Age of Exposure, Task Stimulus, and Type of Second Language Knowledge

Muhammad Asif Qureshi
Zayed University, Abu Dhabi, UAE.

Introduction

Although grammaticality judgment tasks have been extensively used in second language (L2) research, the type of knowledge they elicit has been debated. The two major points discussed with respect to the type of L2 knowledge that GJTs examine include stimulus type and learner’s age. For stimulus type, it has been argued that grammatical sentences measure implicit knowledge (IK), while ungrammatical sentences gauge explicit knowledge (EK) (e.g., Loewen, 2009). For learner’s age, the previous research suggests that younger learners perform better on the measure of IK, while older learners score higher on instruments assessing EK (Gutiérrez, 2013). If the stimulus type argument is accurate, one important claim would follow that younger learners would perform better on the grammatical items that evoke IK than ungrammatical items that elicit EK. This would be reversed for adult learners, who would perform better on the ungrammatical items than the grammatical items; hence, they would perform better in EK than IK. However, previous research has reported varying conclusions about the type of L2 knowledge GJTs invoke, e.g., studies reporting that grammatical sentences measure IK and ungrammatical sentences tap into EK (Ellis, 2009), and those reporting that stimulus types do not render GJTs as either IK or EK factors (Vafee, Suzuki, & Kachisnke, 2016). As the controversy persists, this report, therefore, investigates the relationship between age and type of L2 knowledge as demonstrated by the grammatical and ungrammatical halves in the GJT as compared to an editing task by early and late L2 learners.

Literature Review

Because GJTs are consistently used in L2 research, it is crucial to investigate the type of knowledge they explore (Douglas, 2001; Purpura, 2004). Opinions on the type of L2 knowledge that GJTs elicit range from some degree of IK (Bialystok, 1979) to probable use of IK (Schachter & Yip, 1990) to a combination of the two (Han, 2000; Han & Ellis, 1998). However, GJTs might evoke/elicit both: implicit and explicit knowledge, depending on the features of task design, such as: (a) time constraints – whether tests are speeded (Bialystok, 1979; Ellis, 1991), (b) task stimulus – whether the items investigated are grammatical or ungrammatical (Ellis, 2009; Loewen, 2009), and (c) task modality – whether tests are aural or written (Granena, 2012). Generally, it has been proposed that aural and speeded tasks and grammatical items on GJTs elicit IK, while untimed and written tasks and ungrammatical items produce EK.
Regarding task stimulus – which is the focus of the current study, Ellis (2004) proposes a three-tier process that learners go through while performing a GJT. These include (a) semantic processing, (b) noticing, and (c) reflecting. During semantic processing, L2 learners attempt to understand the meaning of a given sentence; during noticing, learners try to find if a sentence contains any errors; and during reflection, they consider what is incorrect in a sentence, and why it is erroneous. Loewen (2009) argues that the first two steps might be equally applicable to both grammatical and ungrammatical sentences, because in both cases L2 learners need to understand the meaning of a sentence and see if there are any errors in the sentence. However, Loewen (2009) states that reflection may not happen at all if L2 learners notice a sentence as grammatical because they might stop thinking about the sentence. In contrast, for ungrammatical sentences, learners need to reflect to find inaccuracies in the sentence; hence, they reflect and use explicit knowledge.

To confirm the relationship between task stimulus (i.e., grammatical and ungrammatical) and type of L2 knowledge (i.e., implicit and explicit), several investigations have been carried out, but their findings remain inconclusive. Bialystok (1979) found that learners used IK for both task stimuli, while Ellis (2005) and Ellis and Loewen (2007) found that ungrammatical knowledge loaded more strongly on the EK factor in an untimed test.

As task stimulus and age of exposure to L2 are intricately linked, researchers have also attempted to investigate whether early or late learners would perform better on one of the two types of L2 knowledge. Ellis (2005) and Gutiérrez (2013) posit that implicit knowledge is limited to only younger learners, but explicit knowledge can be learnt at any age. Moreover, Bley-Vroman’s (1988) Fundamental Difference Hypothesis (FDH) posits that adults cannot access an innate mechanism for learning a language implicitly; they must use their problem solving skills (DeKeyser, 2000). Considering the FDH and the theory that explicit knowledge can be learned at any age, it can be hypothesized that late L2 learners would perform better on ungrammatical items than on grammatical items.

One way to check such a hypothesis is to compare younger and late L2 learners’ scores on the grammatical and ungrammatical halves of a GJT with their performance on another task (Loewen, 2009). Han and Ellis (1998), for instance, used a timed and a delayed GJT, an oral production task, and a metalinguage task. This study uses an editing task along with a loosely-timed GJT to examine early and late L2 learners’ grammatical proficiency in relation to type of L2 knowledge.

The Current Study

This study is based on the hypothesis that the grammatical items in the GJT (G-GJT) will reflect IK, whereas items in the editing task and the ungrammatical sentences in the GJT (U-GJT) will represent EK. This assumption is based on the findings that posit that the test types/items requiring “analysis” and “conscious awareness” elicit EK (Ellis, 2005). It is, therefore, assumed that the U-GJT will encourage analysis for error identification. Similarly, items in the ET will also require participants to first recognize the error/s, and then correct those. This process of error identification and correction will require L2 learners to explicitly understand the errors. On the other hand, the G-GJT may stimulate a more spontaneous response, because the absence of an error may not impose extra demands for identifying an error and then correcting it (i.e., Loewen, 2009). Thus, inclusion of an ET complements the GJT in a way that provides an opportunity to examine different types of knowledge by comparing participants’ responses to the ungrammatical items in the two tasks.

Moreover, participants’ performance on the two tasks may also highlight performance differences in relation to their age of exposure. As it has been suggested that adult learners lack access to the innate mechanism needed for learning a language and that explicit language learning is not constrained by age, it is predicted that late learners will perform better on the U-GJT and ET than on the G-GJT. In summary, a comparison between participants’ performance on the ET and the two halves of GJT will help to illustrate possible age effects for different types of L2 knowledge (i.e., implicit and explicit). More specifically, the
current study explores the following research question:

To what extent does early and late learners’ performance on the grammatical and the ungrammatical halves of the GJT vary from their performance on the editing task?

Method

Participants

Three hundred and thirty-five foreign language learners from two universities in Pakistan participated in this research. Early starters received English medium instruction (EMI) from the first grade, while the late learners had EMI (i.e., all the course materials, instruction, and assessment in English) from grade eleven. The mean age at testing (MAaT) was (22.70, SD = 2.61) for early learners, and for the late learners, the MAaT was (22.42, SD = 2.90).

Materials

The materials used for data collection included a background questionnaire (BQ), a grammaticality judgment task, and an editing task. The BQ contained 23 questions about participants’ educational history, number of languages known, age of exposure to the L2, age at testing, and their self-rated proficiency in English.

The GJT covered twelve grammatical rules considered problematic for L2 learners. It was adapted from DeKeyser (2000), and was shortened to minimize fatigue effects. The current GJT contained 114 sentences, and the participants had to select each sentence as either grammatical or ungrammatical. The lexile measure for the GJT was 10L to 270L. An example of the items is provided below. The task was loosely timed as the participants were given five minutes more than the time needed to complete the task in the pilot study.

Example:

| Practice Items | Correct | Incorrect |
|----------------|---------|-----------|
| 1. A snake bit she on the leg. | ✔️ | |
| 2. Susan is making some cookies for us. | ✔️ | |
| 3. The baby bird has fall from the oak tree. | ✔️ | |
| 4. The little boy was counting all his pennies last night. | ✔️ | |

In addition to a GJT, an editing task was developed. Previous research has used editing tasks (ET) for examining L2 learners’ ability to identify and correct errors (e.g., Francis, 2012; Ferris & Roberts, 2001). To compare L2 learners’ grammatical proficiency on the GJT with another task, it was decided to create an ET. The ET contained 229 words and was piloted with L1 English and L2 speakers before using it in the actual study. The task contained violations of the same twelve grammatical rules as listed in the GJT. Two instances of error for each morphosyntactic feature contained in the GJT were inserted in the ET. Overall, the ET was comprised of 24 morphosyntactic violations. The participants were supposed to read the passage, find the errors and provide a correction. For example, in the following sentence, “To see the real defects, we have to climb the bridge over”, the participants were supposed to cross the word “over” from its current position and place it immediately after the verb climb.

Data Collection Procedures

Data were collected from two universities in Pakistan. All directions were provided in English because in
Pakistan the medium of instruction at higher education level is English. Procedures were explained to the participants on a projector. First, participants were provided the informed consent form. After the informed consent process was completed, the participants were shown four sample items for each, the GJT and the ET, on a projector. Finally, participants were asked to fill in the GJT first, then, the editing task, and lastly, the background questionnaire. During data collection, three other teachers assisted in distribution and collection of instruments and in monitoring participants.

**Data Coding**

On the GJT, all the items correctly identified as grammatical or ungrammatical received 1 point; all the sentences incorrectly identified received a 0. The obtained reliability measure for the GJT was .91. In contrast to the GJT, participants provided multiple correct responses for the same error/s on the ET. Hence, any response that rectified the erroneous target structure was scored 1. To ensure reliability, one-third (approximately 100/409) of the ET was coded by two coders. All the correct responses were assigned 1, and all the incorrect responses given a 0. The ET obtained an inter-rater Kappa reliability of .98 and internal consistency of KR-20 = .83.

**Results**

To answer the question, between-group and within-group analysis were computed. First early and late learners’ performance on task stimuli (i.e., grammatical and ungrammatical items) in the GJT was examined. No significant difference between the early and late learners on either one of the two stimuli was observed. Results for the *t*-test analysis are provided in Table 1.

**TABLE 1**

|                         | n     | Mean | SD  | t      | df  | p       | d    |
|-------------------------|-------|------|-----|--------|-----|---------|------|
| **Grammatical**         |       |      |     |        |     |         |      |
| Early L2 Learners       | 225   | 39.32| 6.46| -1.87  | 333 | .06     | -.02 |
| Late L2 Learners        | 110   | 40.75| 6.86|        |     |         |      |
| **Ungrammatical**       |       |      |     |        |     |         |      |
| Early L2 Learners       | 225   | 39.50| 6.71| -1.24  | 333 | .21     | -.14 |
| Late L2 Learners        | 110   | 40.49| 7.00|        |     |         |      |

*Note.* Total possible maximum points for each stimulus = 57

Table 1 shows no significant difference between the two groups on the G-GJT (*t* = -1.83, *p* = .06, *d* = -.02) and the U-GJT (*t* = -1.24, *p* = .21, *d* = -.14); however, there is a very small and negative effect size (*d* = -.14) for ungrammatical items, which means late learners obtained slightly higher scores than the early learners on the ungrammatical stimulus.

To explore the within-group performance by early and late learners on the task stimuli in the GJT as compared to their performance on the editing task, a paired sample *t* test was computed. Table 2 provides summary results for the paired sample *t* test.
TABLE 2
Paired Sample t-test for G-GJT, U-GJT, and the Editing Task (Based on Percentage Correct)

| Morphosyntactic features | n   | M     | SD   | t    | df  | p   | d   |
|---------------------------|-----|-------|------|------|-----|-----|-----|
| **Early L2 Learners**     |     |       |      |      |     |     |     |
| Grammatical               | 225 | 68.93 | 11.30| -6.38| 224 | .52 | -.04|
| Ungrammatical             | 225 | 69.24 | 11.76|      |     |     |     |
| Grammatical               | 225 | 68.93 | 11.30| 21.93| 224 | .00 | 1.46|
| Editing task              | 225 | 36.08 | 19.21|      |     |     |     |
| Ungrammatical             | 225 | 69.24 | 11.76| 21.73| 224 | .00 | 1.44|
| Editing task              | 225 | 36.08 | 19.21|      |     |     |     |
| **Late L2 Learners**      |     |       |      |      |     |     |     |
| Grammatical               | 110 | 71.4  | 11.99| .60  | 109 | .54 | .05 |
| Ungrammatical             | 110 | 70.99 | 12.25|      |     |     |     |
| Grammatical               | 110 | 70.44 | 12.13| 10.14| 85  | .00 | 1.09|
| Editing task              | 86  | 45.38 | 20.94|      |     |     |     |
| Ungrammatical             | 110 | 70.31 | 12.01| 10.38| 85  | .00 | 1.11|
| Editing task              | 86  | 45.34 | 20.94|      |     |     |     |

For both early and late learners, a significant difference was found in their scores when grammatical and ungrammatical stimuli were compared with the ET. However, no significant difference was detected when the two stimuli were compared with each other. Mean scores in Table 2 show that both learner-groups scored higher on the grammatical and ungrammatical stimuli of the GJT than on the ET. To further understand the relationship between the two stimuli and the editing task, a bivariate Pearson correlation analysis was computed. Results of the correlations are provided in Table 3.

TABLE 3
Intercorrelations, Means, and Standard Deviations for G-GJT, U-GJT, and the ET for Early (n = 225) and Late Learners (n = 86) (Based on Percentage Correct)

| Tests         | GJT-U | EDIT  |
|---------------|-------|-------|
| **Early Learners** |       |       |
| GJT-G         | .80*  | -.01  |
| GJT-U         | --    | -.03  |
| EDIT          | --    | --    |
| **Late Learners** |       |       |
| GJT-G         | .75*  | .11   |
| GJT-U         | --    | .17   |
| EDIT          | --    | --    |

Results of the correlation analysis reveal a very strong and positive correlation between the grammatical and ungrammatical halves of the GJT for both early and late learner groups, $r(223) = .80, p < .01$, and $r(84) = .75, p < .01$, respectively. This means learners who had relatively high scores on the G-GJT had high scores on the U-GJT as well. However, a similar correlation between the two stimuli on the GJT and participants’ scores on the ET was not observed. Figure 1 provides a comparison of early and late learners’ performance across the G-GJT, U-GJT, and the ET.
In Figure 1, circles represent grammatical stimuli in the GJT, squares symbolize items on the editing task, and diamonds indicate ungrammatical items in the GJT. As shown in Figure 1, late learners achieved higher percentages on ungrammatical stimuli and the editing task, as compared to early learners. Early learners secured higher percentages on grammatical items than ungrammatical items in the GJT or the editing task within-group. Visually, a wider gap between the two stimuli and the ET can be observed for the early learners, while a smaller gap is noticeable for the late learners.

**Discussion and Conclusion**

The current study explored age effects for two different task stimuli (i.e., grammatical and ungrammatical) in the GJT and then compared the two stimuli with the ET. To control for time constraint and modality, the GJT was presented visually and was loosely timed. Findings of the study reveal no significant difference between the early and late learners based on either the G-GJT and U-GJT. This finding has implications for the arguments presented for the impact of stimulus types in the GJT and their hypothesized relation with AoE to L2. First, no significant difference was observed between the G-GJT and U-GJT. This finding indicates a lack of dissonance between the two stimuli for this particular group of learners; hence, it does not confirm previous results that report G-GJT and U-GJT measuring different types of L2 knowledge (i.e., IK & EK; Loewen, 2009). In contrast, the strong correlation found between the G-GJT and U-GJT may imply that both measured a similar type of L2 knowledge. This result supports recent findings that stimulus type does not render a GJT into an IK or EK assessing factors (e.g., Vafee, Suzuki, & Kachisnke, 2016). Moreover, the results also did not reveal any sensitivity to learners’ AoE; rather, both groups obtained very similar mean scores on the grammatical and ungrammatical stimuli on the GJT.

However, based on the GJT results only, it might be difficult to draw any conclusions as to which types of L2 knowledge (i.e., IK or EK) each stimulus measured. Some might argue that because the G-GJT and U-GJT were visual and loosely timed, they did not tap into IK, but rather explored the EK only. To confirm this, learners’ performance on the two stimuli in the GJT was compared with their scores on the ET. The ET
was also visual and loosely timed, but it required participants to critically read the text, identify errors, and then provide corrections. As participants were requested to find errors in the text and provide corrections, it can be claimed that they used their EK during the task completion. While performing the ET, it is very likely that participants applied semantic processing (i.e., understanding the passage), noticing (i.e., identifying errors), and reflecting (i.e., finding solutions), which are the processes recommended for explicit language use (i.e., Ellis, 2004). The significant differences observed between the two stimuli in the GJT in relation to the ET indicate a possible difference between the types of L2 knowledge they measure. This is further heightened by a strong correlation between the two stimuli in the GJT in contrast to their lack of association with the ET. As the ET required error identification and correction, it might be argued that it demanded more explicit evidence of grammatical proficiency as compared to the two stimuli in the GJT. Hence, the G-GJT and U-GJT both measured IK.

The Author

Muhammad Asif Qureshi is an assistant professor at Zayed University, Abu Dhabi. He currently teaches composition to undergraduate students. His research interests include examining Second/Foreign language acquisition with a particular focus on age effects and types of L2 knowledge. In addition, he investigates vocabulary in L2 writing.

English and Writing Studies Program
Zayed University, P.O. Box 144534
Abu Dhabi, United Arab
E-mail: Asif.qureshi@zu.ac.ae

References

Bialystok, E. (1979). Explicit and implicit judgments of L2 grammaticality. Language Learning, 29, 81-103.
Bley-Vroman, R. (1988). The fundamental character of foreign language learning. In W. Rutherford & M. Sharwood-Smith (Eds.), Grammar and second language teaching: A book of readings (pp. 19-30). New York: Newbury House.
DeKeyser, R. (2000). The robustness of critical period effects in second language acquisition. Studies in Second Language Acquisition, 22, 499-533.
Douglas, D. (2001) Performance consistency in second language acquisition and language testing: A conceptual gap. Second Language Research, 17, 442-456.
Ellis, R. (1991). Grammaticality judgments and learner variability. In H. Burmeister & P. Rounds (Eds.), Variability in second language acquisition: Proceedings of the 10th meeting of the Second Language Acquisition Forum (pp. 25–60). Eugene, OR: Department of Linguistics, University of Oregon.
Ellis, R. (2004). The definition and measurement of explicit knowledge. Language Learning, 54, 227–275
Ellis, R. (2005). Measuring implicit and explicit knowledge of a second language. A psychometric study. Studies in Second Language Acquisition, 27, 141 -172.
Ellis, R. (2009). Implicit and explicit knowledge in second language learning, testing, and teaching (Vol. 42). Bristol, UK: 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, 119 - 126.
Ferris, D., & Roberts, B. (2001). Error feedback in L2 writing classes. How explicit does it need to be? Journal of Second Language Writing, 10, 161–184.
Francis, N. (2012). Bilingual competence and bilingual proficiency in child development. Cambridge, MA: MIT Press.
Granena, G. (2012). Age differences and cognitive aptitudes for implicit and explicit learning in ultimate second language attainment (Doctoral dissertation). University of Maryland, College Park. Retrieved from http://libproxy.nau.edu/docview

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

Han, Y. (2000) Grammaticality judgment tests: How reliable and valid are they? *Applied Language Learning, 11*, 177-204.

Han, Y., & Ellis, R. (1998) Implicit knowledge, explicit knowledge and general language proficiency. *Language Teaching Research, 2*, 1-23.

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). Bristol, UK: Multilingual Matters.

Purpura, J. (2004). *Assessing grammar*. Cambridge: Cambridge University Press.

Schachter, J., & Yip, V. (1990). Grammaticality judgments: why does anyone object to subject extraction? *Studies in Second Language Acquisition 12*, 379-392.

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.