

Understanding the Cognitive Processes Underlying Performance in the IELTS Listening Comprehension Test

*Mohammad Ghahramanlou, Zahra Zohoorian and Purya Baghaei **

English Department, Islamic Azad University, Mashhad Branch, Ostad Yusofi St., Mashhad 91871, Iran

Correspondence: puryabaghaei@gmail.com or pbaghaei@mshdiau.ac.ir; Tel.: +98-5136-6350-64/5

Abstract: The purpose of this study is to examine the cognitive processes underlying the listening comprehension section of IELTS and to investigate if they vary in terms of difficulty. For this purpose, a checklist of possible cognitive operations was prepared based on the literature and the candidates' feedback. The checklist consisted of six cognitive operations. A sample of IELTS listening test was given to 310 upper intermediate and advanced students of English. Linear logistic test model was employed to analyse the data. Findings showed that keeping up with the pace of the speaker and understanding reduced forms were the most difficult operations for the listeners. Altogether, the six operations explained 72% of the variance in item difficulty estimates. Implications of the study for the testing and teaching of listening comprehension are discussed.

Keywords: listening comprehension; subskills; validity

Introduction

There is a consensus among listening researchers that listening comprehension includes a variety of subskills; though there is not yet a general agreement on their precise nature or exact quantity [1]. As Valette [2] maintains, the development of cognitive skills taxonomy was intended to be the foundation for designing listening tests to measure levels of ability in listening. Though a controversial concept, subskills divisibility has long helped language test developers in distinguishing levels of understanding and determining item difficulty.

Moreover, the results of such analyses can be beneficial not only to syllabus designers for planning appropriate syllabi but also for the development of test items.

Thus, it should be possible to define listening as complex skill operations which involve a number of subskills or constituent elements [3]. While focusing on the product, a listening test score can disclose nothing about the processes based on which a listener has selected the right answer or the reason why comprehension might have broken down. As a result, lack of evidence in such cases has invoked researchers to conduct studies in order to investigate the operations needed for a listening test taker comprehension. Also, there needs to be clear explanations through empirical research on how the subskills for listening combine to generate comprehension [3].

According to Rayner and Clifton [4] listening is divided into four levels. These include the channel level which consists of attending to the vocalizations and gestures. The next level is the signal level which embraces the identification of the speaker's signals. The third level, known as the intention level, is comprised of understanding the meaning by those signals. The last and the fourth level is the project level whereby the consideration of the joint proposed projects is involved. Though many of the studies on the listening skill have been conducted at the signal and intention levels, several researchers have attempted to investigate those many processes through which listening occurs. In fact there have been several efforts to define listening comprehension in terms of its subskills relying on the fact that listening involves the application of different separate skills. Field [5] mentions two possibilities offered through breaking the listening skill into its subskills. Firstly, it can follow a diagnostic approach by providing a checklist to show understanding breakdowns. Secondly, it provides the possibility of material preparation for a set of short sessions to practice those subskills for the later integration of lengthier listening exercises.

Thus, psycholinguists have tried to perceive listening as a creative imposition of meaning for the language signals received [6]. Having focused on the extraction of listening subskills, several taxonomies and classifications are proposed based on theoretical speculation. One of the initial taxonomies based its essence on a two-stage process of extracting basic linguistic information and utilizing the information for a communicative purpose. This view recurred in literature under various titles such as “construction and utilization processes” or “recognition and selection levels” [7]. In Oakeshott-Tylors’ 1977 study a set of macro and micro comprehension features were identified. Rost [8] proposed the subskills related to perception, interpretation, and information transfer arguing that there are some overlapping or parallel orientations [9]. However, apart from this parallel model, other psycholinguists preferred to conceptualize the hierarchical model. In such a model, the underlying processes are conceived at various levels that are built on each other either from top to bottom or from bottom to top [10]. Yet, other researchers mention the interactive model in which various levels interact simultaneously where general knowledge, contextual expectations, and predictions play a focal role [11].

Richards [12] determined a detailed communicative taxonomy including 18 micro skills for listening comprehension focusing on short-term memory including sounds, stress and intonation patterns, reduced forms, word boundaries, speech rate process, performance variables of speech, rates of delivery, recognition of grammatical structures and systems, major and minor sentence constituents, recognition of meaning for various grammatical forms, recognition of cohesive devices, recognition of communicative functions, using real world knowledge to infer goals, situations and participants, implementing such processes like prediction, inferencing, deduction, or exemplification, distinguishing literal from implied meaning, using nonverbal clues, and using strategies for comprehension. Other

researchers such as Grognet and Va Duzer [13] focused on the active quality of listening and mentioned such constituents for listening as reason determination, using short-term memory, organizing information by events, information prediction, recalling background knowledge, assigning meaning, checking understanding, and transferring the message to long-term memory. Nonetheless, Aitken [14] proposed a four item category including the comprehension of syntax and vocabulary, identification of sound patterns, attitudes, intentions, and reference making. Anderson and Lynch [15], drawing their conclusions from discourse analysis studies, proposed the following three skills:

- Recognizing the conversation topic counting on a native speaker's statements
- Making predictions for the probable topics for which a response is needed
- Recognizing and signalling for the lack of input

Yet, Barta [7] used the verbal report methodology to find the subskills and strategies implemented by the test takers processing thoughts for listening. Relying on a qualitative method, Barta compiled a taxonomy of 27 items which was based on theoretical framework of the ability in listening. These items basically related to strategic listening competences of cognitive, metacognitive and compensatory. In the same vein, a number of other researchers [16];[17];[12];[18].have also proposed several constituent elements and subskills for the listening skill among which Munby's 1978 list is believed to be the lengthiest including 250 subskills. Having analyzed the commonalities among the proposed subskills lists, the researchers sought a common and comprehensive framework. For the purpose of the present study the following taxonomy is proposed:

1. Using syntactic knowledge
2. Using semantic knowledge
3. Understanding details and explicit information

4. Understanding reduced forms
5. Keeping up with the pace of the speaker
6. Making inferences

The significance of syntactic knowledge mostly depends on the fact that most of the misunderstandings occur at this level [19]. Moreover, many of the function words are known to be unintelligible out of context while the speakers normally speak quickly [10]. Similarly, the semantic knowledge is deemed substantial as it deals closely with content words. Studies such as those conducted by [20] also confirm the point as it was found out that as an important predictor, vocabulary knowledge explained 14% of the listening ability. Furthermore, it is undeniable that the ability to elicit details explicitly contributes to the listening ability. Being able to elicit small pieces of information such as numbers and names significantly impacts the ability in listening comprehension [21]. Reduced forms recognition as the fourth subskill refers to the ability of understanding the processes through which the explicit markedness or the perceptual saliency of the morphemes is reduced [22] which considerably affects comprehension. The fifth subskill mainly deals with the ability to grasp the constant flow of speech and words. Keeping up with the speaker's pace is known as an obstacle for many second and foreign language learners [23]. Finally, inference making ability entails the listeners' use of pragmatic knowledge to determine an implied meaning [24]. It involves the understanding of meaning beyond the literal context of the utterance through the application of the knowledge about the speakers' intentions [25].

Taking into consideration the various studies concerning the sub skills of listening, it is observed that there is a similar framework analysing the relevant subskills; that is, most of the studies have employed correlation and regression techniques for analysis. Since analysing the subskills requires the investigation of psychological operations as well, the researchers take a different approach for the analysis which is the Linear Logistic Testing

Model [26]. The model's strength is in testing the hypotheses which identify the cognitive operations considered for psychological units. The model enables the researchers to empirically test hypotheses related to the processes and to establish substantive psychological theories. By identifying the cognitive processes which are needed to solve the items, the construct validity of the items can be demonstrated and new items for testing specific cognitive processes be written. Despite being a very powerful model in understanding the cognitive processes underlying test performance and providing validity evidence, LLTM has not received enough attention in cognitive psychology and education.

Method

Participants

Three-hundred and ten upper intermediate and advanced students of English as a foreign language participated in this study. The participants were both male (37.1%) and female (62.9%) and aged from 18 to 55 ($M=25.32$, $SD=5.65$). Participants were told that they were subjects of a study and were assured that their personal information would be kept confidential. The purpose of the study and the procedure was also explained to them.

Instrument

A sample IELTS listening test was administered to elicit the data needed for the study. The test consisted of 40 items in four different sections. In the first section, the students were asked to listen to a crime committed recently, and answer 10 questions. The task required to read a map and locate places and choose the correct answer among three alternatives related to what happened to the eye witness and the victim of the crime, and write down some pieces of information about the physical appearance of the two thieves involved in the crime.

There were five fill-in-the-gap and five multiple choice questions. For one of the multiple choice questions the participants were supposed to choose two correct answers.

In the second section, the students listened to some information about a hotel, presented by a tour guide followed by an information line. The section comprised 10 questions, five multiple-choice and five fill-in-the-gap. The students were supposed to answer questions on the location of the hotel, the facilities provided, and the price of the hotel.

In the third section, students listened to three friends talking about an exam they had. There were nine fill-in-the-gap items and one multiple-choice question. They had to write down two pieces of information about one of the speakers and answer eight questions about the advice she gave her two friends. Finally, in the fourth section, the students listened to a talk given by an assistant lecturer on a type of migratory bird. They answered 10 fill-in-the-gap questions about the birds.

Analyses

Q-Matrix Specification

The test items were examined by four experts in the field to build the Q-matrix involving the cognitive operations at work for each item. Among them, was an educational supervisor with about 25 years of experience in teaching English in different language institutes in Iran, an English teacher with an M.A degree in TESOL with five years of experience in teaching English, and two IELTS teachers with 10 years of experience in teaching general English and three years of teaching IELTS. This process was done with each expert individually; that is, the researchers played the listening file for each expert, gave a list of cognitive operations prepared in advance, and asked them to tick the operation they think might be involved in answering each question. The list consisted of the following six hypothesized operations discussed earlier:

1. Using syntactic knowledge
2. Using semantic knowledge
3. Understanding details and explicit information
4. Understanding reduced forms
5. Keeping up with the pace of the speaker
6. Making inferences

Results

Since LLTM is an extension of the Rasch model, the standard Rasch model (Rasch, 1960/1980) should fit the data first [26]; [27];[28]. Andersen's 1973 likelihood ratio (LR) test with the median of raw scores as a partitioning criterion showed that the test with 40 items does not fit the Rasch model: $\chi^2=237.721$, $df=39$, $p<0.001$.

The graphical model check (Figure 1) revealed that 17 items lie far from the 45 degree line. If the data fit the Rasch model, it is expected that the items fall close to a 45 degree line. Items falling off the line are misfitting items [29].

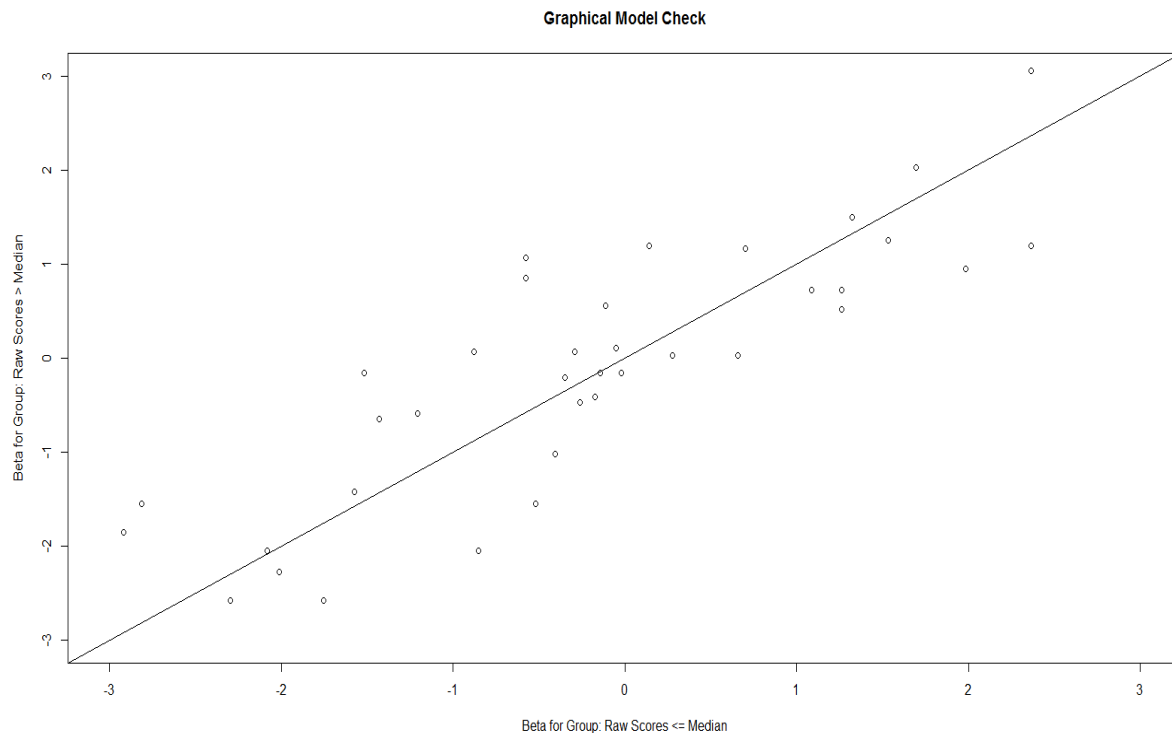


Figure 1: Graphical model check

After deleting 17 items which fall far from the 45 degree line, Rasch model was fitted again. Andersen's LR test with the median of scores as a partitioning criterion showed that the 23 remaining items fit the Rasch model: $\chi^2 = 36.755$, $df = 22$, $p = 0.025$.

4. 2. LLTM Analysis

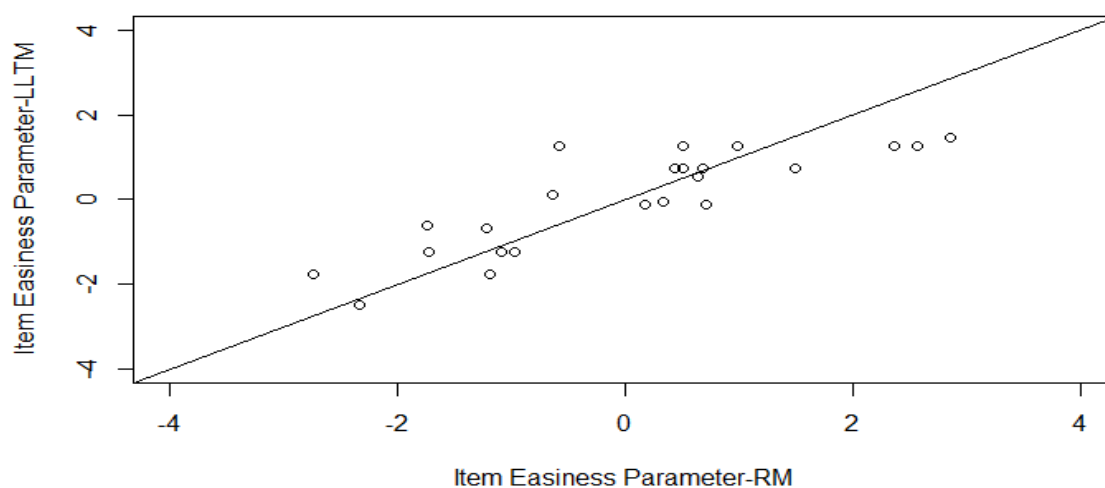
The 23 Rasch model fitting items and the Q-matrix were subjected to LLTM analysis using eRm [30] package in R version 3.11 [31]. Table 2 shows the easiness parameter of the six cognitive operations, their standard errors, and their 95% confidence intervals.

Table 2: Easiness of the basic parameters, standard errors and 95% confidence intervals

Operation	Estimate	Error	lower CI	upper CI
Syntax	-0.716	0.100	-0.912	-0.521
Semantics	-0.516	0.067	-0.648	-0.384

Details & Explicit	-0.727	0.103	-0.928	-0.525
Info.	-1.133	0.090	-1.310	-0.956
Reduced Forms	-1.361	0.083	-1.523	-1.199
Pace	-0.792	0.118	-1.024	-0.561
Inferences				

As explained before, LLTM imposes a linear constraint on the difficulty parameter. That is, we should be able to reconstruct Rasch model-based item parameters by adding the difficulty of the operations needed to solve each item. Figure 2 displays the plot of the relationship between the parameter estimates yielded by the Rasch model and those reconstructed by LLTM using the difficulty of the six cognitive operations. The correlation coefficient between the two sets of parameters was .85.



Comparing the fit of LLTM and the Rasch model with the likelihood ratio test showed that the Rasch model fits significantly better than LLTM, $\chi^2=582$, $df=16$, $p<0.001$. Nevertheless, the correlation between Rasch model-based items estimates and LLTM-reconstructed item

estimates was 0.85; that is, we managed to explain 72% of the variance in item difficulties with the six cognitive operations.

Discussion

The assessment and teaching of foreign language skills has been informed by the concept of subskills, i.e., a set of cognitive abilities or processes that learners need to master and activate to perform on language tasks and communicate successfully. Accordingly, the focus of teaching methods and curricula have largely been on the development of these subskills as applied linguists started to propose lists and taxonomies of subskills to aid language teachers and material developers [32]; [12].in fostering the subskills. However, the psychological reality of these subskills, their independent existence, and discriminability, which altogether boil down to their empirical verifiability has always been a question. Correlational research including regression analysis and exploratory and confirmatory factor analysis have yielded contradictory results.

In the present study, we aimed at finding the cognitive operations or subskills underlying listening comprehension in English as a foreign language. These operations were hypothesized to determine the item difficulty of the individual listening items. After reading the related literature and establishing a common framework consisting of six possible operations, an IELTS listening test was given to 310 upper intermediate and advanced students of English. The test consisted of 40 questions, 17 of which were discarded due to lack of fit to the Rasch model.

A Q-matrix consisting of six cognitive operations was taken as our theoretical listening model. The following operations were considered to be the most influential operations in determining the item difficulties: (a) using syntactic knowledge, (b) using semantic knowledge, (c) understanding details and explicit information, (d) understanding reduced forms, (e) keeping up with the pace of the speaker, and (f) making inferences.

'Keeping up with the pace of the speaker' was found to be the hardest operation, that is, students found it very difficult to process the constant flow of words produced by the speaker when they was delivered with a fast pace.

'Understanding reduced forms' closely followed the 'keeping up with the pace' operation. Due to the discrepancy between the spoken form of English and its written form, understanding what the complete form of the word is, seems to be a demanding task for foreign language learners of English. English people tend to use a lot of short forms in their informal speech. The listener might face contractions such as he's, he'll, etc. Or she might hear some abbreviations or acronyms such as Mr. UNESCO, etc. Another common reduced form that the students might face is clustering.

Clusters can be extremely difficult to tackle with. After talking to the candidates who had taken the test, we noticed a common complaint among a majority of the students about not distinguishing the boundaries of words; that is, where a word starts and where it finishes. Thus, there is difficulty in distinguishing words ending and starting with sounds having the same place of articulation. Words which end and start with sounds of the same place of articulation, after clustering, can hardly be distinguished.

The third most challenging operation was inference making. Students are required, in some stages of listening, to relate some pieces of information to either the previously stated information or their background knowledge or information outside the text. In some cases they make an inference by referring to a previously mentioned piece of information in the text or they might be expected to make a prediction based on the preceding elements of the text.

The fourth place belongs to 'understanding details and explicit information'. It is the ability to find small pieces of information directly stated in the text, including names,

numbers, etc. In this case, the way that students approach the distractors can make a big difference, owing to the fact that some other pieces of information, similar to the requested, might be given which function as distractors.

'Using syntactic knowledge' stood in the fifth place. Students are required to find which grammatical class the word belongs to although this operation does not seem to impose a heavy weight on the item difficulty.

The easiest operation is using semantic knowledge. It seems that upper intermediate and advanced students of English do not seem to have major problems in dealing with the meaning of the words; in other words, the lexis used in the test given did not impose a heavy semantic load on the students at these levels.

There has been a common trend among listening instructors and learners to elicit the words or phrases out of the text and analyze them in isolation. The findings of the current study show that the pace of the speaker and reduced forms play an important role in listening comprehension. Therefore, it is highly recommended to pay a careful attention to the natural pace of speaking and also using the reduced forms. For the given purpose, listening to authentic materials is strongly advised. The importance of 'keeping up with the pace of the speaker' shows that lack of exposure to the real and natural pace of English being spoken can be a major source of problem in listening, which the students might face. The listening materials developed for educational purposes vary significantly with authentic materials regarding the pace or the use of reduced forms. Educational textbooks and educational audio-visual materials aim at transferring the standard type of language to learners, while what really takes place in the real context of communication is completely different. Moreover, teaching the different short forms to the students is recommended as well.

Dearth of research regarding the underlying cognitive operations of listening was the driving force for conducting this study. The findings indicate that 72% of the variance in item difficulty can be explained by the six operations studied. Well-designed prospective studies may ultimately shed light on the other operations that explain the remaining 28% of variance in IELTS listening item difficulties. Apart from the operations, the test format is another influential factor. The next step would be to study the impact of test formats on the students' performance.

References

1. Buck, G., & Tatsuoka, K "Application of the rule-space procedure to language testing: Examining attributes of a free response listening test". *Language Testing*, 15(2), 119–157.
2. Valette, R.M. "Modern language testing (2nd ed.)". New York: Harcourt Brace Jovanovich. (1977).
3. Goh, C. C. M., & Aryadoust, V. "Examining the notion of listening subskill divisibility and its implications for second language listening." *International journal of listening*(2014): 28, 1-
4. Rayner, K., & Clifton, Jr. Charles. "In D. Medin (Ed), *Stevens' handbook of experimental psychology, memory and cognitive processes*. " New York: John Wiley & Sons.
5. Field, J. "Skills and strategies: towards a new methodology for listening. " *ELT Journal*(1998): 52(2), 110-118.
6. Rivers, Wilga M. "Speaking in Many Tongues (Expanded 2nd Edition) ". Rowley, Mass.: Newbury House(1976).
7. Barta, E. " Test takers' listening comprehension sub-skills and strategies" . *WoPaLP* (2010):4, 59-85.
8. Rost, M. " Listening in language learning" . Longman: London(1990).
9. Rost, M. "Teaching and researching listening " . London: Longman. (2002).

10. Clark, H. H. & Clark, E. V. " Psychology and Language " , Harcourt Brace Jovanovich, INC. New York. (1977): p.43.
11. Osgood, C. E " Mondatok megértése és létrehozása. In Cs. " Pléh (Ed.), Szöveggyűjtemény a pszicholingvisztika tanulmányozásához .Budapest: Tankönyvkiadó . (1986): (pp.136-168).
12. Richards, Jack. C. "Listening Comprehension: Approach, Design, Procedure," TESOL Quarterly, Vol. 17, No. 2, (1983):pp. 219-240 .
13. Grognet. A., Van Duzer. C. "Listening Skills in the Workplace. Spring Inst. for International Studies," Denver, CO. (2002).
14. Aitken, K. G. " Measuring Listening Comprehension. English as a Second language. "TEAL Occasional Papers, Vol. 2, Vancouver: British Columbia Association Of teachers of English as an Additional language. ERIC Document No. ED 155 945. (1978).
15. Anderson, A., & Lynch, T. "Listening (11th ed.)." Oxford: Oxford University Press. (2003).
16. Carroll, J. B. "Defining Language Comprehension: Some speculations. " In Language(1972).
17. Willis, J. "Teaching English through English," London: Longman. (1981).
18. Boyle, J. " Factors affecting listening comprehension" . ELT Journal(1984):38, pp. 34-38.
19. Field, J. " Notes on listening: Variability and assimilation ". Modern English Teacher. (1997).
20. McCarthy. J. " Learning to Listen: Teaching an Active Listening Strategy to Preservice Education Professionals" . Topics in Early Childhood Special Education (2000) : 27(4), 223-231.
21. Field, J. " Listening in the language classroom. Cambridge, England:" Cambridge University Press. (2008).
22. Henrichsen, L.E. " Sandhi-variation: A filter of input for learners of ESL" . Language Learning (1984) :34(3), 103-126.

23. Kim, J. "Affective Reactions to Foreign Language Listening Retrospective Interviews with Korean EFL Students" . *Language Research* (2002) :38, 117-151.
24. Vandergift, L. "Recent Development in Second Language Listening Comprehension Research". In Graeme Porte (Ed.), *Language Teaching: Surveys and Studies*. (2007): 210-291.
25. Rose K. R. & G. Kasper (eds.) "Pragmatics in language teaching" . Cambridge: Cambridge University Press. (2001).
26. Fischer, G. H. "The linear logistic test model as an instrument in educational research." *Acta Psychologica* (1973): 37, 359-374.
27. Baghaei, P. & Kubinger, K. D. "Linear logistic test modeling with R" . *Practical Assessment, Research & Evaluation* (2015):20, 1-11.
28. Baghaei, P., & Ravand, H. " A cognitive processing model of reading comprehension in English as a foreign language using the linear logistic test model" . *Learning and Individual Differences*(2015) :43, 100-105.
29. Kubinger, K. D. " Psychological test calibration using the Rasch model: some critical suggestions on traditional approaches." *International Journal of Testing* (2005) : 5(4), 377-394.
30. Mair, P., Hatzinger, R., & Mair, M. J. "*eRm: extended Rasch modeling* [Computer software]. R package version 0.15-4. " (2014). <http://CRAN.R-project.org/package=eRm>.
31. R Core Development Team. "R: A language and environment for statistical computing. "R Foundation for Statistical Computing, Vienna, Austria. (2015).
32. Munby, John. " Communicative Syllabus Design, " Cambridge: CUP. (1983).
33. Alderson, J. C., & Lukmani, Y. " Cognition and reading: Cognitive levels as embodied in test questions." *Reading in a Foreign Language*(1989) : 5, 253–270.
34. Benson C. P. & Hjelt. C. " Listening Competence: A Prerequisite to Communication." Southern Illinois University. (1978).

35. Bond, T. & Fox, C. " Applying the Rasch model: Fundamental measurement in the human sciences (2nd Ed.)" . Mahwah, NJ: LEA. (2007).
36. Morley, J. "Current Perspectives on Improving Aural Comprehension." (1999). Retrieved from <http://www.eslmag.com/MorleyAuralStory.htm>.
37. Pléh, Cs. " A mondatmeg ért és a magyar nyelvben." Budapest: Osiris Kiadó (1998).
38. Richards, J. C. & Schmidt, R. " Longman dictionary of language teaching and applied linguistics". New York, NY, USA. (1985).



©2016 by the authors; licensee *Preprints*, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons by Attribution (CC-BY) license (<http://creativecommons.org/licenses/by/4.0/>).