven. In sum, care was taken to ensure that participants could only make correct judgments in the testing phase if they were able to identify instances of correct and incorrect case marking. Examples of the testing set can be found in Appendix B.

Procedure

As noted above, the experiment consisted of an exposure phase, a testing phase, and a debriefing session.[4] The experimental group completed all phases; the control group only took part in the testing phase and the debriefing session. The exposure and testing phases of the experiment were delivered via the stimulus presentation software Superlab 4.5 (Cedrus Corporation, 2012). Following the testing phase, participants completed a written debriefing questionnaire, followed by an oral interview.

During training, participants were exposed to the stimulus material under incidental learning conditions. This experiment operationalized incidental learning as an experimental condition in which participants are not informed that they will later be tested (Hulstijn, 2003, 2013). Furthermore, following common practice in recent studies of incidental learning (e.g., Grey et al., 2014; Hamrick & Rebuschat, 2012, 2013; Rebuschat & Williams, 2012), the exposure phase was deliberately designed to disguise the real purpose of the training task. Participants were told that they were going to take part in a study on learning foreign language vocabulary. Their task was to listen to a sentence, repeat the sentence aloud, then repeat the foreign word aloud, and finally to match the meaning of the foreign (Czech) word (e.g., *britva*) to one of two pictures (see description of images above) displayed on the monitor (e.g., a razor or a broom). Participants were given no feedback on the accuracy of their decision. Figure 1 illustrates the training procedure.

At no point during the exposure phase were participants informed that the Czech nouns were inflected for case nor that they would be tested afterward. Participants listened to all 144 sentences without a break. These sentences were presented in randomized order for each participant. The entire exposure phase took an average of about 25 min to complete.

Testing phase. Following the exposure phase, participants in the experimental group were informed that the sentences in the previous section were not arbitrary but were part of a complex system. They were then told that they would listen to 48 new sentences, half of which belonged to the complex system, half of which did not. For each test sentence, participants had to decide as quickly as possible if the sentence belonged to the system or not. No feedback was provided on the accuracy of the participants' decisions. In addition, for each grammaticality judgment, participants were also asked to report their overall level of confidence (no confidence, somewhat confident, very confident, or absolutely certain) and the basis of their judgment (guess, intuition, memory, or rule). Participants were provided with instructions and examples of how to select the appropriate confidence and source categories, which is especially important with regard to the distinction between "guess and "intuition." See Figure 2 for an illustration of the testing procedure.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

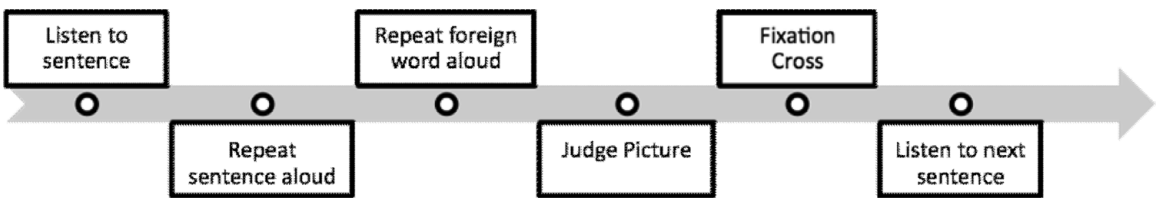


Figure 1. Procedure of training phase.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

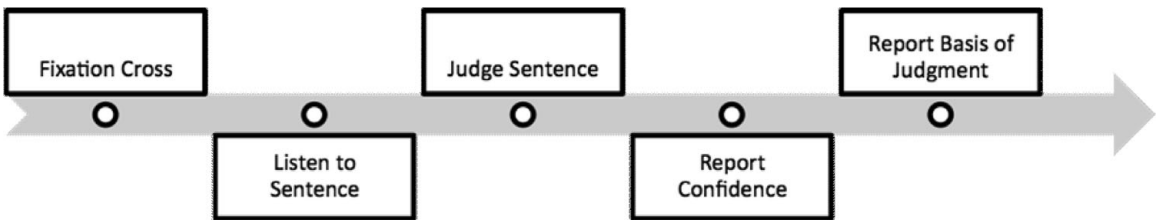


Figure 2. Procedure of testing phase.

Debriefing session. At the end of the experiment, participants were prompted to describe any rules or patterns they might have noticed and to specify at which point in the experiment (e.g., training or testing) this had occurred. In the oral interview, the researcher asked the candidates to elaborate on their responses in the written questionnaire. Finally, the researcher explained the underlying rule system and asked the participants again if they had figured this out or had any intuition about it at any point during the experiment.

Statistical analyses

Performance on the GJT was analyzed using both *d*-prime scores and mean accuracy rates. The *d*-prime scores are based on signal detection theory (for an overview, see Macmillan & Creelman, 2005) and are reported to be a more accurate measure than overall mean accuracy in that they correct for response bias. In the case of the present experiment, a *d*-prime score of 0 can be interpreted as chance performance. A *d*-prime score significantly higher than 0 indicates that the participants are able to discriminate between grammatical and ungrammatical items, with higher *d*-prime scores indicating superior performance.

Before carrying out each analysis, the distribution of scores were plotted and examined for skewness and kurtosis, and Shapiro–Wilk tests of normality were run via SPSS. Unless otherwise noted, the results revealed the data to be normally distributed. Means, standard deviations, exact *p* values, 95% confidence ratings, and effect sizes are reported for all tests of significance.

Before examining the subjective measures of awareness to determine the degree to which the exposure phase resulted in implicit and/or explicit knowledge on the part of the participants, it must first be established that the incidental training conditions produced a significant learning effect. This is necessary because subjective measures of awareness are designed to disentangle whether knowledge acquired during an experiment is implicit or explicit in nature. If the results of the experiment do not indicate a significant learning effect, then it can be assumed that no knowledge, implicit or explicit, was acquired as a result of incidental exposure (see Rebuschat, 2013).

RESULTS

Performance on the GJT served as the measure of learning. Confidence ratings, source attributions, and retrospective verbal reports served to distinguish implicit and explicit knowledge.

Performance on the grammaticality judgment task

The results revealed that the experimental group ($M = 55.44\%$, $SD = 7.00\%$) significantly outperformed the untrained controls ($M = 49.71\%$, $SD = 5.80\%$) on the GJT, $t\ (40) = 3.166$, $p = .003$, 95% confidence interval (CI) = (2.20%, 10.20%), $d = 0.89$. The average *d*-prime score on the GJT for the experimental group was

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Table 3. *Descriptive statistics (A-prime) and results of independent samples t tests for experimental and control groups across nominative, accusative, S-V-O, and O-V-S test items*

Item Type	A'			Results				
Group	*M*	*SD*	*SE*	*t*	*df*	*p*	95% CI	*d*
Nominative								
Experimental	0.530	0.203	0.044	1.338	40	.19	−0.055, 0.223	0.44
Control	0.446	0.180	0.039					
Accusative								
Experimental	0.640**	0.200	0.044	2.012	40	.05	0.011, 0.212	0.67
Control	0.529	0.125	0.027					
Type 1 S-V-O								
Experimental	0.555	0.190	0.041	1.10	40	.28	−0.044, 0.222	0.39
Control	0.466	0.265	0.058					
Type 2 O-V-S								
Experimental	0.614**	0.145	0.031	2.338	40	.02	0.013, 0.201	0.73
Control	0.507	0.147	0.031					

Note: S-V-O, Subject–verb–object; O-V-S, object–verb–subject.
*$p < .05$. **$p < .01$. The asterisks refer to the results of a one-sample *t* test (performance against chance levels).

0.314 ($SD = 0.369$), which was significantly higher than the average *d*-prime score of the control group, −0.035 ($SD = 0.494$), $t\,(40) = 2.587$, $p = .013$, 95% CI = (0.08, 0.62), $d = 0.80$. In addition, the *d*-prime analysis revealed several instances in which the number of false alarms exceeded the number of hits in the performance of individual participants, thus calling into question the appropriateness of the parametric *d*-prime analysis. To address this problem, nonparametric *A*-prime sensitivity scores were calculated, which have been argued to be a more sensitive measure under such circumstances (see, e.g., Grier, 1971; Stanislaw & Todorov, 1999; Zhang & Mueller, 2005). The average *A*-prime score for the experimental group was 0.585 ($SD = 0.097$), which was significantly different from the control group, 0.487 ($SD = 0.139$), $t\,(40) = 2.645$, $p = .012$, 95% CI = (0.02, 0.17), $d = 0.82$. Thus, the results of the *A*-prime analysis corroborated the results from the previous parametric analyses in indicating that participants were able to successfully discriminate between grammatical and ungrammatical case endings in the testing set.

As Table 3 demonstrates, further *A*-prime analyses uncovered that while the experimental group performed significantly above the control group on accusative markers, the scores of the two groups did not significantly differ on nominative items (see Table 3). Finally, participants in the experimental group demonstrated superior performance on items written in syntactic Pattern 2 (object–verb–subject), but did not perform significantly better than the control group on test items following syntactic Pattern 1 (subject–verb–object).

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Because the data above point toward differential performances within categories related to morphology (nominative vs. accusative) as well as within syntax (Pattern 1 vs. Pattern 2), further analyses were carried out in order to control for item effects and participant effects in the data. To this end, a logit mixed-effect regression model (Jaeger, 2008; for an overview of mixed-effect models in L2 research, see Cunnings, 2012) was carried out using R (R Core Team, 2014) and lme4 (Bates, Maechler, Bolker, & Walker, 2014) to examine the relationship between group (experimental vs. control) and performance. Within this model, performance was modeled as a binary outcome, "group" as a fixed-effect and "item" and "subject" as crossed random intercepts. A null model, which incorporated only the random effects for the intercepts of item and subject, was computed. Following this, a model was calculated that included group as a fixed effect. A likelihood ratio test was carried out to compare the null model to the model incorporating the effects of group. The results of this test indicated that the latter model led to a significant improvement in fit in comparison to the null model, χ^2 (1) = 6.78, $p < .001$, $\varphi = 0.08$. These results confirm a significant difference in performance between the experimental and control groups, while statistically controlling for the effects of subject and item variance.

Retrospective verbal reports

The analysis of retrospective verbal reports indicated that all participants (21/21) reported noticing the morphological inflections at the end of the foreign words during the training phase of the experiment. Comments such as "the foreign words always ended in vowels" or, more specifically, "all the words ended in *–u* or *–a*" as well as "the endings of the words changed" were taken as evidence that the learners had attended to and subsequently become aware of the surface features of the morphological markers. Furthermore, these comments also suggest that the learners had developed some awareness beyond the surface forms, specifically that these inflections were part of a pattern or behaved in a certain way. Despite this, none of the participants were able to demonstrate understanding of the underlying rule system by verbalizing the rules governing the use of the inflections. When prompted to guess, 13 participants stated that the affixes might have represented noun class, such as gender. Two other subjects mentioned that they thought that the inflection might be connected in some manner with the position of the foreign word in the sentence, but they could not explain the nature of this connection. At the very end of the experiment, when the rules were explained to the participants, none of the participants claimed that they had thought of the rules at any point during the experiment. In sum, the analysis of the retrospective verbal reports thus suggests that participants exhibited awareness at the level of noticing, but below the threshold of verbalization (e.g., Schmidt, 1990).

Subjective measures of awareness

Source attributions. As can be seen in Table 4, an analysis of the source attributions revealed that participants largely relied on their intuition rather than

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Table 4. *Accuracy (%) and proportion (%) of responses across confidence ratings and source attributions*

Confidence Rating	Accuracy (%)	Proportion (%)	*p*
No confidence	51.9	15	.70
Somewhat confident	56.3	50	.02
Very confident	57.1	31	.04
Absolutely certain	59.3	4	.35
Source Attribution			
Guess	50.5	15	.92
Intuition	56.3	50	.02
Memory	57.7	19	.08
Rule	58.8	16	.06

on explicit rule knowledge in their judgments. In addition, the accuracy of the participants' judgments was only significantly above chance levels when they indicated that they had made their judgment based on intuition. This means that the *guessing criterion* of implicit structural knowledge was not met. However, there are two pieces of evidence that lend support to the claim that the participants developed some unconscious grammatical knowledge. Performance of the experimental group was significant against chance on intuition judgments, which has been argued to be reflective of some unconscious structural knowledge (see, e.g., Dienes, 2010, 2013). This is in contrast to participants' nonsignificant performance when attributing decisions to rule, thus providing evidence that participants had not acquired explicit rule knowledge.

Confidence ratings. As shown in Table 4, participants' performance was significantly better than chance levels when indicating that they were somewhat confident or very confident in their decisions. While this strongly suggests that the participants were partially conscious of their judgment knowledge, an additional logit mixed-effect model with crossed random effects of subjects and items (Jaeger, 2008) was performed in order to analyze if there was any underlying interaction between confidence and accuracy. A likelihood ratio test revealed that the level of confidence and accuracy was not significantly related, χ^2 (3) = 1.50, $p = .68$. Furthermore, an effect size estimation, using the Cramer φ, indicated a very weak association between confidence and accuracy ($\varphi = 0.04$). This result satisfies the requirements for the *zero-correlation criterion* (Dienes & Scott, 2005), and provides further evidence that the participants' judgment knowledge of the target case-markers was, at least partially, unconscious in nature.

DISCUSSION

The results of this experiment demonstrated that participants can develop some knowledge of L2 case markings after a minimal amount of exposure, without

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

feedback, and as a result of incidental exposure. Concerning the relationship between learning and awareness in this study, an analysis of the retrospective verbal reports indicates that this learning effect was realized in the absence of verbalizable rule knowledge, but in the presence of low levels of awareness; in this case, the participants reported becoming aware of the surface features of the inflected forms during the training phase of the experiment. Furthermore, traditional analyses of the subjective measures of awareness (guessing criterion and zero-correlation criterion) suggest that the exposure phase had resulted in some implicit knowledge of L2 morphology among the experimental subjects.

Research Question 1: To what extent can L2 case marking be acquired under incidental learning conditions?

One finding of particular interest was that the performance of the experimental group was significantly better than that of the control group on accusative case items, but not on nominative items. As noted above, this differential performance accords with the findings of Robinson's (2002, 2005) study, which reported improved performance for only the locative marker, but not for the noun–incorporation and ergative markers in the miniature Samoan system he utilized to examine the acquisition of L2 morphology. Robinson (2002) speculated that the superior performance and concomitant awareness of the form–function mapping associated with the locative marker could be attributed to the amount of processing required by the locative training items. As Robinson (2002) notes, in order to answer the comprehension question related to the locative case items, participants had to attend to the information conveyed by the case marker. In contrast, comprehension questions targeting the ergative case items, which showed the lowest levels of learning, could be solved without attending to the ergative marker due to the opacity and communicative redundancy of this construction. Unlike in Robinson's research, a differential amount of semantic processing does not seem to offer an explanation for the patterns observed in the present study. Here, it does not appear that the participants necessarily needed to attend to the semantic information conveyed by either the nominative or the accusative case-markers during training. The training task required participants to match the meaning of the foreign word to one of two pictures, which could, in theory, be performed by relying solely on the lexical clues available within the sentence.

What then could account for the differential performance on the nominative versus accusative items? A possible explanation could be attributed to what has been labeled in the cognitive psychological literature as the "bizarreness effect" (e.g., Riefer & Rouder, 1992), specifically, the finding that distinctive stimuli are better recalled than more common stimuli. In the case of the present study, this effect could potentially be connected with the phonological characteristics of the morphemes used to mark the two cases. The phoneme associated with the nominative marker, *–a* (/*a*/), was probably perceived by English speakers as more natural in a word-final position than the phoneme used to realize the accusative marker *–u* (/*u*/), given the phonotactic rules of the English language. Thus, the

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

"oddity" of the accusative morpheme /*u*/ might have resulted in increased salience, leading to more noticing of the accusative marker during the training phrase. This interpretation is also lent support by participants in the experimental group were able to perform significantly better on accusative items than their counterparts in the control group on items with an object–verb–subject word order (a nontypical English word order), but not on items with a subject–verb–object word order (the prototypical word order in English). In this case, it also seems possible that the "foreign-ness" of the object–verb–subject items might have resulted in increased attention during exposure.

Finally, given the multifariousness of the training conditions and stimulus materials, it is not possible to pinpoint which component, or combination thereof, is responsible for the incidental learning that was observed in this study. It seems likely that the repetition of the exemplars in the training set (both as part of the sentence and individually) as well as elicited imitations as part of the training condition were likely to have played a role in achieving a learning effect, given the role that frequency arguably plays in L2 acquisition (e.g., Ellis 2002, 2008; N. Ellis, 2005). The procedure of repeating the sentence aloud has also been used in previous experiments operating within the AGL paradigm (see, e.g., Rebuschat & Williams, 2012). It could be argued that repetition led to an increased amount of rehearsal within short-term memory, creating additional opportunities for noticing to occur at the time of encoding.

One criticism that could be levied against the results of this experiment is that the performance of the experimental group was only slightly above chance levels (~56%), thus the advantage observed for the experimental group could possibly have been the result of a sampling error. While this criticism is best addressed through replication, these results should not be surprising given the difficulty in acquiring inflectional morphology (e.g., Larsen-Freeman, 2010) as well as the minimal amount of exposure the participants received in the present study (approximately 30 min). Furthermore, it is worth comparing the results of this experiment to those of previous research operating within the AGL paradigm as well as within L2 research. Artificial grammar learning studies typically result in a 55% to 80% performance on GJTs (DeKeyser, 2003). In addition, the results obtained here appear similar to those in Grey et al. (2014), the study that is most directly comparable to the current research. In Grey et al. the mean accuracy (M = 56.3%) on the delayed picture-matching task roughly corresponds to the mean average observed on the GJT in the present study (M = 55.4%). Although caution needs to be exercised when making comparisons across studies due to differences in their designs, participants, and contexts, these trends indicate that the findings of the present study are not atypical within the AGL paradigm and in L2 research investigating the acquisition of L2 case marking following incidental exposure.

Despite the positive results reported here, there are a number of limitations in relation to measurement of learning in the present study. One such limitation stems from the use of the GJT as the sole outcome measure. The rationale for employing grammaticality judgments was that this is one of the most commonly used tasks in implicit learning research in both cognitive psychology and L2 research. However, utilizing grammaticality judgments as a sole assessment of learning

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

has been criticized because they only assess participants' ability to recognize the target construction, and provide no information about their ability to produce it (e.g., Hama & Leow, 2010; Leow & Hama, 2013). When we consider the slight learning effect observed in the present study, it would seem unlikely that any evidence of learning would have emerged on assessments requiring the participants to produce the target morphological case endings, given the argument that production requires a higher order of knowledge than recognition (e.g., Bloom, 1956; Keenen & MacWhinney, 1987). In addition, asking participants to produce output might promote the development of conscious knowledge during the testing phase. However, along with including delayed measures to analyze the durability of learning (cf. de Graaf, 1997), future research would undoubtedly benefit from employing alternate or multiple measures in order to gain a fuller picture of the durability, quality, and quantity of learning that has taken place. Including behavioral measures, such as reaction times, in incidental learning experiments would seem particularly fruitful because behavioral techniques allow for more sensitive assessment of implicit knowledge (Hulstijn, 2013; Leung & Williams, 2012; Williams & Rebuschat, 2012) than mean accuracy rates on traditional test such as the GJT.

In addition, because English is not a morphologically rich language, the English L1 background of the participants in this study might have influenced their attention to and processing of the inflectional morphemes. Previous studies (Ellis & Sagarra, 2011; Kempe & Brooks, 2013; Rast, 2010; Williams, 2005) have suggested that recognition and processing of specific grammatical features might be facilitated by prior exposure to similar constructions in an L1 or L2. Concerning the acquisition of morphology, Ellis (e.g., 2006a, 2006b, see also Ellis & Sagarra 2010a, 2010b, 2011) has used morpheme acquisition order data in the context of connectionist models to argue that the acquisition of grammatical features can be moderated through the process of learned, selective attention as well as via a variety of other transfer phenomena. Given that the participants in this study did not have a background in morphologically rich languages, it would appear that they might not have been primed to develop an awareness beyond the level of noticing (Schmidt, 1990) of the morphological inflections present in the input. Because the results from previous research (Williams, 2004, 2005) has suggested a relationship between linguistic background and the degree that participants are able to learn in the absence of awareness, future research would benefit from directly exploring whether an L1 or L2 background in a language with rich morphology influences the processing and subsequent acquisition of L2 case markings under incidental learning conditions.

Research Question 2: What type of knowledge is acquired, implicit or explicit?

In regard to the second research question, the results of the retrospective verbal reports indicated that none of the participants had become aware of the morphological rule system. The lack of verbalizable knowledge shows that subjects can develop knowledge of L2 morphology without being able to verbally account

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

for their performance. This interpretation should be viewed with caution, given the lack of concurrent data, such as might be made available from think-aloud protocols (e.g., Hama & Leow, 2010) to support this claim. Although none of the participants in this study were able to verbalize the underlying rule system, all of the participants reported noticing, or becoming aware of, the changing inflections during the training phase of the experiment. This report of noticing is indicative of a low level of awareness of the target linguistic forms (Schmidt, 1990).

While the results of the retrospective verbal reports point to a low level of awareness on the part of the learners, the results of the subjective measures of awareness paint a slightly more complex picture as to the nature of the knowledge acquired in this experiment. Concerning the confidence ratings, the immediate results suggest that participants developed both conscious and unconscious knowledge: that the knowledge was partly conscious, as demonstrated by an increased accuracy when participants indicated some confidence in their decision, but also partly unconscious, in that there was not a significant relationship between confidence and accuracy, thus satisfying the zero-correlation criterion, which holds that a lack of correlation between confidence and accuracy is indicative of implicit knowledge (cf. Dienes & Scott, 2005). Supporting the interpretation of the presence of implicit knowledge, an analysis of the data from the source attributions indicated that participants performed at above chance levels when making judgments based on intuition, which has also been argued to be reflective of unconscious judgment knowledge (e.g., Dienes, 2008). In sum, traditional analyses of the measures of awareness demonstrate that the incidental exposure in the present study resulted in both conscious and unconscious knowledge on the part of the participants.

Assuming that these results do reflect implicit knowledge, this finding partly runs counter to the results from Robinson (2002, 2005) and Grey et al. (2014), where the detected learning effects were found to be largely driven by explicit knowledge. This divergence from previous findings could be attributed to the smaller amount of input to which the participants were exposed in the training phase in other studies. In this study, participants received 72 exposures to each of the two case markers in the training set, a total of 144 exposures. In contrast, in Robinson's (2002, 2005) study, participants received 450 total exposures, 150 for each of the three case endings. It seems plausible that such a difference in total amount of input could result in increased opportunities to notice the morphological forms during the training phase and develop greater awareness of the underlying morphological rule system.

In Grey et al. the training phase consisted of only 128 exposures presented auditorily, which is comparable to the type and amount of exposure in the present study. Note that the stimulus material in Grey et al.'s study comprised English words with Japanese case markers added to the end (*Vet-ga injection-o gave*). A number of studies have demonstrated that enhancing the target structure in the input, for example, by underlining the target form within a reading text or presenting the target structures in bold, can promote the noticing, and subsequent acquisition, of this target linguistic structure (e.g., Alanen, 1995; Leow, 2002). Given the results of these studies, and the argument that salience is a contributing

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

factor to whether particular linguistic structures are noticed in the input (e.g., Ellis, 2006a; Peters, 1985; Schmidt, 2001; Slobin, 1985), it follows that attaching the morphological marker to the end of the English word could have increased the salience of the case markings, thus leading to heightened levels of awareness throughout the exposure phase.

However, at this point, an alternative interpretation of the patterns observed above is worth offering by questioning the assumption that the results of the confidence ratings and the source attributions are necessarily indicative of unconscious knowledge. While the zero-correlation criterion was met by the data from the confidence ratings, the lack of a relationship between confidence and accuracy could be a relic of the slight overall learning effect. This interpretation of the results is in line with the view that it is probably not possible to develop grammatical, native-speaker-like intuition as a result of a single laboratory training session (e.g., Leow & Hama, 2013) and a recent suggestion by Serafini (2013) that source attribution data could be interpreted differently. According to Serafini, the source attributions could represent "degrees of explicitness," with above chance performance on intuition responses indicating low levels of awareness or knowledge that has been noticed, and above chance responses on rule responses reflecting awareness at the higher level of understanding. This interpretation runs counter to the traditional interpretation (e.g., Dienes, 2008; see also Rebuschat & Williams, 2012) regarding the phenomenology of intuition reflecting native-speaker-like implicit knowledge. Serafini's (2013) interpretation would appear to sit well with the results of the present study, in particular the reports of noticing on the retrospective verbal reports. Taken together then, the results from both subjective measures and retrospective verbal reports in the present study would point toward low levels of awareness among the participants in this study. Subsequently, the results here would fit well with previous research (e.g., Brooks & Kempe, 2013; Grey et al., 2014) that has found an important role for awareness in the acquisition of inflectional morphology.

Finally, when reflecting critically on the measures of awareness included in the present study, a limitation of this research lies in the lack of information with respect to the strategies adopted by participants when dealing with the stimuli in the exposure phase. Incidental learning conditions as operationalized here provide no guarantee that learners actually learn incidentally, that is, without intention. For example, Robinson (2002, 2005) reported that his participants engaged in rule-search behavior under an experimental condition that was designed to create opportunities for incidental learning. To address this shortcoming, future studies, like earlier research in this area (e.g., DeKeyser, 1995; Robinson 1997, 2002, 2005), could include retrospective questions asking participants if they were actively involved in searching for rules. This information could increase the construct validity of incidental learning studies. In particular, it could reveal whether the intended incidental learning conditions succeeded in orienting learners' attention toward the meaning of the stimuli and in preventing them from utilizing explicit learning strategies. Furthermore, this point underscores the potential benefits of eliciting concurrent data, for example, through think-aloud protocols (e.g., Hama & Leow, 2010), as a means of shedding light on the processes that occur at the time of encoding.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Conclusion

This study set out to examine the degree to which L2 case markings can be acquired as a result of incidental exposure and investigate whether the resulting knowledge was implicit or explicit in nature. In regard to the first research question, the experiment described here provided some evidence that participants can develop partial knowledge of an artificial L2 case marking system under incidental learning conditions, without feedback, and after a very limited amount of exposure. These results are in line with previous research investigating the incidental learning of L2 case marking (Grey et al., 2014; Robinson, 2002, 2005). Like previous research, the present study suggests that there are limits to the extent to which L2 case markers can be acquired incidentally. Similarly to Grey et al.'s findings, only a slight learning effect was observed, and this was driven by performance on only one of the morphological categories as in Robinson's (2002, 2005) work.

Thus, taken together, these results lend support to the notion that the acquisition of L2 morphology through incidental exposure is not as straightforward as the learning of other grammatical areas, such as L2 word order. Previous research that has investigated the acquisition of L2 syntax under incidental learning conditions (e.g., Rebuschat & Williams, 2012; Williams & Kuribara, 2008) has demonstrated clearer learning effects as a result of similar training and after a comparable amount of total exposure to the target linguistic constructions. One question worth exploring is just how much exposure is needed for L2 learners to acquire inflectional forms. In addition, the one study that has examined the acquisition of both L2 word order and case marking (Grey et al., 2014) reported greater effects for word order than for morphology. The research design in the present study did not allow for comparing the relative ease of acquiring case marking and other areas of grammar. Future research could further explore the difficulty involved in acquiring L2 morphology incidentally as opposed to other types of constructions. Another question worth exploring might be whether the learning of inflectional morphology simply requires more exposure than other areas of grammar, such as word order. If this is the case, an interesting avenue might be to examine the time course of learning in order to gain a clearer picture of how much exposure is needed for learners to acquire these forms, as well as the impact that individual difference variables, such as nonverbal IQ (e.g., Brooks & Kempe, 2013; Brooks et al., 2006; Kempe & Brooks, 2008), have on the trajectory of learning.

Finally, as has been noted by many others (e.g., Hulstijn, 1997), caution must be exercised when attempting to extrapolate laboratory studies to naturalistic language acquisition, in particular when these experiments utilize artificial grammars, whose generalizability to natural languages has long been challenged (e.g., Schmidt, 1994b; see Rogers, Révész, & Rebuschat, 2015, for a brief overview). The use of artificial languages stems from the quantitative tradition, which stresses "reliability over validity and control over context" (Hulstijn et al., 2014, p. 365). In this sense, the advantages of artificial languages are clear: they allow the experimenter complete control over the input present in the experiment, and allow the experiment to effectively control for other confounding variables, such as prior

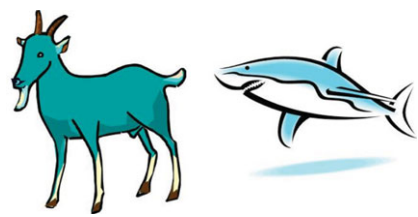
https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

knowledge on the part of the participants. However, the limitations of artificial languages, such as the system employed in the present study, are also apparent. These limitations include that artificial languages represent much simplified versions of natural language systems and often completely ignore important aspects of language, such as pragmatics.[5] Despite these reservations, it is hoped that this experiment can join the list of previous research that has utilized artificial or semiartificial languages in their experimental design, and that has undoubtedly contributed to our understanding of L2 acquisition (e.g., de Graaf, 1997; DeKeyser, 1995, 1997; Grey et al., 2014; Hama & Leow, 2010; Leung & Williams, 2011, 2012, 2014; Rebuschat & Williams, 2012; Robinson, 1997, 2010; Williams, 2005).

In conclusion, despite the limitations of the present study, it is believed that this research will contribute to several areas within L2 research. First, this study adds to the growing body of research on the incidental learning of various aspects of L2 grammar, in particular to the degree that L2 morphology can be acquired solely as the result of incidental exposure. Second, through the use of retrospective verbal reports and subjective measures of awareness, this experiment represents a methodological contribution toward the measurement of implicit and explicit knowledge, thus adding to the ongoing, contentious debate regarding the role of awareness in L2 acquisition.

APPENDIX A

Sample items and pictures presented during the exposure phase for nominative, objective, and instrumental items (color online)

Morphological Category	Example Sentence and Pictures
Nominative	Last summer the grass ate the koza in the field.  Last year the studna supplied water to the town. 

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

APPENDIX A (*cont.*)

Morphological Category	Example Sentence and Pictures
Accusative	Last month the kasu opened Patrick with the key.   Today Peter used a britvu in the bathroom.  

APPENDIX B

Examples of grammatical and ungrammatical patterns used in the acceptability judgment testing sets

Pattern	Grammatical
Nominative	In the evening the bunda warmed him in the park.
Accusative	This morning she changed the plenu in the bathroom.
	Ungrammatical
Nominative (–u*)	In summer the lisku ate the eggs in the field.
Accusative (–a*)	Last weekend he baked the kachna in the oven.

ACKNOWLEDGMENTS

We thank the anonymous reviewers for their constructive comments on previous versions of this article. All errors remain our own.

NOTES

1. Although Krashen differentiates between conscious "learning" and unconscious "acquisition," the terms learning and acquisition will be used interchangeably in this article to refer to the development of L2 knowledge.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

2. However, in an early paper (Schmidt, 1994b), Schmidt presents a weak form of the noticing hypothesis that states that awareness at the level of noticing, like understanding, is facilitative but perhaps not necessary.
3. As suggested by an anonymous reviewer, it is important to stress that incidental learning conditions do not control or guarantee that learners will necessarily learn incidentally (e.g., Schmidt, 1993). Numerous studies have demonstrated that some learners may employ explicit learning strategies even under incidental learning conditions (e.g., DeKeyser, 1995).
4. Note that the order of the phases in the experiment, specifically that the debriefing session follows the testing phase, can impact on participants' reported levels of awareness. For example, Rebuschat (2008) found that participants developed conscious awareness during the testing phase of the experiment.
5. We thank an anonymous reviewer for this suggestion.

REFERENCES

Alanen, R. (1995). Input enhancement and rule presentation in second language acquisition. In R. Schmidt (Ed.), *Attention and awareness in foreign language learning and teaching* (pp. 259–302). Honolulu, HI: University of Hawaii Press.

Andringa, S., & Rebuschat, P. (Eds.) (2015). New directions in implicit and explicit language learning [special issue]. *Studies in Second Language Acquisition*, *37*.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2014). lme4: Linear mixed-effects models using Eigen and S4. R package version1.0–7 [Computer software]. Retrieved January 23, 2015, from http://CRAN.R-project.org/package=lme4

Bialystok, E. (1982). On the relationship between knowing and using forms. *Applied Linguistics*, *3*, 181–206.

Bloom, B. S. (1956). *Taxonomy of educational objectives: Handbook 1. The cognitive domain*. New York: David McKay.

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.

Brooks, P. J., Kempe, V., & Donachie, A. (2011). Second language learning benefits from similarity in word endings: Evidence from Russian. *Language Learning*, *61*, 1142–1172.

Brooks, P. J., Kempe, V., & Sionov, A. (2006). The role of learner and input variables in learning inflectional morphology. *Applied Psycholinguistics*, *27*, 185–209.

Cedrus Corporation. (2012). SuperLab Version 4.5 [Computer software]. San Pedro, CA: Author.

Clahsen, H., Felser, C., Neubauer, K., Sato, M., & Silva, R. (2010). Morphological structure in native and nonnative language processing. *Language Learning*, *60*, 21–43.

Cobb, T. (2012). *Web Vocabprofile*. Retrieved November 14, 2012, from http://www.lextutor.ca/vp/

Cunnings, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, *28*, 369–382.

de Graaf, R. (1997). The eXperanto experiment: Effects of explicit instruction on second language acquisition. *Studies in Second Language Acquisition*, *19*, 249–297.

DeKeyser, R. M. (1995). Learning second language grammar rules. *Studies in Second Language Acquisition*, *17*, 379–410.

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

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. Doughty & M. Long (Eds.), *Handbook of second language acquisition* (pp. 313–348). Oxford: Blackwell.

DeKeyser, R. M. (2005). What makes learning second-language grammar difficult? A review of issues. *Language Learning*, *55*, 1–25.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

DeKeyser, R. M. (2009). Cognitive–psychological processes in second language learning. In M. Long & C. Doughty (Eds.), *Handbook of second language teaching* (pp. 119–138). Oxford: Blackwell.
Dienes, Z. (2004). Assumptions of subjective measures of unconscious mental states: Higher order thoughts and bias. *Journal of Consciousness Studies*, *11*, 25–45.
Dienes, Z. (2008). Subjective measures of unconscious knowledge. *Progress in Brain Research*, *168*, 49–64.
Dienes, Z. (2010). Measuring the conscious status of knowledge. In L. A. P. Miranda & A. I. Madariaga (Eds.), *Advances in cognitive science: Learning, evolution and social action* (pp. 113–128). Bilbao, Spain: University of the Basque Country Press.
Dienes, Z. (2013). Conscious versus unconscious learning of structure. In P. Rebuschat & J. N. Williams (Eds.), *Statistical learning and language acquisition* (pp. 337–364). Berlin: De Gruyter.
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.
Dörnyei, Z. (2009). *The psychology of second language acquisition*. Oxford: Oxford University Press.
Ellis, N. C. (2002). Frequency effects in language acquisition: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, *24*, 143–188.
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. (2006a). Selective attention and transfer phenomena in L2 acquisition: Contingency, cue competition, salience, interference, overshadowing, blocking, and perceptual learning. *Applied Linguistics*, *27*, 164–194.
Ellis, N. C. (2006b). Language acquisition as rational contingency learning. *Applied Linguistics*, *27*, 1–24.
Ellis, N. C. (2008). Implicit and explicit knowledge about language. In J. Cenoz & N. H. Hornberger (Eds.), *Encyclopedia of language and education: Vol. 6. Knowledge about language* (pp. 119–132). New York: Springer.
Ellis, N. C. (2011). Implicit and explicit SLA and their interface. In C. Sanz & R. Leow (Eds.), *Implicit and explicit language learning: Conditions, processes, and knowledge in SLA & bilingualism* (pp. 35–47). Washington, DC: Georgetown University Press.
Ellis, N. C., & Sagarra, N. (2010a). The bounds of adult language acquisition: Blocking and learned attention. *Studies in Second Language Acquisition*, *32*, 553–580.
Ellis, N. C., & Sagarra, N. (2010b). Learned attention effects in L2 temporal reference: The first hour and the next eight semesters. *Language Learning*, *60*, 85–108.
Ellis, N. C., & Sagarra, N. (2011). Learned attention in adult language acquisition: A replication and generalization study and meta-analysis. *Studies in Second Language Acquisition*, *33*, 589–624.
Ellis, N. C., & Schmidt, R. (1997). Morphology and longer-distance dependencies: Laboratory research illuminating the A in SLA. *Studies in Second Language Acquisition*, *19*, 145–171.
Ellis, N. C., & Schmidt, R. (1998). Rules or associations in the acquisition of morphology? The frequency by regularity interaction in human and PDP learning of morphosyntax. *Language and Cognitive Processes*, *13*, 307–336.
Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language: A psychometric study. *Studies in Second Language Acquisition*, *27*, 141–171.
Gor, K. (2010). Beyond the obvious: Do second language learners process inflectional morphology? *Language Learning*, *60*, 1–20.
Grey, S., Williams, J. N., & Rebuschat, P. (2014). Incidental exposure and L3 learning of morphosyntax. *Studies in Second Language Acquisition*, *36*, 1–34.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Grey, S., Williams, J. N., & Rebuschat, P. (2015). Individual differences in incidental vocabulary learning: Phonological working memory, learning styles, and personality. *Learning and Individual Differences*, *38*, 44–53.

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. *Studies in Second Language Acquisition*, *32*, 465–491.

Hamrick, P. (2013). *Development of conscious knowledge during early incidental learning of L2 syntax*. Unpublished doctoral dissertation, Georgetown University.

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

Hamrick, P., & Rebuschat, P. (2013). Frequency effects, learning conditions, and the development of implicit and explicit lexical knowledge. In J. Connor-Linton & L. Amoroso (Eds.), *Measured language: Quantitative approaches to acquisition, assessment, processing, and variation*. Washington, DC: Georgetown University Press.

Hinz, J., Krause, C., Rast, R., Shoemaker, E., & Watorek, M. (2013). Initial processing of morphological marking in nonnative language acquisition: Evidence from French and German learners of Polish. In L. Roberts, A. Ewert, M. Pawlak, & M. Wrembel (Eds.), *EUROSLA yearbook 13* (pp. 139–175). Amsterdam: John Benjamins.

Hopp, H. (2013). The development of L2 morphology. *Second Language Research*, *29*, 3–6.

Hulstijn, J. H. (1997). Second language acquisition research in the laboratory: Possibilities and limitations. *Studies in Second Language Acquisition*, *19*, 131–143.

Hulstijn, J. H. (2003). Incidental and intentional learning. In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 349–381). Malden, MA: Blackwell.

Hulstijn, J. H. (2007). Psycholinguistic perspectives on second language acquisition. In J. Cummins & C. Davison (Eds.), *The international handbook on English language teaching: Vol. 2. Current trends and issues* (pp. 701–713). Norwell, MA: Springer.

Hulstijn, J. H. (2013). Incidental learning in second language acquisition. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Oxford: Wiley–Blackwell.

Hulstijn, J. H., & Ellis, R. (Eds.) (2005). Implicit and explicit second-language learning [special issue]. *Studies in Second Language Acquisition*, *27*.

Hulstijn, J. H., Young, R. F., Ortega, L., Bigelow, M., DeKeyser, R., Ellis, N. C., et al. (2014). Bridging the gap: Cognitive and social approaches to research in second language learning and teaching. *Studies in Second Language Acquisition*, *36*, 361–421. doi:10.1017/S0272263114000035

Jaeger, T. F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, *59*, 434–446.

Jiang, N. (2004). Morphological insensitivity in second language processing. *Applied Psycholinguistics*, *25*, 603–634.

Jiang, N. (2007). Selective integration of linguistic knowledge in adult second language learning. *Language Learning*, *57*, 1–33.

Keenan, J. M., & MacWhinney, B. (1987). Understanding the relationship between comprehension and production. In H. W. Dechert & M Raupach (Eds.), *Psycholinguistic models of production* (pp. 149–155). Norwood, NJ: Ablex.

Kempe, V., & Brooks, P. J. (2008). Second language learning of complex inflectional systems. *Language Learning*, *58*, 703–746.

Kempe, V., Brooks, P. J., & Kharkhurin, A. (2010). Cognitive predictors of generalization of Russian grammatical gender categories. *Language Learning*, *60*, 127–153.

Kempe, V., & MacWhinney, B. (1998). The acquisition of case-marking by adult learners of Russian and German. *Studies in Second Language Acquisition*, *20*, 543–587.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon Press.

Larsen-Freeman, D. (2010). Not so fast: A discussion of L2 morpheme processing and acquisition. *Language Learning*, *60*, 221–230.

Leow, R. P. (2002). Models, attention, and awareness in SLA: A response to Simard and Wong's (2001) "Alertness, orientation, and detection: The conceptualization of attentional functions." *Studies in Second Language Acquisition*, *24*, 113–119.

Leow, R. P. (2015). Implicit learning in SLA: Of processes and products. In P. Rebuschat (Ed.), *Implicit and explicit learning of languages* (pp. 47–65). Amsterdam: John Benjamins.

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

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.

Macmillan, N., & Creelman, C. (2005). *Detection theory: A user's guide*. 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.

Peters, A. (1985). Language segmentation: Operating principles for the perception and analysis of language. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 2, pp. 1029–1067). Hillsdale, NJ: Erlbaum.

Rast, R. (2008). *Foreign language input: Initial processing*. Clevedon: Multilingual Matters.

Rast, R. (2010). The use of prior linguistic knowledge in the early stages of L3 acquisition. *IRAL, International Review of Applied Linguistics in Language Teaching*, *48*, 159–183.

R Core Team. (2014). *R: A language and environment for statistical computing [Computer software]*. Vienna: R Foundation for Statistical Computing. Retrieved January 20, 2015, from http://www.R-project.org/

Reber, A. S. (1967). Implicit learning of artificial grammars. *Journal of Verbal Learning and Verbal Behavior*, *6*, 855–863.

Reber, A. S. (1993). *Implicit learning and tacit knowledge: An essay on the cognitive unconscious*. London: Clarendon Press.

Rebuschat, P. (2008). *Implicit learning of natural language syntax*. Unpublished doctoral dissertation, University of Cambridge.

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

Rebuschat, P. (Ed.). (2015). *Implicit and explicit learning of languages*. Amsterdam: John Benjamins.

Rebuschat, P., Hamrick, P., Riestenberg, K., Sachs, R., & Ziegler, N. (2015). Triangulating measures of awareness. *Studies in Second Language Acquisition*, *37*, 299–334.

Rebuschat, P., Hamrick, P., Sachs, R., Riestenberg, K., & Ziegler, N. (2013) Implicit and explicit knowledge of form–meaning connections: Evidence from subjective measures of awareness. In J. Bergsleithner, S. Frota, & J. K. Yoshioka (Eds.), *Noticing: L2 studies and essays in honor of Dick Schmidt* (pp. 255–275). Honolulu, HI: University of Hawaii Press.

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

Riefer, D. M., & Rouder, J. N. (1992). A multinomial modeling analysis of the mnemonic benefits of bizarre imagery. *Memory & Cognition*, *20*, 601–611.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

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. (2002). Effects of individual differences in intelligence, aptitude and working memory on adult incidental SLA: A replication and extension of Reber, Walkenfeld and Hernstadt, 1991. In P. Robinson (Ed.), *Individual differences and instructed language learning* (pp. 211–266). Amsterdam: Benjamins.

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.

Robinson, P. (2008). Attention and awareness. In J. Cenoz & N. H. Hornberger (Eds.), *Knowledge about language: Encyclopedia of language and education* (2nd ed., Vol. 6, pp. 1–10). New York: Springer.

Robinson, P. (2010). Implicit artificial grammar and incidental natural second language learning: How comparable are they? *Language Learning*, *60*, 245–263.

Robinson, P., Mackey, A., Gass, S., & Schmidt, R. (2012). Attention and awareness in second language acquisition. In S. Gass & A. Mackey (Eds.), *The Routledge handbook of second language acquisition* (pp. 247–267). 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. 273–298). Amsterdam: John Benjamins.

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

Schmidt, R. (1993). Awareness and second language acquisition. *Annual Review of Applied Linguistics*, *13*, 206–226.

Schmidt, R. (1994a). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review—Revue de l'AILA*, *11*, 11–26.

Schmidt, R. (1994b). Implicit learning and the cognitive unconscious: Of artificial grammars and SLA. In N. Ellis (Ed.), *Implicit and explicit learning of languages* (pp. 165–209). London: Academic Press.

Schmidt, R. W. (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 Hawaii Press.

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

Schmidt, R. (2010). Attention, awareness, and individual differences in language learning. In W. M. Chan, S. Chi, K. N. Cin, J. Instanto, M. Nagami, J. W. Sew, et al. (Eds.), *Proceedings of CLaSIC 2010* (pp. 721–37). Singapore: National University of Singapore, Centre for Language Studies.

Serafini, E. (2013). *Cognitive and psychological factors in the long-term development of implicit and explicit second language knowledge in adult learners of Spanish at increasing proficiency*. Unpublished doctoral dissertation. Georgetown University.

Shanks, D. R., & John, M. F. (1994). Characteristics of dissociable human learning system. *Behavioral and Brain Sciences*, *17*, 367–447.

Slobin, D. (1985). Crosslinguistic evidence for the language-making capacity. In D. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 2, pp. 1157–1256). Hillsdale, NJ: Erlbaum.

Stanislaw, H., & Todorov, N. (1999). Calculation of signal detection theory measures. *Behavior Research Methods, Instruments, and Computers*, *31*, 137–149.

Tagarelli, K., Borges Mota, M., & Rebuschat, P. (2011). The role of working memory in the implicit and explicit learning of languages. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings*

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press

of the 33rd Annual Conference of the Cognitive Science Society (pp. 2061–2066). Austin, TX: Cognitive Science Society.

White, L. (2009). Grammatical theory: Interfaces and L2 knowledge. In W. Ritchie & T. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 49–68). Leeds: Emerald Group Publishing.

Williams, J. N. (2004). Implicit learning of form-meaning connections. In J. Williams, B. VanPatten, S. Rott, & M. Overstreet (Eds.), *Form-meaning connections in second language acquisition* (pp. 203–218). Mahwah, NJ: Erlbaum.

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. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319–344). Bingley: Emerald Group Publishing.

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

Williams, J., & Rebuschat, P. (2012). Statistical learning and syntax: What can be learned, and what difference does meaning make. In P. Rebuschat & J. Williams (Eds.), *Statistical learning and language acquisition* (pp. 237–264). Berlin: Mouton deGruyter.

Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and nonparametric estimate of sensitivity. *Psychometrika*, *70*, 203–212.

https://doi.org/10.1017/S0142716415000247 Published online by Cambridge University Press