



Iranian Journal of Applied Linguistics (IJAL)

Vol. 25, No. 1, March 2022, 61-77

---

## The Effect of Online Assessment on Speaking Complexity, Accuracy, and Fluency of Iranian Intermediate EFL Learners

Malihe Maldar

*Tabaran Institute of Higher Education, Mashhad, Iran*

---

### Abstract

---

Evaluating the oral abilities of foreign language learners through online assessments attitudes a significant challenge for both teachers and students. Conducting skill evaluations in person, through direct interaction between the examiner and the learner, enables more authentic communication. However, shifting to a digital format and incorporating technological tools may cause discomfort among students and take away the friendly atmosphere of the assessment. This study aimed to investigate the impact of online assessment on the speaking complexity, accuracy, and fluency of intermediate English as a foreign language (EFL) learners. Initially, 50 learners took the Oxford Placement Test, with only 44 scoring between 50 and 60, placing them at level B1. These students were then divided into an experimental group and a control group. The Test of English for International Communication (TOEIC) was administered as a pretest and posttest for both groups, alongside a speaking pretest to assess complexity, accuracy, and fluency (CAF) scores. The experimental group underwent online assessment (OA) via the Adobe Connect platform, while the control group engaged in book exercises without any specific treatment. Following the intervention, both groups took a speaking posttest (TOEIC). Data analysis involved Pearson Correlation Coefficient, kurtosis, skewness, descriptive statistics, and sample t-tests. The findings revealed that online assessment significantly improved participants' speaking accuracy and fluency, although it did not have a significant impact on speaking complexity.

**Keywords:** Accuracy, Complexity, Fluency, Online assessment

---

*Corresponding author:* English Department, Tabaran Institute of Higher Education, Mashhad, Iran  
*Email address:* malihe\_maldar@yahoo.com



## **1. Introduction**

Internet technologies have become part of professional, academic, individual, and commercial life. Students consider that using collaborative technologies helps them grow learning productivity, apply a suitable method to learning, and enhance their understanding of course content (Kember et al., 2010). Quarto et al. (2018) mentioned that online education is developing in various ways worldwide due to the emergence of new technologies, the global spread of the Internet, and the increasing need for human resources to receive regular training for improvement. Online assessment tools offer a convenient way for students to practice and demonstrate their speaking skills in a simulated environment, preparing them for real-world communication situations (Luoma, 2004).

In the field of English language education, the utilization of technology in assessment practices has significantly transformed the evaluation of students' speaking skills (Brown, 2014). With the advent of online assessment tools, educators are exploring the impact of these digital platforms on the development of speaking complexity, accuracy, and fluency in English as a foreign language (EFL) contexts (Chapelle, 2016). Luoma (2004) stated that online assessment on students' linguistic abilities, focusing on how these innovative tools can enhance the teaching and learning of spoken English. The integration of online speaking assessments allows educators to provide more immediate and personalized feedback to students, facilitating targeted language development (Fulcher, 2014). Technology-enhanced speaking assessments can increase student engagement and motivation, leading to improved language proficiency outcomes (Chapelle, 2003). Shohamy (2001) believed that as educators navigate the complexities of incorporating technology into language assessment practices, collaboration between researchers, practitioners, and technology developers is essential to ensure the effectiveness and validity of online speaking assessments.

## **2. Literature review**

### *2.1. Online assessment*

In recent years, the integration of technology into language assessment practices has gained significant attention in the field of English language education. Online speaking assessments, in particular, have emerged as a valuable tool for evaluating students' oral proficiency and promoting language development. Several studies have highlighted the advantages of incorporating online speaking assessments into language education. Fulcher (2014) emphasizes that these assessments enable educators to provide immediate and personalized feedback to students, leading to targeted

language development. So, this constant feedback supports scaffolding learning which causes interaction effectively among students and educators (Ludwig-Hardman & Dunclap, 2003). Additionally, technology-enhanced speaking assessments have been shown to increase student engagement and motivation, ultimately improving language proficiency outcomes (Chapelle, 2003). Luoma (2004) notes that online assessment tools offer a convenient way for students to practice and demonstrate their speaking skills in a simulated environment, preparing them for real-world communication situations.

According to Maley and Kiss (2018), online assessment is a new technique to evaluate students' learning process in a web-based situation or through the Internet. Online assessment should be viewed as a system for evaluating students' academic achievement. The teachers should enlarge the evaluation measures utilized throughout the educational conveyance of the internet-based course to turn it into a thriving online assessment (Robles & Braathen, 2002). Also Baleghizadeh and Saeedi (2021) believed that the use of trained teachers who can effectively teach and assess students in online classes positively impacts students' learning and contributes to their satisfaction with the classroom experience. If online assessment methods are not taught to teachers, they evaluate learners using the same old methods by which the themselves were assessed (Taghizadeh, Mazdayasan and Mahdavi-rad, 2020). According to their research educators lack adequate skills for online assessment, and they should be trained in effective methods. This article attempts to utilize an approach for online assessment.

## 2.2. *Speaking CAF*

In today's digital age, the ability to communicate effectively in a foreign language is becoming increasingly important. Research has shown that speaking online assessments can provide a more authentic and interactive environment for language learners to demonstrate their speaking abilities (Bachman & Palmer, 2010). It is assumed that second/foreign language presentation could be described by three dimensions of CAF (Ellis, 2008; Larsen-Freeman, 2009). Michel (2017) described complexity as the amounts, difficulties, richness, and diversity of the student's performance

Michel (2017) described complexity as the amounts, difficulties, richness, and diversity of the student's performance. In this study, complexity is the average number of clauses per T-unit

(Larsen-Freeman, 2006, p. 597). T-unit is an independent clause and clauses attached to or embedded in it.

Fluency is distinct from a student's speed and amount of speech without dysfluency markers such as self-correction, functionless repetitions, and false starts – in "coping with real-time processing" (Wolfe-Quintero et al., 1998, p. 14). In this study, fluency is the average number of words in each T-unit (Larsen-Freeman, 2006, p. 597).

Accuracy is designated as the rank to which English students' production is created on the instruction method of the target language (Ellis & Barkhuizen, 2005). In this research, accuracy is the proportion of error-free T-units (Larsen-Freeman, 2006, p. 597).

Reviewing the related literature, there have been some studies examining the effects of online learning on speaking CAF. Although previous studies have researched the evolution of online speaking, they have not assessed the dimensions (CAF) of speaking in an online context. So, this research planned to investigate the effects of online assessment on Iranian EFL learners' speaking CAF by formulating the following questions:

1. Does the online assessment have any significant effect on the speaking fluency, accuracy complexity, of Iranian intermediate EFL learners?

### **3. Materials and Method**

#### *3.1. Participants and Setting*

The participants of this research were 44 Iranian EFL learners who were chosen from 50 students from a private institute. And all of them were at the intermediate level. None of the learners lived in an English –Speaking country, and they spoke Persian as their mother tongue. They were studying at the institute, located in Iran. In this study, participants were selected based on convenient non-random sampling. This research was conducted in 2021, spring, and summer semesters. Twenty-two students were in the experimental group 9 of them were male, 13 were female, and 22 students were in the control group, 11 females and 13 males. Their age ranged from 16 to 20 years old. They were in level B1.

#### *3.2. Instruments*

*Oxford Placement Test.* The Oxford Placement Test (OPT) which comprises 100 questions divided into two sections, was administered to assess the uniformity of the students' proficiency

levels (see Appendix 1). This assessment aligns with the Common European Framework of Reference (CEFR) standards, categorizing students who achieved scores between 50-60 as being at the B1 level.

*Pre-Test and Post-Test.* The Test of English for International Communication (TOEIC) was used as a pre-test and post-test in this research. It is an international standardized test of English language proficiency for non-native speakers and it is intentionally designed to measure the everyday English skills of people working in an international environment. The test-retest reliability of scores of the TOEIC Speaking was reported to be approximately .80, based on the data of 16,867 test-takers.

*Larsen Freeman's Rubric.* In the present study, the rubric of Larsen Freeman (2006, p. 295) was used to measure speaking complexity, accuracy, and fluency. "Fluency: "Average number of words in each T-unit." complexity: "Average number of clauses in each T-unit." Accuracy: "The proportion of error-free T-units."

### 3.3. Materials

*Speak Now Book 3.* For this study, Richards and Bohlke (2012) used the book *Speak Now three*.

*Pictures.* In this research, for measuring CAF, 4 specific pictures were used as a treatment for online classes. So, the pictures were selected according to participants' levels.

### 3.4. Procedure

Initially, 50 Persian language students learning English at the Institute in Mashhad were tested to determine their homogeneity. English was their foreign language, and their mother tongue was Persian. OPT test consisted of two parts. In the first part, the students answered 50 questions. The second part consisted of 50 questions ranging from moderate to complex and ten single-choice questions. Only 44 Students received scores between 50 and 60, which put them in level B1. Then they were divided into two groups. Both groups had two sessions in a week, one hour and a half sessions; the semesters included 14 sessions in 7 weeks. The experimental group was held on an online platform, and the control group class was held in a face-to-face class. At the beginning of the semester, a pre-test was administered to both groups.

The pre-test consisted of several parts. Part one involved reading a text aloud, describing a picture, responding to some questions, proposing a solution, and expressing an opinion. In Part 2, learners should describe the picture on the screen in as much detail as possible. They had 30

seconds to prepare their responses. In Part 3, learners responded to 3 questions immediately without preparation. In Part 4, learners answered three questions based on the information provided. In Part 5, learners were presented with a problem and asked to propose a solution. In Part 6, learners gave their opinions about a specific topic.

Students' voices were recorded, and the teacher calculated the pre-test and post-test scores based on the speaking complexity, accuracy, and fluency according to the rubric of Larsen Freeman (2016, p. 579). The experimental group received treatment every three sessions. The treatment given to the experimental group consisted of 4 different photos. These pictures had different subjects, and Teacher gave the some words as a clue then students had to describe them and explain the details. If something were left, the teacher asked them some questions. The students did not have time to think and describe the pictures. The online class was held on the Adobe Connect platform. On days when the assessment was done, the teacher gave specified time to students. And students entered the class in turn. And the teacher showed them one of 4 pictures randomly, each student talked for about 4 to 5 minutes for a particular picture. Students received feedback as a mark. This treatment was repeated in the third, sixth, ninth, and twelfth sessions.

The control group did not receive any treatment. They were 22 students in the class, they did not receive any treatment, and they were involved with exercises that were in the book. The number and length of the sessions were the same for both groups. And both groups passed the "Speak now 3" book. At the end of the semester, the post-test was taken. This test was also used as a final exam. In the post-test, the same method was used to record the students' voices, and then the teacher calculated their CAF scores.

Meanwhile, the students' recorded voices were given to two raters to get reliability. One of the raters had a Ph.D. in Applied Linguistics. Moreover, the other was M.A in TEFL. They used Larsen Freeman's Rubric too, which was applied for this research to score the CAF. All the Tables related to the raters' reliability tests are listed.

### *3.5. Data Analysis*

Data were analyzed using SPSS software, version 26. To show inter-rater reliability, Pearson Correlation Coefficient was used. For normality of data, Kurtosis and skewness were applied. Descriptive statistics and independent samples t-tests were used to answer the research questions. For inter-group analysis, independent samples t-test was conducted to compare the performance groups in the pre-test and post-test.

#### 4. Results

Before conducting any analyses of the pretest, and post-test, it was essential to check the normality of the distributions. Thus, the Kolmogorov-Smirnov test of normality was used to check the normality. For the data to be normal Skewness and Kurtosis should be in +2 and -2. The result of pre-test normality for the first question showed that skewness is .54 and Kurtosis is -.90. And the post-test normality for the first question was Skewness of .54 and Kurtosis -1.1. Therefore the data for the post-test and pretest for the first question was normal. the pretest normality for the second question showed that Skewness was .36, and Kurtosis was .35. and the Skewness for the posttest normality was .16 and Kurtosis was -1.6. Therefore the data for the pretest and post-test for the second question was normal. The result of pretest normality for the third question showed that the Skewness was .39 and Kurtosis was -.11. Also the Skewness for the post-test normality was .27 and Kurtosis was -.31. So, the data of pretest and post-test for the third question was normal.

Table 1. *Descriptive statistics of the Proficiency Test*

|                | N  | Minimum | Maximum | Sum     | Mean    | Std. Deviation | Variance |
|----------------|----|---------|---------|---------|---------|----------------|----------|
| Placement test | 44 | 52.00   | 60.00   | 2422.00 | 55.0455 | 2.24079        | 5.021    |

Descriptive statistics were used to define the main features of the data collected from the placement. Table 1 organized the main analysis of data about the placement test. It shows information such as mean, standard deviation, and the number of participants. The pretest scores were 52 to 60 and the mean scores ( $M=55.0455$ ) show that 44 students in this research were at the intermediate level. Then, 44 participants were randomly distributed into two groups. 22 learners were in the experimental group and 22 learners in the control group.

Tables 2 show the results of the pre-test and post-test of fluency. To check the null hypotheses of the study, the independent sample was applied for the first question.



Table 2. Results of the Pre-test and post-test of Speaking Fluency in the Experimental and Control Groups

| Group        | N  | Mean     | Std.     | T      | df | MD     | sig(two-tailed) | 95%CI   |         |
|--------------|----|----------|----------|--------|----|--------|-----------------|---------|---------|
|              |    |          |          |        |    |        |                 | Low     | upper   |
| C.G.pretest  | 22 | 89.5455  | 6.11577  | -.359  | 42 | -.636  | .722            | 4.21740 | -2.9446 |
| E.G.pretest  | 22 | 90.1818  | 5.64536  |        |    |        |                 |         |         |
| C.G.posttest | 22 | 91.1818  | 5.81162  | 10.708 | 42 | -45.77 | .000            | -54.399 | -37.146 |
| E.G.posttest | 22 | 136.9545 | 19.18947 |        |    |        |                 |         |         |

This table shows that the mean score for students in the control group pretest was 89.5, and for the experimental group, the pretest was 90.1. The standard deviations in the table show that the variation in the data is a bit wider for the control group pretest (SD=6.1) than the experimental group pretest (SD=5.6). By looking at Mean scores, it can be indicated that, on average, students in the experimental group were better than the students in the control group. The mean difference in the pretest was -6363. And the "t" value was -.359. The p-value is more than (0.05) so it designated that there was no significant difference between the two groups in the pre-test of speaking fluency. In other words, the level of participants' fluency in both groups was equal. The outcomes of the post-test of speaking fluency show that the mean score for students in the control group post-test was 91.1, and for the experimental group, the post-test was 136.9. As well, the standard deviation for the experimental group post-test (SD=19.1) is wider than the control group post-test (SD=5.8). From the mean scores can be identified that students in the experimental group achieved better than the students of the control group in overall speaking fluency. The results of the independent samples t-test for the experimental and control groups post-test display that the mean difference was -45.7 and the "t" value was 10.7. The p-value indicated in the "sig (2 tailed)" is .00 which is less than (0.05) and it showed that there was a significant difference between the post-test of both groups. So, the outcomes indicated that using pictures as a treatment for the experimental group, had a significant effect on Iranian Intermediate EFL learners' speaking fluency.

Table 3 shows the results of the pretest and posttest of speaking accuracy in the experimental and control groups. To check the null-hypotheses 2, of the study, independent samples were applied.

Table 3. *Results of the Pre-test and post-test of Speaking Accuracy in the Experimental and Control Groups*

| Group        | N  | Mean     | Std.     | T      | df | MD       | sig(two-tailed) | 95%CI    |         |
|--------------|----|----------|----------|--------|----|----------|-----------------|----------|---------|
|              |    |          |          |        |    |          |                 | Low      | upper   |
| C.G.pretest  | 22 | 98.1818  | 6.08383  | .914   | 42 | - 1.7727 | .366            | - 5.6871 | -2.1417 |
| E.G.pretest  | 22 | 99.9545  | 6.76459  |        |    |          |                 |          |         |
| C.G.posttest | 22 | 99.5455  | 8.49497  | 15.491 | 42 | -52.545  | .000            | 59.3905  | 45.700  |
| E.G.posttest | 22 | 152.0909 | 13.45153 |        |    |          |                 |          |         |

This table displays that the pretest means score for students in the control groups was 98.1, and for the experimental group was 99.9. The standard deviations of the pretest show that the variation in the data is a bit wider for the experimental group (SD=6.7) than for the control group (SD=6.0). Therefore the results from the independent samples t-test of the pretest indicate that the "t" value for is .91 which is far away from 0. The p-value, in the "sig (2 tailed) is .36 which is more than 0.05 and it displayed that, there was no significant difference between the pretest of both groups. In other words, the level of participants' fluency in both groups was equal. The outcomes of the speaking accuracy post-test of the groups designate that the mean score for students in the control groups was 99.54, and for the experimental group was 152.09. In addition, the standard deviation for the experimental group (SD=13.451) is wider than the control group (SD=8.49). From the mean scores can be stated that students in the experimental group achieved better than the students of the control group in overall speaking accuracy. To understand the significant differences between the performance of the students of the control and experimental group, look at the results of the independent samples t-test. It shows the mean difference was -5254. And the "t" value was 15.491. The p-value is less than (0.05) so it indicated that there was a significant difference between the post-test of both groups. So, we can determine that the observed difference between

the two means was statistically significant. Therefore, the results displayed that online assessment had a significant impact on Iranian Intermediate EFL learners' speaking accuracy.

Table 4 shows the results of the pretest and post-test of speaking in the complexity in the experimental and control groups. To investigate the null hypotheses for research question 3 of this study, independent samples were applied.

Table 4. *Results of the Pre-test and Post-test of Speaking Complexity in the Experimental and Control Groups*

| Group        | N  | Mean    | Std.    | T     | df | MD      | sig(two-tailed) | 95%CI   |          |
|--------------|----|---------|---------|-------|----|---------|-----------------|---------|----------|
|              |    |         |         |       |    |         |                 | Low     | upper    |
| C.G.pretest  | 22 | 100.590 | 7.0620  | -.270 | 42 | -.5909  | .789            | -5.0113 | 3.8294   |
| E.G.pretest  | 22 | 101.181 | 7.46188 |       |    |         |                 |         |          |
| C.G.posttest | 22 | 101.136 | 6.92336 | -.677 | 42 | -1.5000 | .502            | 5.97301 | -2.97301 |
| E.G.posttest | 22 | 102.636 | 7.75546 |       |    |         |                 |         |          |

This table displays that the mean score for students in the control group pretest was 100.59, and the experimental group pretest was 101.18. The standard deviations in the table above show that the variation in the data is for the experimental group (SD=7.062) and the control group (SD=7.461). The results from the independent samples t-test indicate that the "t" value is -.270 which is far away from 0. The p-value, in the "sig (2 tailed) is .789 which is more than 0.05 and it displayed that, there was no significant difference between the pretest of both groups. In other words, the level of participants' complexity in both groups was equal. The outcomes of the post-test of speaking complexity show that the mean score for students in the control group post-test was 101.13, and for the experimental group, the post-test was 102.636. Also, the standard deviation for the experimental group post-test (SD=7.755) is a bit wider than the control group (SD=6.9233). From the mean scores can be stated that students' performance in the experimental group is a little better than the control group in overall speaking complexity. To understand the significant differences between the performance of the students of the control and experimental group in the

posttest notice the results of the independent samples t-test. The mean difference was -1.5000 and the "t" value was .677. The p-value is more than (0.05) so it indicated that there was no significant difference between the post-test of both groups. So, we can determine that the observed difference between the two means was not statistically significant. Therefore, the results displayed that online assessment didn't have any significant impact on Iranian Intermediate EFL learners' speaking complexity. So the null hypothesis was accepted.

## 5. Discussion

The researchers found that students who underwent online assessment showed significant improvements in speaking ,accuracy, and fluency compared to those who were assessed through traditional methods. This suggests that online assessment can be a valuable tool for enhancing speaking skills in English language learners. So the first hypothesis was accepted. Hsu-chon (2019) found that using task repetition and post-task transcribing have a significant effect on speaking fluency and accuracy. The findings of the current study stated that there was no significant effect on the speaking complexity of Iran Intermediate EFL Learners.

One possible explanation for these results is that online assessment allows for more opportunities for practice and feedback, which can help students improve their speaking skills over time. Additionally, the use of technology in assessment may motivate students to engage more actively in speaking tasks, leading to better performance. Since online classes are tedious for most students, they have to be managed in new and innovative ways. Up-to-date methods help students to participate more in the classes. Studies show that using teaching aids such as games, videos, and pictures motivates students to participate in online classes. Also, students' interaction with teachers and their classmates increases their motivation.

The findings of this study have important implications for language educators and curriculum designers. Incorporating online assessment into language learning programs can provide students with valuable opportunities to practice and improve their speaking skills. It also highlights the importance of integrating technology into language teaching to enhance student learning outcomes. Also This research shows that speaking complexity is more complex than speaking accuracy and fluency. Therefore, assessing speaking complexity is not successful just by showing pictures. Teachers can take the initiative to use the assessment methods that are common in face-to-face classes and online classes.

## 6. Conclusion

In conclusion, the research suggests that online assessment can be a valid and reliable method for evaluating speaking skills, but it may present challenges that can affect the complexity of language produced. It demonstrates that online assessment can have a positive impact on the development of speaking skills in English language learners. The modality of the assessment (synchronous vs. asynchronous) and the type of assessment (dynamic vs. diagnostic) can influence the outcomes in terms of speaking accuracy and fluency. By providing students with more opportunities for practice, feedback, and engagement through technology, online assessment can help facilitate the development of speaking skills in English as a foreign language. Additionally, factors such as the examiner's behavior, technical efficiency of the platform, and the learners' anxiety levels play a role in the effectiveness of online speaking assessments.

This research had several limitations; the study may have a limited generalizability due to the small sample size of participants. They were limited to 44 intermediates EFL learners selected non-randomly from about 28 institutes Mashhad.

The study may have relied on a limited set of assessment tools to evaluate speaking complexity, accuracy, and fluency. Using a more comprehensive range of assessment measures could provide a more nuanced understanding of the effects of online assessment. Only four pictures were used in this research. In this study, speaking complexity, accuracy, and fluency were measured according to Larsen Freeman; that researcher could use another score rating according to other experts.

## 6. References

- Baleghzadeh, S., & Saeedi, M. (2021). A Qualitative Evaluation of an Online Teacher Training Program in Iran: EFL Teachers' Perceptions. *Iranian Journal of Applied Linguistics*, Vol. 24, No. 2, September 2021, 1-28.
- Brown, H. D. (2014). *Principles of language learning and teaching: A course in second language acquisition*. Pearson.
- Chapelle, C. A., & Voss, E. (2016). 20 years of technology and language assessment in language learning & technology. *Language Learning & Technology*, 20(2), 116-128.

- Ellis, R. (2008). *The study of second language acquisition (2nd edition)*. Oxford: Oxford University Press.
- Ellis, R. and G. Barkhuizen (2005). *Analysing Learner Language*. Oxford: Oxford University Press.
- Goodwin, C. 1981. *Conversational Organization: Interaction Between Speakers and Hearers*. New York: Academic Press.
- Fulcher, G.(2013). *Practical language testing*. Routledge
- Hsu, H. C. (2019). The combined effect of task repetition and post-task transcribing on L2 speaking complexity, accuracy, and fluency. *The Language Learning Journal*, 47(2), 172-187.
- Kember, D., McNaught, C., Chong, F. C., Lam, P., & Cheng, K. F. (2010). Understanding the ways in which design features of educational websites impact upon student learning outcomes in blended learning environments. *Computers & Education*, 55(3), 1183-1192.
- Larsen-Freeman, D. (2006). The emergence of complexity, fluency, and accuracy in the oral and written production of five Chinese learners of English. *Applied linguistics*, 27(4), 590-619.
- Luoma, S. (2004). *Assessing speaking*. Cambridge, UK: Cambridge University Press
- Ludwig-Hardman, S., & Dunclap, J. C. (2003). Learner support services for online students: scaffolding for success. *International Review of Research in Open & Distance Learning*, 4(1): 1–15.
- Maley, A., & Kiss, T. (2018). Becoming a Creative Teacher. In *Creativity and English Language Teaching* (pp. 181-201). Palgrave Macmillan, London.
- Michel, M. (2017). Complexity, accuracy and fluency (CAF). In Shawn Loewen & Masatoshi Sato (Eds.), *The Routledge Handbook of Instructed Second Language Acquisition*. London: Routledge.
- Quarto, N., Shailendra, S., Meyer, N. P., Menon, S., Renda, A., & Longaker, M. T. (2018). Twist1-haploinsufficiency selectively enhances the osteoskeletal capacity of mesoderm-derived parietal bone through downregulation of Fgf23. *Frontiers in physiology*, 9, 1426.
- Robles, M., & Braathen, S. (2002). Online assessment techniques. *Delta Pi Epsilon Journal*, 44(1), 39-49.
- Shohamy, E. (2001). *The power of tests: A critical perspective on the uses of language tests*. Essex, UK: Pearson.
- Taghizadeh, M., Mazdayasna, G., & Mahdavi-rad, F. (2020). Language Assessment Courses at Iranian State Universities: Are they Comprehensive Enough to Develop Valid Language

Assessment Literacy (LAL) among EFL Students?. *Iranian Journal of Applied Linguistics*, Vol. 23, No. 2, September 2020, 106-145.

Wolfe-Quintero, K. (1998). The connection between verbs and argument structures: Native speaker production of the double object dative. *Applied Psycholinguistics*, 19(2), 225-257.

## Appendix

### OXFORD PLACEMENT TEST 2 GRAMMAR TEST PART 1

Name \_\_\_\_\_  
Total Listening \_\_\_\_\_ / 100      Total Grammar Part 2 \_\_\_\_\_ / 50  
Total Grammar Part 1 \_\_\_\_\_ / 50      Grand total \_\_\_\_\_ / 200

Look at these examples. The correct answer is indicated in bold.

- A In warm climates people **like** likes are liking sitting outside in the sun.  
B If it is very hot, they sit **at** in under the shade.

Now the test will begin. Tick the correct answers.

- 1 Water **is freezing** ~~be freezing~~ ~~freezes~~ at a temperature of 0° C. 1 \_\_\_  
2 In some countries **there is** ~~is~~ ~~it is~~ dark all the time in winter. 2 \_\_\_  
3 In hot countries people wear light clothes **for keeping** ~~to keep~~ ~~for to keep~~ cool. 3 \_\_\_  
4 In Madeira they have **the good** ~~good~~ ~~a good~~ weather almost all year. 4 \_\_\_  
5 Most Mediterranean countries are **more warm** ~~the more warm~~ ~~warmer~~ in October than in April. 5 \_\_\_  
6 Parts of Australia don't have **the** ~~some~~ ~~any~~ rain for long periods. 6 \_\_\_  
7 In the Arctic and Antarctic **it is** ~~there is~~ ~~it has~~ a lot of snow. 7 \_\_\_  
8 Climate is very important in **most** ~~of most~~ ~~the most~~ people's lives. 8 \_\_\_  
9 Even now there is **little** ~~few~~ ~~less~~ we can do to control the weather. 9 \_\_\_  
10 In the future **we'll need** ~~we are needing~~ ~~we can need~~ to get a lot of power from the sun and the wind. 10 \_\_\_  
11 Pele is still perhaps **most** ~~the most~~ ~~the more~~ famous footballer in the world. 11 \_\_\_  
12 He **had been** ~~is~~ ~~was~~ born in 1940. 12 \_\_\_  
13 His mother **not want** ~~wasn't wanting~~ ~~didn't want~~ him to be a footballer. 13 \_\_\_  
14 But he **used** ~~ough~~ ~~has used~~ to watch his father play. 14 \_\_\_  
15 His father **made him to** ~~made him~~ ~~would make him to~~ practise every day. 15 \_\_\_

- 16 He learned to use or his left foot or and his left foot and both his left foot and his right. 16 \_\_\_
- 17 He got the name Pele when he had only ten years was only ten was only ten years. 17 \_\_\_
- 18 By 1956 he has joined joined had joined Santos and had scored in his first game. 18 \_\_\_
- 19 In 1957 he has been picked was picked was picking for the Brazilian national team. 19 \_\_\_
- 20 The World Cup Finals were in 1958 and Pele was looking forward to play to playing to be playing. 20 \_\_\_
- 21 But he hurt this the his knee in a game in Brazil. 21 \_\_\_
- 22 He thought he isn't going to couldn't wasn't going to be able to play in the finals in Sweden. 22 \_\_\_
- 23 If he hadn't been weren't wouldn't be so important to the team, he would have been left behind. 23 \_\_\_
- 24 But he was a such such a a so brilliant player, they took him anyway. 24 \_\_\_
- 25 And even though even so in spite of he was injured he helped Brazil to win the final. 25 \_\_\_

- The history of the World Cup is quite a a quite quite short one. 26 \_\_\_
- Football has been is being was played for 27 \_\_\_
- above over more than a hundred years, but the first World Cup 28 \_\_\_
- competition did not be was not was not being held until 29 \_\_\_
1930. Uruguay could win were winning had won the Olympic football 30 \_\_\_
- final in 1924 and 1928 and wanted be being to be World Champions for the third time. 31 \_\_\_
- Four teams entered from Europe, but with a little few little success. 32 \_\_\_
- It was the first time which that when professional teams 33 \_\_\_
- are playing would play had played for a world title. 34 \_\_\_
- It wasn't until four years later more further that a 35 \_\_\_
- European team succeeded to win in winning at winning 36 \_\_\_
- for the a its first time. The 1934 World Cup was 37 \_\_\_
- again won by a the one home team, 38 \_\_\_
- what this which has been the case several times since 39 \_\_\_
- then. The 1934 final was among between against two 40 \_\_\_
- European teams, Czechoslovakia and Italy, which that who won, 41 \_\_\_
- Went on to win winning to have won the 1938 final. Winning 42 \_\_\_
- successive finals is something that is not was not has not been achieved 43 \_\_\_
- again until Brazil did thesd them it in 1958 and 1962. If Brazil 44 \_\_\_
- would have won would win had won in 1966 then the 45 \_\_\_



authorities would have needed to    the original World Cup replaced. 46 \_\_\_\_

But England stopped the Brazilians    a third successive win. An England player, 47 \_\_\_\_

Geoff Hurst, scored three goals in the final and won it almost    48 \_\_\_\_

1966 proved    the last year that England 49 \_\_\_\_

even qualify for the finals till 1982, though they got in as winners in 1970. 50 \_\_\_\_

**GRAMMAR TEST Part 2**

- 51 Many  persons  people  peoples nowadays believe that everyone should learn to use computers. 51\_\_
- 52 The majority of children in the UK  have  has  are having access to a micro-computer. 52\_\_
- 53 There are more computers per head in England than  anywhere else  somewhere else  anywhere other in the world. 53\_\_
- 54 Learning a computer language is not the same  as  like  than learning a real language. 54\_\_
- 55 Most people start off with 'Basic',  who  what  which is the easiest to learn. 55\_\_
- 56 Children seem to find computers to easy, but many adults aren't used to  work  the work  working 56\_\_  
with microtechnology.
- 57 There aren't  no  any  some easy ways of learning to program a computer. 57\_\_
- 58 The only way to become really proficient is to practise a lot  on your own  by your own  on your self. 58\_\_
- 59 You can pick up the basics quite quickly if you  want to  would  are willing to make en effort. 59\_\_
- 60 Most adults feel it would be easier if only they  would have started  would start  had started 60\_\_  
computer studies earlier.
- 61 Some people would just  rather  prefer  better not have anything to do with computers at all. 61\_\_
- 62 A lot have resigned themselves to never even  know  known  knowing how a computer works. 62\_\_
- 63 Microtechnology is moving so fast that hardly  anybody  nobody  no one can keep up with it all. 63\_\_
- 64 It's no use  in trying  to try  trying to learn about computers just by reading books. 64\_\_
- 65 Everyone has  difficulty in learning  difficulties to learn  it difficult to learn 65\_\_  
if they can't get 'hands-on' experience.

Below is a letter written to the 'advice' column of a daily newspaper. Tick the correct answers.

Dear Marge,

- I am writing  I will write  I should write to you because I 66\_\_
- am not knowing  don't know  know not what to do. I'm twenty-six and a teacher at 67\_\_
- a primary school in Norwich where  I'm working  I've worked  I work for the last five years. 68\_\_
- When I  was  have been  had been there for a couple of years, one of the older members of staff 69\_\_
- would leave  left  had been leaving and a new teacher 70\_\_
- would be  became  was appointed to work in the same department as me. 71\_\_
- We  worked  have worked  should work together with the same classes during her first year 72\_\_
- and had the  opportunity for building  possibilities to build  chance to build up a good professional 73\_\_
- relationship. Then, about eighteen months after  she has arrived  to have arrived  arriving 74\_\_
- in Norwich, she decided to buy  her own  herself  her a house. 75

She was tired of  to live  live  living in rented accommodation and wanted a place 76 \_\_\_\_  
 by her own  of her own  of herself. At about the same time, I 77 \_\_\_\_  
 was given  have been given  gave notice by the landlord of the flat 78 \_\_\_\_  
 what I was living  that I had lived  I was living in 79 \_\_\_\_  
 and she asked me if I  liked  had liked  would like to live 80 \_\_\_\_  
 with her. She  said  told  explained me that by the time she 81 \_\_\_\_  
 would pay  would have paid  had paid the mortgage 82 \_\_\_\_  
 and the bills  it  there  they wouldn't be 83 \_\_\_\_  
 a lot  many  few left to live on. She suggested 84 \_\_\_\_  
 us to  we should  we may share the house and share the costs. 85 \_\_\_\_  
 It seemed like a good idea, so after  we'd agreed  we could agree  we agreed with all the details 86 \_\_\_\_  
 what  that  who needed to be sorted out, we moved into the new house together. 87 \_\_\_\_  
 At the end of this month  we have lived  we have been living  we'll have been living 88 \_\_\_\_  
 together for a year and a half. It's the first time  I live  I'm living  I've lived with anybody before, but 89 \_\_\_\_  
 I should guess  I might have guessed  I'd have guessed what would happen. I've fallen in love 90 \_\_\_\_  
 with her and now she's been offered another job 200 miles away and is going to move. I don't know what  
 to  
 Do. Please give me some advice.  
 Yours in shy desperation,  
 Steve

Look at the following examples of question tags in English. The correct form of the tag is ticked.

- A He's getting the 9.15 train,  isn't he  hasn't he  wasn't he?  
 B She works in a library,  isn't she  doesn't she  doesn't he?  
 C Tom didn't tell you,  hasn't he  didn't he  did he?  
 D Someone's forgotten to switch off the gas,  didn't one  didn't they  haven't they?

Now tick the correct question tag in the following 10 items:

- 91 Steve's off to China,  has he  hasn't he  isn't he? 91 \_\_\_\_  
 92 It'll be a year before we see him again,  won't it  won't we  shan't it? 92 \_\_\_\_  
 93 I believe he's given up smoking,  isn't he  don't I  hasn't he? 93 \_\_\_\_  
 94 I'm next on the list to go out there,  am not I  are I  aren't I? 94 \_\_\_\_  
 95 No doubt you'd rather he didn't stay abroad too long,  shouldn't you  wouldn't you  hadn't you? 95 \_\_\_\_  
 96 He's rarely been away for this long before,  is he  hasn't he  has he? 96 \_\_\_\_

- 97 So you think he'll be back before November,  shall he  will he  do you? 97 \_\_\_\_  
 98 Nobody's disagreed with the latest proposals,  did he  has he  have they? 98 \_\_\_\_  
 99 We'd better not delay reading this any longer,  should we  did we  had we? 99 \_\_\_\_  
 100 Now's hardly the time to tell me you didn't need a test at all,  did you  is it  isn't it? 100 \_\_\_\_