



INDIVIDUAL VS. PAIRED FORMATS OF ORAL PROFICIENCY EXAM: DOES IT MATTER IN TERMS OF SCORES AND PREFERENCE?

Zafer Susoy¹

Anadolu University,
Eskişehir, Turkey

Abstract

The purpose of this preliminary study is to examine whether there is a relationship between the achievement scores and two different test formats used in speaking assessment – paired or individual. The study also attempts to explore the test takers' preferences pertaining to the test formats employed and the reasons behind. The participants of this study are 25 freshman students studying in the English Language Teaching department of a well-known state university in Turkey. The data used in the quantitative analysis comes from the grades scored by the same cohort of students in two oral proficiency exams conducted within the scope of course assessment during the regular exam period of the program. Secondly, the qualitative data was obtained from the semi structured interviews carried out with 14 of the 25 students immediately after the paired speaking test. The interview questions specifically sought to reveal how the participants felt about the paired format of speaking test, how they would compare this format to the individual format, and if they were given the chance which format they would prefer and why. We have come to see that students got better scores in the paired speaking test format that we designed in the form of paired discussion in comparison to the individual format where the test takers' interlocutor is the course lecturer. The difference in the mean scores obtained in the two different formatted exams was found statistically significant. Our qualitative findings showed that the majority of the students liked and supported the practice of being paired.

Keywords: speaking assessment, oral proficiency, assessment formats

1. Introduction

Testing of oral performance and proficiency in a second language (L2) has attracted a widespread attention as a difficult area of empiric investigation. Oral proficiency interviews have been the traditional means to rate and assess oral performance in a L2.

¹ Correspondence: email zafersusoy@gmail.com

Silverman (1976, as cited in van Lier, 1989) described some noticeable characteristics of an interview conducted to rate one's oral proficiency and posed some validity problems of oral proficiency interviews (OPIs). According to Silverman (1976) the interviews has got a pre-planned structure, and an unbalanced distribution of talk, the interviewer being superior and making all the interactional decisions such as initiating, ending or bring up a novel topic etc.. As Grice's (1975) maxims suggested, however, a real life conversation is changeable during the course in terms of consequences and allocates more or less the same responsibilities for the parties to build the talk.

Test designers often strive to produce 'real-life' test conditions so that the test scores can extend and depict authentic and non-test situations (Bachman, 1990; Bachman & Palmer, 1996). An example to these allegedly 'real-life' speaking tests are those where more than one test-taker fulfill a task together through collaborative communication and where raters score their proficiency by examining the oral performance (Van Moere, 2013). Particularly, paired speaking format where the test takers interact in dyads has had a wide coverage in the field of oral proficiency testing. Since 1980's, the paired format has been adopted to test speaking abilities of test takers in four major tests of English famous worldwide; they, namely are, PET, KET, FCE and CAE (Birijandi and Bagherkazemi, 2011).

One of the benefits of paired speaking testing proposed in the literature is concerning its comparatively high capacity to reflect the common classroom practices of speaking as Brooks (2009) writes; "*paired testing more closely mirrors the type of oral interaction the students would likely encounter at university and it reflects the type of speaking tasks commonly used in the classroom*" (p. 324). Likewise, Van Moere (2013) mentions *the positive washback* of the paired format. Another benefit of the paired speaking test format cited in the literature is regarding the language characteristics produced by the test takers. The test takers' performances on a paired format were analyzed in a relatively recent study by Marochi (2008, as cited in Kanga, 2012) and the results revealed three scopes of language on which the test takers performed better than they did on an individual testing format where they were asked to communicate only with the instructor. These scopes were;

- the load of language generated by the test takers;
- a more complicated level of grammar achieved;
- a wider variety of vocabulary.

These better and high quality aspects of language that could be produced in the paired oral examination format can consequently lead to higher grades achieved by the test takers (Brooks, 2009; Davis, 2009). Brooks (2009) found that 12 out of his total 16 participants achieved higher scores on paired format in comparison with the individual format. Likewise, Davis (2009) also showed that learners on a paired speaking format indicated a wider range of reciprocal communicative action and thus got higher scores. Similarly, Paired format of speaking testing is claimed to offer oppurtunities for the test takers to generate more conversational patterns (Galaczi, 2008, 2014) .With regard to how the communication occurs in these types of oral testing situations, Galaczi (2008)

addressed to three types of interactional design; collaborative, parallel and asymmetric and the study showed that the test takers scored higher while mutually structuring the interaction in the collaborative mode. One reason cited in the literature explaining these asymmetrical communicative patterns is the asymmetrical power relation between the test taker and the rater in individual test designs (Van Lier, 1989).

Együd and Glover (2001) claimed that students largely prefer being paired in oral examination and that students could manage to generate better English when paired compared to the individual assessment format as mentioned earlier in this section. The preferences and performance across the two types of speaking test, however, also vary across studies. In Kanga's (2012) study, students justify an inclination towards the individual type of test although they were found less successful in terms of performance in individual test compared to the paired format. The study indicated that there lies several factors behind the students' inclination to individual test type despite low performance index. These reasons include students' feeling more relaxed and confident and the test's being administered more easily. Some researchers asserted that interacting in pairs is lowering the anxiety of partners (Ikeda, 1998; Norton, 2005)

In spite of the recently documented advantages, Foot (1999), more than a decade ago, posed some precautions against the paired speaking test format some of which are currently still valid. Among the Foot's (1999) warnings were;

- there is a little body of practical support based on observation and experimentation to prompt the paired format
- the attendance of a scorer in the testing environment also over shade the everyday speech format
- candidates might be at different levels of language proficiency and anxiety and they might possess different personality traits and accents, which, in the end, endangers the use of the paired testing format.
- the quantity of the allocated test time for the dyads is generally the same allocated for one individual in individual test format, which seems quite short for pairs to engage in a genuine interaction.

Foot's (1999) caveats has been under investigation since. There are some considerations regarding the effect of the partner's level of language proficiency over the other partner in a pair and the issue of proficiency is far away from being plain. Davis (2009), for example, conducted a study to investigate the partners' proficiency on one another's performance and found no significant influence. Davis's results (2009) also showed that lower-ability test takers performed better when paired with a higher-ability partner, however, the study overall concludes that the proficiency differences between partners make no obstacle for the administration of paired speaking tests. However, Iwashita (1998) found that the interlocutor proficiency mattered more when choosing the partners to be paired. The study showed that both low and high proficient test takers attained better grades generating more linguistic output with a high proficient partner.

Different personality traits of the partners have also been examined to find out whether it effects the communication and the overall oral results in group discussion performances (Berry, 2004; Bonk & Van Moere, 2004). Berry (2004) found out that extroversion as a personality trait is a positive contribution to the oral performance. Likewise, Bonk & Van Moere (2004) first controlled the proficiency levels and then compared the oral assessment scores of a huge number of shy and outgoing students. The results revealed that the outgoing students somewhat excelled the shyer ones in their oral performances. O' Sullivan (2002) stressed a slight effect of interlocutor familiarity in paired speaking performances. In the study, however, even this small scale effect is limited to the gender of the partner since the effect was only found significant when a female partner interacted with an unfamiliar partner. As can be seen from the modestly reviewed brief literature, testing the oral language proficiency is a highly complex issue with both alleged advantages (Davis, 2009; Van Moere, 2013), shortcomings (Foot, 1999) and preferences or disfavours of test takers (Együd and Glover, 2001; Kanga, 2012). Therefore, the purpose of this preliminary study is to examine whether there is a relationship between the achievement scores and two different test formats – paired or individual. The study also attempts to explore the test takers' preferences pertaining to the test formats employed and the reasons behind. To achieve the purpose, the current study seeks to answer two research questions;

1. Do test-takers' speaking examination scores differ depending on whether the interlocutor is an examiner or a peer?
2. Which of the speaking test format do test-takers prefer and why?

2. Methodology

2.1 Participants and the Research Context

The participants of this study are 25 freshman students studying in the English Language Teaching department of a well-known state university in Turkey. The participants' ages vary between 18 and 20. A class of freshman was chosen based on the convenience sampling method (Creswell, 2012). The participants were all assured that the participation in the study is voluntary and would never lead to any bias over their course grades. The participants of the current study supposedly have high English proficiency since they are foreign language majors selected through a University Entrance Exam which has got a demanding foreign language test. On being chosen based on the scores of this exam, the university test office also test the English proficiency of the freshman and offer a year-long preparatory language training if necessary before starting the four-year degree program of English language teacher training. In the first year of the degree program, the students take several basic language skill courses one of which is the 'Oral Communication Skills I and II' offered in two subsequent semesters. In this oral communication skill course, the students follow a speaking course book full of discussion topics and these topics are discussed in the classroom with the guidance and prompts of the course instructor. The assessment

of the course is done through oral interviews conducted twice a year. In these interviews, the individual students pick up one of the offered topics and after preparing within several minutes start to talk about the chosen topic. The instructor obviously interferes with the process of the learners' talk and engages in conversation with, questions, commentaries, support or any kind of prompt and mimicry. While the course instructor interacts with the candidate, he/she also rates the candidates' performance following a rating rubric adopted by the department. An inter-rater, a colleague of the instructor, also observes the interaction and assesses the performance. A final score is achieved by the mean of the two separate raters.

2.2. Data Collection and Analysis Procedure

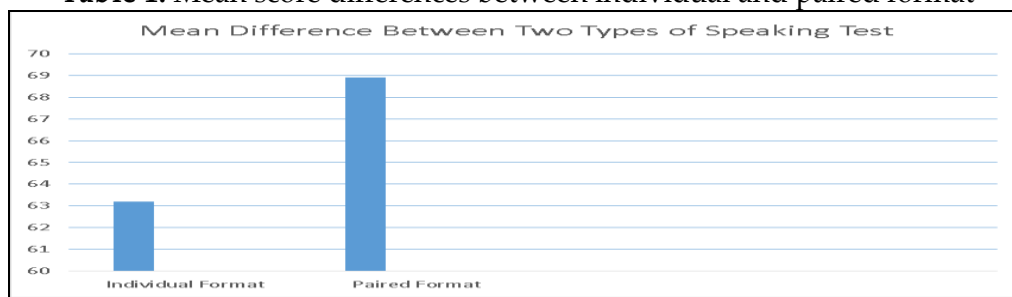
The data used in the quantitative analysis comes from the grades scored by the students in two oral proficiency exams conducted within the scope of course assessment during the regular exam period of the program. The midterm exam was traditionally conducted, namely in the form of individual or one-to-one interaction with the course instructor, put differently, the interlocutor of the students was the course instructor. As a difference, the second exam during the finals was carried out in the paired speaking format. In this second paired exam, the students were asked to pick up a partner and discuss on a given topic together. In both exams, there was a co-rater for interrater reliability and the discussions topics were prepared by the researcher and the assisting course instructor. However, the topics of the paired speaking exam were more provocative in nature to prompt a discussion. These topics were specifically chosen from a book titled *Taboos and Issues* (MacAndrew & Martinez, 2001) (see Appendix A). A paired samples t-test was computed using SPSS statistical software to see whether there is a significant difference between two distinct formats of speaking test – individual vs. paired.

Secondly, the qualitative data was obtained from the semi structured interviews carried out with 14 of the 25 students immediately after the paired speaking test. The interview questions specifically sought to reveal how the participants felt about the paired format of speaking test, how they would compare this format to the individual format, and if they were given the chance which format they would prefer and why (see Appendix B). The interviews were conducted in Turkish, the participants' mother tongue, to maximize the ease of administration. The interviews were audio-recorded and transcribed by the researcher. To analyze the interview data, we performed a thematic content analysis as suggested by Weber (1990) and Creswell (2012). To reveal the codes and categories out of large pieces of qualitative data, we through read the data, noting on the margins and coloring the related, consistent, recurring unit of ideas so that they would later on come together to create codes. To depict the categories, verbatim quotations from the data were used.

3. Results

In this section, the results of the both statistical computation and the qualitative analysis are presented. The paired samples t-test, which was computed to find out whether there is a significant difference between test formats, revealed that the scores that student obtained separately from two distinct testing format are statistically and significantly different, ($M=63,2$, $SD=12,3$) and ($M=68,9$, $SD=14,2$), individual and paired format respectively, $t(24)=-2,57$, $p<.05$. In other words, the students got better scores in the paired speaking test where they discussed a given topic compared to the individual testing of oral proficiency where they interacted with the course lecturer. There is mean difference of 5,7 in the total success rates of two exams, with paired format's mean being higher and statistically significant as can be seen from Table 1 below:

Table 1: Mean score differences between individual and paired format



The qualitative analysis of the transcriptions revealed that 10 out of 14 interviewees preferred paired speaking test over the individual one. 10 students noted that they would chose to sit in the individual type of oral examination if they were given a change to choose. Reasons of this preference vary. First of all, there are affective reasons pertaining to how they feel during the speaking tests. All of those who favored the paired format stated that they felt quite more comfortable, relaxed and found the paired format far less anxiety provoking. A majority of the participants attributed the comfortable feelings associated with the paired format to their partners, namely to the fact that the partners are already well acquainted friends whom they themselves chose, who are familiar and similar to themselves in terms of proficiency and general attitudes as exemplified below;

"I feel so relaxed because he is my friend, he speaks almost as well as I speak. I easily guess what he will say on the topic as I have got to know him well". (Student 1)

"I can plan what I am going to say while my friend is speaking, just like having a normal conversation. She is a very close friend of mine." (Student 4)

"You know your friend that's why you feel comfortable because we are nearly at the same level" (Students 7)

“When you speak with the lecturer, you get stressed because he speaks like a native. But my friend’s pronunciation is easier to understand” (Student 13)

According to our analysis, those who favored the traditional one-to-one or individual test design stated the opposite, claiming the paired format to be demoralizing making them feel desperate and nervous. The reason why they feel in such a negative way about the paired format is that they, as a dyad, could not maintain the conversation and could not support each other, and once one of them had difficulty in talking, the other one also remained silent and they thought this led to their failure in the exam. For this reason, they highlighted that in the individual format, the help of the lecturer in the form of oral prompts, clues or guiding questions saved them from absolute silence. This idea can be clearly seen in the quotations below;

“The instructor can make me talk and think more by asking questions. But we always hush with my friend” (Student 5)

“When the instructor sees that we are having difficulty, he paraphrases his questions and this helps our speech diversify” (Student 6)

“I kept quite, my friend kept quite and we just stood there silent. If I had talked to the instructor, he would have done something to keep the conversation going” (Student 8)

While some students keep silent especially when paired to discuss, some students told that speaking to a friend in the paired discussion format encouraged them to easily generate more ideas about the topic without getting nervous. Some students also said that they could manage to provide interactional help for one another when needed;

“When you talk alone, you also think by yourself, but when you get paired you share the responsibility of the topic. He poses an idea then I pose another idea and we either agree or disagree and this means lots of talk together” (Student 12)

“For example, while I am talking about A, my friend talks about B and this makes me think about C” (Student 9)

“When my friend is stopped up, I immediately interfered. Especially, if we disagree on the topic, that produces more talk” (Student 4)

We also asked in the interview whether the participants had any recommendation to enhance the speaking exams offered in the department or what characteristics an individually ideal speaking test should carry. We get only two recommendations, one from a paired format advocate and one from individual format

advocate. The rest of the students did not propose any suggestion apart from highlighting their preference, either paired or individual format. The recommendations are verbatim quoted below:

“There may be voice record, perhaps in the lecturer’s office or in a classroom but without the lecturer himself in the room” (Student 8).

“The person sitting and watching (co-rater) beside is always taking notes and I am very disturbed of his/her presence. This makes me highly stressed.” (Student 2)

4. Conclusion

As noted earlier, the question of assessing speaking ability in a L2 and deciding on the test design is one of the most challenging areas of the foreign language testing area (Birjandi & Bagherkazemi, 2011). In our preliminary analysis, we have come to see that students got better scores in the paired speaking test format that we designed in the form of paired discussion in comparison to the individual format where the test takers’ interlocutor is the course lecturer. The difference in the mean scores obtained in the two different formatted exams was found statistically significant. The better performance which resulted in the higher scores in the paired format was verified by some previous studies with which our results are in tune (Davis, 2009; Van Moere, 2013). Our purpose was twofold one of which was to uncover the preferences. For this purpose, we conducted a semi-structured interview and our interview findings showed that the majority of the students liked and supported the practice of being paired as also stressed by (Együd and Glover, 2001; Kanga, 2012). However, the participants who favored the individual test design (n=4) felt anxious due to the poor performance of their interlocutors. This kind of disfavor was verified in several previous studies (Iwashita, 1998; Foot, 1999). As the number of the participants in the current study is quite limited, the findings are confined to the research setting and not generalizable beyond. Studies which also control the personality and proficiency variables with higher number of participants are highly needed to investigate the effect of the paired speaking test format, which is very scarce.

References

1. Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford, UK: Oxford University Press.
2. Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford, UK: Oxford University Press.

3. Berry, V. (2004). A study of the interaction between individual personality differences and oral performance test facets. Unpublished doctoral dissertation. King's College, University of London, UK.
4. Birjandi, P., & Bagherkazemi, M. (2011). From Face-to-Face to Paired Oral Proficiency Interviews: The Nut is Yet to be Cracked. *English Language Teaching*, 4(2), p169.
5. Bonk, W. J., & Van Moere, A. (2004). L2 group oral testing: The influence of shyness/outgoingness, match of interlocutors' proficiency level, and gender on individual scores. Paper presented at the annual meeting of the Language Testing Research Colloquium, Temecula, California
6. Brooks, L. (2009). Interacting in pairs in a test of oral proficiency: Co-constructing a better performance. *Language Testing*, 26, 341-366.
7. Creswell, J. W. (2012). *Educational research: Planning, conducting, and evaluating quantitative*. Prentice Hall
8. Davis, L. (2009). The influence of interlocutor proficiency in a paired oral assessment. *Language Testing*, 26(3), 367-396.
9. Egyud, G., & Glover, P. (2001). Readers respond. Oral testing in pairs-secondary school perspective. *ELT journal*, 55(1), 70-76.
10. Foot, M. C. (1999). Relaxing in pairs. *ELT Journal*, 53(1), 70-76.
11. Galaczi, E. D. (2008). Peer-peer interaction in a speaking test: The case of the First Certificate in English examination. *Language Assessment Quarterly*, 5, 89-119.
12. Galaczi, E. (2014). Interactional competence across proficiency levels: How do learners manage interaction in paired speaking tests? *Applied Linguistics*, 35(5), 553-574.
13. Grice, H. P. (1975). Logical conversation. In P. Cole & J. Morgan (Eds.), *Syntax and semantics 3: Speech acts* (pp. 41-58). New York: Academic Press
14. Ikeda, K. (1998). The paired learner interview: A preliminary investigation applying Vygotskian insights. *Language, Culture and Curriculum*, 11(1), 71-96.
15. Iwashita, N. (1998). The validity of the paired interview in oral performance assessment *Melbourne Papers in Language Testing*, 5(2), 51-65.
16. Kanga, K. N. (2012). Individual and paired oral proficiency testing: A study of learners' preference (Order No. 1517174). Available from ProQuest Dissertations & Theses Global. (1039138663). Retrieved from <http://search.proquest.com/docview/1039138663?accountid=7181>
17. MacAndrew, R., & Martínez, R. (2001). *Taboos and issues*. Language Teaching Publications.
18. Norton, J. (2005). The paired format in the Cambridge speaking tests. *ELT*, 59(4), 287-297.
19. O'Sullivan, B. (2002). Learner acquaintanceship and oral proficiency test pair-task performance. *Language Testing*, 19(3), 277-295.
20. van Lier, L. (1989). Reeling, writhing, drawling, stretching and fainting in coils: Oral proficiency interviews as conversation. *TESOL Quarterly*, 23(3), 489-508.

21. Van Moere, A. (2013). Paired and group oral assessment. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics* (pp. 1-4). Oxford, England: Wiley-Blackwell
22. Weber, R. P. (1990). *Basic content analysis* (2nd ed.). Thousand Oaks, CA: Sage Publications

Appendix A: Discussion topics used in the paired speaking test

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"In certain cases, police and army officers can and should use torture to get the information they urgently need to prevent death, injury, any loss or damage of other people."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"Even if you are the only witness and telling the truth in a criminal case, it may be okay to accept the large sum of money offered to you by the defendants to change your testimony. Remember that you immensely need money."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"Animals can have no actual right at all. No one should be cruel to them but there is no point in objecting to eating them or doing medical and cosmetic experiments on them if we want to use these products safely."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"When a girl marries, she exchanges the attention of many men for the inattention of one."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"I don't mind my body being cut up and my organs taken out either to help sick people or just for the sake of science. After all, I am dead!"

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"Cancer research is more important than HIV research since people with HIV deserve what they have, and they can wait."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"It is unfair of the government to spend my taxes to support the people of other countries. I will never get anything from their country, will I? So, refugees should be sent back to their home country."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"It is not always better for people to go out with or marry people with similar ages. Age should be out of question when it comes to love and relationships."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"It is hard to understand that women always earn less than men in sport. The men's champion of Wimbledon wins £477, 500 but the women's champion of the same tournament wins 430.000."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"I can forgive but I cannot forget" is only another way of saying "I cannot forgive."

Pair Discussion: Discuss the following topic with your pair. You may take either 'against' or 'for' stance, or you may both agree or disagree with the topic and support each other's opinion. Defend your point through your reasons, or negotiate through conversation.

"It is acceptable for people to bury their loved ones or relatives somewhere around their houses, for example in their garden."

Appendix B: (Semi-structured interview questions)

- 1) Az önce bir konuşma sınavından çıktınız ve arkadaşınızla konuştunuz. Bu formatla ilgili olarak ne düşünüyorsunuz?
- 2) Sınavda arkadaşınızla konuşmak size kendinizi nasıl hissettirdi?
- 3) Vize sınavını hatırlayacak olursanız, orada hocayla konuşmuşunuz. Vize sınavınızı az önceki sınavla nasıl kıyaslıyorsunuz?
- 4) Bir kez daha sınava girme şansınız olsaydı, hangi formatı tercih ederdingiz? Neden?
- 5) Son olarak idealinizde bir konuşma sınavı nasıldır ve/veya nasıl olmalıdır? Getirebileceğiniz bir öneri var mı?

Zafer Susoy
INDIVIDUAL VS. PAIRED FORMATS OF ORAL PROFICIENCY EXAM:
DOES IT MATTER IN TERMS OF SCORES AND PREFERENCE?

Creative Commons licensing terms

Author(s) will retain the copyright of their published articles agreeing that a Creative Commons Attribution 4.0 International License (CC BY 4.0) terms will be applied to their work. Under the terms of this license, no permission is required from the author(s) or publisher for members of the community to copy, distribute, transmit or adapt the article content, providing a proper, prominent and unambiguous attribution to the authors in a manner that makes clear that the materials are being reused under permission of a Creative Commons License. Views, opinions and conclusions expressed in this research article are views, opinions and conclusions of the author(s). Open Access Publishing Group and European Journal of Foreign Language Teaching shall not be responsible or answerable for any loss, damage or liability caused in relation to/arising out of conflicts of interest, copyright violations and inappropriate or inaccurate use of any kind content related or integrated into the research work. All the published works are meeting the Open Access Publishing requirements and can be freely accessed, shared, modified, distributed and used in educational, commercial and non-commercial purposes under a [Creative Commons Attribution 4.0 International License \(CC BY 4.0\)](https://creativecommons.org/licenses/by/4.0/).