System 72 (2018) 215–225

Contents lists available at ScienceDirect

System

journal homepage: www.elsevier.com/locate/system

Exploring second language learners' grammaticality judgment performance in relation to task design features

Li-Ju Shiu [a, *], Şebnem Yalçın [c], Nina Spada [b]

[a] National Chi Nan University, 1 University Rd. Puli, 545 Nantou County, Taiwan
[b] University of Toronto, 252 Bloor St W, Toronto, ON M5S 1V6, Canada
[c] Boğaziçi University Faculty of Education, Department of Foreign Language Education, Bebek 34342 İstanbul, Turkey

ARTICLE INFO

Article history:
Received 12 January 2017
Received in revised form 14 December 2017
Accepted 14 December 2017

Keywords:
Grammaticality judgment tasks
Task modality
Target structure difficulty
Task design features

ABSTRACT

This paper reports on an investigation of how second language (L2) learners' grammaticality judgment task (GJT) performance varies according to time constraints, task modality, and task stimulus in relation to two target features. One hundred and twenty EFL students were asked to judge items as grammatical or ungrammatical on four computer-based GJTs – two differing along the timed/untimed dimension and two differing along the aural/written dimension. Each GJT consists of 60 items (30 grammatical and 30 ungrammatical) focusing on two grammatical features in English, the passive voice and the past progressive, which were hypothesized to differ in terms of their learning difficulty. The results indicated that time constraints, task modality and task stimulus played a significant role in affecting L2 learners' GJT performance. Furthermore, although the learners performed better on the past progressive items, their GJT performance indicated similar patterns in relation to task design features across both target structures.

1. Introduction

Grammaticality judgment tests (GJTs)[1] have long been used to provide information about second language (L2) learning, including, for example, the investigation of adult L2 learners' access to Universal Grammar (e.g., Bley-Vroman, Felix, & Ioup, 1988), the critical period hypothesis in second language acquisition (SLA) (e.g., Johnson & Newport, 1989; Johnson, 1992), and L2 learners' use of different types of L2 knowledge (Bialystok, 1979; Bowles, 2011; Ellis & Loewen, 2007; Ellis, 2005; Godfroid, Loewen, Jung, Park, & Gass, 2015; Han & Ellis, 1998; Kim & Nam, 2016; Vafaee, Suzuki, & Kachisnke, 2017; Zhang, 2015). The extensive use of GJTs in SLA research derives from the hypothesis that they serve as promising measures of learners underlying linguistic competence. As a result, learners' GJT performance has been used to argue for and against different theoretical positions and empirical findings (e.g., Birdsong, 1989; Ellis, 1991; Han & Ellis, 1998; Hedgcock, 1993). The popularity of using GJTs in SLA research is also in part due to the fact that they are comparatively easy to administer to a large number of participants and they can assess knowledge of target features that are difficult to elicit in learners' production (Loewen, 2009).

* Corresponding author. 353 Wunsin First Rd. Nantun District, Taichung City 408, Taiwan.
E-mail address: ljshiu@ncnu.edu.tw (L.-J. Shiu).

[1] GJTs have been referred as "acceptability judgment tasks" in some studies because "acceptability judgment" was considered to be a more appropriate term to refer to learners' task performance (Ionin & Zyzik, 2014).

https://doi.org/10.1016/j.system.2017.12.004
0346-251X/

GJTs have different forms, including asking test takers to make grammaticality judgments, and/or to identify, correct and/or explain erroneous forms (Chaudron, 1983; Loewen, 2009). Among these task requirements, those asking L2 learners to judge the overall grammaticality of sentences have received more attention, particularly in recent studies exploring learners' use of different types of L2 knowledge (e.g., Ellis, 2009; Loewen, 2009; Godfroid et al., 2015; Gutierrez, 2013; Kim & Nam, 2016; Vafaee et al., 2017). It has been generally observed that L2 learners' performance on GJTs varies with learner-related factors (e.g., L2 proficiency level), linguistic features (e.g., target structures), and GJT task design variables (e.g., time constraints, task stimulus, and modality) (Bialystok, 1979; Ellis, 1991; Hedgcock, 1993; Godfroid et al., 2015; Gutierrez, 2013; Loewen, 2009; Murphy, 1997). Among the task design features, modality (i.e. visual versus aural) and target feature (i.e. more/less difficult to learn) have received less attention.

Due to the popularity of GJTs as research instruments to measure learners' L2 knowledge in the field of SLA (Loewen, 2009), it is essential to continue to explore how design features contribute to leaners' GJT performance. The current study builds on previous GJT research, exploring whether, and if so, to what extent the four variables—time constraints, task stimulus, task modality, and target features—affect L2 learners' grammaticality judgments.

2. Literature review

2.1. GJT design features and L2 learners' GJT performance

Below we review some of the GJT research indicating that learners' performance may vary with time constraints, task stimulus, task modality, and target features.

2.1.1. Time constraints

GJT research has generally found that L2 learners perform better on untimed GJTs than timed ones (Bowles, 2011; Ellis, 2005; Godfroid et al., 2015; Gutierrez, 2013; Han & Ellis, 1998; Han, 2000; Loewen, 2009; Mandell, 1999; Zhang, 2015). One often-given explanation for this finding is that with unlimited time, L2 learners, especially those receiving extensive amounts of classroom instruction, can take advantage of their L2 explicit knowledge while making judgments (e.g., Ellis, 2009; Loewen, 2009).

Ellis (2004) has proposed that when making grammaticality judgments, leaners are likely to go through three steps. First, they have to process semantically to understand the sentence (semantic processing). Then they need to detect if there is anything ungrammatical (noticing). If there is no grammatical error, they can make their judgment at this point. However, if learners notice something ungrammatical, they may reflect upon what (or maybe why) is not correct to confirm their initial detection of the ungrammatical element (reflecting). If learners are given enough time, they may go through the three steps before making a judgment. Accordingly, their GJT performance is better given the greater time allowance.

2.1.2. Task stimulus

Research using GJTs has also reported that L2 learners differ when judging grammatical and ungrammatical sentences. The majority of the studies employing GJTs have found that L2 learners perform better on grammatical rather than ungrammatical items (e.g., Bialystok, 1979, 1986; Gutierrez, 2013; Murphy, 1997; Kim & Nam, 2016; Loewen, 2009; Vafaee et al., 2017). However, a small number of studies using GJTs (e.g., Bley-Vroman et al., 1988; Gass, 1983) have found the opposite.

Reviewing a number of GJT studies, Hedgcock (1993) noted several possible factors that might affect learners' performance in judging grammatical and ungrammatical sentences, including, for example, the syntactic or semantic complexity of the test sentences, and influence of learning experience. To illustrate, errors may be easier to identify in sentences with simple structures than in sentences with complex structures because in the former, the errors may be more salient. Errors that occur in sentences with more complex semantic meanings might be more likely to be unnoticed because learners' attention might be more focused on the *meaning* rather than the *form* of the sentences. Erroneous forms of the features that have been extensively practiced might also be easier to detect than those that have been less frequently practiced because the former might have been overtly corrected.

Other views as to what factors might contribute to GJT judgments include Birdsong (1989) who argues that learners might tend to reject a grammatical sentence when unsure about its grammaticality. Ellis (1991) proposed that learners might consider an ungrammatical sentence to be grammatical due to lack of sufficient L2 knowledge. Gutierrez (2013) argued that L2 learners might resort to different types of L2 knowledge to respond to grammatical versus ungrammatical items; they might use explicit knowledge to respond to ungrammatical items and implicit knowledge in response to grammatical items.

2.1.3. Task modality

Researchers (e.g., Johnson, 1992; Penney, 1989; Wong, 2001) have assumed that task modality plays a role in influencing learners' GJT performance. In a comprehensive review of psychological research on modality differences, Penney (1989) argued that aural and visual verbal materials are processed in different parts of the memory system and by different mechanisms. McDonald (2000) also argued that decoding phonological stimuli is more demanding than decoding written stimuli because the former imposes more of a processing load. Wong (2001) explored whether modality affected how learners processed linguistic input, finding that learners had difficulty simultaneously attending to both form and meaning when the input was presented in an aural mode but not in a written mode.