Learning and Individual Differences 43 (2015) 100–105

Contents lists available at ScienceDirect

Learning and Individual Differences

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

A cognitive processing model of reading comprehension in English as a foreign language using the linear logistic test model

Purya Baghaei [a,*], Hamdollah Ravand [b,c]

[a] English Department, Islamic Azad University, Mashhad Branch, Mashhad, Iran
[b] Vali-e-Asr University of Rafsanjan, Iran
[c] University of Jiroft, Iran

ARTICLE INFO

Article history:
Received 29 January 2015
Received in revised form 2 September 2015
Accepted 5 September 2015

Keywords:
Reading comprehension
Cognitive processing
Linear logistic test model
Q-matrix

ABSTRACT

Reading comprehension in a foreign language (FL) is a complex process with many underlying cognitive components. Many second language researchers have tried to explain reading comprehension in terms of taxonomies of subskills and processes. However, the nature of these components is not yet known. Previous research using exploratory and confirmatory factor analysis has yielded contradictory results. The purpose of this study is to investigate the underlying cognitive components and processes of FL reading comprehension using linear logistic test model (LLTM; Fischer, 1973). For the purpose of the present study, the data of 400 applicants taking an advanced high-stakes reading comprehension test were used. The cognitive processes underlying the test were derived. The derived processes were reading for details, making inferences, reading for main idea, syntax, and vocabulary. Linear logistic test model showed that making inferences is the hardest process to employ and vocabulary the easiest. The implications for teaching and testing reading comprehension are discussed.

1. Introduction

Reading comprehension in a foreign language is a complex process with many underlying cognitive components. Many researchers have tried to explain reading comprehension in terms of taxonomies of subskills (Alderson & Lukmani, 1989; Hughes, 2003; Jang, 2009; Lumley, 1993; Munby, 1978). However, the nature of these subskills is not yet known. We still do not know 'whether separable comprehension subskills exist, and what such subskills might consist of and how they might be classified' (Alderson, 2000, p. 10). Evidence for empirical separability of subskills supports the commonplace practice of dividing language skills into smaller units as a basis for syllabus design and test development and validation.

Some researchers argue that reading is a single global construct (Carver, 1992; Rost, 1993). Alderson (2005), based on the results obtained from the piloting of different DIALANG components, concluded that subskills do not contribute to distinguishing between different levels of reading language proficiency hence they are inseparable. Schedl, Gordon, Carey, and Tang (1996) studied the dimensionality of the reading section of the TOEFL (Test of English as Foreign Language) and concluded that different items which were supposed to measure different reading subskills do not explain differences in item difficulty or test dimensionality.

Qualitative studies which seek to record content experts' agreement on the subskills that are needed to answer reading items have also failed to show consensus among experts (Alderson & Lukmani, 1989). Conversely, there are some other researchers who found evidence and argued for the divisibility of reading comprehension (e.g., Anderson, Bachman, Perkins, & Cohen, 1991, Bachman, Davidson, & Milanovic, 1996, Carr & Levy, 1990, Davis, 1968, Drum, Calfee, & Cook, 1981, Grabe, 1991, 2000, Koda, 2007, Lumley, 1993, Nevo, 1989, Weir, Huizhong, & Yan, 2000).

Lennon (1962) argues that it is only possible to reliably measure four subskills in reading ability: word knowledge, comprehension of explicitly stated meaning, comprehension of implicit/inferential meaning, and appreciation. However, Carroll's (1993) factor analytic studies of cognitive tests identified four factors in reading: (general) reading comprehension, special reading comprehension, reading decoding and reading speed. Hudson (1996) suggested that reading comprehension involves processing skills such as local textual comprehension, global textual comprehension, and inference-making. In much the same vein, Weir and Porter (1996) conceptualized reading comprehension as having four categories of processes and skills: (1) local careful reading, (2) global careful reading, (3) local expeditious reading, and (4) global expeditious reading.

Freedle and Kostin (1991, 1992, 1993), using multiple regression, identified a set of item and text characteristics that predicted the difficulty of reading comprehension items of Scholastic Aptitude Test (SAT), Graduate Record Examination (GRE), and the Test of English as a Foreign Language (TOEFL). Freedle and Kostin (1991) found that eight variables predicted 58% of item difficulty variance on the SAT. In their next study (1992) they found that seven variables explained 41% of the variation in item difficulties on GRE. Finally, they found that 11

* Corresponding author.
E-mail addresses: pbaghaei@mshdiau.ac.ir (P. Baghaei), Ravand@vru.ac.ir (H. Ravand).

http://dx.doi.org/10.1016/j.lindif.2015.09.001
1041-6080/

item characteristics accounted for 58% of the variance in item difficulties in the TOEFL (Freedle & Kostin, 1993).

The current general idea among applied linguists is that reading comprehension in a foreign language is composed of several distinct subskills but there is no consensus on the nature of these subskills as different researchers have identified different subskills for second language reading comprehension with or without empirical evidence (Alderson & Lukmani, 1989; Hughes, 2003; Lumley, 1993; Munby, 1978).

Previous studies on the cognitive processes underlying reading comprehension have mostly been either with qualitative or correlational methods (e.g., factor analysis or multiple regression). Correlational approaches such as multiple regression or factor analysis are, according to Gorin (2005) and Sonnleitner (2008), limited in the following ways: (1) high correlations between cognitive processes and item difficulty does not guarantee causality, and (2) correlational methods are strongly affected by the range of item difficulties. For example, with items of similar difficulties the cognitive processes in question would not account for the small differences in item difficulties whereas item sets of wide range of difficulty would lead to high correlations with the difficulty-causing cognitive processes. Multiple regression studies, as noted by Buck, Tatsuoka, and Kostin (1997), are further limited in that these studies use item scores to predict item difficulties according to the processes/subskills involved in answering the items. The problem is that the predictors are too many and the items being coded for each predictor (i.e., process) are too few. Cohen and Cohen (1983) suggest about 30 items should be coded for each predictor for stable results in multiple regression. In a study with 10 item predictors, for example, there should be at least 300 items while in the multiple regression studies, which often involve more than 10 attributes, rarely are there more than 100 items. On the contrary, IRT-based methods such as LLTM enable researchers to model cognitive operations and estimate item difficulties in a non-correlative manner.

A neglected methodological approach in foreign language reading research for understanding reading comprehension is cognitive processing models such as LLTM and cognitive diagnostic modeling (CDM). In these models the sources of item difficulty are identified and parameterized to test hypotheses about the cognitive components that underlie item solving. Determining sources of item difficulty not only explicates validity of uses and interpretation of scores at the item level but also is useful for rule-based item generation and prediction of item difficulty on the basis of its cognitive features.

Latent component cognitive processing models such as the LLTM or component latent trait model (CLTM, Embretson, 1984) have sporadically been used to empirically study sources of item difficulty in first and FL reading (Embretson & Wetzel, 1987; Gorin, 2005; Sonnleitner, 2008). The need to break down learning concepts into smaller manageable units or 'learning quanta' for optimal teaching and learning has long been recognized in education (Fischer, 1973; Taber, 2004). Parameterizing these smaller chunks helps in understanding areas of difficulty and in devising remedial programs to help struggling learners. One psychometric model to accomplish this goal is Fischer's (1973) LLTM. LLTM is an extension of the Rasch model (Rasch, 1960/1980) which imposes linear constraints on the difficulty parameter. The model assumes that the overall difficulty of an item, estimated with the standard Rasch model, is the sum of the difficulty of the cognitive processes that are needed to solve the item. The model estimates the difficulty of those processes and tells us whether we have been successful in reconstructing item difficulty parameters (estimated by the standard Rasch model) with the difficulty of the cognitive processes (Baghaei & Kubinger, 2015).

Identifying cognitive operations and processes which are sources of item difficulty through LLTM, allows researchers to: (1) understand the nature of the construct and its underlying components, (2) construct new items with predetermined item difficulty parameters without the need to pilot them to estimate their difficulties (Fischer & Pendl, 1980), (3) predict the difficulty of items which have unique combinations of the estimated cognitive components. This is particularly helpful in item banking and adaptive testing, and (4) develop appropriate teaching activities that deal with the specific cognitive operations and processes.

It must be noted that the concept of subskills or processes under LLTM is different from that under factor analysis. While in factor analysis the subskills are hypothesized to be separate dimensions of ability this is not the case in LLTM as a prerequisite for LLTM is that the unidimensional Rasch model should hold first. Therefore, cognitive operations in LLTM are not independent dimensions of ability, unlike those recognized in factor analysis; otherwise the unidimensional Rasch model would not have fitted the data in the first place. Bejar (1983) states that unidimensionality does not mean that a single operation is at work when test-takers answer the items. In fact, there might be several operations and processes involved but as long as they work in unison unidimensionality would hold.

The processes under LLTM, while being psychologically separate, work in unison and are heavily interconnected. Research in the area of foreign language abilities shows that correlational methods fail to distinguish such processes as the final conclusion of most such studies is that language subskills are not discriminable. For instance, in a recent study Goh and Aryadoust (2014) in the context of listening comprehension in English as a second language using CFA found that their postulated listening subskills were highly correlated "making the models inadmissible because discriminant validity has been violated" (p. 16). Only when the higher-order aggregate-level CFA was employed they managed to show the separability of the postulated subskills. They further continue that "….subskills … were empirically divisible, should the right modeling approach be used" (p. 19). It is, therefore, emphasized that the theoretical status of subskills under LLTM is that of simultaneous parallel processes which are highly interdependent possibly, as suggested by Goh and Aryadoust (2014), due to the existence of a higher-order general factor. We argue that the model is superior to correlational methods in the study of foreign language subskills since language processing entails parallel processing of interconnected or associative neural networks in the brain (Bechtel & Abrahamsen, 1991). We believe that LLTM is an answer to the long-running debate in the field of language testing on the dimensionality of language constructs (McNamara, 1996).

2. Previous applications of the LLTM

LLTM has been applied to educational and psychological contexts to identify two categories of processes which might contribute to item difficulty: construct-related processes and construct-irrelevant processes. Kubinger (2008, & 2009) pointed to some possible applications of LLTM to identify the effect of some construct-irrelevant processes such as item position effects, speeded presentation of items, content-specific learning, and item response format.

Embretson and Wetzel (1987) employed LLTM to test a model of multiple-choice (MC) paragraph comprehension items. They proposed a model with two processing stages for paragraph comprehension: (1) a text representation process which referred to textual characteristics such as the number of proposition and arguments in passages and (2) a decision process with three events: translating the visual stimuli of the alternatives into a meaningful representation, locating the relevant text for evaluating alternatives, and falsifying or confirming alternatives. The results of their study showed that decision processes contribute to item difficulty considerably more than text representation processes and, therefore, reading comprehension item difficulty depends more on response decision processes than the paragraph. They conclude that two uncorrelated factors are involved in MC reading items: verbal ability, i.e., understanding the text and reasoning ability, i.e., the ability to select the correct alternative. Gorin (2005) investigated variations in item difficulty due to experimentally manipulating four