Assessing L2 task performance: Understanding effects of task design

Parvaneh Tavakoli *

Applied Linguistics, London Metropolitan University, North Campus Room TM1-78, Tower Building,
166-220 Holloway Road, London N7 8DB, UK

Abstract

The overarching aim of the research reported here was to investigate the effects of task structure and storyline complexity of oral narrative tasks on second language task performance. Participants were 60 Iranian language learners of English who performed six narrative tasks of varying degree of structure and storyline complexity in an assessment setting. A number of analytic detailed measures were employed to examine whether there were any differences in the participants' performances elicited by the different tasks in terms of their accuracy, fluency, syntactic complexity and lexical diversity. Results of the data analysis showed that performance in the more structured tasks was more accurate and to a great extent more fluent than that in the less structured tasks. The results further revealed that syntactic complexity of L2 performance was related to the storyline complexity, i.e. more syntactic complexity was associated with narratives that had both foreground and background storylines. These findings strongly suggest that there is some unsystematic variance in the participants' performance triggered by the different aspects of task design.
© 2009 Elsevier Ltd. All rights reserved.

Keywords: L2 performance; Storyline complexity; Task design; Task structure

1. Introduction

Over the past two decades, studying task design and performance conditions has become a burgeoning area of research within task-based language teaching and assessment. Research on Second Language Acquisition (SLA) has concentrated on tasks and investigated task difficulty, task design and performance conditions and the effect they have on language learning and language performance (Bygate, 2001; Crookes, 1989; Foster and Skehan, 1996; Norris et al., 2002; Samuda and Bygate, 2008). Although research on language teaching has focused on different characteristics and performance conditions of tasks that would influence task difficulty and language performance in instructional settings, researching task difficulty and task design in a language testing (LT) context is a more recent topic of investigation. The main reason is that, for a long time, tasks in

---

* Tel.: +44 (0) 20 7133 2468.
E-mail address: P.Tavakoli@londonmet.ac.uk

LT were taken to be of equal difficulty (Fulcher and Marquez Reiter, 2003). In an early attempt to explore task difficulty, Pollitt (1991) challenged the assumption in LT that all tasks are of equal difficulty, and suggested that performance tests need to be constructed of tasks with increasing difficulty. Iwashita, McNamara, and Elder (2001) maintain that defining task difficulty has been a great challenge for test developers and language educators because tasks do not lend themselves readily to categorization for test purposes. Skehan (2001) has pointed out that in the current models of language proficiency, the underlying and generalized competences do not easily predict performance across different tasks and performance conditions. Reconceptualizing task difficulty, Fulcher (2003) argues that the framework of test method facets as proposed by Bachman (1990)[1] cannot be used to investigate task difficulty because:

(a) it is difficult to get agreement on precisely what each characteristic means
(b) there is no information on how or when method effects might influence scores, and
(c) as an 'unordered checklist' the Bachman model would be difficult to use in research or task design" (p. 63).

LT research now acknowledges that tasks are of different difficulty levels and appreciates the fact that a hierarchy of task difficulty has yet to be established (Bachman, 2002; Iwashita, et al., 2001). Bachman (2002) contends that understanding the effects of tasks on language performance and how test-takers interact with tasks is "the most pressing issue facing language performance assessment" (p. 471). The above discussion strongly suggests that identifying different aspects of task design that determine task difficulty is now considered an important current challenge for LT researchers since an index of task difficulty will be essential in selecting appropriate tasks, in providing a more reliable assessment of oral ability and in improving the validity of the interpretations and uses that are made based on the test results.

This article attempts to address the question of whether different aspects of oral narrative task design affect task difficulty and L2 performance. It will briefly examine the findings of task-based research, the tenets of a cognitive approach to task difficulty and the contribution such research could have for LT. The current research study in which the effects of task structure and storyline complexity on language performance are investigated will then be reported and the implications the findings of the study have for LT will be discussed.

2. Task design and second language performance

Since the mid-1980s, design and performance conditions of L2 speaking tasks have been widely investigated by SLA researchers. In pedagogy, tasks have been mainly investigated from two different perspectives to language teaching: an interactional perspective and a cognitive information processing perspective. Within the interactional perspective to task-based research, some researchers have focused on the role of interaction in the development of L2 particularly with respect to the negotiation of meaning (Long, 1989; Pica and Doughty, 1985). The main focus of the research in this interactional perspective has been exploring the effect of pair and group-based interaction on the development of L2 learner interlanguage. A second group of researchers, also adopting the interactional perspective, have attempted to explore how learners co-construct meaning while they are engaged in interaction (Duff, 1993; Van Lier and Matsu, 2000). The main interest of this group has been to allow participants to shape the task to meet their own needs and to build meanings collaboratively. Although the interactional approach to tasks has shed light on current understandings of task design and task variability in SLA, LT research has not been able to employ the findings of this body of research mainly because this approach is more related to classroom and pedagogy settings than to LT contexts. Moreover, it is clear that some of the valuable evidence provided by the interactional perspective, e.g. variability in performance on the same task on different occasions (Duff, 1993), questions the feasibility of analyzing task design in any static way or predicting variability of language performance on tasks.

---

[1] Bachman's (1990) framework of test method facets includes five major categories: (1) the testing environment; (2) the test rubric; (3) the nature of the input the test taker receives; (4) the nature of the expected response to that input, and (5) the relationship between input and response.