Using Google Podcasts to Develop Prospective EFL Teachers’ Listening Comprehension Skills

Marwa A. Naeem
Faculty of Education, Kafr El–Sheikh University, Kafr El–Sheikh, 2021

Abstract:
Attempting to investigate the impact of using Google Podcasts in developing prospective EFL teachers’ listening comprehension skills, this study was conducted during the first semester of the academic year 2021/2022. One hundred and seventy participants enrolled in the Fourth Year, English Major listened to a Google podcast based on Charles Dickens’ *Hard Times*. Methodologically, the one–group pretest–posttest design was adopted. A valid and reliable listening comprehension test consisting of fifty items was administered to the participants. The experiment lasted for seven weeks in which participants received extensive listening sessions. After the posttest, data were statistically analyzed by the Paired Sample *t*-Test. Findings revealed that podcasts developed prospective EFL teachers’ recognition of literal meaning, purpose determination, reasoning analysis and conclusion drawing. Two listening comprehension skills were not significantly developed; namely: inference making and evidence finding. Thorough implications were provided and recommendations for further studies were suggested.

Keywords: Google Podcasts, Listening Comprehension, Aural Skill Development.
استخدام مدونات جوجل الصوتية لتنمية مهارات الفهم السمعي لدى الطلاب المعلمين دارسي اللغة الإنجليزية كلغة أجنبية

مروه أحمد رفعت نعيم
أستاذ مناهج المغة الإنجميزية وطريق تدريسها المساعdirty الكلية التربية، جامعة كفر الشيخ

مُستخلص

في محاولة لاستقصاء تأثير استخدام مدونات جوجل الصوتية في تطوير مهارات الفهم السمعي لدى الطلاب المعلمين دارسي اللغة الإنجليزية كلغة أجنبية، أجريت هذه الدراسة خلال الفصل الدراسي الأول من العام الجامعي 2021/2022. حيث استمع مائتا وسبعون مشاركًا إلى مدونة صوتية تستند إلى رواية "Hard Times" لـ تشارلز ديكنز. و من الناحية المنهجية، فقد اعتمدت الدراسة الحالية على التصميم البحثي القائم على الاختبار القلبي والبعدي لمجموعة واحدة. تم تطبيق اختباراً للفهم السمعي مكون من خمسين بندًا على المشاركين بعد التحقق من صدقته و ثباته. استمرت التجربة لمدة سبعة أسابيع حيث تلقى المشاركون جلسات استماع مكثفة. و عقب الاختبار البعدي، تم تحليل البيانات إحصائيًا بواسطة اختبار ت للعينة المرتبطة. توصلت النتائج إلى أن مدونات جوجل الصوتية نمت الفهم السمعي لدى الطلاب المعلمين دارسي اللغة الإنجليزية كلغة أجنبية و ذلك في مهارات المعنى الحرفي، وتحديد الغرض، وتحقيق الهدف. و استخلاص النتائج، إن النتائج أظهرت أيضاً أنه لم يتم تنمية مهارات من مهارات الفهم السمعي و هي مهارة الاستدلال ون مهارة التوصيل إلى الأدلة. و وفقاً لنتائج الدراسة، تم تقديم مقترحات و توصيات لمزيد من الدراسات.

الكلمات المفتاحية: مدونات جوجل الصوتية، الفهم السمعي، تنمية المهارات السمعية
Introduction:

Although Egyptian prospective EFL teachers study both explicit and implicit academic courses that are supposed to develop their listening skill, they do not often promote a satisfactory level of listening comprehension. Lack of practice is believed to be a major cause for such a problem. Egyptian prospective EFL teachers do not have neither intensive nor extensive exposure to listening materials. Their encounter with listening is confined to limited doses to achieve certain course tasks and pass the official exam. A suggested solution to this problem is giving prospective EFL teachers a chance to listen extensively to native English through Google Podcasts.

In general, the word podcast is the result of combining two words: *pod* that refers to Apple’s portable audio player, *iPod*, and *cast* which is a contraction of *broadcast*. To define it comprehensively, a podcast is a voice recording that deals with an idea. It has relatively short time that does not usually exceed thirty minutes. A podcast is often a part of a series that is stored on a website for easy access with the possibility of constant update. Although attention is directed to podcasting recently, it is not a new idea. A podcast is a radio show that can be downloaded to a computer or listened to on the Internet (Khadimally, 2018; Simonson, Zvacek & Smaldino, 2019).

Podcasts, as a type of Mobile-Assisted Language Learning (MALL), were proposed as an effective technological tool that can aid EFL learners develop their listening and speaking skills. Mobile phones, the primary medium for listening to podcasts, have become a significant
part of students' lives. Interestingly, it was reported that college students spend an average of nine hours on their smartphones daily (Kang, 2016).

Tidal (2021) illustrated that there are various types of podcasts. A solo style is the simplest type of podcast in which a single speaker discusses a topic and there is no need for complicated equipment or other speakers. In addition, an interview is a popular type of podcast through which a host speaker asks a series of questions to a number of guests. Besides, there are panel podcasts which encompass three or more speakers discussing a topic. On a technical basis, a panel podcast is more complicated than a solo one. However, panel podcasts are more informative and vivid. On another track, multiple-host or conversational podcasts involve two or more people in a debate about a certain issue. The difference between a panel podcast and a conversational one is that the speakers in the latter adopt contradictory viewpoints and each defends what he / she believes. Aiming at teaching the listener something or developing his / her skills, an instructional podcast usually requires careful planning and scripting. In most cases, it is accompanied with a supporting material like a PDF or a website. The last type is the narrative podcast in which a story is told whose nature and length determine the number of episodes.

Podcasts have many pros. Manglik (2020) pointed out that podcasts allow listeners to involve in multitasks while listening, i.e., driving, cooking or working out. Furthermore, podcasts are creative in the sense that they paint pictures in listeners’ minds and set their
imagination free. In terms of access and speed, audio material has a more straightforward consumption than videos. Technically, podcasts do not consume much battery and data of mobiles as they do not require listeners to glue to their mobile screens all the time. To add, sound effects can be employed in podcasts to increase listeners’ imaginative power, provide them with more details and keep them away from boredom.

However, some educators mentioned a number of challenges of using podcasts in educational contexts. Podcasts are believed to be unable to hold students’ attention. They warned against the potential distraction podcasts may cause specially when they are not well-planned. Moreover, podcasts may lead to teachers’ complacency in the sense that they replace real teaching with podcasts. There is also the fear that students may abandon reading books for the sake of podcasts (Zitouni et al., 2021).

In respect of listening comprehension, it is simply defined as a language learner’s ability to “understand native speech at normal speed in unstructured situations” (Turel, 2021: xxii). It has certain aspects that differ in their learning than others. The most problematic aspects that face a foreign language learner in listening comprehension are hearing the sounds, understanding intonation and stress, coping with redundancy and noise, fatigue, understanding different accents and using visual and environmental clues (Ur, 1984).

Hamada (2016) identified eight key listening comprehension skills; namely: recognizing literal meaning, understanding vocabulary, making
inferences, identifying main idea, determining purpose, analyzing reasoning, drawing conclusions and finding evidence. He recommended that learners’ training should start with the phoneme level. Phoneme perception is believed to be a prerequisite to all listening comprehension skills.

A worthwhile remark was provided by Poelmans (2003) about the effective methods for teaching listening comprehension. She explained that the methods used to train learners on listening comprehension should encompass a sort of overlapping between the training materials and evaluation ones. It was also asserted that acquiring listening comprehension skills in a foreign language differs from that in a learner’s native language.

The recent studies that investigated the relationship between listening to podcasts and developing listening comprehension are somehow limited. Among such studies was that led by Mirza (2021) in which he investigated the podcasting impact on Iranian senior undergraduate TEFL students’ listening comprehension enhancement and motivation increase. An experimental research design that involved two groups: an experimental group and a control one was used. Sixty undergraduates took part in the study. Statistically, the Kolmogorov–Smirnov test of normality was adopted to confirm group homogeneity. Moreover, $t$-test for independent samples was used to analyze the data of the pre and posttests. It was found out that podcasts had a positive effect on improving TEFL undergraduates’ listening comprehension. In addition, participants’ motivation witnessed a notable increase towards
using podcasts as a learning tool according to the results of a motivation questionnaire.

On the same track, Gonulal (2020) explored how extensive listening through podcasts and vodcasts improved FL listening skills. The sample for that study consisted of forty EFL college students and the experiment lasted for one year. Listening log assignments, listening progress tests, proficiency tests and a listening log questionnaire represented the tools of the study. Both quantitative and qualitative analyses of data were provided. Findings revealed that students spent approximately one hour per week on extensive listening with podcasts or vodcasts outside the classroom. Students less commonly preferred vodcasts and they were more attracted to podcasts. In addition, students’ achievement was notably increased. Affectively, students reported that extensive listening practice with podcasts and vodcasts was highly effective. It helped them improve their pronunciation abilities and knowledge of words and phrases. Nonetheless, there was an occasional frustration with the pace of speech in podcasts and vodcasts. Generally, it was concluded that digital technologies can boost language skills; specially listening.

Abdulrahman, Basalama and Widodo (2018) scrutinized EFL learners’ listening comprehension after using podcasts. The participants in their study were sixty Indonesian high school students divided equally into an experimental group and a control one. Cluster random sampling was employed and quasi–experimental design with the post–test only was adopted. Moreover, the researchers used a survey questionnaire
for the experimental group so as to elicit their impression on studying via podcasts. Statistically, one way ANOVA was utilized. Results proved the positive impact of podcasts on listening comprehension. Furthermore, the questionnaire asserted EFL students’ positive attitude towards the use of podcast in a listening classroom. It was reported that podcasts offered authentic materials, interesting activities including listening exercises and meaningful tasks for students, therefore, they felt more motivated to learn English. It was concluded that podcasts are recommended to teach EFL listening.

Evaluating the incorporation of self-regulated and awareness strategies with podcast activities and their effectiveness in developing ESL listening and speaking, Kang (2016) had a worthy dissertation that came up with enlightening results. The sample for the study included twenty intermediate ESL learners enrolled in an intensive English program at a southwestern American university. The experiment lasted for twelve weeks, and quasi-experimental research design was utilized. The study encompassed two independent variables; namely: time (within-subjects factor) and group (between-subjects factor). As for dependent variables, they were six: listening skills, speaking skills, attitudes, self-regulation, self-awareness, and use of podcasts. The researcher employed mixed-method comparisons so as to analyze two groups’ performances on the pre-, post- and delayed post-measures. Furthermore, there was an analysis of participants’ listening logs and speaking performances. Findings showed that both groups improved significantly on listening and speaking achievement tests. In addition,
the experimental group made a slightly better progress on listening comprehension. However, the experimental group produced slightly slower and more comprehensible speech. Both groups had positive attitudes towards podcasts. It was concluded that podcasts are crucial for today’s ESL classrooms.

Aiming at scrutinizing the effect of mobile learning on EFL learners’ listening anxiety and listening comprehension, Rahimi and Soleymani (2015) selected fifty intermediate students and sampled them equally in two groups: an experimental group and a control one. Two instruments were employed: a foreign language listening anxiety questionnaire to identify learners’ listening anxiety level, and the Key English Test (KET) to assess their listening comprehension. The experiment took a semester in which the experimental group listened to podcasts through their mobile phones or portable digital media players. Simultaneously, the control group utilized their desktop computers to do their listening activities. Findings revealed that listening anxiety of the podcasting group reduced significantly after the experiment. In addition, a significant difference on listening comprehension in favor of the podcasting group was found.

In Japan, Cross (2014) made a case study in which he explored the effect of BBC’s podcasts on promoting an EFL learner’s autonomous listening. Metacognitive instruction was utilized, and the learner was introduced to metatextual skills and activities. The experiment took nine weeks; the learner listened to journal entries and interviews. Regular guidance and continuous feedback were provided by
the researcher. Findings showed that extensive listening to podcasts remarkably enhanced the learner’s autonomous listening and sophisticated metatextual skills.

In the Saudi context, Al Qasim and Al Fadda (2013) referred to the scarcity of research works that handled employing technology in developing FL listening. Consequently, they made a quantitative study in which they investigated the impact of podcasts on listening comprehension. Forty-six female higher education Saudi EFL students participated in the study. They were randomly chosen and divided into two groups: a control group and an experimental one as an experimental research design was used. Treatment took six weeks, then data were statistically analyzed by \( t \)-test. Results pointed out that podcasts prominently enhanced college students’ listening comprehension.

The above review of related studies referred to a positive relationship between employing podcasts and listening development. However, there is a scarcity of studying this relationship in the Egyptian educational environment. Consequently, the current study attempted to fill this gap in the body of EFL literature. Theoretically, the researcher utilized blended learning to carry out her experiment. Participants listened to the podcasts online outside the classroom. Weekly face-to-face discussions were held about what they had listened to. While participants maintained their learning pace, the researcher was just a facilitator for the learning process.
To elaborate the problem of the study, the following main question was raised: What is the effect of using Google Podcasts on developing prospective EFL teachers’ listening comprehension skills?

Some sub-questions were derived from the main question as follows:

1 – What is the effect of using Google Podcasts on developing prospective EFL teachers’ recognition of literal meaning?
2 – What is the effect of using Google Podcasts on developing prospective EFL teachers’ inference making?
3 – What is the effect of using Google Podcasts on developing prospective EFL teachers’ purpose determination?
4 – What is the effect of using Google Podcasts on developing prospective EFL teachers’ reasoning analysis?
5 – What is the effect of using Google Podcasts on developing prospective EFL teachers’ conclusion drawing?
6 – What is the effect of using Google Podcasts on developing prospective EFL teachers’ evidence finding?
7 – What is the effect of using Google Podcasts on developing prospective EFL teachers’ listening comprehension skills as a whole?

Answering these questions entailed turning them into the following null hypotheses that could be statistically tested:

1 – There is no significant difference between the pretest and the posttest mean scores on the recognition of literal meaning at the (0.01) level of significance.
2 – There is no significant difference between the pretest and the posttest mean scores on inference making at the (0.01) level of significance.

3 – There is no significant difference between the pretest and the posttest mean scores on purpose determination at the (0.01) level of significance.

4 – There is no significant difference between the pretest and the posttest mean scores on reasoning analysis at the (0.01) level of significance.

5 – There is no significant difference between the pretest and the posttest mean scores on conclusion drawing at the (0.01) level of significance.

6 – There is no significant difference between the pretest and the posttest mean scores on evidence finding at the (0.01) level of significance.

7 – There is no significant difference between the pretest and the posttest mean scores on listening comprehension skills as a whole at the (0.01) level of significance.

**Method**

**Participants**

One hundred and seventy participants took part in the current study. They were all students enrolled in the Fourth Year, English Major at the Faculty of Education in Kafr El–Sheikh University. Participants listened to the podcasts during the first semester in the academic year 2021/ 2022. Their ages range from 21 to 22 years old. In general,
participants have been studying English for sixteen years. In the last three years of their academic study, they received specialized listening courses developed by their lecturers according to the Faculty of Education, Kafr El–Sheikh University Bylaw through which they were supposed to develop their listening comprehension skills. Technically, the researcher adopted a purposive sampling as it would be more suitable for the nature of the experiment in which an extensive listening material is given to all participants.

**Instruments**

For data collection, the researcher constructed a listening comprehension skill test on Google Forms. The test was divided into six dimensions each of which measured a listening comprehension skill under study. A test tryout on a sample of thirty prospective EFL teachers was made to establish its validity and reliability. Concerning validity, the researcher submitted the test to seven educational experts in the field of TEFL, language testing and psychological measurement. They made some modifications of the form and number of the test items that were taken into account in the final form of the test. As for reliability, the split-half method was adopted. The Spearman–Brown prophecy formula (Ary, Jacobs & Sorensen, 2010: 244) rendered an $r = .89$. Consequently, the test reliability was confirmed. In regard to the optimum time limit for the test, the researcher calculated the means of participants’ spent time on the test tryout. Therefore, the test optimum time was thirty minutes including instructions and answering participants’ questions before administration.
In its final form, the total number of test items was fifty MCQs. The researcher made use of Spark Notes (2021) quizzes at Hard Times Study Guide in test construction. The following table illustrates the distribution of the test items to the listening comprehension skills:

<table>
<thead>
<tr>
<th>No.</th>
<th>Listening Comprehension Skill</th>
<th>Items that Measure the Skill</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Recognition of Literal Meaning</td>
<td>1 – 34</td>
</tr>
<tr>
<td>2</td>
<td>Inference Making</td>
<td>35 – 39</td>
</tr>
<tr>
<td>3</td>
<td>Purpose Determination</td>
<td>40 – 42</td>
</tr>
<tr>
<td>4</td>
<td>Reasoning Analysis</td>
<td>43 – 45</td>
</tr>
<tr>
<td>5</td>
<td>Conclusion Drawing</td>
<td>46 – 48</td>
</tr>
<tr>
<td>6</td>
<td>Evidence Finding</td>
<td>49 – 50</td>
</tr>
</tbody>
</table>

The bulk number of test items was devoted to the recognition of literal meaning. Then, a balanced distribution of items to other skills followed. The logic behind item distribution was twofold. The Spark Notes quiz database imposed concentrating on the recognition of literal meaning. Moreover, this skill is usually stressed in literary material.

**Material**

A ready–made Google podcast based on Charles Dickens’ Hard Times was used as an extensive listening material in this study (Loyal Books, 2021). There was a narrator and several other readers who played the different roles of the novel characters. The podcast total listening time was twelve hours divided into 37 episodes.

**Procedure**

The current study adopted the one–group pretest–posttest design. Three steps were followed. First, a pretest that measured listening
comprehension skills was administered. Second, participants listened to Charles Dickens’ *Hard Times* scheduled podcasts for seven weeks in a row (three sessions per week). Weekly discussions were held, and the researcher provided counsels whenever requested. Third, participants were post tested on listening comprehension skills.

*Table (2): Listening Task Schedule*

<table>
<thead>
<tr>
<th>Session</th>
<th>Content</th>
<th>Duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Episodes: 1 – 3</td>
<td>30 mins</td>
</tr>
<tr>
<td>2</td>
<td>Episodes: 4 – 5</td>
<td>33 mins</td>
</tr>
<tr>
<td>3</td>
<td>Episodes: 6</td>
<td>33 mins</td>
</tr>
<tr>
<td>4</td>
<td>Episodes: 7 – 8</td>
<td>31 mins</td>
</tr>
<tr>
<td>5</td>
<td>Episodes: 9 – 10</td>
<td>31 mins</td>
</tr>
<tr>
<td>6</td>
<td>Episodes: 11 – 12</td>
<td>28 mins</td>
</tr>
<tr>
<td>7</td>
<td>Episodes: 13 – 14</td>
<td>33 mins</td>
</tr>
<tr>
<td>8</td>
<td>Episodes: 15 – 16</td>
<td>35 mins</td>
</tr>
<tr>
<td>9</td>
<td>Episodes: 17</td>
<td>34 mins</td>
</tr>
<tr>
<td>10</td>
<td>Episodes: 18 – 19</td>
<td>34 mins</td>
</tr>
<tr>
<td>11</td>
<td>Episodes: 20 – 21</td>
<td>38 mins</td>
</tr>
<tr>
<td>12</td>
<td>Episodes: 22</td>
<td>31 mins</td>
</tr>
<tr>
<td>13</td>
<td>Episodes: 23</td>
<td>31 mins</td>
</tr>
<tr>
<td>14</td>
<td>Episodes: 24</td>
<td>32 mins</td>
</tr>
<tr>
<td>15</td>
<td>Episodes: 25 – 26</td>
<td>33 mins</td>
</tr>
</tbody>
</table>
Results

In terms of statistics, the Paired Sample $t$-Test was the adequate technique to analyze data according to the nature of the study design. Using the Paired Sample $t$-Test formula (Rajaretnam, 2015: 137), the study null hypotheses were tested.

**Hypothesis One**: There is no significant difference between the pretest and the posttest mean scores on the recognition of literal meaning at the (0.01) level of significance.

To test this hypothesis, the Paired Sample $t$-Test formula was adopted basing on the data rendered by the pre- and posttests. The following table sums up the attained results:

*Table (3): Paired Sample $t$-Test of the Recognition of Literal Meaning*

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>Episodes: 27 – 28</td>
<td>33 mins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>Episodes: 29 – 30</td>
<td>39 mins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>Episodes: 31 – 32</td>
<td>47 mins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Episodes: 33 – 34</td>
<td>50 mins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>Episodes: 35</td>
<td>30 mins</td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Episodes: 36 – 37</td>
<td>30 mins</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Total Listening Time</strong></td>
<td></td>
<td></td>
<td></td>
<td>12 Hrs</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>D</th>
<th>$s$</th>
<th>$n$</th>
<th>df</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>15.818</td>
<td>72.890</td>
<td>170</td>
<td>169</td>
<td>9.929*</td>
</tr>
</tbody>
</table>

(* ) Significant at the (0.01) level
Results indicated that the obtained $t$ was 9.929. This obtained value was greater than the critical one which equals 2.626 when the degree of freedom is 169 and the hypothesis is null. Consequently, there is a significant difference between the pretest and the posttest mean scores on the recognition of literal meaning at the (0.01) level of significance; therefore, Hypothesis One is rejected.

**Hypothesis Two**: There is no significant difference between the pretest and the posttest mean scores on inference making at the (0.01) level of significance.

The Paired Sample $t$-Test was used to test this hypothesis. The required test data are presented in Table (4):

*Table (4): Paired Sample $t$-Test of Inference Making*

<table>
<thead>
<tr>
<th>D</th>
<th>s</th>
<th>n</th>
<th>df</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.571</td>
<td>10.661</td>
<td>170</td>
<td>169</td>
<td>0.004</td>
</tr>
</tbody>
</table>

The calculated $t$ value was 0.004. Compared to the critical $t$, which equals 2.626 when the degree of freedom is 169 and the test is two-tailed, the second null hypothesis was accepted. It was concluded that there was no significant difference between the pretest and the posttest mean scores on inference making at the (0.01) level of significance.

**Hypothesis Three**: There is no significant difference between the pretest and the posttest mean scores on purpose determination at the (0.01) level of significance.

Testing this hypothesis entailed utilizing the Paired Sample $t$-Test. The test data were collected and summed up in Table (5):

*Table (5): Paired Sample $t$-Test of Purpose Determination*

<table>
<thead>
<tr>
<th>D</th>
<th>s</th>
<th>n</th>
<th>df</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.388</td>
<td>4.560</td>
<td>170</td>
<td>169</td>
<td>3.487*</td>
</tr>
</tbody>
</table>

(*) Significant at the (0.01) level
Since the critical $t$ equals 2.626 at a degree of freedom of 169 and the test is two-tailed and the obtained $t$ equaled 3.487, the third null hypothesis is rejected. Thus, there is a significant difference between the pretest and the posttest mean scores on purpose determination at the (0.01) level of significance.

**Hypothesis Four**: There is no significant difference between the pretest and the posttest mean scores on reasoning analysis at the (0.01) level of significance.

Like the previous three hypotheses, Hypothesis Four was tested by the Paired Sample $t$-Test. Table (6) displays the collected data and the reached results:

*Table (6): Paired Sample $t$-Test of Reasoning Analysis*

<table>
<thead>
<tr>
<th>D</th>
<th>$s$</th>
<th>n</th>
<th>df</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.059</td>
<td>3.779</td>
<td>170</td>
<td>169</td>
<td>2.917*</td>
</tr>
</tbody>
</table>

(*) Significant at the (0.01) level

According to $t$ value statistical table, the critical $t$ equals 2.626 when the degree of freedom is 169 and the test is two-tailed. The obtained $t$ that equaled 2.917 was greater than the critical value. Therefore, it was concluded that there was a significant difference between the pretest and the posttest mean scores on reasoning analysis at the (0.01) level of significance. Hence, Hypothesis Four was rejected.

**Hypothesis Five**: There is no significant difference between the pretest and the posttest mean scores on conclusion drawing at the (0.01) level of significance.
Statistically, the hypothesis was tested by the Paired Sample $t$-Test. The necessary data for testing this hypothesis are presented in Table (7):

**Table (7): Paired Sample $t$-Test of Conclusion Drawing**

<table>
<thead>
<tr>
<th>$D$</th>
<th>$s$</th>
<th>$n$</th>
<th>$df$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.724</td>
<td>4.868</td>
<td>170</td>
<td>169</td>
<td>4.184*</td>
</tr>
</tbody>
</table>

(*) Significant at the (0.01) level

Reviewing the statistical tables, the critical $t$ value that equals 2.626 was less than the obtained $t$ that equaled 4.184 when the degree of freedom is 169 and the test is two-tailed. Accordingly, the fifth null hypothesis was rejected and the result was that there was a significant difference between the pretest and the posttest mean scores on conclusion drawing at the (0.01) level of significance.

**Hypothesis Six**: There is no significant difference between the pretest and the posttest mean scores on evidence finding at the (0.01) level of significance.

The Paired Sample $t$-Test was used to test this hypothesis. The fundamental data for testing this hypothesis are illustrated in Table (8):

**Table (8): Paired Sample $t$-Test of Evidence Finding**

<table>
<thead>
<tr>
<th>$D$</th>
<th>$s$</th>
<th>$n$</th>
<th>$df$</th>
<th>$t$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.706</td>
<td>3.199</td>
<td>170</td>
<td>169</td>
<td>2.114</td>
</tr>
</tbody>
</table>

The calculated $t$ ($= 2.114$) was less than the critical one ($= 2.626$) when the degree of freedom is 169 and the test is two-tailed. In this case, the null hypothesis was accepted. It was concluded that there was
no significant difference between the pretest and the posttest mean scores on evidence finding at the (0.01) level of significance.

**Hypothesis Seven**

**There is no significant difference between the pretest and the posttest mean scores on listening comprehension skills as a whole at the (0.01) level of significance.**

This hypothesis was tested using the Paired Sample \( t \)-Test. Table (9) sums up the collected data and the attained results:

*Table (9): Paired Sample \( t \)-Test of Listening Comprehension Skills*

<table>
<thead>
<tr>
<th>( D )</th>
<th>( s )</th>
<th>( n )</th>
<th>( df )</th>
<th>( t )</th>
</tr>
</thead>
<tbody>
<tr>
<td>23.265</td>
<td>99.960</td>
<td>170</td>
<td>169</td>
<td>12.476*</td>
</tr>
</tbody>
</table>

(*): Significant at the (0.01) level

Reviewing the statistical tables, the critical \( t \) value that equals 2.626 was far much less than the obtained \( t \) that equaled 12.476 when the degree of freedom is 169 and the test is two-tailed. Hence, the seventh null hypothesis was rejected, and it was concluded that there was a significant difference between the pretest and the posttest mean scores on listening comprehension skills as a whole at the (0.01) level of significance.

**Discussion**

The majority of listening comprehension skills was notably developed by regular listening to podcasts. Specifically, EFL learners’ recognition of literal meaning obviously improved. This improvement can be attributed to the fact that podcasts are easy to retrieve several times. Consequently,
EFL learners could maintain a good deal of specific details and literal information retention. Such an explanation accords with its counterparts provided by Gonulal (2020); Abdi & Makiabadi (2019), and Shahid & Ali (2017).

Interestingly, EFL learners’ inference making did not seem to be improved by listening to podcasts. Inference making is quite associated with higher–order thinking skills. It may be worthy to note that Egyptian EFL learners are not accustomed to employ such skills in their learning. Neither EFL curricula nor exams devote adequate space to teach and test higher–order thinking skills. A plausible reason of this result was provided by Ifedayo, Ziden, and Ismail (2021). They reported that university staff’s as well as students’ low–level behavioural intention negatively impacted podcast acceptance. It was recommended to promote podcast users’ acceptance of podcasts first, then employ them to enhance linguistic skills.

Moreover, podcasts proved to be an enhancer of purpose determination skill. Asmara (2017) attributed such an enhancement to the authentic language used in the podcasts. She pointed out that language authenticity helps students focus on the purposes of what they listen to. To add, the Charles Dickens’ *Hard Times* Podcast was in a British accent. EFL learners remarked on that saying the British accent and its pace helped them focus on what they listened to and gave them adequate time to take notes.

Reasoning analysis was also developed through listening to podcasts. The researcher believed that the weekly face–to–face discussions were an important factor in developing EFL learners’ reasoning analysis. Such discussions opened new horizons about the novel themes, characters and
other dimensions. Exchanging ideas was beneficial to develop the ability to analyze the content of the podcasts. This result is associated with Syahabuddin and Rizqa’s (2021) conclusion in which they highlighted the importance of the face-to-face component in learning via podcasts.

EFL learners also developed their conclusion drawing skill. This listening comprehension skill was based on and connected to other skills such as purpose determination and reasoning analysis. The researcher thought that all these three skills could be considered as one package. That is to say that developing one of these skills logically contributed to the development of another. This explanation lined with that provided by Albán (2021); Harahap (2020), and Alshaikhi & Madini (2016). Nevertheless, Lovett (2019) believed that nonfictional podcasts served many pragmatic purposes that positively affect listeners’ daily lives. Nonfictional podcasts could be used for prompting responses. They were also useful in composing texts and engaging in activities. In addition, nonfictional podcasts were regarded as a means for conveying knowledge. Fictional podcasts were not recommended for the previous purposes.

However, podcasts failed to develop EFL learners’ evidence finding skill. This result seemed odd and unexplainable. Logically, evidence finding is a higher-order thinking skill like purpose determination, reasoning analysis and conclusion drawing. While all these skills were significantly developed, it was bewildering that evidence finding was not. Literature and previous studies provided no adequate explanation either. On the contrary, many such studies as Gachago, Livingston and Ivala (2016); McKinney, Dyck and Luber (2009), and Lee and Chan (2005) concluded that podcasts were very effective in developing higher education
students’ listening comprehension skills including evidence finding. The researcher viewed that more studies should investigate this skill meticulously.

In general, podcasts had a considerable fruitful contribution to developing EFL learners’ listening comprehension skills. König (2021) asserted the importance of enthusiasm in podcasts. Enthusiastic podcasts – that were recorded with vivid rather than monotonous tones – were reported to be more attractive and interesting to learners. In the current case, Charles Dickens’ Hard Times Google podcast offered an enthusiastic material with a narrator’s and many characters’ speech. Therefore, listeners’ motivation and imagination were both stimulated. Probing into the positive and meaningful impact of listening to radio news on developing listening comprehension skills, Persichitte, Suparman and Spector (2017) concluded that language learners need to listen for at least twenty days in a row to get their listening comprehension skills improved.

During the experiment, the researcher noted the importance of the weekly face-to-face discussions in which several questions related to the finished listening sessions were raised. Students benefited from peer answers and exchanging experiences through discussion. The researcher provided insightful directions and hints as well.

Drawing attention to the difficulties embedded in the way of relying on podcasts in university courses, Ratledge, Dalporto and Lewy (2020) argued that a sort of adaptation had to be made before integrating such tools as podcasts into higher education. The fact that a large proportion of university students live in rural areas that hardly enjoy technological facilities had to be considered. For the desirable adaptation, they
suggested a number of procedures. Courses should be mobile–friendly, i.e., colleges ought to update course modules to be handled by smartphones. In addition, the government with collaboration with educational institutions should make plans to provide families and students in rural areas with adequate devices to learn online. Connecting rural–dweller university students to the internet is another governmental mission. Affectively, curricula should impart and reinforce the value of community from afar.

**Conclusion**

Google Podcasts can effectively develop prospective EFL teachers’ listening comprehension skills. Listening extensively to a novel and intensively for almost 35 minutes per day is probably fruitful for EFL college learners. Nonetheless, not all listening comprehension skills seem to be developed by listening to podcasts. Further studies are recommended to investigate the causes behind this selective development. The face–to–face guidance is believed to be necessary even if it is used at long intervals. Besides, it is advised to generalize the current study results with caution, specially when dealing with university students, specially those who live in rural areas.
References


Poelmans, P. (2003). Developing second-language listening comprehension: effects of training lower-order skills versus higher-order strategy. Amsterdam: LOT.


Spark Notes (2021). Hard times study guide. Retrieved from https://www.sparknotes.com/lit/hardtimes/#Quick%20Quizzes on Friday, October 1, 2021 at 5:00 p.m.


