

The Correlation between Syntactic Knowledge and Reading Comprehension

Novita Pratiwi

English Language Teaching-Universitas Negeri Malang
Jl. Semarang No.5, Malang 65145, Indonesia. E-mail: novita@sahabatpulau.org

Abstract: This study aimed to find out the precise degree of relationship between syntactic knowledge (represented by grammar scores) and reading comprehension within EPT context. The data collected were in form of examinees' individual scores of EPT taken from two language institutes. The data representing 160 examinees were divided into two sets of scores: grammar and reading. Utilizing statistical analysis software, the researcher analyzed the data using Pearson's correlation coefficient. The result shows that syntactic knowledge and reading comprehension have strong positive correlation. It indicates that examinees' grammar scores might predict their reading scores within EPT.

Key Words: syntactic knowledge, reading comprehension

Abstrak: Penelitian ini bertujuan untuk mengetahui nilai pasti dari hubungan antara pengetahuan sintaktis (diwakili oleh nilai tata bahasa) dengan pemahaman membaca dalam konteks EPT. Data yang terkumpul berbentuk nilai individu peserta EPT yang diperoleh dari dua lembaga bahasa. Data yang mewakili 160 peserta dibagi menjadi dua: nilai tata bahasa membaca. Dengan memanfaatkan aplikasi analisis statistik, peneliti menganalisis data menggunakan koefisien korelasi Pearson. Hasilnya menunjukkan bahwa pengetahuan sintaksis dan pemahaman membaca memiliki korelasi positif yang kuat. Hal tersebut mengindikasikan bahwa nilai tata bahasa peserta mungkin dapat memprediksi nilai membaca mereka dalam EPT.

Kata kunci: pengetahuan sintaksis, pemahaman membaca

INTRODUCTION

In working on EPT, the test takers are expected to master two language skills which are reading and listening comprehension, as well as one language component which is structure or grammar. Numerous studies had been conducted to discover ways, techniques, methods or media which are qualified to help them in mastering all three of them. Nevertheless, most of those studies implied that the test-takers have problem in mastering language skills and components because they lack in vocabulary and have limited syntactic knowledge. Readers might have difficulties in developing comprehension if they do not understand the words meaning (Chen, 2014). Additionally, their syntactic knowledge may influence their comprehension as well and they need syntactic knowledge in integrating their background knowledge and word meaning (Koda, 2005).

A similar study to the current one was carried out by Morvay (2012) in which she conducted her research to 64 native Hungarian students who were learning English as a foreign language in their 12th grade in Slovakia. The aim of her study was to discover the relationship between their ability to process complex syntax and their foreign language reading comprehension. The analysis of her study showed that syntactic knowledge is a statistically significant estimator for foreign language reading comprehension. The study provides the evidence that processing ability of complex syntactic structures in a foreign language do contribute to learners' efficient reading comprehension in that language.

Syntactic knowledge (also syntactic competence or syntactic awareness), after all, refers to grammatical knowledge which, according to Kumaravadivelu (2006), is a syntactic system that deals with the rules of grammar. To boot, Randall (2007) mentioned that

grammatical knowledge is closely associated with our knowledge of word and the knowledge of how to connect the content words together to make meaningful sentences. He also mentioned that it is one of the four types of knowledge which are traditionally associated with knowing a word, beside phonological, morphological, and semantic knowledge. Additionally, Cain (2007) refers the grammatical or syntactic awareness as the ability to manipulate and to reflect on the language's grammatical structure. Thus, she further explained that one of the tasks that measure this ability is grammar correction task where a sentence having grammatical or morphological anomaly must be corrected. If it is so, then it proven that most questions in grammar section of EPT is associated with syntactic knowledge.

Regarding to the relation between syntactic knowledge and reading comprehension, plenty of studies have been conducted to see the effects of syntactic knowledge on language skills, especially reading. Not only do researchers study the relationship between them, but they also described how syntactic knowledge affects reading. Nation and Snowling in Cain (2007) mentioned that syntactic knowledge had a fundamental role in reading comprehension. On that account, Cain (2007) expressed that syntactic awareness has been hypothesized to specifically relate to reading comprehension.

Additionally, Linde-Lopez (2008) elaborated the difference of reading comprehension and grammar test by saying that the former one is receptive in nature as what the students need to do is only connecting the text and comprehending the meaning of what the information is about within the paragraphs; while the latter one is productive in nature as the learners have to decide whether the sentences are correct or not, correct them if they are not, or answer the questions. Jung (2009), however, stated that reading comprehension is still affected either directly or indirectly by grammar knowledge even though it is mostly conceptual. It is so because the role of grammar becomes more complex when the issue turns to second language reading since learners must learn how to construct phrases, even more to construct ones in a new language. In addition, a study by Dwaik in Chen (2014) showed that grammar structures are utilized more than vocabulary by second language readers when they try to comprehend a text. Therefore, mainly improving their syntactic knowledge can also improve their lexical knowledge (vocabulary) indirectly.

Gathering the information from the numerous studies mentioned above, the researcher tried to re-

late them with the reality. Based on the preliminary study that has been done by the researcher, she discovered that generally, people who have an experience in taking EPT find it difficult to work on either in one or two of its three sections. Anyhow, after digging the information deeper, she found out that they mostly got problems in grammar section for the non-English department students. Meanwhile, English department students generally have problems in reading and listening sections.

Related to the theories mentioned in the beginning, the problems which the test takers deal with are mostly caused by the lack of vocabulary and limited syntactic knowledge. Most of the test takers, however, utilize their syntactic knowledge more than vocabulary when they try to comprehend a text in reading comprehension section. This urges the researcher to find out the correlation between syntactic knowledge and reading comprehension. Therefore, the present study was intended to measure the degree of relationship between syntactic knowledge and reading comprehension.

Relating to the background of the study, the researcher formulated a problem of this research as the following: *Will the higher syntactic knowledge the examinees have be the higher scores they get on reading comprehension section?* And based on the research problem, the hypothesis of this study was as follows: *The higher syntactic knowledge the examinees have, the higher scores they will get on reading comprehension section EPT*

METHOD

As a quantitative research, correlational study offers insight into narrowly defined questions which help to identify the relationships among variables (Creswell, 2012). Since this study was aimed to investigate the precise degree of relationship between syntactic knowledge with reading and listening comprehension, the researcher will conduct a correlational study which is intended to measure the relationship between two or more continuous variables—the variables that show gradational differences (Creswell, 2012; Latief, 2017). Moreover, Creswell (2014) also states that correlational design is another non-experimental form of research in which the researchers make use of the correlational statistics to explain and calculate the degree of relationship (or association) between two or more variables or set of scores. In line with that, Lodico, Spaulding, and Voegtle (2010)

implied that the purpose of correlational research is to determine two or more variables and to examine whether there are any relationships between them. Therefore, this study was intended to obtain the information concerning the correlation between syntactic knowledge and reading comprehension. The correlation of both, however, is in the context of EPT.

In a correlational study, the researcher needs to use predictors to predict the outcome in one variable. Therefore, there are two kinds of variable in this study: predictor variable and criterion variable. First, the predictor variable means to give prediction of the result, while the criterion variable is the predicted outcome. As Lodico, Spaulding, and Voegtle (2010) explained that predictor variable is used to predict something, while the variable that is being predicted is called the criterion variable. Hence, the predictor variable in this study is syntactic knowledge, while the criterion variable is reading comprehension. The design can be illustrated can see in Figure 1.

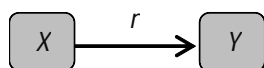


Figure 1. Correlational Research Design.

As can be seen from Figure 1, X represents the predictor variable which is syntactic knowledge. Meanwhile, Y represents the criterion variable which is reading comprehension. Lastly, r represents the correlation coefficient. This study was conducted by using quantitative approach since the researcher tried to describe the relationship between the examinees' syntactic knowledge to their reading comprehension. It deals more with number rather than the forms of words (Ary et. al, 1979). This involved data collection from the English Proficiency Test (EPT).

Population and Sample

Population is the group of individuals having the same characteristics that researchers want to generalize their findings to (Muijs, 2004; Creswell, 2012). In deciding the population, however, there are three steps in identifying population that will be taken into account by the researcher: general, target, and accessible populations can see in Figure 2. Banerjee & Chaudhury (2010) defined that general population is a whole group of participants which some characteristics are needed to be confirmed. According to Asiamah, Mensah, and Oteng-Abayie (2017), it is universally known by researchers. That is why in general

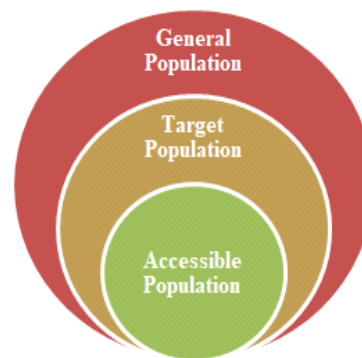


Figure 2. The Conceptualization of the Relationship between General, Target and Accesible Population

population, participants must partake at least one aspect of interest (Bartlett, Kotrlik, & Higgins, 2001; Creswell, 2003). Hence, it is this aspect that makes them eligible as the members of population (Asiamah et al., 2017).

However, not every population member can be involved in the study since the general population is typically rough in the sense that it often include the ones whose addition in the study would intervene goal, assumption, and/or context of the research (Asiamah et al., 2017). Therefore, the researcher needs to refine the general population by eliminating several member of the population. For a large population, selecting participants by using a set of criteria without defining both target and accessible population might result in neglecting the most qualified and appropriate group of participants and might unable the researcher to obtain 'the most suitable' samples (Asiamah et al., 2017).

Additionally, they also mentioned that the condition of target as well as accessible populations is essential for the large population. Hence, the part left of general population after its filtering process is called target population. It is described as the group consisting participants or individuals with certain characteristics (Bartlett et al., 2001; Creswell, 2003). Target population is more filtered than general population as it does not consist any characteristic that contradicts a research assumption, context, or goal (Asiamah et al., 2017).

Meanwhile, the accessible population is obtained after discarding all participants from target population which probably will or will not be involved or which is inaccessible at the period of the study (Bartlett et al., 2001). Besides, Sedibe (2006) agreed that accessible population is mostly affected by time and other extrinsic variables such as resources, distance, money,

and safety. For this reason, Asiamah et al. (2017) concluded that accessible population is the final group of the participants. The data will be obtained from this group, either from all of its participants or from a sample extracted from it.

All the above explanations lead to the statement that the general, target, and accessible populations of the present study can be seen on Figure 3. As the general population needs to cover the entire group that has at least one same attribute which is serving certified TOEFL ITP and providing similar EPT, therefore the general population of the current study was all certified institutions providing TOEFL ITP in Indonesia. The researcher took this consideration into account since the institutions serving TOEFL ITP develop and re-design the EPT by themselves in referring to the actual TOEFL ITP which makes the EPT reliable. Despite all the qualified institutions which share the same attribute, not all of them may participate in this study, requiring the researcher to eliminate some of them. Hence, the target population in this study was all certified institutions providing TOEFL ITP and similar EPT in student cities (the cities in which there are many students coming from different provinces in Indonesia) since there are a big number of students who need TOEFL ITP and similar EPT certificate for various purposes as applying for scholarships or for job vacancies. Hence, the data collected were expected to be representative enough to be generalized for there were a bunch of them.

Since it was not possible to reach all student cities in Indonesia, which was the target population, the researcher must define that part of the accessible population. Factors such as expense, time, and accessibility prevented the researcher from obtaining the data from all of the target population. Consequently, the accessible population in this research was all certified

institutions providing TOEFL ITP and similar EPT in Malang. By collecting data from certified institutes, the researcher expected them to be valid and reliable. Moreover, in order to have representative data to be generalized, she used total sampling technique to obtain as much data as possible for her study. Therefore, the researcher noted that the certified institutes in Malang are Balai Bahasa dan Budaya Universitas Negeri Malang, Pusat Bahasa Universitas Brawijaya, Azet Language Center, and English First Malang. However, there were only two institutions which granted the data access to the researcher: Balai Bahasa dan Budaya Universitas Negeri Malang and Pusat Bahasa Universitas Brawijaya.

Data Collection

Data collecting is the process of gathering and measuring information from variables related to the study which enables the researcher to answer stated research questions. The data, being quantitative in nature, will lead to statistical analysis and mathematical computation and be typically illustrated in charts or graphs. In the current study, the data gathered were in the form of EPT scores provided by certified institutes in Malang. Since most of the institutions administer EPT at least once a week, the researcher only needed to gather the score records for the last two or three months as it depends on the availability of the records accessed. By collecting the existing records of EPT scores provided by the mentioned certified institutions, the data collected were expected to be reliable and the results of analysis are expected to be valid. Ultimately, there were two sets of data in this study: scores of grammar and reading sections, and scores of grammar and listening sections.

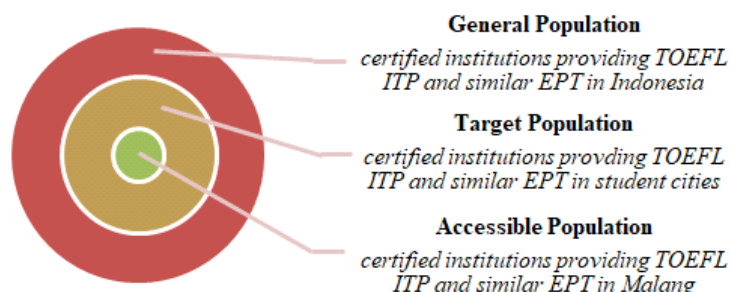


Figure 3. The Conceptualization of the Relationship between General, Target and Accessible Population of the Present Study (adoption from Asimah. et. all, 2017)

Data Analysis

In analyzing the data, the first thing that the researcher did was to begin the analysis process by organizing the data that were collected into three sets of data (grammar, reading, and listening scores) which made it easier to be analyzed. The second step was encoding the data gathered, including applying them into the scatter diagram which aims to show the illustration of the correlation among variables that are presented visually. The third stage was using the formula of Pearson's correlation coefficient which is given below or simply using statistical analysis software. It is used to analyze the relationship between two or more continuous variables (Muijs, 2004). Having finished all the data analysis, the researcher formulated the conclusion based on the research questions stated in Chapter I. Since the correlation coefficient proposed by Pearson as cited in Muijs (2004) is ranged between +1 to -1, the researcher used the categorization or guidelines of correlation coefficient (r) as can be seen in Table 1.

Table 1. Guidelines for Intrepeting Pearson Correlation Coefficient

| r Value | Meaning |
|---------|--|
| 0 | No correlation |
| <0./-1 | Weak positive or negative correlation |
| <0./-3 | Modest positive or negative correlation |
| <0./-5 | Moderate positive or negative correlation |
| <0./-8 | Strong positive or negative correlation |
| <0./-8 | Very strong positive or negative correlation |
| +1 | Perfect positive correlation |
| -1 | Perfect negative correlation |

Description of Statistical Data

For the data analysis purpose, the descriptive statistics regarding the basic source of the data collected was presented after using SPSS software as the overview of the examinees' overall achievement on grammar and reading within English Proficiency Test. The descriptions of each variable (one predictor and one criterion) are shows in Table 2.

Table 2. Descriptive Statistics for Grammar, Reading, and Listening Scores

| | Grammar (X) | Reading (Y ₁) |
|----------------|-------------|---------------------------|
| N | 160 | 160 |
| Mean | 42.35 | 44.49 |
| Std. Deviation | 6.234 | 6.149 |
| Standart error | .439 | .486 |
| Minimum | 24 | 31 |
| Maximum | 67 | 59 |

Table 2 shows, the achievement of 160 examinees in reading was slightly better than that of grammar by having mean value of 44.49. Meanwhile, the mean value of grammar is 42.35. However, the values of standard deviation show the wide variation that is higher for grammar score (6.234) and slightly lower for reading score (6.149). In other words, despite the variation of both scores, their means and standard deviation are pretty much similar with only little differences. As additional information, the table shows that from 160 examinees, the grammar scores range from 24 to 67, while the reading scores are at range from 31 to 59.

Based on the result of multiple comparisons in Table 2, it can be seen that there are differences of means and p-values for grammar to reading. The mean of grammar is compared to the mean of reading, resulting the mean difference of -2.137. It means that the mean score of grammar is two points lower than the mean score of reading. Furthermore, Table shows the p-value of .004. The p-value presented is less than .05 of the significance level. Therefore, it is likely that both of them differ from one another.

The Correlation between Syntactic Knowledge and Reading Comprehension

To analyze the extent of correlation between both grammar and reading sets of scores, it is essentially needed to confirm the availability of a linear correlation between the both variables. As can be seen from Figure 4, there seems to be a positive linear correlation between both variables which means that the higher the examinees' scores on grammar, the higher scores they will get on reading section in EPT.

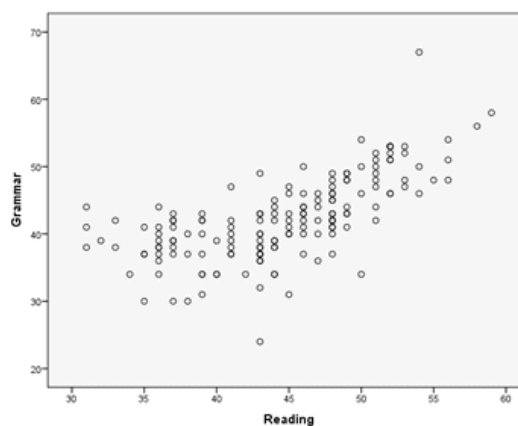


Figure 4. Scattered Diagram for the Correlation between Grammar and Reading Scores

Figure 4, the relationship between grammar and reading scores was analyzed using correlation coefficient of Pearson product-moment shown in Table 2. From the calculation of SPSS showed on Table 3, it provides three important information: (1) the Pearson correlation coefficient, (2) the significance level, and (3) the number of cases.

Table 3. The Correlations between Grammar and Reading Score

| | | Grammar | Reading |
|---------|---------------------|---------|---------|
| Grammar | Pearson Correlation | 1 | .673** |
| | Sig. (2-tailed) | | .000 |
| | N | 160 | 160 |
| Reading | Pearson Correlation | .673** | 1 |
| | Sig. (2-tailed) | .000 | |
| | N | 160 | 160 |

** . Correlation is significant at the 0.01 level (2-tailed)

The researcher found that the correlation coefficient between examinees' grammar and reading scores within English Proficiency Test (EPT) was .673. The coefficient correlation was greater than the critical value of significance level at .01 (.673 > .01). This means that there was an almost high, positive correlation between both variables, with the high grammar scores the examinees' have associated with their higher reading scores, which leads to the rejection of null hypothesis.

DISCUSSION

Within this study, the researcher hypothesized that the higher syntactic knowledge the examinees have, the higher scores they will get on reading comprehension section. It also means that syntactic knowledge and reading comprehension has positive correlation. As has been previously mentioned in the background, syntactic knowledge refers to grammar rules understanding and sentences construction. No wonder, some experts consider it the same as grammar knowledge or syntactic awareness.

In the current study, the first thing that the researcher tried to investigate was the correlation between syntactic knowledge and reading comprehension within English Proficiency Test (EPT). Regarding the relationship of them, numerous scholars had conducted similar research. In their study about the factors that influence skills of syntactic awareness in good readers and poor-comprehension ones, for instance, Nation and Snowling in Cain (2007) discover the fact that a strong correlation was found between

syntactic awareness skill and reading comprehension. This finding, however, was related to children as their population and samples, which is different from the present study that focused on EPT examinees who are adults.

Since the researcher gathered the data for this study from two language institutions in Malang which most of the participants or examinees are adults which range from university students to employees, she tried to relate her study to the previous one which was carried out by Guo, Roehrig, and Williams (2011) who studied adults in their research. They concluded that compared to beginner learners (children), adults' reading ability relies more on morphological awareness and syntactic awareness. They also elaborated that the relations between syntactic or morphological awareness and reading development are complicated. Moreover, they might be intervened by vocabulary skills. Particularly, syntactic awareness results in the increasing of breadth and depth of word knowledge, which also influence reading comprehension. This result of theirs shows the similarity to the present study in which the reading comprehension of examinees might be predicted by their grammar scores.

After analyzing the data for this study, the researcher found that within EPT context, there is a strong positive correlation between syntactic knowledge and reading comprehension. It means that the examinees' scores in grammar section could predict their scores in reading section. To continue the evidence in this research, the findings indicate that by having high syntactic awareness when working on the grammar section, the examinees are also assisted by this awareness when they working on the reading section. According to Herriman in Guo et al. (2011), there are three aspects included in syntactic awareness: the first is aware that sentence is a basic unit of written language, the second is aware about acceptability of grammar and how well it is formed as associated to the strings of word or sentences, and the third is judging the relations between syntactic structure and sentences semantic properties. These aspects are quite helpful when working on both grammar and reading section as there are numerous grammar-related questions in reading section as well. Moreover, syntactic knowledge shows students ability to internalize and use structures in English that are grammatical (Norris in Hashemin & Mahmoudi, 2016). Both Herriman and Norris show how far syntactic knowledge matter to reading comprehension.

Not only do grammar-related questions show up in reading section, but among varied types of pre-reading activities there is also the pre-teaching of grammar (Hashemin & Mahmoudi, 2016). This indicates that both have similar characteristics when it comes to written language. Furthermore, Jung (2009) implied that, reading comprehension is still influenced by the grammar knowledge either directly or indirectly even though is mostly conceptual. When the issue turns to send language reading, the role of grammar becomes more complex as learners must learn how phrases are constructed.

In their study, Fatemi and Shadman (2012) concluded that the higher the scores on structure test, the better the performance on the reading comprehension test. It can be interpreted that having knowledge of syntax might result in having better score in reading comprehension. Furthermore, Guo et al. (2011) stated one of the major findings in their study that syntactic awareness is positively related to reading comprehension. Particularly, learners acquire new vocabulary by utilizing syntactic knowledge through the use of context clues. In this case, they employ both syntactic and contextual information to guess new words meaning. Nevertheless, the idea of deciding the scope of grammar is a prickly issue mainly since the knowledge of grammar is overlapped with vocabulary. Hence, more empirical studies are crucial to generate the precise nature of the relationship between both grammar and second language reading.

Related to that, the researcher composed that syntactic knowledge which is covered in grammar section of EPT might contribute to reading comprehension in various ways. For instance, higher level of syntactic knowledge is associated to greater accuracy and fluency in solving the questions related to the incorrect words based on their role in sentence structures, which would affect reading comprehension. Specifically, Guo et al. (2011) claimed that because adult learners have mastered high-level syntactic awareness, they might also use this skill in the reading process. However, having the fact that this study was a correlational one, the researcher was unable to provide a clear description of developmental shift in syntactic knowledge contribution to reading comprehension. Thus, further experimental studies or other designs are necessary to examine the role of syntactic knowledge in predicting examinees' reading comprehension within EPT.

CONCLUSION

From the result of the study, the researcher drew some conclusions which can answer the three research questions proposed based on research problems. The research question was about the correlation between syntactic knowledge and reading comprehension within EPT. Based on the analysis of the data, syntactic knowledge—which is represented by grammar scores within EPT—has high positive correlation to reading comprehension. This means that H01 is rejected and signifies that the higher scores the examinees get on grammar section, the higher scores they get on reading section.

REFERENCES

- Asiamah, N., Mensah, H. K., & Oteng-Abayie, E. F. (2017). General, target, and accessible population: Demystifying the concepts for effective sampling. *The Qualitative Report*, 22(6), 1607.
- Banerjee, A., & Chaudhury, S. (2010). Statistics without tears: Populations and samples. *Industrial Psychiatry Journal*, 19(1), 60–65. doi: 10.4103/0972-6748.77642.
- Bartlett, J. E., Kotrlik, J. W., & Higgins, C. C. (2001). Organizational research: Determining appropriate sample size in survey research appropriate sample size in survey research. *Information Technology, Learning, and Performance Journal*, 19(1), 43–50. doi: 10.1109/LPT.2009.2020494.
- Cain, K. (2007). Syntactic awareness and reading ability: Is there any evidence for a special relationship? *Applied Psycholinguistics*, 28(4), 679–694. doi: 10.1017/S0142716407070361.
- Chen, K. Y. (2014). Vocabulary, syntactic knowledge and reading comprehension: the perspective of college EFL Students. *An Interdisciplinary Journal*, 39–52.
- Hosseini Fatemi, A. (2012). The predictive power of vocabulary breadth & depth, and syntactic knowledge in reading comprehension performance of Iranian advanced EFL students. *Studies in Applied Linguistics*, 2(2), 1–5.
- Guo, Y., Roehrig, A. D., & Williams, R. S. (2011). The relation of morphological awareness and syntactic awareness to adults' reading comprehension: Is vocabulary knowledge a mediating variable. *Journal of Literacy Research*, 43(2), 159–183. doi: 10.1177/1086296X11403086.

- Hashemin, K., & Mahmoudi, A. (2016). Effects of STM span, syntactic awareness, and vocabulary knowledge on Iranian EFL learners' reading comprehension. *Theory and Practice in Language Studies*, 6(12), 2297–2303. doi: 10.17507/tpls.0612.09.
- Jung, J. (2009). Second language reading and the role of grammar. *Studies in Applied Linguistics and TESOL*, 9(2).
- Lopez, A. L. (2008). To what extent does grammar knowledge account for competence in FL reading comprehension in university students?. *Revista española de lingüística aplicada*, 21, 181–200.
- Morvay, G. (2012). The relationship between syntactic knowledge and reading comprehension in EFL learners. *Studies in Second Language Learning and Teaching*, 2(3), 415–438.
- Sedibe, M. (2006). *A comparative study of the variables contributing towards the establishment of a learning culture in schools* (Unpublished doctoral dissertation). University of Pretoria, South Africa.