8th ICEEPSY 2017
The International Conference on Education and Educational Psychology

MEASURING ENGLISH GRAMMAR TEST: A RASCH ANALYSIS APPROACH

Fatimah Ali (a)*, Nor Suhardiliana Sahar (b), Noor Raha Mohd Radzuan (c), Zuraina Ali (d),
Afifuddin Husairi Hussain (e)
*Corresponding author

(a) Universiti Malaysia Pahang, Pekan, Pahang, Malaysia, tym@ump.edu.my
(b) Universiti Malaysia Pahang, Pekan, Pahang, Malaysia, liana@ump.edu.my
(c) Universiti Malaysia Pahang, Pekan, Pahang, Malaysia, nraha@ump.edu.my
(d) Universiti Malaysia Pahang, Pekan, Pahang, Malaysia, zuraina@ump.edu.my
(e) MyRASCH (Rasch Measurement Model Malaysia Chapter, Malaysia, afifuddinhusairi@gmail.com

Abstract

Students' mastery of English language has always been the concern of many stakeholders especially the employers. Various standard language tests are used to measure the level of spoken and written language competency in general. However, not many language tests assess grammatical competency among the ESL learners. The grammar test which was developed aims to identify the students' strengths and weaknesses in all 7 important grammatical elements such as Part of Speech, Tenses, Subject verb agreement, Relative Clauses, Conditionals, Passive and Verb Forms. The purpose of this study was to measure the reliability and validity of the instrument (UMP-EPT Grammar) using RASCH. The instrument which consists of 60 items was administered to 1694 first-year engineering students at Universiti Malaysia Pahang, Malaysia. Data collected were analyzed using WINSTEPS version 3.80.1. Findings revealed that item reliability and item separation was 1.0 and 20.53 respectively, while for person reliability and person separation, the results were .80 and 2.03 respectively. The study implies that the grammar test developed has great potential to be used for any high-stake test.

1. Introduction

Increasing proficiency in English language among students is one of the main priorities of Malaysia Educational Blueprint 2013-2025 (MEB). However, based on the MEB report, achievement in the English language subject is significantly low (MoE Malaysia, 2013). According to Souriyavongsa, Rany, Zainol Abidin and Leong (2013), one of the factors that contributes to students' low achievement in English language is fear of making mistakes because lack of basic grammatical knowledge. Yau (2014) reveals that in Malaysian context, the most common grammatical errors made by tertiary level students in written form are: singular/plural, articles, prepositions, adjective/noun, subject-verb agreement (SVA), and tenses.

To help realise MEB aim of increasing students' English language proficiency, Universiti Malaysia Pahang (UMP) Strategic Plan 2011-2015 in KRA 1A requires undergraduate students to achieve Malaysian University English Test (MUET) Band 3/IELTS 5.5/TOEFL 550, and postgraduate students at IELTS 6.0/ TOEFL 570. However, only 52% of UMP students registered for undergraduate programmes in 2013 achieved Band 3 or above for their MUET. This indicates a huge gap between students' English language proficiency level and the target of KRA 1A. Hence, to achieve the target of KRA 1A, Universiti Malaysia Pahang-English Proficiency Test (UMP-EPT) has been introduced as the benchmark to identify students' level of English proficiency and to stream them into appropriate level on English language courses. UMP-EPT consists of three components in which one of them is a grammar test.

1.1. Historical background of English language in Malaysia

The English language use in Malaysia is very much tied up with the historical and education background of the country. In particular, the influence of English can be traced as early as nineteen century when the British Empire widened its seach for gold and glory to the South East of Asia including Malaysia (or previously known as Malay Peninsular). They had first landed in Penang for trading. And in trading they communicated with the locals using sign language and English (Mohd Faisal Hanafiah, 2004).

With the introduction of the Resident System in the 1870's, the use of English language spread to the locals. The British recruited the local people who were able to understand some basic of conversational English. Since commerce and trading activities expanded especially in town areas, the mastery of the language became better. Besides trading and commerce that contributed to the use of English in the country, education also contributed to its usage among the locals. By 1950's several types of schools were introduced. This included high schools and convents that used English as the medium of instructions.

Subsequently when Malaysia was granted its independence, despite the use of the national language, English continued to be important due the legacy of the colonial era. The language was used as the only language in the post World War II that symbolised the modern nation at that particular time. However, after 1967, English was not given any official status due to the changes in the government policy in that it shifted the English-medium school to Malay-medium ones – a process which was completed nationwide by 1983 (Asmah Haji Omar & Noor Ein Mohd Noor, 1981). In schools, the

language is taught as a compulsory subject in primary and secondary schools throughout the country. For students in primary schools, they learn the language for six years that is from the age of seven (7) to twelve (12). The learning of the language continues in their secondary schools in which they are exposed to another five years of learning English. English is accepted as the second language, that is second importance in the ranking of languages in Malaysia (Asmah, 1976; Dumanig, David, & Symaco, 2017). In addition, other languages that are commonly use in the country are the Chinese dialects and Tamil.

With the role of English as a second language, its recognition in the education planning and policy is only secondary since the emphasise is on Bahasa Melayu (Asmah, 1981). English, however, is made a compulsory subject in the Malaysian schools. The students in Malaysia are required to sit for Pentaksiran Tingkatan Tiga (PT3 – Form Three Assessment) and Sijil Pelajaran Malaysia (SPM –Form Five Malaysian Certificate of Education). Meanwhile, at the tertiary level education, despite the evolution of Bahasa Melayu in the Malaysian education system, English is paramount until today in the country (Darmi & Albion, 2013).

2. Problem Statement

The literature for the past five (5) years was sought to gauge the area of studies pertaining to grammar in the Malaysian context. A study conducted by Fauziah Hassan and Nita Fauzee Selamat (2017) showed the fact that there was no 'synergy' between what was learnt in the class with the ones that were tested in the students' exams, made learners were not able to be proficient in the English language. In their study, teachers that were employed as respondents argued that the books used and the official syllabus focused on topic or themes while in the exams, students were assessed in relation to their four skills in language apart from grammar.

To date, most of the studies of grammatical errors in Malaysia focus on the writing or speaking skills and not on grammar test per se. Kaur's (2013) study investigated learners' capability in lexical knowledge, in particular grammar, found that students at one of the universities in Malaysia lacked in some areas of the grammar knowledge. Majority of them had limited knowledge on grammar rules for instance plural verbs, auxiliary verbs and infinitive "to". Also, they had difficulties in using tenses in their writing exercises. It was troublesome for them as they were not sure whether to use "ed", or "s" in the sentences they wrote. Similar findings were found by Singh, C. K. S., Singh, A. K. J., Razak, N. Q. A., & Ravinthar (2017) in their study of grammar errors made by ESL tertiary students in writing.

In regard to speaking proficiency, Mohammad Azannee Haji Saad & Murad Hassan Mohammed Sawalmeh (2014) analysed errors made by less proficient L2 Malaysian learners in role-play presentations. They found that the most frequent mistakes made by students were verb form and word form on top of SVA.

The present study employed Rasch measurement to measure the reliability and validity of the instrument used in this case the UMP-EPT Grammar test paper. Rasch measurement has been used in language learning to guage the effectiveness of a classroom assessment. A study conducted by Vogel and Engelhard (2011) revealed that Rasch model enables the researchers to obtain that there was a significance increase with regards to students' marks in their pre and post-tests learning French grammatical structures. Although only a small number of students i.e. 44 of them were involved in the

study, Rasch measurement showed that the use of guided inductive approach had resulted in students' learning of French grammatical structures to be more proficient.

Another use of Rasch measurement in language testing was employed by Howard (2012). His aim of using the measurement was to identify whether or not objective testing was suitable for assessing grammar skills, in particular; internal punctuation (commas, semicolons, and colons) and identification of syntactical structures (phrases and clauses). Although the results of the Rasch analysis in the researcher's study showed that the reliability estimates were lower for Punctuation Items, it was however, decided that revision was not necessary. Such was due to the fact that the values for test dimensionality and estimate reliability for global MNSQ infit was .98 (ZSTD = .1) and outfit was .91 (ZSTD = 0.0). These figures showed that the items for Punctuation felt within acceptable ranges to be reliable in assessing grammar skills among secondary school pupils.

In summary, it is imperative to conduct a study of the different type of grammatical errors made by Malaysian students in a grammar test particularly at the tertiary level to consolidate the findings of previous studies. The findings will also inform course developers to focus on the students' weak areas which in turn will facilitate language learning and the mastery of English language grammar.

3. Research Questions

This study aims to address two (2) following questions:

    i)   Is the instrument able to evaluate students' grammatical competency?
    ii)   Which grammatical items are found difficult by the students?

4. Purpose of the Study

The purpose of this study is to measure the reliability and validity of the instrument (UMP-EPT Grammar) using RASCH.

5. Research Methods

This study employed one-shot design and the analysis was conducted by using Rasch Measurement Model.

5.1. Rasch Measurement Analysis

The data were analysed by using Rasch Measurement Analysis. Nor Irvaoni and Mohd Saidfudin (2012) describe Rasch as 'an analysis of probabilistic and inferential and it focuses on the pattern of item responses that stipulates the interaction between a person and an item based on a mutual latent trait'. It predicts the likelihood of how a person of different ability level for a particular trait should respond to an item of a certain level of difficulty. The difference between the ability of the person and the difficulty of the item determines the probability of success (Bond & Fox, 2007). Rasch can transform ordinal data into

interval and it assumes that the item difficulty is the attribute that is influencing the person responses while the person ability is the attribute that is influencing the item difficulty estimates (Linacre, 2010).

There are two fundamental expectations in Rasch Model theorem assumption. The first expectation is that a person who is more competent has a greater chance to answer all items correctly; and secondly, an easy item is more likely to be answered correctly by all persons. Rasch can identify the relationship between item difficulty and person's ability, as such it is possible to analyse the item fit and the person fit (Nor Irvoni Mohd Ishar & Rosmimah Mohd Roslin, 2016). The item fit denotes an index which indicates the functionality of the item while the person fit refers to an index which describes the responses of an individual. For the data to fit the Rasch Model, the expected values of the standardised values are 0 and the mean square fit indices are 1. A misfit item, on the other hand, means the item fails to measure the intended trait because of its unsuitability, too easy or too difficult for a person to do. Therefore, the following goodness of fit criteria must be fulfilled in order to verify the fit and misfit for persons or items (Nor Irvoni & Mohd Saidfudin, 2012) :

i. Point Measure Correlation (PMC), $0.4 < x < 0.8$

ii. Mean Square (Infit & Outfit), $0.5 < y < 1.5$

iii. Z standard (Infit & Outfit), $-2.0 < Z < +2.0$

5.2. Participants

The participants in the present study were recruited from1694 first-year undergraduate students at Universiti Malaysia Pahang in Malaysia. The students had enrolled into nine different faculties. All participants had different first languages and English was learnt either as a first, second, third or foreign language. Their age range was between 20 and 22 years old.

5.3. Instrument

A set of UMP English Proficiency Test (UMP-EPT) Grammar question paper was used to investigate grammatical errors made by Malaysian language learners. The question paper consists of forty multiple-choice questions (Section A) and twenty open-ended questions (Sections B and C). The grammatical items tested were:

- part of speech
- subject-verb agreement
- tenses
- verb forms
- relative clauses
- conditionals
- passive structure

6. Findings

6.1. Reliability Indices

The Rasch item-person statistics in Table 1 are obtained from the data to examine the fit of the data to the Rasch model. 60 items were analyzed using WINSTEPS version 3.80.1 software. By using Rasch, we can measure how well the items fit within the underlying construct. The test raw score of Cronbach's Alpha marke a strong reliability of 0.82, which indicates that the items in the grammar test were able to measure the students' grammatical accuracy.

Table 1. Summary of Fit Statistics for Measuring English Grammar Test

(N = 1694, i = 60)	Item – Person Interaction			
	Item Location		Person Location	
Mean	0.00		-0.31	
SD	1.44		0.72	
MnSq (Infit & Outfit)	1.00	0.99	1.00	0.99
z-std (Infit & Outfit)	-0.2	-0.2	0.0	0.0
Maximum Measure	2.90		2.62	
Minimum Measure	-4.82		-2.26	
Reliability Indices				
Separation	20.53		2.03	
Reliability	1.00		0.80	
Model error	0.06		0.31	
Cronbach's Alpha		0.82		

In terms of item reliability, the instrument used had a very strong reliability at +1.0 logits on a 0 – 1 scale similar to interpreting Cronbach's Alpha, and the item separation index of 20.53, indicating a good item range. Therefore, it can be assumed that the findings from the instrument is replicable across comparable cohorts. The mean of –0.31 logits in Table 1 indicates that some of the items were comparatively difficult for the respondents to endorse. The instrument is capable of yielding a good person separation of 2.03, indicating there is an acceptable separation of measures along the scale, compared with the errors of measurement, which are comparatively smaller (.31). This implies that the power of the tests of fit of the model was very good. The Person Infit MnSq value is at the ideal 1, and z-std value that is 0, gives an indication of the goodness of fit of the instrument, and that it measured what it was supposed to measure.

6.2. Construct Unidimentionality and Item Dependency

Rasch uses Unidimensionality to measure construct validity by conducting the Principal Component Analysis (PCA) as shown in Table 2.

Table 2. Standardize Residual Variance (in eigenvalue units)

Description	Empirical (%)	Model (%)
Variance explained by measures	29.0%	29.1%
Unexplained variance	71.0%	70.9%
Unexplained variance in 1st contrast	3.0%	

To fulfill Unidimensionality, items in the instrument must measure the same composite of abilities – the students' grammatical English language proficiency. As indicated in Table 2, the Rasch PCA of residuals yields the raw variance explained by measure of 29.0% which is very close to the variance expected by the model (29.1%). However, it is far from the minimum Unidimensionality threshold of 40% is an indication of a strong measurement dimension (Conrad, Conrad, Dennis, Riley, & Funk, 2011).

6.3. Item Analysis

From Figure 1, the location of the mean item is above the mean person. It shows that majority of the students were unable to endorse most of the difficult items. The easiest item in the grammar test paper is question number 25 which is on passive structure while the most difficult item is question number 46 which is on verb form. This finding is comparable with the findings of a study conducted by Mohammad Azannee Haji Saad & Murad Hassan Mohammed Sawalmeh (2014). Relatively, items number 25, 29, 31, and 54 seemed to be too easy for the students to endorse. Eighteen (18) items are somewhat easy with seventeen (17) of them are below the mean person or below the student's ability. However, item 21 stars to challenge the students' ability.

Eight items belong to difficult items i.e. Levels 3 and 4. Seven (7) or 88% of the items are from dichotomous scale in Sections B and C. They are relatively difficult for the students to answer because that have to fill in the blanks using their own words and they also have to analyse five sentences to find a grammatical error in each sentence. The most difficult item (Item 46) is also from Section B. From the analysis, items on tenses, SVA, relative clauses and conditional are considered difficult for the students.
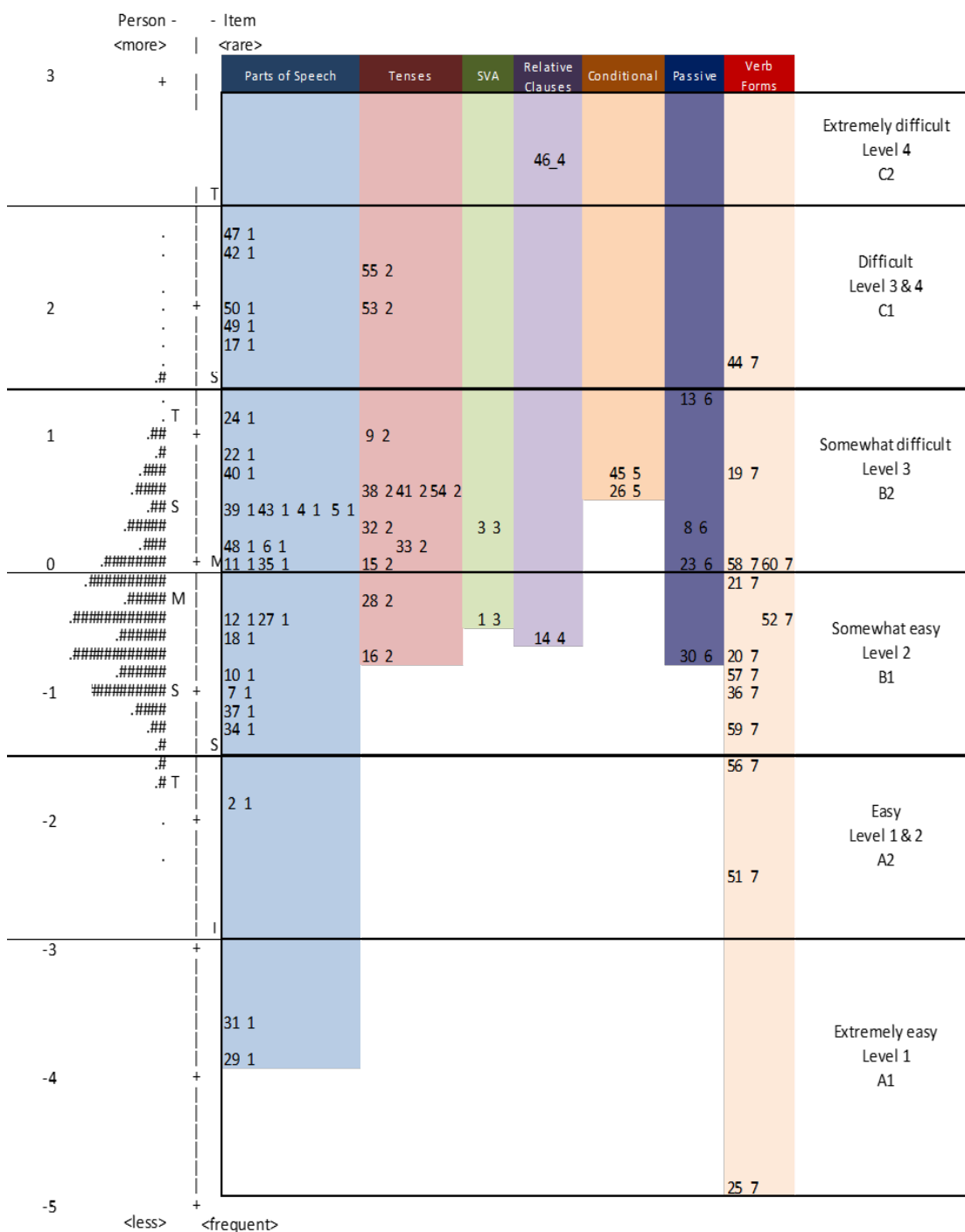
Figure 01. Person Item Distribution Map

7. Conclusion

This study is a preliminary attempt to measure the reliability and validity of the grammar test paper developed using Rasch. As a high-stake test which enables students to receive credit exemption on the basis of excellence, the test paper must be able to accurately measure students' level of grammatical proficiency. The findings from the present study revealed that although the instrument has a good reliability, further analysis is needed to ensure each item is working well to measure the construct. More work need to be done to develop new reliable items and rigorous analysis must be conducted to finally be successful in developing a good grammar test instrument.

Acknowledgments

References

Asmah Haji Omar (1976), The Teaching of Bahasa Malaysia in the Context of National Language Planning. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Asmah Haji Omar., & Noor Ein Mohd Noor. (1981). National language as medium of instruction. Kuala Lumpur: Dewan Bahasa dan Pustaka.

Bond, T. G., & Fox, C. M. (2007). Applying The Rasch Model: Fundamental Measurement in the Human Sciences. Lawrence Erlbaum Associates.

Conrad, K. J., Conrad, K. M., Dennis, M. L., Riley, B. B., & Funk, R. (2011). Validation of the Substance Problem Scale to the Rasch Measurement Model, GAIN Methods Report 1.2.

Darmi, R., & Albion, P. (2013, July). English language in the Malaysian education system: its existence and implications. In Third Malaysian Postgraduate Conference (MPC) 2013, p. 175.

Dumanig, F., David, M., & Symaco, L. (2017). Competing Roles of the National Language and English in Malaysia and the Philippines: Planning, Policy and Use. Journal Of International And Comparative Education (JICE), 104-115. doi:10.14425/00.45.77

Fauziah Hassan & Nita Fauzee Selamat (2017). Why Aren't Students Proficient in ESL: The Teachers' perspective. The English Teacher, 17.

Howard, P. S. V. H. (2012). Development of a measure assessing knowledge and use of internal punctuation to signal syntactic relationships (Doctoral dissertation, University of Denver).

Kaur, N. (2013). A case study of tertiary learners' capability in lexical knowledge. GEMA: Online Journal of Language Studies, 13(1), 113-126.

Linacre, J. M. (2010). Predicting responses from Rasch measures. Journal of Applied Measurement, 11(1), 1–10.

Ministry of Education Malaysia. (2013). Malaysia Education Blueprint 2013-2025: Preschool to Post-Secondary Education. Putrajaya, Malaysia: Ministry of Education Malaysia

Mohd Faisal Hanafiah (2004). English language and the language of development: A Malaysian Perspective, Jurnal Kemanusiaan, p.106-120

Nor Irvoni Mohd Ishar & Rosmimah Mohd Roslin. (2016). Customer Voice Retaliation (CVR) Construct Verification: A Rasch Analysis Approach. Procedia Economics and Finance 37 (2016), p 214 – 220

Nor Irvoni, M. I., & Mohd Saidfudin, M. (2012). Students' Perception towards Quality Library Service Using Rasch Measurement Model. In 2012 International Conference of Innovation, Management and Technology Research (ICIMTR2012) (pp. 668 – 672).

Singh, C. K. S., Singh, A. K. J., Razak, N. Q. A., & Ravinthar, T. (2017). Grammar Errors Made by ESL Tertiary Students in Writing. English Language Teaching, 10(5), 16.

Souriyavongsa, T., Rany, S., Abidin, MJZ. & Leong, LM. (2013). Factors causes students low English language learning: A case study in the National University of Laos. International Journal of English Language Education, 1(1), p 179-192.

UniKl, B. M. I., Mohammad Azannee Haji Saad & Murad Hassan Mohammed Sawalmeh. (2014). Error Analysis in Role-play Presentations among Less Proficient L2 Malaysian Learners. International Journal of English and Education, 3(3), p 346 – 355

Vogel, S. P., & Engelhard Jr, G. (2011). Using Rasch measurement theory to examine two instructional approaches for teaching and learning of French grammar. The Journal of Educational Research, 104(4), 267-282.

Yau, A. (2014). A Case Study of Grammatical Errors Made by Malaysian Students. International Journal of Science Commerce and Humanities, 2(5), p 154 – 160.