ATU PRESS

ISSUES IN LANGUAGE TEACHING (ILT)
VOL. 13, NO. 1, 1-33, June 2024
https://doi.org/10.22054/ilt.2024.78148.833 Document Type: Research Paper

Examining L2 Learners' Performance on Implicit and Explicit Measures: The Role of Grammaticality and Structure Difficulty

Sepideh Mehraein*
Ph.D. Candidate of Applied Linguistics, University of Tehran, Iran

Hamideh Marefat
Professor of Applied Linguistics, University of Tehran, Iran

Received: February 07, 2024**; Accepted:** June 16, 2024

Abstract
Extensive research exists on the effects of task design features on measuring L2 learners' implicit and explicit knowledge. However, the role of structure difficulty has received limited attention. Additionally, the use of fine-grained measures of implicit knowledge has remained underexplored. To address these gaps, utilizing objective criteria to select easy (plural -s) and difficult (third-person -s) structures, a total of 256 experimental items, equally divided into grammatical and ungrammatical, as well as easy and difficult structures, were developed and administered to 32 advanced L2 learners. A word monitoring task (WMT) assessed their implicit knowledge through reaction time (RT) and grammaticality sensitivity index (GSI), while a timed grammaticality judgment test (TGJT) measured their automatized explicit knowledge through accuracy scores. The WMT results showed longer RTs for ungrammatical items and larger GSI for the plural -s items, revealing participants' more implicit knowledge of the easy structure. The results of the TGJT revealed that L2 learners judged grammatical items more accurately than ungrammatical ones and the plural -s items more accurately than third-person -s ones, showing participants' more automatized explicit knowledge of the easy and grammatical structures. The findings highlight the influence of grammaticality and structure difficulty on knowledge retrieval and suggest that advanced L2 learners exhibit stronger implicit and automatized explicit knowledge of the easy structure. These findings underscore the need for tailored instructional approaches to address difficult structures and emphasize the importance of using real-time psycholinguistic measures to examine L2 learners' implicit knowledge.

Keywords: Implicit Knowledge, Explicit Knowledge, Automatized Explicit Knowledge, Grammaticality, Structure Difficulty

***Corresponding author's email:** sepideh.mehraein@ut.ac.ir

INTRODUCTION

The constructs of implicit and explicit knowledge of the second language (L2) are central to the field of second language acquisition (SLA). Implicit knowledge refers to linguistic knowledge without awareness, whereas explicit knowledge involves conscious linguistic knowledge (Andringa & Rebuschat, 2015; DeKeyser, 2003; Hulstijn, 2005; Williams, 2009). As previous studies have empirically demonstrated, explicit and implicit knowledge are discrete constructs that can be measured using various tests (e.g., Bowles, 2011; R. Ellis, 2005; Gutiérrez, 2013; Suzuki & DeKeyser, 2017; Vafaee et al., 2017; Zhang, 2015). These studies have shown that in measuring implicit and explicit knowledge, the extent to which these two types of knowledge are involved depends on several task design features. According to some studies, the factor that differentiates implicit and explicit knowledge is time pressure (Bowles, 2011; R. Ellis, 2005; R. Ellis & Loewen, 2007; Zhang, 2015). However, more recent studies revealed that time pressure does not guarantee the retrieval of implicit knowledge and that time-pressured tasks that direct learners' attention to form (e.g., grammaticality judgment test (GJT)) cannot be considered as measures of implicit knowledge (Mehraein et al., 2022; Suzuki & DeKeyser, 2015; Vafaee et al., 2017). Some scholars highlighted the role of grammaticality as another task feature in distinguishing and measuring implicit and explicit knowledge (e.g., Bley-Vroman et al., 1988; Gutiérrez, 2013; Kim & Nam, 2017; Loewen, 2009; Vafaee et al., 2017). Nevertheless, different results have been found concerning the differences between grammatical and ungrammatical items. Another task feature, namely structure difficulty, was also shown to be crucial in measuring implicit and explicit knowledge (e.g., R. Ellis, 2006; Mehraein et al., 2022; Shiu et al., 2018). However, this last feature needs to be systematically and objectively explored. Previous studies have not consistently used objective criteria to categorize target structures as easy or difficult, nor have they employed fine-grained measures of implicit knowledge. Our motivation is to address these gaps using rigorous selection

criteria and psycholinguistic measures of implicit knowledge. One such method is the word monitoring task (WMT), which indirectly measures L2 learners' grammatical sensitivity in milliseconds (ms) while they are reading/listening for comprehension (Granena, 2013; Mehraein et al., 2022; Suzuki, 2017; Suzuki & DeKeyser, 2015, 2017).

Therefore, this study attempted to investigate the role of grammaticality (i.e., grammatical/ungrammatical) and structure difficulty (i.e., easy/difficult) in measuring L2 learners' implicit and automatized explicit knowledge. To do so, plural -s and third-person -s were selected as easy and difficult structures, respectively, based on objective selection criteria (frequency, saliency, functional value, metalanguage, early/late acquisition), and experimental items were divided into grammatical and ungrammatical. Then, learners' implicit knowledge was assessed through WMT, a real-time comprehension measure that utilized reaction time (RT) and involved no awareness of linguistic features, and their automatized explicit knowledge was measured through timed GJT, which required awareness and directed test takers' attention to form.

LITERATURE REVIEW

Implicit and Explicit Knowledge

There are numerous definitions available for implicit and explicit knowledge. In the context of SLA, it has been argued that implicit knowledge is *knowledge of language*; it is the intuitive understanding of language, encompassing an innate sense of correctness and acceptability beyond awareness (Sharwood Smith, 1981). On the other hand, explicit knowledge refers to *knowledge about language*, involving awareness of L2 items and structures through metalinguistic awareness (R. Ellis, 1997). More comprehensively, R. Ellis (2005) suggests that implicit and explicit knowledge of an L2 can be distinguished in terms of seven principal dimensions: degree of awareness, focus of attention, time available, certainty, systematicity, metalanguage, and learnability. According to him, of these characteristics, (1) degree of awareness, (2) focus of attention, (3) time

available, and (4) metalanguage are more central and should be taken into account when distinguishing implicit and explicit knowledge. Therefore, implicit knowledge is defined as knowledge that requires no awareness, can be drawn on under time pressure, draws L2 learners' attention to meaning, and involves no metalinguistic knowledge. In contrast, explicit knowledge requires conscious awareness, can be accessed without time pressure, draws learners' attention to form, and makes them use their metalinguistic knowledge. He asserts that when designing a particular test, these criteria would make the test a measure of implicit or explicit knowledge. However, some scholars believe that among the above-mentioned criteria, awareness can be considered the primary criterion for the distinction between implicit and explicit knowledge and that time pressure cannot be a distinguishing factor between knowledge types (Andringa & Rebuschat, 2015; DeKeyser, 2009; Mehraein et al., 2022; Suzuki & DeKeyser, 2015, 2017; Vafaee et al., 2017; Williams, 2009). Therefore, a further distinction is made in the literature between implicit knowledge and automatized explicit knowledge. Both implicit knowledge and automatized explicit knowledge involve rapid access to linguistic knowledge, yet they are differentiated by the awareness criterion (DeKeyser, 2009). While implicit knowledge requires no awareness, automatized explicit knowledge involves consciousness about linguistic forms even under time pressure. Irrespective of these differences among the knowledge types, most studies have shown that there are distinct linguistic knowledge types that can be measured through different tests (Bowles, 2011; R. Ellis, 2005; R. Ellis & Loewen, 2007; Gutiérrez, 2013; Isbell & Rogers, 2021; Kim & Nam, 2017; Mehraein et al, 2022; Suzuki & DeKeyser, 2015, 2017; Vafaee et al., 2017; Zhang, 2015).

The Measurement of Implicit and Explicit Knowledge

In the last two decades, ongoing attempts have been made to construct measures of implicit and explicit knowledge. In an initial attempt, taking into account the features that characterize implicit and explicit knowledge, R. Ellis (2005) developed a test battery consisting of an oral narrative test (ONT), a

timed GJT, an elicited imitation task (EIT), an untimed GJT, and a metalinguistic knowledge test (MKT). Results of exploratory factor analysis (EFA) showed that the first three tests (i.e., ONT, EI, and timed GJT), which were time-pressured tests, loaded on the implicit knowledge factor whereas the last two tests (i.e., untimed GJT and MKT) loaded on the explicit knowledge factor. However, Isemonger (2007) criticized the use of EFA and suggested conducting a confirmatory factor analysis (CFA) on the same data due to the fact that R. Ellis (2005) tested a hypothesis based on an understanding of the constructs and the model. Isemonger (2007) also recommended testing some rival models from the original EFA as he states, "It is important that rival models are tested because the fit of a particular model does not preclude the possibility that other untested models fit better" (p. 109). Responding to Isemonger (2007), R. Ellis and Loewen (2007) reanalyzed the R. Ellis' (2005) data by conducting CFAs. They tested the implicit/explicit model from the original EFA against a rival model. Their results confirmed the implicit/explicit model but showed no model fit for the rival model. Later, R. Ellis's (2005) study was replicated by Bowles (2011), who tested Spanish heritage learners and L2 Spanish learners. The results of CFA showed that test scores loaded on a two-factor model, as in R. Ellis (2005). The CFA findings of another replication study by Zhang (2015) also provided empirical support for the construct validity of the R. Ellis' (2005) test battery with a different learner population and language. Following these investigations, numerous studies have utilized the tests from R. Ellis's (2005) study with different task design features to measure the effect of implicit/explicit instruction and implicit/explicit assessment (e.g., Akakura, 2012; Baleghizadeh & Derakhshesh, 2016; Ghahari & Piruznejad, 2016; Ghorbani & Atai, 2012; Godfroid, 2016; Rezaei & Mehraein, 2019).

Task Design Features

Time Pressure

All of the above-mentioned studies have revealed that implicit and explicit knowledge are distinct constructs that can be measured with different tests,

and time pressure is the crucial feature in distinguishing the two knowledge types. It was believed that imposing time pressure allows L2 learners to access their implicit knowledge, while with unlimited time, they can take advantage of their explicit knowledge. However, more recent studies have cast doubt on the contributing role of time pressure in implicit and explicit measures (e.g., Mehraein et al., 2022; Suzuki & DeKeyser, 2015; Vafaee et al., 2017). These studies have shown that imposing time pressure cannot restrict access to explicit knowledge sufficiently to guarantee the utilization of implicit knowledge. Proficient L2 learners may still access their automatized explicit knowledge with awareness even under time constraints, which differs from implicit knowledge requiring no awareness. Put differently, though implicit and automatized explicit knowledge are accessed quickly, the two are distinguished as the former is tapped into without awareness while the latter is applied with awareness. Thus, time pressure cannot be considered a pivotal factor in measuring different knowledge types, and time-pressured tasks (e.g., timed GJTs) are not necessarily measures of implicit knowledge. Based on this evidence, in the current study, the timed GJT is considered to measure automatized explicit knowledge, not implicit knowledge.

Grammaticality

Another line of research has pursued the role of grammaticality as another distinguishing factor in measuring implicit and explicit knowledge. Gutiérrez (2013) examined this factor in the context of Spanish L2 learners by employing a test battery consisting of a timed GJT, an untimed GJT, and an MKT, calculating GJT scores for grammatical items and ungrammatical items separately. Two hypothesized models were compared using CFAs: The grammatical/ungrammatical model and the timed/untimed model. In the first model, both the grammatical and ungrammatical items of timed GJT loaded on the construct of implicit knowledge, and the grammatical and ungrammatical items of untimed GJT and MKT loaded on the construct of explicit knowledge. In the second model, the grammatical items of both timed

and untimed GJTs loaded on the construct of implicit knowledge, and the ungrammatical items of both types of GJTs and MKT loaded on explicit knowledge. The results showed that the grammatical/ungrammatical model was a better fit for the data, indicating that grammatical items tap into implicit knowledge, while ungrammatical items tap into explicit knowledge. Moreover, the results of repeated-measures ANOVA indicated that although both time pressure and grammaticality had significant effects on the GJT scores, the mean differences between grammatical and ungrammatical items were considerably larger than those between timed and untimed tests. As such, Gutierrez concluded that grammaticality is the key factor distinguishing between the use of implicit and explicit knowledge. Regarding the role of grammatically, Vafaee et al. (2017) found different results. They tested several CFA models, and their results revealed that only the ungrammatical sentences in both timed and untimed GJTs loaded together with scores on the MKT on a factor they labeled explicit knowledge, but the grammatical sentences did not tap into implicit knowledge. Thus, they concluded that manipulating grammaticality in form-focused tasks such as GJTs cannot make them measures of implicit knowledge. Research has also reported that L2 learners differ when judging grammatical and ungrammatical sentences. Most of the studies have demonstrated L2 learners' better performance on grammatical items compared to ungrammatical items (e.g., Bialystok, 1979, 1986; R. Ellis, 2005; Gutierrez, 2013; Kim & Nam, 2017; Loewen, 2009; Mehraein et al., 2022; Shiu et al., 2018; Vafaee et al., 2017). As can be inferred from these studies, better performance on grammatical items may indicate L2 learners' stronger implicit knowledge of the structures. However, few studies have found the opposite (Bley-Vroman et al., 1988; Gass, 1983).

Structure Difficulty

In addition to time pressure and grammaticality, a few studies (e.g., Bialystok, 1979; R. Ellis, 2006; Mehraein et al., 2022; Shiu et al., 2018) have posited that the variations in the performance of L2 learners in the implicit and explicit tasks may be attributed to another task design feature, namely

structure difficulty. In an early study, Bialystok (1979) scrutinized the performance of L2 learners in relation to target structures by administering aural GJTs: one required the participants to respond within 3 seconds, while the other allowed a 15-second response time. These GJTs comprised grammatical and ungrammatical items, targeting adjectives, object pronouns, and verbs. Each target structure encompassed three governing grammar rules that were categorized as easy, intermediate, or difficult, based on the subjective judgments of Bialystok and native speakers. The results revealed that learners' performance on GJTs varied with the difficulty of target structures, with better performance on easy structures under extended response times. In another study, R. Ellis (2006) employed R. Ellis' (2005) test battery to explore the relationship between the learning difficulty of 17 grammatical structures and the two knowledge types. Although the selection of the structures was based on several criteria, they were not classified into easy and difficult structures. The findings revealed a striking trend: what may be easy in terms of implicit knowledge may pose difficulties in terms of explicit knowledge, and vice versa. Shiu et al. (2018) investigated the influence of task design features (i.e., time constraints, task modality, task stimulus, and target features) on L2 learners' performance, focusing on past progressive and passive voice structures, which were hypothesized to differ in terms of difficulty level. The results revealed that learners performed better on the past progressive items (i.e., easy structure) compared to the passive items (i.e., difficult structure). Although the structure difficulty in Shiu et al.'s (2018) study was defined in terms of more objective criteria (i.e., structure formation, input frequency, phonological saliency, and early/late acquisition), only GJTs (measures of automatized explicit knowledge) were employed without including fine-grained measures of implicit knowledge. More recently, Mehraein et al. (2022) investigated how structure difficulty affects the types of cognitive representations that L2 learners utilize in various tests of implicit and explicit knowledge. To achieve this goal, WMT, EIT, timed GJT, untimed GJT, and MKT were employed, consisting of an equal number of grammatical and ungrammatical structures classified as either easy or

difficult. Regarding the classification of the target structures into easy and difficult, a holistic rating was used (Housen & Simoens, 2016) as experienced EFL teachers rated the target structures through a questionnaire based on R. Ellis's (2005) list of 17 structures problematic to learners. Based on their judgments, 8 structures (4 easy and 4 difficult) were chosen. The results revealed that the easy structures of the GJTs, regardless of the time conditions, primarily tap into automatized explicit knowledge, and the difficult structures in the EI, timed GJT, untimed GJT, and MKT tap into explicit knowledge. Moreover, in all of the tests, there were significant performance differences between the easy and difficult structures and also between grammatical and ungrammatical items.

The results of these studies indicate that structure difficulty can be a contributing task design feature in learners' performance on implicit and explicit tasks. However, this aspect has yet to be systematically and objectively investigated as the above-mentioned studies were not consistent in using selection criteria for dividing the target structure into easy and difficult. Bialystock (1979) and Mehraein et al. (2022) selected the structures based on subjective criteria, and R. Ellis (2006) did not classify target structures into easy/difficult. The only study that employed objective criteria to divide items into easy and difficult was Shui et al. (2018), who only validated the GJT tests without referring to the role of structure difficulty on implicit tests. To the best of our knowledge, no study has comprehensively examined the impact of structure difficulty on fine-grained measures of implicit and automatized explicit knowledge based on objective criteria. Therefore, it is essential to review the objective criteria that determine the ease or difficulty of learning structures, both implicitly and explicitly, focusing on the target structures used in this study.

Objective Criteria for Determining Structure Difficulty

The literature is replete with criteria that can objectively determine the ease/difficulty of learning grammatical structures implicitly or explicitly. Of

these criteria, frequency, saliency, and functional value can explain the ease or difficulty of implicitly acquiring different grammatical features, and metalanguage pertains to structure difficulty in terms of explicit knowledge (N. Ellis 1996, 2002; R. Ellis, 2005, 2006; Goldschneider & DeKeyser, 2005). The challenge of structure difficulty can also be attributed to another factor, which relates to both implicit and explicit knowledge; it pertains to the fact that certain structures are acquired at an early stage in the process of L2 learning, while others are acquired at a later stage (R. Ellis, 2005). These criteria will be explained as follows, highlighting the differences between the target structures in this study.

Frequency

N. Ellis (1996) suggests that implicit acquisition is influenced by the relative frequencies in the linguistic input. Frequent features are generally easier to acquire. Regarding the target structures used in the current study, plural -s is widely frequent in English, appearing in many contexts (Bybee, 2007). In comparison, third-person -s is less frequent in the linguistic input. Thus, frequent exposure to plural -s facilitates its acquisition, whereas less frequent exposure to third-person -s can challenge learners.

Saliency

Goldschneider and DeKeyser (2005) analyzed the order of acquisition of morphemes and found that the sequence is affected by the concept of saliency. Saliency refers to how noticeable a grammatical feature is in the input, with factors like perceptual salience, phonological salience, syntactic category, morphophonological regularity, and frequency contributing to a morpheme's saliency. Morphemes with higher saliency are acquired more easily through implicit processing. Plural -s is salient (easy to notice) and noticeable as it follows a consistent pattern, making it easier for learners to spot and process. Third-person -s is less conspicuous in the input, making it a challenge for learners to notice and process.

Functional Value

Grammatical forms often serve discoursal, semantic, or pragmatic functions. Forms that realize a single, non-redundant function are typically easier to learn compared to forms that serve multiple functions or are often redundant (Nichols, 1992). For example, plural -s holds high functional value as it marks noun plurality. Its clear and distinct function facilitates its acquisition as it aligns with the communicative One-to-One Principle (Andersen, 1984). However, third-person -s expresses present tense, person, and number. This structure, while important for subject-verb agreement, is argued to be redundant for communication. In cases where its omission does not lead to difficulty or misunderstanding, learners might not attach the same level of functional value to this grammatical feature.

Metalanguage

Metalanguage can vary in its level of technicality, ranging from 'semi-technical' to 'technical.' In general, the more technical the metalanguage required to formulate a rule, the more difficult that rule will be to learn (R. Ellis, 2006). According to Hawkins (1999), for plural -s, metalinguistic knowledge requirements are relatively low. The concept (pluralization) is straightforward, and learners can easily understand and articulate the rule without advanced linguistic terminology. Third-person -s may require a higher degree of metalinguistic knowledge. To fully understand this structure, learners may need to grasp concepts related to subject-verb agreement and verb conjugation. This involves more technical metalanguage. As such, learners may face a greater challenge in accessing their explicit knowledge of this grammatical feature.

Early/Late Acquisition

Early-acquired structures are typically those linguistic features or rules that L2 learners tend to learn in their early stages of language development. These structures are often considered "easy" because they are typically acquired with relative ease and processed intuitively. On the contrary, late-acquired

structures are those linguistic features that individuals typically learn at later stages of their language-learning journey. These structures are often considered "difficult" because they may involve complexity, irregularity, or less common usage. Plural -s, a frequently occurring morphological feature in English, is acquired at an earlier stage in L2 development. In contrast, third-person -s is considered a relatively late-acquired structure in L2 learners' language development, and their implicit and explicit knowledge of this structure may not be as well-established as their knowledge of plural -s.

New Implicit Knowledge Measures: Real-Time Comprehension Tests

As previously discussed, Suzuki (2017), Suzuki and DeKeyser (2015), and Vafaee et al. (2017) have questioned the construct validity of the commonly administered tests of implicit knowledge such as timed GJT. In search of measures to substitute old tests of implicit knowledge, these studies have included timed GJT and EIT along with new psycholinguistic measures (WMT, SPRT, and visual world task) and found the latter measures to be purer measures of implicit knowledge. Suzuki and DeKeyser (2015) compared WMT and EIT with MKT, a measure of explicit knowledge. Their results showed that EI performance was associated with MKT scores and that only online error detection tests (e.g., WMT) could index implicit knowledge. Also, Vafaee et al. (2017) hypothesized that imposing time constraints or manipulating grammaticality does not lead GJTs to be measures of implicit knowledge because they draw attention to form. They employed SPRT, WMT, timed GJT, untimed GJT, and MKT and tested 20 CFA models, including new models, previous models (e.g., timed/untimed, or grammatical/ungrammatical), and some rival models. Their results revealed that WMT and SPRT are more fine-grained measures of implicit knowledge, and the different types of GJTs measure different levels of explicit knowledge. Similarly, Suzuki (2017) and Suzuki and DeKeyser (2017) employed three timed form-focused tasks (i.e., a fill-in-the-blank test, an auditory GJT, and a written GJT) in addition to three real-time comprehension

tasks (i.e., a visual-world task, a WMT, and an SPRT). They also considered the role of automatized explicit knowledge. The results of CFA and multitrait-multimethod (MTMM) analyses provided evidence that the two GJTs and fill-in-the-blank test loaded on the factor of automatized explicit knowledge, and the visual-world task, WMT, and SPRT loaded on the factor of implicit knowledge. All of these studies indicated that real-time psycholinguistic measures can indirectly measure learners' grammatical sensitivity without explicitly prompting them to make grammaticality judgments. WMT, as one of the best online processing measures, examines whether L2 speakers are sensitive to grammatical errors while they are reading/listening for comprehension and provides RT (Granena, 2013; Mehraein et al., 2022; Rezaei & Mehraein, 2019; Suzuki, 2017; Suzuki & DeKeyser, 2015, 2017; Suzuki et al., 2023). Therefore, the present study employed WMT, a fine-grained measure of implicit knowledge, as part of the data collection procedure.

RESEARCH QUESTIONS

The following research questions were formulated and examined for the purpose of the study.

1. How do grammaticality (grammatical vs. ungrammatical) and structure difficulty (easy vs. difficult) affect the L2 learners' performance on WMT?
2. How do grammaticality (grammatical vs. ungrammatical) and structure difficulty (easy vs. difficult) affect the L2 learners' performance on timed GJT?

METHOD

Participants

Seventy-four learners of English as a foreign language (EFL) (48 female, 26 male) were recruited in Tehran, Iran. The selection of the participants was based on their proficiency level using C-tests adapted from Ishihara et al.

(2003). Advanced learners were chosen as they were proficient enough to perform the tasks in the study and most probably had implicit and automatized explicit knowledge of the target structures. Therefore, the requirement for participation was a minimum score of 90 out of 100 on the C-tests. Based on this criterion, 41 EFL learners were selected and asked on Telegram to arrange an appointment with the experimenter to attend the lab. Of these 41 students, only 33 (18 females and 15 males) in the age range of 18 to 39 completed the tasks. They consisted of undergraduates (n = 21) majoring in English literature, M.A. students/holders (n = 9), and Ph.D. candidates/holders (n = 3) majoring in teaching English as a foreign language (TEFL). All participants received an incentive upon completion of the tasks, either as a course credit or a small fee. Each participant filled in written informed consent prior to the experiment.

Target Structures

This study targeted two English structures - plural -s and third-person -s. The reasons for selecting these structures were three-fold. First, as the current study is part of a larger experiment conducted in an EEG laboratory, these two structures could easily be incorporated into psycholinguistic measures (e.g., WMT) in an EEG task. Second, R. Ellis (2005, 2009) suggested that plural -s and third-person -s are among the most problematic structures for L2 learners. Third, these target structures were hypothesized to differ in terms of their structure difficulty based on the objective criteria reviewed in the review of the literature section (i.e., frequency, saliency, functional value, metalanguage, and early/late acquisition).

Adopting all these criteria, we considered plural -s structure as the easy structure and third-person -s as the difficult one, with plural -s benefitting from the features that facilitate its implicit and explicit knowledge acquisition. Therefore, it was hypothesized that L2 learners possess higher levels of implicit and explicit knowledge for the easy structure. Having selected the structures, different items were written, half of them grammatical

and half ungrammatical. Two counterbalanced lists were created. These sets were similar in terms of item length (8 to 10 words). In List 1, half of the items were grammatical, and half ungrammatical. The grammaticality of the sentences was reversed in List 2 to make sure no target item appeared twice in one list.

Instruments

English Proficiency Test

C-tests adapted from Ishihara et al. (2003) were used to establish the participants' general English proficiency. There is a growing body of evidence supporting the validity of C-tests as predictors of L2 learners' general proficiency (Dörnyei & Katona, 1992; Eckes & Baghaei, 2015; Eckes & Grotjahn, 2006). To answer these tests, linguistic knowledge, textual knowledge, discourse structure, reading skills and strategies, knowledge of the world, and expectancy grammar are required. Therefore, cloze tests could be used to predict overall language ability (Heaton, 1988). In the current study, three C-test passages were used. In these passages, the second half of every fifth or sixth word was deleted for the test takers to complete. Each C-test comprised a passage of 112 to 143 words with 50 partially filled blanks in total. The maximum score was 100 points, with two points allocated for each correct response. C-tests were selected due to their quick administration time (10 –15 minutes). The results of the sampled population's performance (N = 33) on the C-tests indicated that scores ranged from 90 to 99 with a mean score of 93.84 and standard deviation of 2.48, suggesting that the participants were homogenous and proficient enough (advanced L2 learners) to perform the tasks in the study.

Word Monitoring Task

The WMT served as a tool for measuring participants' sensitivity to grammatical errors while engaging in online sentence reading for comprehension. The WMT imposed dual-task conditions: word monitoring

and sentence comprehension. In this task, participants were presented with a monitoring word in a rectangular box in the center of the screen for 2 seconds and were asked to read a subsequent sentence appearing on the monitor. The words in the sentence were displayed automatically one at a time for 450 ms with interstimulus intervals of 200 ms. The participants were required to press the space button immediately when they identified the monitoring word that appeared after the relevant target structure in a sentence. RTs were measured from the onset of the monitoring word to the point they pushed the button. To ensure that the participants processed the sentences for meaning, they were presented with follow-up true/false comprehension statements after each item and were requested to indicate its truth or falsehood by pressing two fixed keys on the keyboard (i.e., up key for true, and down key for false). The ratio between the true and false answers was kept equal. Each comprehension statement remained on the screen for 5 seconds. This dual-task paradigm minimizes the application of (automatized) explicit knowledge during real-time comprehension. This is achieved by time-locking the utilization of grammar knowledge to hundreds of milliseconds (Suzuki & DeKeyser, 2015, 2017). The WMT is, thus, arguably a pure measure of implicit knowledge (Suzuki, 2017; Suzuki & DeKeyser, 2015; Vafaee et al., 2017).

This task included 128 experimental items (64 grammatical and 64 ungrammatical - 64 targeting the easy structure and 64 the difficult structure) with a sentence length of 8 to 10 words. The monitoring words consistently emerged as either the fifth or sixth word in all experimental items, making the critical regions the fourth or fifth word accordingly. Sixty-four filler sentences (half grammatical and half ungrammatical) targeting other grammatical structures were also used. The monitoring word in the fillers appeared in different regions so as to prevent participants from predicting where this word would appear. Participants familiarized themselves with the procedure by first completing eight practice trials (4 grammatical and 4 ungrammatical – 4 easy items and 4 difficult) before the experiment. The following are two samples of the experimental items in WMT, including the

monitoring words (underlined), target structures (bolded), and the follow-up statements.

(1) **Easy Structure (Plural -s)**

Grammatical: We have heard several **employees** witnessed the accident.

Ungrammatical: We have heard several **employee** witnessed the accident.

Follow-up: Several employees witnessed the accident. (True)

Grammatical: We realized few **students** attended the graduation ceremony.

Ungrammatical: We realized few **student** attended the graduation ceremony.

Follow-up: All the students attended the ceremony. (False)

(2) **Difficult Structure (Third-Person -s)**

Grammatical: Our experienced manager **stays** calm in all complicated situations.

Ungrammatical: Our experienced manager **stay** calm in all complicated situations.

Follow-up: The manager gets angry very soon. (False)

Grammatical: The computer program **corrects** spelling and grammatical mistakes.

Ungrammatical: The computer program **correct** spelling and grammatical mistakes.

Follow-up: The program can correct mistakes. (True)

WMT was used to investigate structure difficulty and grammaticality in terms of implicit knowledge as it calls for a primary focus on meaning and does not require learners to utilize metalinguistic knowledge. The rationale behind this task is that the difference in the RTs to target word identification between grammatical and ungrammatical items provides the index for online grammatical sensitivity as participants are expected to unconsciously slow down whenever they face an ungrammaticality. It is hypothesized that there is a delay in response to a monitoring word that follows an ungrammatical

form, compared to a grammatical form, if they have implicit knowledge of the target features. By contrast, if a participant's implicit knowledge is not well-developed, they may not perceive the ungrammaticality of the preceding ungrammatical form, leading to minimal impact on his or her RTs (Godfroid, 2016). With regard to the effect of structure difficulty, it is hypothesized that the L2 learners demonstrate more sensitivity to the violations in the easy structure compared to the difficult one, as they are expected to have more implicit knowledge of the easy structure. These hypotheses can be checked by computing the participants' grammaticality sensitivity index (GSI) (Godfroid, 2016; Granena, 2013; Suzuki & DeKeyser, 2015). GSIs are computed by subtracting the RTs to grammatical items from the RTs to ungrammatical items for both easy and difficult structures. A positive high GSI suggests a slowdown in processing ungrammatical items, reflecting grammatical sensitivity and implicit knowledge (Godfroid, 2016). Negative GSIs and GSIs close to zero imply a lack of sensitivity, which points to a lack of implicit knowledge (Godfroid, 2016). Furthermore, to ensure participants' involvement in dual-task processing, they were required to achieve a 75% comprehension accuracy threshold for inclusion in the data analyses (Suzuki, 2017). Following this, one participant was excluded from further analysis due to his high error rate. Reliability indexed by Split-half reliability for this task was .82.

Timed Grammaticality Judgment Test

In this computer-delivered TGJT, the items were automatically presented in the center of the screen word by word. Each word was displayed for 450 ms with interstimulus intervals of 200 ms, and at the end of each sentence, a question mark was presented on the screen for 1000 ms. The participants were asked to make a judgment as quickly and accurately as possible as to whether the sentence they read was grammatical or ungrammatical using two fixed keys on the keyboard (up key for grammatical and down key for ungrammatical). They were told that they could press the buttons while

reading the items or after seeing the question mark. In TGJT, participants' attention was solely on the form, prompting them to consciously and rapidly draw upon their automatized explicit knowledge, which is different from the requirements of real-time comprehension tasks such as WMT (DeKeyser, 2003, 2009; Mehraein et al., 2022; Suzuki & DeKeyser, 2017; Vafaee et al., 2017). According to Paradis (2009) and Suzuki and DeKeyser (2015, 2017), L2 learners (mostly advanced learners) access their (automatized) explicit knowledge consciously and quickly, even when pressured for time.

The stimulus sentences comprised 128 sentences, equally divided into grammatical and ungrammatical, as well as easy and difficult structures. The sentences differed from those employed in the WMT, and no filler sentences were included. Before the experiment, the participants were given eight practice sentences (4 grammatical and 4 ungrammatical – 4 easy and 4 difficult) to familiarize them with the fast-paced nature of the test. Provided below is a set of two experimental items for easy and difficult structures, with highlighted sections representing target structures.

(3) **Easy Structure (Plural -s)**

Grammatical: Almost all of these **experiments** have serious flaws.

Ungrammatical: Almost all of these **experiment** have serious flaws.

(4) **Difficult Structure (Third-Person -s)**

Grammatical: My intelligent friend **persuades** his father to lose weight.

Ungrammatical: My intelligent friend **persuade** his father to lose weight.

Although this test is not a pure measure of explicit knowledge, it was hypothesized that as paying attention to form inevitably raises L2 learners' awareness and prompts them to access their metalinguistic knowledge, TGJT serves as a measure of their automatized explicit knowledge of the easy and difficult structures. It is predicted that they score higher on grammatical sentences and also on easy items compared to the ungrammatical and difficult items.

TGJT items were scored dichotomously (i.e., correct and incorrect), and unanswered items were counted as incorrect. Furthermore, participants with an error rate exceeding 25% were eliminated from the analysis. Two individuals scored below 75%, leading to their exclusion from subsequent analysis. Then, the accuracy scores were computed for the grammatical and ungrammatical items as well as the easy and difficult sentences based on the responses of the remaining participants. The maximum total score for the task was 128, with each condition having a maximum score of 32. The reliability indexed by Cronbach's alpha for TGJT was .79.

Procedure

Prior to the data collection session, participants were selected based on their performance in C-tests adapted from Ishihara et al. (2003) to ensure all participants were advanced learners of English. The study was conducted in one session, with each participant individually tested in a quiet EEG lab room. Upon arrival, participants received an overview of the study and were asked to fill out a written informed consent form. The experimental session began with a brief familiarization phase. Participants were introduced to the tasks through completing eight practice trials for each test (four grammatical and four ungrammatical) to become accustomed to the procedure. Following the practice phase, participants completed the two main tasks, progressing from implicit to explicit: the WMT, followed by the timed GJT. Given the substantial number of items (196 in the WMT and 128 in the TGJT), scheduled breaks were provided to maintain engagement and accuracy. Participants received one 5-minute break during the TGJT and two 5-minute breaks during the WMT, during which they were offered snacks. This allowed participants to rest and remain focused throughout the session. Both tasks were programmed and delivered using MATLAB to ensure precise timing and data collection. The collected data were analyzed to compute RTs and accuracy scores. The entire session, including consent, practice, task execution, and breaks, lasted approximately 2 hours.

RESULTS

RQ1. How do grammaticality (grammatical vs. ungrammatical) and structure difficulty (easy vs. difficult) affect the L2 learners' performance on WMT?

Table 1 presents the descriptive statistics for WMT for the 32 participants. As the table indicates, the mean RTs for the grammatical and ungrammatical sentences targeting the easy structure (plural -s) were 588 (SD = 78) and 609 (SD = 66), respectively, and the mean RTs were 574 (SD = 70) and 585 (SD = 60) for grammatical and ungrammatical items of the difficult structure (third-person -s). This shows that the learners were more sensitive to the easy structure; however, irrespective of structure difficulty, they showed more slowdowns in ungrammatical items. Following Godfroid (2016) and Granena (2013), GSIs were also computed by subtracting the RTs to grammatical items from the RTs to ungrammatical items. The magnitude of this sensitivity index served as a measure of L2 learners' implicit knowledge (Suzuki & DeKeyser, 2015, 2017). A higher GSI for the structure reveals more implicit knowledge of that structure. Therefore, as Table 1 shows, the participants had larger GSI in the case of easy items, indicating stronger implicit knowledge of plural -s compared to the third-person -s across conditions in the WMT.

Table 1
Descriptive Statistics for RTs in the WMT

	Grammatical *M (SD)*	Ungrammatical *M (SD)*	GSI
Easy Structure (Plural -s)	588 (78)	609 (66)	21
Difficult Structure (Third-Person -s)	574 (70)	585 (60)	11

To check the normality of the distribution of RTs, the Kolmogorov-Smirnov test of normality was conducted, which confirmed the normality assumption of the data from WMT. Following this, a repeated measures ANOVA with grammaticality and structure type as within-participant factors was conducted. The results showed a significant main effect for grammaticality ($F(1, 31) = 5.961$, $p = .021$, $\eta_p^2 = .161$), a main effect for structure difficulty

(F (1, 31) = 11.891, p = .002, η_p^2 =.277), and no interaction between the two variables (F (1, 31) = 1.646, p = .209, η_p^2 =.050). Regarding the effect sizes, the guideline for interpreting the magnitude of partial eta squared effect sizes (i.e., .01 small; .06 medium; .14 large) (Gray & Kinnear, 2012) was used. Both structure difficulty and grammaticality had large effect sizes, with the former being even larger. These results, as can be seen in Figure 1, suggest that the participants indicated larger slowdowns in responding to the monitoring word in (a) ungrammatical sentences of both structure types compared to their grammatical counterparts, and (b) easy items compared to difficult ones.

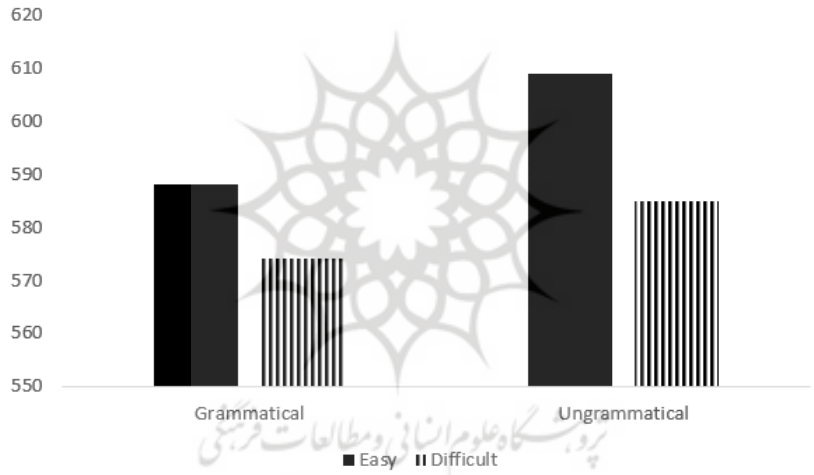

Figure 1.
Mean RTs (ms) to Easy and Difficult Items Across Conditions in the WMT

RQ2. How do grammaticality (grammatical vs. ungrammatical) and structure difficulty (easy vs. difficult) affect the L2 learners' performance on timed GJT?

Table 2 shows the mean accuracy scores (and standard deviations (SD)) for easy and difficult items across the different grammaticality conditions in the timed GJT.

Table 2
Descriptive Statistics for Accuracy Scores in the Timed GJT

	Grammatical *M (SD)*	Ungrammatical *M (SD)*
Easy Structure (Plural -s)	29.77 (2.12)	29.38 (2.23)
Difficult Structure (Third-Person -s)	30.00 (1.91)	27.87 (2.47)

Prior to conducting the repeated measures ANOVA, the Kolmogorov-Smirnov test confirmed the normality of the data for the ungrammatical items, but this was not the case for the grammatical conditions. Nevertheless, it should be noted that repeated measures ANOVA is generally robust to non-normality when the sphericity assumption is met, which was the case in this study (Blanca et al., 2023). The results of the repeated measures ANOVA for GJT items with grammaticality and structure difficulty as within-participant factors showed a significant main effect for grammaticality, ($F (1, 30) = 8.390$, $p = .007$, $\eta_p^2 = .219$), structure difficulty, ($F (1, 30) = 4.419$, $p = .044$, $\eta_p^2 = .128$), and an interaction between grammaticality and structure difficulty ($F (1, 30) = 8.205$, $p = .008$, $\eta_p^2 = .215$). Concerning the effect sizes, grammaticality had a large effect size, and the structure difficulty effect tended to be approaching large. There was also a large effect size for their interaction, showing that the effect of grammaticality varied depending on the structure difficulty. As Figure 2 indicates, participants performed better on grammatical items for both target structures, and their performance was significantly better in plural -s structure than in third person -s structure only in ungrammatical items.

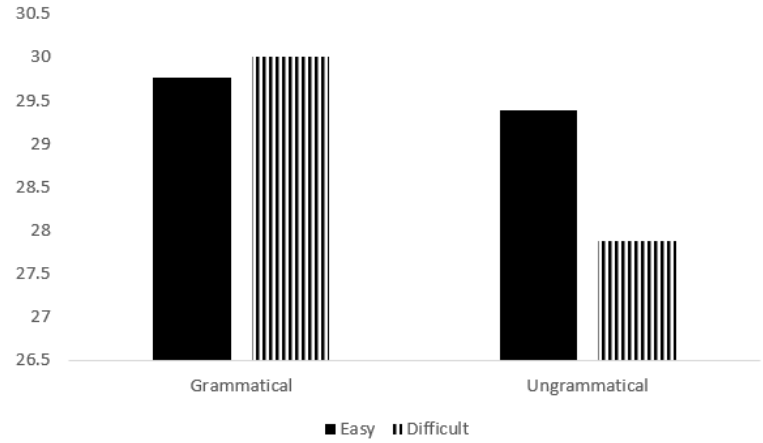

Figure 2.
Mean Accuracy Scores of Easy and Difficult Items Across Conditions in the Timed GJT

When comparing participants' performance on easy and difficult items across both tasks, it was observed that learners performed significantly better on and demonstrated greater sensitivity to the easy items compared to the difficult ones. This lends support to the hypothesis that the third-person -s was more difficult than the plural -s, even for the advanced participants.

DISCUSSION

The prime goal of this study was to find out whether L2 learners' performance on target structures hypothesized to differ in terms of difficulty (i.e., third-person -s more difficult than plural -s) can be interpreted as a contributing factor in measuring implicit and explicit knowledge. More specifically, the present study aimed to examine the effects of grammaticality and structure difficulty on L2 learners' performance on tests of implicit and automatized explicit knowledge (i.e., WMT and TGJT). The descriptive statistics of RTs and accuracy scores and the results of the repeated-measures ANOVAs provide answers to this purpose. The mean RTs and accuracy scores on the grammatical items were significantly faster and more accurate than those on the ungrammatical ones, and the mean RTs and accuracy scores on the easy

(i.e., plural -s) items were significantly slower and more accurate than those on the difficult (i.e., third-person -s) items. The RT data in WMT indicated more slowdowns in ungrammatical items and also in easy sentences, demonstrating L2 learners' sensitivity to these items. In addition, the RTs showed larger GSI for plural -s items, revealing participants' more implicit knowledge of this structure compared to the third-person -s, as anticipated. As the results of timed GJT show, the L2 learners in this study judged grammatical items more accurately than ungrammatical ones and the plural -s items more accurately than third-person -s ones. This also indicates that although L2 learners have grammatical knowledge of both structures, they exhibited a greater degree of automatized explicit knowledge of the easy structure. Overall, these results indicate that the L2 learners processed grammatical vs. ungrammatical and easy vs. difficult sentences differently when performing different implicit and explicit knowledge tests and that they showed greater degrees of implicit and explicit knowledge of the easy structure.

The finding that L2 learners' RTs to the ungrammatical items in WMT were longer than those to the grammatical ones is in line with the results of previous studies (Mehraein et al., 2022; Suzuki & DeKeyser, 2015; Vafaee et al., 2017). A plausible interpretation of this finding is that in the case of grammatical violations, the learners used the information of the linguistic structures based on their implicit knowledge of the structures; therefore, they spent more time processing the errors because these violations were opposed to their expectations. As a result, RTs to the target words that appeared after ungrammatical sentences were slower.

The observed effect of structure difficulty based on RTs in WMT contradicts the results of Mehraein et al.'s (2022) study, which revealed that the RTs to difficult structures were significantly longer than those to easy structures in WMT. The results of the current study revealed that advanced learners showed more slowdowns and higher GSI in the easy items than in the difficult ones. This implies that the L2 learners showed more sensitivity to the easy items, revealing their more implicit knowledge of plural -s

compared to the difficult items (Godfroid, 2016). The controversy between the two studies may arise from variations in the target structures and the criteria used for selecting those structures. To select easy and difficult structures, Mehraein et al. (2022) employed a subjective, holistic rating based on experienced EFL teachers' judgments on R. Ellis's (2005) list of problematic structures and chose 8 structures (i.e., 4 easy, 4 difficult), quite different from the ones in the current study. In the present study, only two structures (i.e., plural -s and third-person -s) were chosen as easy and difficult structures, respectively, based on more objective criteria: frequency, saliency, functional value, metalanguage, and early/late acquisition of structures.

Regarding the timed GJT, in line with the results of numerous GJT studies (e.g., Bialystok, 1979, 1986; R. Ellis, 2005; Gutierrez, 2013; Kim & Nam, 2017; Loewen, 2009; Mehraein et al., 2022; Shiu et al., 2018; Vafaee et al., 2017), our findings revealed that the L2 learners exhibited significantly better performance on the grammatical than on the ungrammatical items. One potential rationale for this is that ungrammatical items require more analysis of linguistic knowledge (even for advanced learners), and this can be problematic in a timed task. However, the finding regarding the role of grammaticality contradicts that reported in Bley-Vroman et al. (1988), where learners exhibited greater accuracy in judging ungrammatical sentences compared to grammatical ones. A possible reason for this discrepancy may be related to task design features (e.g., time pressure or number of experimental items).

Concerning the effect of structure difficulty shown in TGJT, the results of the current study are in line with those of Bialystok (1979), Mehraein et al. (2022), and Shiu et al. (2018) that all showed L2 learners performed better and scored higher in easy structures. This reveals that the participants had more automatized explicit knowledge of the easy items (or plural -s). The findings also conform to those of R. Ellis (2006), in which learners scored higher in plural -s structure than third-person -s in both implicit and explicit tasks. However, in R. Ellis's (2006) study, no psycholinguistic (pure) measure of implicit knowledge was used, and timed

GJT was treated as a measure of implicit knowledge. Moreover, the plural -s items in his study differed from those used in the current study based on number and type. R. Ellis (2006) used only four items (two grammatical and two ungrammatical) to measure learners' knowledge of plural -s. One set (grammatical and ungrammatical) was similar to the items used in our study, but the other set measured the learners' knowledge of the plural -s using coordination (e.g., **Martin sold a few old coins and stamp to a shop*). Thus, the similarity between the findings of the two studies can be related to the inherent difficulty of the plural -s and third-person -s. These results suggest that, irrespective of the number or type of items used or the specific linguistics tasks employed, L2 learners have a stronger linguistic knowledge, whether implicit or automatized explicit, of the plural -s compared to the third-person structure.

The results demonstrated that advanced L2 learners possess more implicit and automatized explicit knowledge of the structures that are frequent and salient in the input, hold high functional value, require less metalinguistic knowledge, and are early-acquired. This result is partially supported by Shiu et al. (2018), who chose past progressive as the easy and passive as the difficult structures based on saliency, frequency, inherent structural complexity, and early/late acquisition. L2 learners' performance on GJTs indicated that the learners performed significantly better on the easy items that were frequent, salient, and early-acquired than on the passive ones. Although the selection criteria were not the same in the two studies, both investigations yielded similar findings showing a higher level of learners' automatized explicit knowledge pertaining to the easy structures.

CONCLUSION

The present study set out to explore the role of grammaticality and structure difficulty in measuring advanced L2 learners' implicit and explicit knowledge using WMT and timed GJT. Based on the criteria of frequency, saliency, functional value, metalanguage, and early/late acquisition, plural -s, and

third-person -s were chosen as easy and difficult structures, respectively, and then, grammatical and ungrammatical items were written. The data analyses showed that both grammaticality and structure difficulty contributed to the retrieval of implicit and automatized explicit knowledge. Furthermore, although L2 learners possessed implicit and explicit knowledge of both target structures, they indicated stronger implicit and explicit knowledge of the easy structure, that is plural -s, as expected.

The findings of this study have significant implications for the field of SLA, particularly in understanding how different types of structures are processed by advanced learners. The study suggests that frequent, salient, and functionally valuable structures, like the plural -s, are acquired more readily and are more deeply ingrained in learners' implicit knowledge systems. This supports the notion that language instruction should emphasize such features to enhance learning efficiency and effectiveness. Additionally, the study highlights the importance of structure difficulty in L2 learning. The differential performance on the plural -s and third-person -s structures underscores the need for tailored instructional approaches that address the specific challenges posed by more difficult grammatical forms. Educators and curriculum developers should consider incorporating strategies that increase the saliency of difficult structures using input enhancement strategies, thereby facilitating their acquisition. This approach can provide valuable insights into the intricacies of language acquisition and inform more effective pedagogical practices. Moreover, the findings advocate for the use of psycholinguistic measures in SLA research to obtain a nuanced understanding of learners' grammatical sensitivity and knowledge types.

However, our study has a few limitations that should be borne in mind when interpreting its results. The first limitation lies in the fact that the test scores were derived from only two target structures — one easy (plural -s) and one difficult (third-person -s). The limited number of target structures examined in the present study could restrict the generalizability of our findings. Therefore, it is imperative that researchers conduct further studies on a variety of target structures. The second limitation has to do with the

relatively limited sample size of this study. Moreover, there was a lack of variation in terms of the participants' language proficiency (i.e., they were advanced L2 learners). Thus, for more reliable results in future studies, it may be worth expanding the current research to a larger sample size and different proficiency levels. Third, in our study, we employed only two measures - WMT for implicit knowledge and timed GJT for automatized explicit knowledge, which might have jeopardized the accuracy of the findings. As such, it is recommended that future studies use more than one fine-grained measure for each knowledge type to more accurately reflect the learners' degree of implicit and explicit knowledge.

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Sepideh Mehraein http://orcid.org/0000-0002-9457-1757
Hamideh Marefat http://orcid.org/0000-0001-6281-4470

References

Akakura, M. (2012). Evaluating the effectiveness of explicit instruction on implicit and explicit L2 knowledge. *Language Teaching Research, 16*(1), 9-37.

Andersen, R. W. (1984). The one to one principle of interlanguage construction. *Language Learning, 34*(4), 77-95.

Andringa, S., & Rebuschat, P. (2015). New directions in the study of implicit and explicit learning: An introduction. *Studies in Second Language Acquisition, 37*(2), 185-196.

Baleghizadeh, S., & Derakhshesh, A. (2017). Measuring the effectiveness of explicit and implicit instruction through explicit and implicit measures. *Iranian Journal of Applied Linguistics, 20*(1), 81-111.

Bialystok, E. (1979). Explicit and implicit judgements of L2 grammaticality. *Language Learning, 29*(1), 81-103.

Bialystok, E. (1986). Factors in the growth of linguistic awareness. *Child Development, 57*(2), 498-510.

Blanca, M. J., Arnau, J., García-Castro, F. J., Alarcón, R., & Bono, R. (2023). Repeated measures ANOVA and adjusted F-tests when sphericity is violated: Which procedure is best? *Frontiers in Psychology, 14*, 1192453.

Bley-Vroman, R., Felix, S., & Loup, G. (1988). The accessibility of universal grammar in adult language learning. *Second Language Research, 4*(1), 1-32.

Bowles, M. A. (2011). Measuring implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition, 33*(2), 247-271.

Bybee, J. L. (2007). *Frequency of use and the organization of language.* Oxford University Press.

DeKeyser, R. M. (2003). Implicit and explicit learning. In C. J. Doughty & H. M. Long (Eds.), *The handbook of second language acquisition* (pp. 312-348). Blackwell Publishers.

DeKeyser, R. (2009). Cognitive-psychological processes in second language learning. In M. H. Long & C. J. Doughty (Eds.), *The handbook of language teaching* (pp. 119–138). Wiley-Blackwell.

Dörnyei, Z., & Katona, L. (1992). Validation of the C-test amongst Hungarian EFL learners. *Language Testing, 9*(2), 187-206.

Eckes, T., & Baghaei, P. (2015). Using testlet response theory to examine local dependence in C-tests. *Applied Measurement in Education, 28*(2), 85-98.

Eckes, T., & Grotjahn, R. (2006). A closer look at the construct validity of C-tests. *Language Testing, 23*(3), 290-325.

Ellis, N. C. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18*(1), 91-126.

Ellis, N. C. (2002). Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition, 24*(2), 143-188.

Ellis, R. (1997). *Second language acquisition.* Oxford University Press.

Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. *Studies in Second Language Acquisition, 27*(2), 141-172.

Ellis, R. (2006). Modelling learning difficulty and second language proficiency: The differential contributions of implicit and explicit knowledge. *Applied Linguistics, 27*(3), 431-463.

Ellis, R. (2009). Measuring implicit and explicit knowledge of a second language. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 31-64). Multilingual Matters.

Ellis, R., & Loewen, S. (2007). Confirming the operational definitions of explicit and implicit knowledge in Ellis (2005): Responding to Isemonger. *Studies in Second Language Acquisition, 29*(1), 119-126.

Gass, S. (1983). The development of L2 intuitions. *TESOL Quarterly, 17*(2), 273-291.

Ghahari, S., & Piruznejad, M. (2016). Recast and explicit feedback to young language learners: Impacts on grammar uptake and willingness to communicate. *Issues in Language Teaching, 5*(2), 209-187.

Ghorbani, M., & Atai, M. R. (2012). Evaluating the effectiveness of explicit and implicit form-focused instruction on explicit and implicit knowledge of EFL learners. *Issues in Language Teaching, 1*(1), 58-92.

Godfroid, A. (2016). The effects of implicit instruction on implicit and explicit knowledge development. *Studies in Second Language Acquisition, 38*(2), 177-215.

Goldschneider, J. M., & DeKeyser, R. M. (2005). Explaining the "Natural Order of L2 Morpheme Acquisition" in English: A meta-analysis of multiple determinants. *Language Learning, 55* (S1), 27-77.

Granena, G. (2013). Individual differences in sequence learning ability and second language acquisition in early childhood and adulthood. *Language Learning, 63*(4), 665–703.

Gray, C., & Kinnear, P. R. (2012). *IBM SPSS statistics 19 made simple*. London: Psychology Press.

Gutiérrez, X. (2013). The construct validity of grammaticality judgment tests as measures of implicit and explicit knowledge. *Studies in Second Language Acquisition, 35*(3), 423-449.

Hawkins, J. A. (1999). Processing complexity and filler-gap dependencies across grammars. *Language 75*(2), 244–285.

Heaton, J. B. (1988). *Writing English language tests*. Longman.

Housen, A., & Simoens, H. (2016). Introduction: Cognitive perspectives on difficulty and complexity in L2 acquisition. *Studies in Second Language Acquisition, 38*(2), 163-175.

Hulstijn, J. H. (2005). Theoretical and empirical issues in the study of implicit and explicit second-language learning: Introduction. *Studies in Second Language Acquisition, 27*(2), 129-140.

Isemonger, I. M. (2007). Operational definitions of explicit and implicit knowledge: Response to R. Ellis (2005) and some recommendations for future research in this area. *Studies in Second Language Acquisition, 29*(1), 101-118.

Isbell, D. R., & Rogers, J. (2021). Measuring implicit and explicit learning and knowledge. In P. Winke & T. Brunfaut (Eds.), *The Routledge handbook of second language acquisition and language testing* (pp. 304–313). Routledge.

Ishihara, K., Hiser, E., & Okada, T. (2003). Modifying C-test for practical purposes. *Doshisha Studies in Language and Culture, 5*(4), 539-568.

Kim, J. E., & Nam, H. (2017). Measures of implicit knowledge revisited: Processing modes, time pressure, and modality. *Studies in Second Language Acquisition, 39*(3), 431-457.

Loewen, S. (2009). Grammaticality judgment tests and the measurement of implicit and explicit L2 knowledge. In R. Ellis, S. Loewen, C. Elder, R. Erlam, J. Philp & H. Reinders (Eds.), *Implicit and explicit knowledge in second language learning, testing and teaching* (pp. 94-112). Multilingual Matters.

Mehraein, S., Marefat, H., & Karami, H. (2022). L2 English knowledge: Implicit, explicit, or automatized? The role of structure difficulty. *Journal of Asia TEFL, 19*(1), 17-34.

Nichols, J. (1992). *Linguistic diversity in space and time.* University of Chicago Press.

Paradis, M. (2009). *Declarative and procedural determinants of second languages*. Philadelphia, PA: John Benjamins Publishing Company.

Rezaei, A., & Mehraein, S. (2019). Implicit and explicit instruction and EFL learners' implicit knowledge development: Evidence from word monitoring task. *Iranian Journal of Applied Linguistics, 22*(1), 116-153.

Sharwood Smith, M. (1981). Consciousness-raising and the second language learner. *Applied Linguistics, 2*(2), 159-169.

Shiu, L. J., Yalçın, Ş., & Spada, N. (2018). Exploring second language learners' grammaticality judgment performance in relation to task design features. *System, 72*, 215-225.

Suzuki, Y. (2017). Validity of new measures of implicit knowledge: Distinguishing implicit knowledge from automatized explicit knowledge. *Applied Psycholinguistics, 38*(5), 1229-1261.

Suzuki, Y., & DeKeyser, R. (2015). Comparing elicited imitation and word monitoring as measures of implicit knowledge. *Language Learning*, *65*(4), 860-895.

Suzuki, Y., & DeKeyser, R. (2017). The interface of explicit and implicit knowledge in a second language: Insights from individual differences in cognitive aptitudes. *Language Learning*, *67*(4), 747-790.

Suzuki, Y., Jeong, H., Cui, H., Okamoto, K., Kawashima, R., & Sugiura, M. (2023). An fMRI validation study of the word-monitoring task as a measure of implicit knowledge: Exploring the role of explicit and implicit aptitudes in behavioral and neural processing. *Studies in Second Language Acquisition, 45*(1), 109-136.

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition, 39*(1), 59-95.

Williams, J. N. (2009). Implicit learning in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *The new handbook of second language acquisition* (pp. 319-53). Emerald Press.

Zhang, R. (2015). Measuring university-level L2 learners' implicit and explicit linguistic knowledge. *Studies in Second Language Acquisition, 37*(03), 457-486.