

The Role of Video-Based Shadowing Practices in L2 Oral Proficiency Development

Sultan MICIK*

Filiz RIZAOĞLU**

ARTICLE INFO

Received: 07.09.2023
Revised form: 30.01.2024
Accepted: 29.02.2024
Doi:10.31464/jlere.1356763

Keywords:

Shadowing
Speech features
Intonation
Second language learning
Pronunciation

ABSTRACT

This mixed-methods study investigated the effect of shadowing practices on oral proficiency and pronunciation, specifically in terms of comprehensibility, pronunciation, intonation, and speech rate in second language learners. To this end, intact freshmen classes in an English language teaching program were assigned as experimental and control groups. The experimental group carried out 11 shadowing tasks bi-weekly in two academic terms and submitted their shadowing recordings to the learning management system to receive feedback from their instructor. After the intervention, oral rating forms, an activity evaluation survey, and a semi-structured interview were used to investigate the outcomes. The pre- and post-treatment oral recordings of the participants were rated online through a 7-point Likert scale by native speakers of English. The ratings of the experimental and control groups were compared through ANOVAs. The descriptive statistics revealed that the experimental group demonstrated a relatively higher but small increase in the oral rating components (comprehensibility, intonation, speech rate) except for the pronunciation of individual sounds. However, the pre- and post-test differences of both groups were not found to be statistically significant in the ANOVA analysis. The 5-point Likert scale activity evaluation questionnaire and the content analysis of the interviews showed that the experimental group participants thought shadowing improves pronunciation and overall, participants had positive attitudes towards the shadowing practice.

Acknowledgments

This study was supported by the Scientific Research Coordination Unit (BAP) of Pamukkale University under the Master Thesis project numbered 2019EĞBE007.

Statement of Publication Ethics

The ethics committee approval has been obtained for the current study: Pamukkale University Social Sciences Ethics Board, Date: 10.04.2019, Decision Number: 68282350/2018/G04

Authors' Contribution Rate

Both authors have contributed equally to the inception, data collection, analysis, and writing-up of the study.

Conflict of Interest

Micik, S., & Rızaoğlu, F. (2024). The role of video-based shadowing practices in L2 oral proficiency development. *Journal of Language Education and Research, 10(1)*, 166-190.

* English language teacher, ORCID ID: <https://orcid.org/0000-0001-5172-5341>, Şuhut Anadolu Lisesi, sultaanugur@gmail.com

** (Corresponding Author) Asst. Prof., ORCID ID: <https://orcid.org/0000-0003-2442-592X>, Pamukkale University, Foreign Language Education Department, frizaoglu@pau.edu.tr

Introduction

In foreign language learning contexts, learners have limited opportunities to be exposed to speech models and to improve their speech features. With the widespread availability of online materials and tools and increasing emphasis on oral skills instruction, novel or updated forms of oral practice have been introduced. One of these is the shadowing technique, whereby the learner listens to a model speaker and repeats the speech as closely as possible, with only a very slight delay (Hamada, 2017). Even though the emergence of shadowing dates back to the 1950s as a psycholinguistic experimental technique (Cherry, 1953), the concept has been transferred to the field of simultaneous interpreter training and more recently to second language (L2) teaching (Foote & McDonough, 2017).

Although it resembles the repetition technique of the Audio-Lingual Method (ALM), shadowing has different properties. The repetition in shadowing is simultaneous with the model speaker, whereas the ALM repetition drills include some pauses between the model speaker and the act of repetition. The lack of a time interval and the higher speed of naturalistic speech in shadowing directs the listener to focus on sound recognition (Sumiyoshi, 2019, p. 6), whereas in the ALM style repetition, due to its offline mode, the learners have more cognitive resources to focus on both the structural/grammatical and semantic aspects of the sentences while repeating them (Kadota, 2007, pp. 29-31, as cited in Hamada, 2016b).

The cognitive underpinnings of the shadowing task are usually explained on the basis of Baddeley's multi-component working memory model (Baddeley, 2003). To elaborate, working memory has four components: a primary attention control system called the central executive, the phonological loop devoted to temporary storage of sound information, the audiovisual sketchpad devoted to processing visual information and an episodic buffer, an interface between the components and the long-term memory. It is proposed that the phonological loop has critical functions that are important for language learning: phonological short-term storage, and subvocal rehearsal before transferring information to long-term memory (Baddeley et al., 1998). Shadowing is argued to activate the subvocal rehearsal function by encouraging overt rehearsal of the heard speech (Kadota, 2007, as cited in Sumiyoshi, 2019). A recent shadowing study lends support to this benefit based on evidence from brain imaging. The study reported improvement in working memory, changes in brain structure, and activity in areas associated with the phonological loop after a one-month intensive shadowing and read-aloud training, (Takeuchi et al., 2021). This improvement in subvocal rehearsal and the phonological loop also reflected itself in the improvements in reading speed and time in silent reading (Kadota et al., 2015, as cited in Kadota, 2021).

In addition, shadowing is considered to be a bottom-up processing activity since it is not possible to focus on top-down features of language in a limited time. The need to focus on specific aspects of speech features in shadowing is argued to activate bottom-up processing, which in turn improves auditory skills in the L2 (Hamada, 2016a, 2017, p. 20). The automatization in decoding words in L2 speech is expected to lower the cognitive

processing load in working memory and gradually leave more resources to the comprehension of content and the recognition of speech features, such as pronunciation and stress (Hamada, 2019, p.5).

Due to the listening component in shadowing, the initial L2 studies traced its potential for enhancing auditory skills (Bovee & Stewart, 2009; Foote & McDonough, 2017; Hsieh et al., 2013; Lin, 2009; Martinsen et al., 2017; Mori, 2011). In recent years, several experimental studies, mainly in Japan, revealed that weekly shadowing practices improve L2 learners' listening comprehension (Hamada, 2016a, 2017; Lin, 2009; Saito et al., 2011; Sumiyoshi, 2019).

Literature Review

There is a growing number of studies that investigate shadowing practices related to L2 pronunciation and other speech features. Most shadowing studies have employed pre-test/post-test designs with two to ten weeks of shadowing practices carried out in university settings in far eastern countries, the US, and Canada (Bovee & Stewart, 2009; Foote & McDonough, 2017; Hsieh et al., 2013; Lin, 2009; Martinsen et al., 2017; Mori, 2011; Teeter, 2017). The weekly shadowing practices may be classroom- or assignment-based and are tracked through participants' self-recordings. The more recent studies incorporate mobile or computer-based training instruments to track the participants. However, most shadowing studies lack a control or comparison group, probably due to the length of the training involved. Similarly, the participants are usually trained in intact L2 classes at universities since it is complicated to maintain participant retention for controlled practices in longitudinal studies. Even in intact classes, not all students complete all of the tasks required for adequate practice. For the same reason, the sample sizes are usually small, limited to 15-20 participants. It is also difficult to find raters who can devote hours to evaluate the participants' pronunciation samples, which might be another reason for the small sample sizes. Some of the shadowing studies have an additional qualitative component, usually a survey and/or interview, which inquire into the participants' first-hand shadowing experiences (Bovee & Stewart, 2009; Foote & McDonough, 2017; Martinsen et al., 2017).

An example of a common shadowing design can be found in an exploratory study, in which a 10-week shadowing intervention was applied to 400 university students in Japan (Bovee & Stewart, 2009). The participants were assigned audio files to shadow and were asked to send their shadowing recordings to their instructor for grading. The pre- and post-shadowing recordings of 21 randomly selected students were rated by eight native speakers of English in terms of overall pronunciation. The majority of the participants received higher ratings for pronunciation in the post-intervention recordings. The low-proficiency students benefited the most from the shadowing practice. A post-intervention participant survey also revealed that the majority had positive experiences and rated the activity as beneficial for improving their pronunciation and intonation. This early longitudinal study is one of the few that considered the proficiency level of the students. However, a major limitation to this study is the lack of a control group, since the improvement in learners' pronunciation might be due to external factors rather than

shadowing. Another limitation is the lack of inferential statistics, without which it is statistically hard to understand whether the improvement is due to chance.

Some studies have used the laboratory context to track learners' shadowing practices in a more controlled way. In a pre-experimental study, 20 Japanese-speaking university students were involved in a 10-week shadowing training in a computer-assisted language learning laboratory (Mori, 2011). The training involved 30-minute practices of shadowing approximately one-minute news videos with and without reading aloud the transcripts and recording the final read-aloud performance. The participants were asked to read aloud a specific passage before and after the training and their performance was analyzed in terms of rhythm, intonation, and stress in an acoustic analysis program. The sound wave analyses indicated improvement in stress, intonation, and rhythm patterns. The use of acoustic analysis provides a more objective form of progress measurement in this study. On the other hand, many shadowing researchers rely on human raters because they represent more naturalistic perceptions of speech.

Shadowing has also been used in the US for developing teaching assistants' oral skills in the L2, which is expected to enhance effective teaching. In a case study, five Chinese graduate students practiced shadowing for two weeks (Mishima & Cheng, 2017). The participants aimed to improve their intelligibility, pronunciation, intonation, rhythm, and fluency to pass an oral proficiency test required for a teaching assistantship. They were asked to choose a TED talk to shadow the first three minutes of the speech and record their voiceover as a video clip using a web-based animation tool. Then they received peer and instructor feedback on their performance. Two certified speaking exam raters evaluated participants' progress by comparing the final shadowed speech samples with their final oral proficiency test scores. Although the ratings indicated that all participants improved their prosodic control and became more intelligible compared to their previous oral proficiency scores, only two of the participants demonstrated enough progress to pass the proficiency test. Also, in an online survey, participants reported that shadowing improved overall speaking skills, fluency, pronunciation, and rhythm. The complementary interview revealed that participants especially liked receiving feedback on their video clips. In this study, the participants chose the model speeches, which is beneficial in terms of learner autonomy; however, this can delimit the generalizability of the research findings because the complexity and the speed of the model speech are not consistent across participants.

As can be seen, most studies use recordings of shadowing or read-aloud performances to measure participants' progress. Several studies have also looked into the effects of shadowing on spontaneous speech. In a longitudinal shadowing study (Martinsen et al., 2017), the oral proficiency progress of 19 L2 learners of French in a US high school was measured via both a read-aloud and a picture description task. The learners received a semester-long training, performing video-based shadowing for five to ten minutes three times a week in the classroom, followed by video-based shadowing exercises with or without subtitles for 30 minutes in a laboratory. The ratings of pre- and post-training read-aloud tasks by three native speakers revealed that shadowing exercises improved students' general accent, intelligibility, sentence and word stress, and overall intonation significantly. However, pre- and post-training picture description task ratings did not differ

significantly, implying that shadowing may result in more significant changes in controlled speech than in spontaneous speech in a 10-week intervention.

In another extemporaneous speaking study, L2 English learners' shadowing ability, comprehensibility, accentedness, and fluency were measured through a shadowing task and a picture dictation task before, during, and after shadowing training at a Canadian university (Foote & McDonough, 2017). Twenty-two participants shadowed short dialogues from TV series, recorded themselves, and sent the recordings to the researchers for eight weeks. Short segments of the picture dictation recordings were rated by 22 native English speakers. The participants were found to improve in terms of comprehensibility, fluency, and imitation, but not in terms of accentedness. Additionally, the post-training interview showed that the participants had positive attitudes toward shadowing and found it useful for improving their pronunciation.

As seen, most of the shadowing studies have used mixed-methods design and have also explored participants' perspectives towards shadowing through surveys and interviews. Across various studies, the first-hand evaluations of shadowing practices have been mainly positive. In the regular surveys collected throughout a semester-long shadowing practice (Martinsen et al., 2017), students reported initial difficulty performing the task, however, the reported task complexity gradually decreased with regular practice. In the same study, ratings of task complexity, authenticity, and speed of speech in the shadowing videos varied among students, yet students overall perceived the task to be beneficial for pronunciation development (Martinsen et al., 2017). It is worth noting that satisfaction with the in-class shadowing activity was much lower than that of the lab-based activity due to the noisy environment. The chance to exercise at one's own pace was expressed as the main strength of the lab activity. The interesting cultural content of the videos also contributed to student motivation according to the survey results. Overall, students also expressed a preference for using subtitles while performing shadowing, which provides a multimodal input advantage. Similar results were obtained in a large-scale mobile-assisted shadowing training with 1000 Japanese freshmen who performed five shadowing tasks weekly for 12 weeks. Participants' linguistic self-confidence, interest in English, attitudes towards communicating in English, perceptions of English ability, and ideal L2 self were found to increase significantly (Teeter, 2017). Feedback is also regarded by participants as a positive practice to accompany shadowing (Mishima & Cheng, 2017; Sumiyoshi & Svetanant, 2017).

So far, only one speech-focused shadowing study has employed an experimental design with a control group. In this study, 14 Taiwanese university students were randomly assigned to control and experimental groups (Hsieh et al., 2013). The control group was assigned traditional repetition-based pronunciation activities for improving English through a computer program. The experimental group, however, received an 8-hour shadowing training to improve their English pronunciation. Before and after the training, both groups were given four texts to read aloud, and their performances were recorded. The experimental group significantly outperformed the control group in terms of pronunciation, fluency, and intonation as assessed by a computer-assisted pronunciation training program. It is argued that the need to remember the target word/sentence for a

longer time in classical repetition practice might prevent the learner from attending to the speech features in the material, which results in lower effectiveness when compared to shadowing. Although the addition of a control group is a strength of this study, studies with larger sample sizes are needed to generalize these results.

Research Aim and Research Questions

Based on this background, shadowing has been found to have the potential for improving L2 learners' auditory skills, however, more evidence is required to understand the effects of shadowing on the improvement of oral skills, more specifically segmental and suprasegmental speech features. A review of the literature pointed to a lack of long-term shadowing experiments involving a control group and Turkish learners of English. Existing studies were also limited in terms of the speech features they focused on. The present study was carried out to investigate to what extent the shadowing technique impacts Turkish EFL learners' speech features and to what extent the participants find shadowing practices beneficial for developing oral proficiency. The research questions were formulated as follows:

1. Following long-term shadowing training, to what extent does the read-aloud performance of L2 learners change in terms of comprehensibility, pronunciation, intonation and rhythm, and speech rate?
2. How do L2 learners evaluate the effectiveness of shadowing practice and its influences on their speech features?

Methodology

Research Design/Model

This study employed a mixed-methods design, as mixing both quantitative and qualitative approaches offers a more inclusive understanding of research questions (Creswell & Creswell, 2017). More specifically, the convergent parallel mixed methods design (Creswell & Creswell, 2017) was adopted. In the quantitative quasi-experimental part, data were collected via pre- and post-test oral recordings and a survey. In the qualitative part, a semi-structured interview was administered with volunteers from the experimental group. The use of both quantitative and qualitative methods aimed at compensating for each other's limitations. Besides, the pre-and post-test design with an experimental and a control group was employed to control for the effect of extraneous factors on the participants' speech features.

Publication Ethics

Before carrying out the study, ethics approval was granted by the Ethics Board of the researchers' university (Pamukkale University Social Sciences Ethics Board, 10.04.2019, No. 68282350/2018/G04).

Participants

The study was conducted in an English Language Teaching (ELT) Program at a state university in southwestern Turkey. Students were from different regions of Türkiye, with the majority residing in the Aegean Region. Freshman students ($n = 90$) enrolled in three intact sections of the speaking course for two terms comprised the sample. Two of the sections were assigned as the experimental group and one section was assigned as the control group. The inclusion criteria for the participants were to have completed at least 10 of the 11 assigned shadowing practices over two terms, to have completed the pre- and post-test recordings, and to be a late sequential Turkish-English bilingual. After the removal of participants who did not meet the criteria ($n = 34$) or had low sound quality recordings ($n = 4$), the experimental group comprised 28 participants (9 male). The control group comprised 24 (13 male) participants from another section. The mean age of the participants was 19, (*Range* = 18 – 25). The participants were at B1 (intermediate) English language proficiency level based on the Oxford Quick Placement Test (Allan, 2004). At the time of data collection, the participants had been learning English as a foreign language for about 10 years, yet the weekly hours and the quality of K12 instruction were limited and varied among the participants.

Instruments

At the outset, a language background form was given to the participants. The second instrument was the read-aloud test which served as the pre- and post-test. A paragraph-long diagnostic speech text from a pronunciation book for L2 learners (Dauer, 1993) was selected. The participants had not seen the paragraph before the pre-test. They were instructed to record themselves individually in a silent meeting room in the faculty while reading the paragraph aloud at a natural speed. They were left alone during the recording so that they felt comfortable and were not permitted to record themselves more than once.

The third instrument was an activity evaluation survey which was administered to the experimental group. Consisting of 40 Likert items and two open-ended items, this survey investigated the experimental group's attitudes toward the shadowing practices. Some items were adapted from the existing surveys (Bovee & Stewart, 2009; Foote & McDonough, 2017; Sumiyoshi & Svetanant, 2017). Two open-ended questions asked the respondents to indicate the positive and negative sides of shadowing practice. These items gave the participants the chance to indicate their perspectives in a more personalized way. The responses to the two items were analyzed through content analysis.

The fourth instrument was the semi-structured interview, which comprised 12 open-ended questions. The items were adapted from the study of Foote and McDonough (2017). The interviews were held face-to-face with volunteering 10 experimental group participants. The interview data were transcribed and analyzed through content analysis.

Data Collection

Before the intervention, the speaking classes taught by the same instructor were divided into experimental and control groups. A pre-test was applied to all participants after the scheduled speaking course hours. In the pre-test, the participants were told to read aloud a specific paragraph individually and record themselves via a voice recorder. This pre-test recording process of all the participants was completed in approximately two weeks.

Before the training, the instructor gave a shadowing tutorial to the experimental group in the classroom. In addition, the instructor introduced the basic speech features of English (intonation, rhythm, connected speech) in a two-hour speaking class. The written training materials and an instruction file were provided in the learning management system (See the Appendix for the instructions).

The experimental group was assigned a shadowing task bi-weekly, in which they were asked to watch an assigned video clip for gist and shadow the model speech at least twice without the subtitles. Then the participants were required to say the sentences by shadowing the speaker while keeping an eye on the subtitles. They were required to consider the prosodic features of the speech while shadowing and recording themselves. Finally, they were asked to review and re-record their final performance till they felt pleased with their recording to submit to the learning management system or the instructor's e-mail address. The completion time for each task was two weeks since the students had a busy schedule in which they had to complete assignments for nine other compulsory courses. The instructor gave feedback to each participant in the learning management system, based on their pronunciation errors, speech rate, connected speech features, and intonation. The participants were told to note their practice time and the number of recording trials for each task in their submission. Varying durations (20 minutes to 2 hours) were reported for these trials, which diminished as they mastered the procedure. The practice time decreased through the end of the semester when participants carried out the procedure with more ease. In the first term, the participants were asked to select 2-minute sections of videos, such as TED talks to shadow or the videos selected by the instructor and send their final practice in a video file. In the second term, the instructor provided some video excerpts from movies or TED talks for the participants to choose from and asked the participants to send their recordings in an audio file.

While performing the tasks was a course requirement for the experiment group, the control group participated in regular speaking classes for two terms. To provide equal opportunities for learners, after the intervention ended, the instructor did shadowing practice with the control group. After receiving 11 shadowing tasks over two terms, the experiment group completed a task evaluation survey. In addition, in the post-test, both the control and experimental group participants recorded their read-aloud performances of the same passage individually. Lastly, one of the researchers interviewed some of the experimental group participants.

Data Analysis

Four (3 females) native speakers of American English were the raters of this study. Three of them were English a second language (ESL) instructors and each had a Ph.D. in ELT and one of them was a retired teacher. They volunteered to contribute to this research. Two rating sessions were arranged on Google Forms. Each session comprised 52 randomly presented recordings of the pre- and post-performances of the participants. The raters went through a training component in each rating session and completed four demo ratings before the actual ratings to reinforce interrater reliability. They were requested to carry out the evaluation in a quiet place with a headset. Rating session 2 was delivered to the raters a few days after they completed rating session 1.

Each audio was evaluated through a 7-point Likert scale in terms of comprehensibility, pronunciation errors, intonation, and speech rate. The rating criteria were adapted from Saito et al. (2015) and Foote and McDonough (2017). The ICC (intra-class correlation) reliability of the ratings was .72 for comprehensibility, .65 for intonation, .61 for pronunciation, and .61 for speech rate. Based on Cicchetti (2001) and Koo & Li (2016), these values indicate moderate to good reliability.

The descriptive statistics of the pre- and post-test scores were calculated. Four 2 (group: control vs. experimental) x 2 (test: pre-test vs. post-test) mixed ANOVAs were performed to understand whether the shadowing practice created a change in the ratings. The assumptions of mixed ANOVA were verified. The normality assumption, skewness, and kurtosis values for each group's speech features were found to be in acceptable ranges (-1.5/+1.5) based on Tabachnick & Fidell (2013). Tests of normality, box plots, Q-Q plots, and histograms showed normality. The equality of variances assumption was also met.

The analysis of the survey responses was performed through bar charts. Inferential statistics were not performed on the survey aggregately, since it was not a unidimensional scale. Also, content analysis was performed on the data gathered from the two open-ended questions. The interview recordings were transcribed manually and analyzed through content analysis. To classify the content elements, common words, and phrases in the transcriptions were coded. Frequency tables were prepared to classify the emerging themes. Additionally, 30% of the data were coded by another coder who was an English language teacher. Based on Miles and Huberman's (1994) formula, the inter-coder reliability was found to be 90%.

Results

The Results of the Read-Aloud Ratings

The first research question aimed at understanding to what extent the assigned shadowing practices lead to differences in L2 learners' comprehensibility, pronunciation, intonation and rhythm, and speech rate, as indicated by read-aloud ratings.

The first research question aimed at understanding to what extent the assigned shadowing practices lead to differences in L2 learners' comprehensibility, pronunciation, intonation and rhythm, and speech rate, as indicated by read-aloud ratings. In the pre-test,

the mean comprehensibility ratings of both groups were found to be similar as shown in Table 1.

Table 1. Descriptive Statistics for Comprehensibility Pre- and Post-tests (7-point Likert scale)

		<i>N</i>	<i>M</i>	<i>Median</i>	<i>Min</i>	<i>Max</i>	<i>SD</i>	<i>95% CI</i>
Pre-Test	Control	24	4.31	4.38	2.75	6.25	.92	[3.93, 4.70]
	Experimental	28	4.3	4.38	2.75	5.75	.95	[3.93, 4.66]
Post-Test	Control	24	4.22	4	2.75	5.5	.75	[3.90, 4.54]
	Experimental	28	4.49	4.5	2.75	5.75	.8	[4.18, 4.80]

Only the experimental group demonstrated a small improvement in terms of comprehensibility in the post-test. A mixed ANOVA revealed no interaction between group and testing time $F(1, 50) = 1.29, p = .26, \eta_p^2 = .025$, as well as no main effects of group $F(1, 50) = .398, p = .53, \eta_p^2 = .008$ and testing time $F(1, 50) = .16, p = .69, \eta_p^2 = .003$. In terms of pronunciation, the descriptive statistics showed little pre- and post-test difference in the mean scores (see Table 2).

Table 2. Descriptive Statistics for Pronunciation Pre- and Post-test Rating Scores (7-point Likert scale)

		<i>N</i>	<i>M</i>	<i>Median</i>	<i>Min</i>	<i>Max</i>	<i>SD</i>	<i>95% CI</i>
Pre-Test	Control	24	4.14	4	3	6	.73	[3.83, 4.44]
	Experimental	28	4.13	4.5	2.75	5.5	.95	[3.75, 4.50]
Post-Test	Control	24	4.26	4.25	3	5.5	.66	[3.98, 4.54]
	Experimental	28	4.24	4.25	2.75	5.5	.79	[3.93, 4.54]

A mixed ANOVA revealed no interaction between group and testing time $F(1, 50) = .001, p = .97, \eta_p^2 = .00$, as well as no main effects of group $F(1, 50) = .001, p = .97, \eta_p^2 = .00$ and testing time $F(1, 50) = .87, p = .36, \eta_p^2 = .00$. Concerning intonation, the pre-test post-test differences were minimal (See Table 3).

Table 3. Descriptive Statistics for Intonation Pre- and Post-test Rating Scores (7-point Likert scale)

		<i>N</i>	<i>M</i>	<i>Median</i>	<i>Min</i>	<i>Max</i>	<i>SD</i>	<i>95% CI</i>
Pre-Test	Control	24	3.99	4	1.75	5.75	.93	[3.78, 4.45]
	Experimental	28	4.12	4.25	2.5	5.75	.87	[2.50, 5.75]
Post-Test	Control	24	4.06	4	2.25	5.75	.81	[3.72, 4.40]
	Experimental	28	4.45	4.75	2.5	5.75	.85	[4.12, 4.77]

A mixed ANOVA revealed no interaction between group and testing time $F(1, 50) = 1.26, p = .27, \eta_p^2 = .03$, as well as no main effects of group $F(1, 50) = 1.46, p = .23, \eta_p^2 = .03$ and testing time $F(1, 50) = 3.09, p = .09, \eta_p^2 = .06$. As demonstrated in Table 4, the mean scores of pre- and post-test for speech ratings were similar.

Table 4. Descriptive Statistics for Speech Rate Pre- and Post-test Rating Scores (7-point Likert scale)

		<i>N</i>	<i>M</i>	<i>Median</i>	<i>Min</i>	<i>Max</i>	<i>SD</i>	<i>95% CI</i>
Pre-Test	Control	24	4.41	4.5	2.5	6	.93	[4.01, 4.80]
	Experimental	28	4.51	4.38	2.75	6	.8	[4.20, 4.82]
Post-Test	Control	24	4.47	4.5	2.75	5.75	.71	[4.17, 4.77]
	Experimental	28	4.86	5	3.25	6.25	.87	[4.86, 5.19]

A mixed ANOVA revealed no interaction between group and testing time $F(1, 50) = 1.11, p = .3, \eta_p^2 = .02$, as well as no main effects of group $F(1, 50) = 1.74, p = .19, \eta_p^2 = .03$ and testing time $F(1, 50) = 2.29, p = .14, \eta_p^2 = .04$.

Results of the Shadowing Task Evaluation Survey

Within the scope of the second research question, the experimental group participants were asked to evaluate the effectiveness of the shadowing practice and its influences on their speech features. Analysis of the items related to the general contributions of shadowing, effects related to suprasegmental features, listening comprehension, evaluations of the shadowing procedures, and impressions related to the potential of long-term shadowing practice are presented in Figures 1-6.

Regarding the general contributions of the shadowing exercises, 81% of the participants agreed (23% strongly) that there was a development in their general speaking skills with the help of shadowing, as seen in Figure 1.

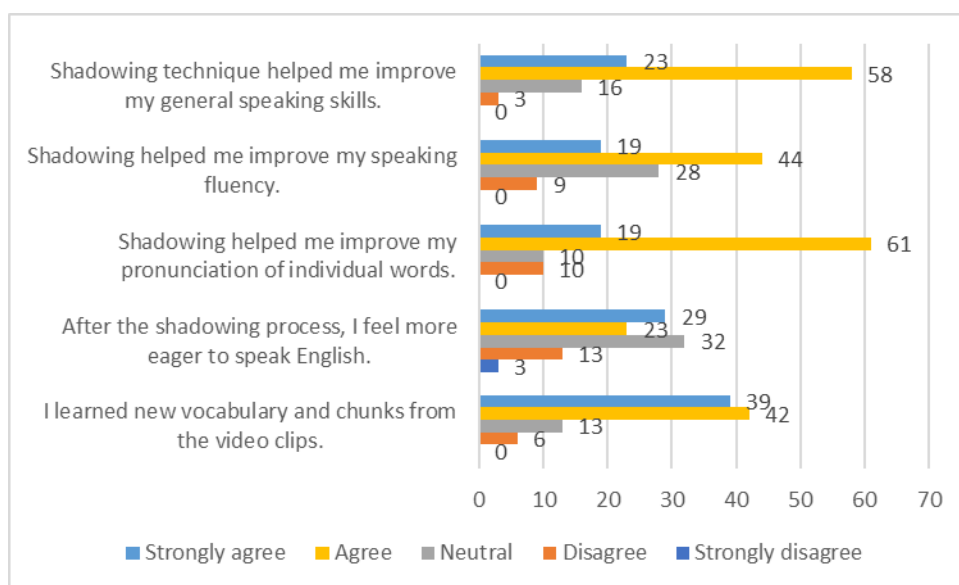


Figure 1. The evaluations of the general contributions of shadowing (%)

Regarding fluency development, the rate of agreement was slightly lower, at 63% (19% strongly), and 28% neither agreed nor disagreed. Most participants agreed that shadowing improved the pronunciation of individual words, as seen in 80% of the responses (19% strong agreement). However, nearly half of the participants (52%) agreed (29% strongly) with the statement that their willingness to speak English increased after the shadowing practice, and 32% neither agreed nor disagreed. As for the other benefits of shadowing, 81% of the participants agreed (39% strongly) that they learned new vocabulary and chunks from the shadowing video clips.

In the survey, 90% of the participants approved (42% strongly) of the importance of pronunciation and intonation skills for them, as shown in Figure 2.

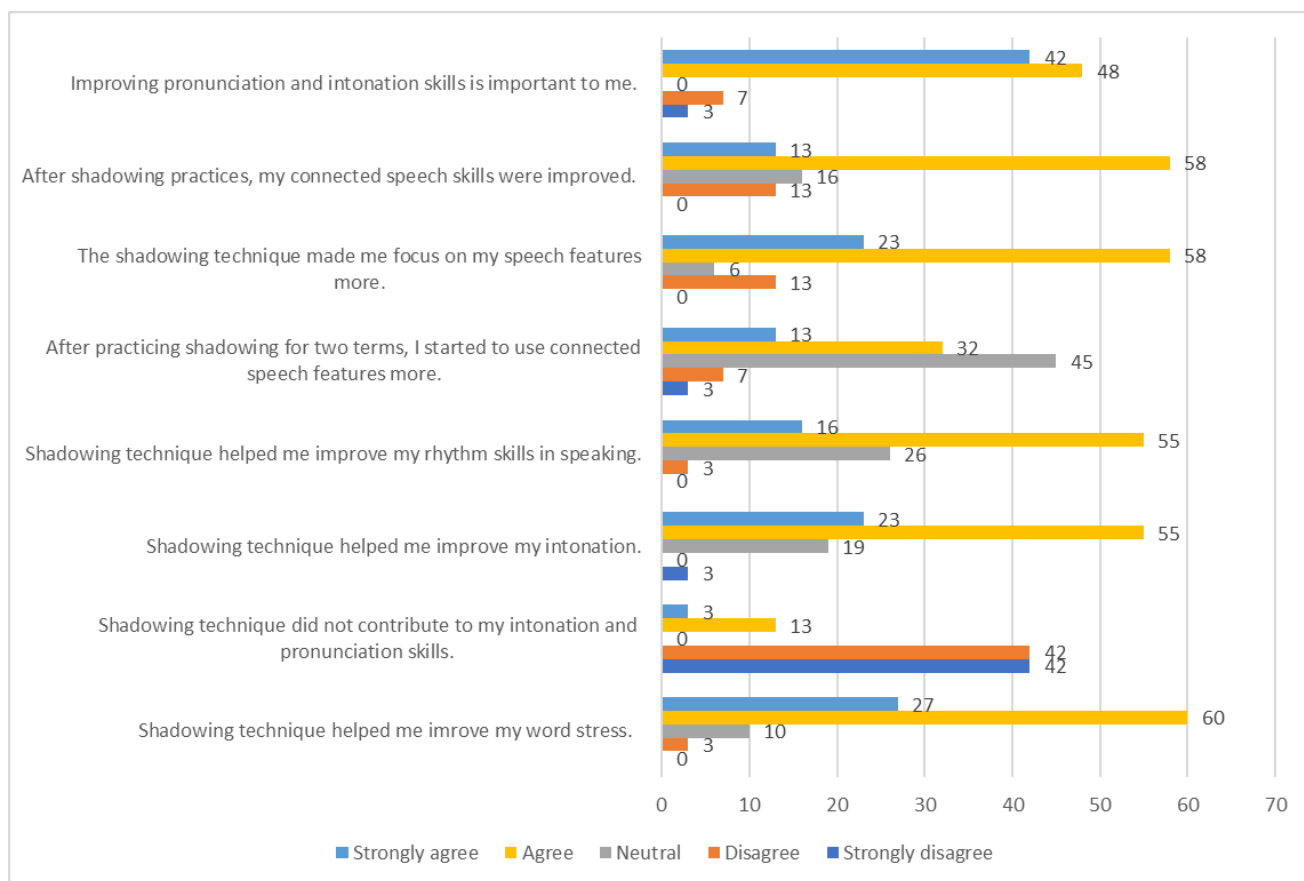


Figure 2. The evaluations of the contributions of shadowing on suprasegmental features (%)

Some of the survey statements specifically probed the contributions of shadowing to the improvement of suprasegmental skills. The majority (71%) agreed (13% strongly) that their connected speech skills improved, and they became more cognizant of their speech features after shadowing (81%), while only 45% agreed that they started to use these skills more, and 45% neither agreed nor disagreed. This suggests that participants had increased awareness of suprasegmentals which did not match the improvement of productive skills in these features. On the other hand, 71% concurred (16% strongly) that shadowing helped them improve rhythm skills while speaking English, while 26% neither agreed nor disagreed. Moreover, 78% of the participants believed (23% strongly) that shadowing improved their intonation and pronunciation skills (19% were neutral) and 84% disagreed with the cross-check statement that shadowing did not contribute to these skills. Similarly, 87% agreed (27% strongly) that shadowing improved word stress.

In terms of listening comprehension, 42% of the participants disagreed (13% strongly) that they had difficulty in listening comprehension when they watched the video clips for the first time, although 32% were undecided and 26% showed agreement with the statement (See Figure 3).

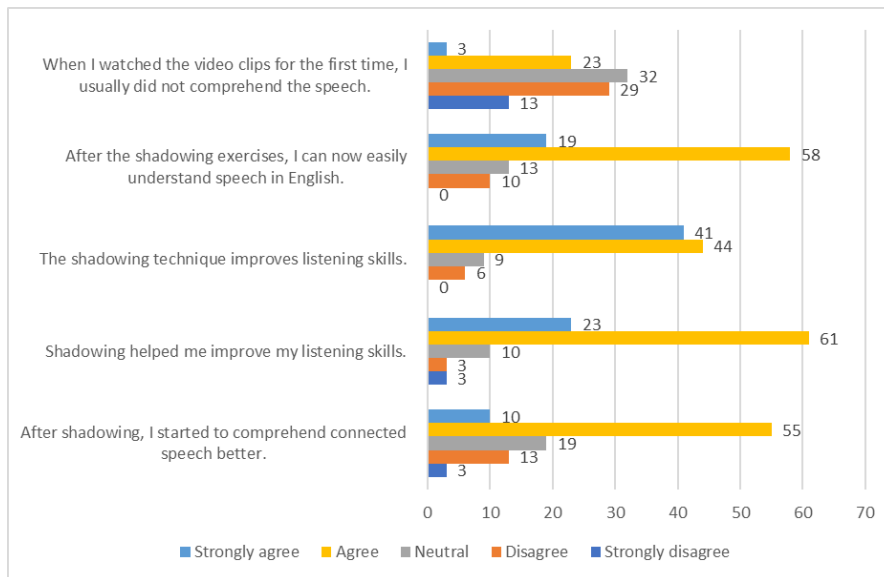


Figure 3. The evaluations of the contributions of shadowing on listening comprehension (%)

The majority (77%) agreed (19% strongly) that it was easier to comprehend English speech after the shadowing practice. Similarly, 85% agreed (41% strongly) that shadowing improves listening comprehension. Considering comprehension of connected speech, 65% reported improvement (10% strong agreement) and 19% were undecided. The response rates for items related to the shadowing procedures are presented in Figure 4.

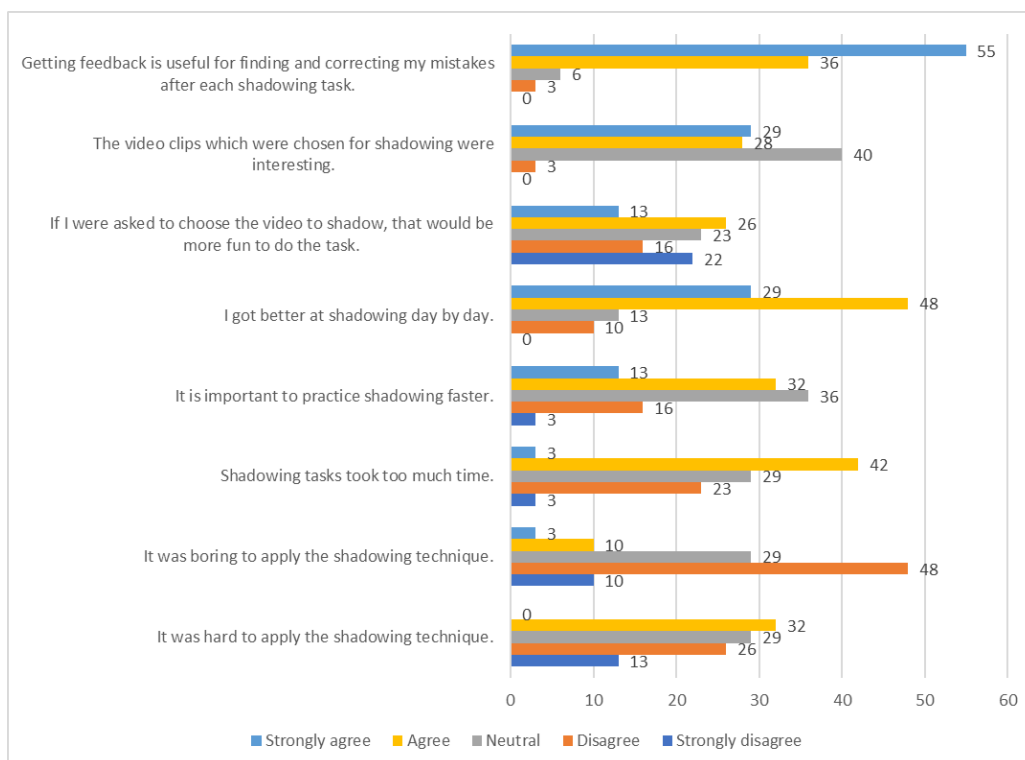


Figure 4. The evaluations of the shadowing procedures (%)

Receiving feedback was found to be the most favored aspect of the shadowing intervention with 91% agreement (55% strong). As for the selection adequacy of the videos, 57% of the students showed agreement (29% strong), while 40% were undecided. This suggests that students have diverse interests, some of which differ from the area of the selected videos. However, when asked if it would be better to choose their videos, only 39% were willing to do so and 38% disagreed. Most students (77%) agreed (29% strongly) that they improved their shadowing skills gradually, which indicates shadowing requires some practice. Concerning the rate of shadowing, 45% agreed (13% strongly) that it should be practiced faster, while 36% neither agreed nor disagreed. Overall, 45% of the students agreed that shadowing tasks took too much time and 29% opted for the neutral option. On the other hand, 58% disagreed that shadowing was boring, while 29% were undecided. Similarly, 39% disagreed that shadowing was hard to practice, 29% were undecided, and 32% agreed with the statement. The participants were also asked to provide additional ratings on some features of the shadowing procedure on a 7-point scale, as shown in Figure 5.

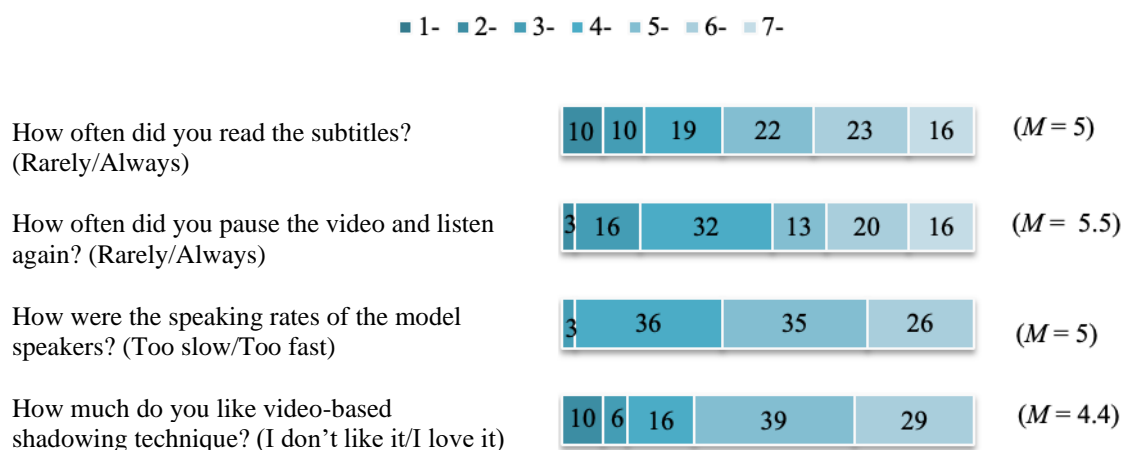


Figure 5. The reported attitudes toward the shadowing technique and information about shadowing practices (%)

The mean rating for the frequency of reading the subtitles while shadowing was 5. Similarly, the mean rating for how often they paused the video to listen again was 5.5. Students rated the speaking rate of the speakers in the videos as 5. Finally, the mean rating for participants' fondness of shadowing was 4.4.

The survey also comprised items related to the potential of shadowing for foreign language learning, as shown in Figure 6.

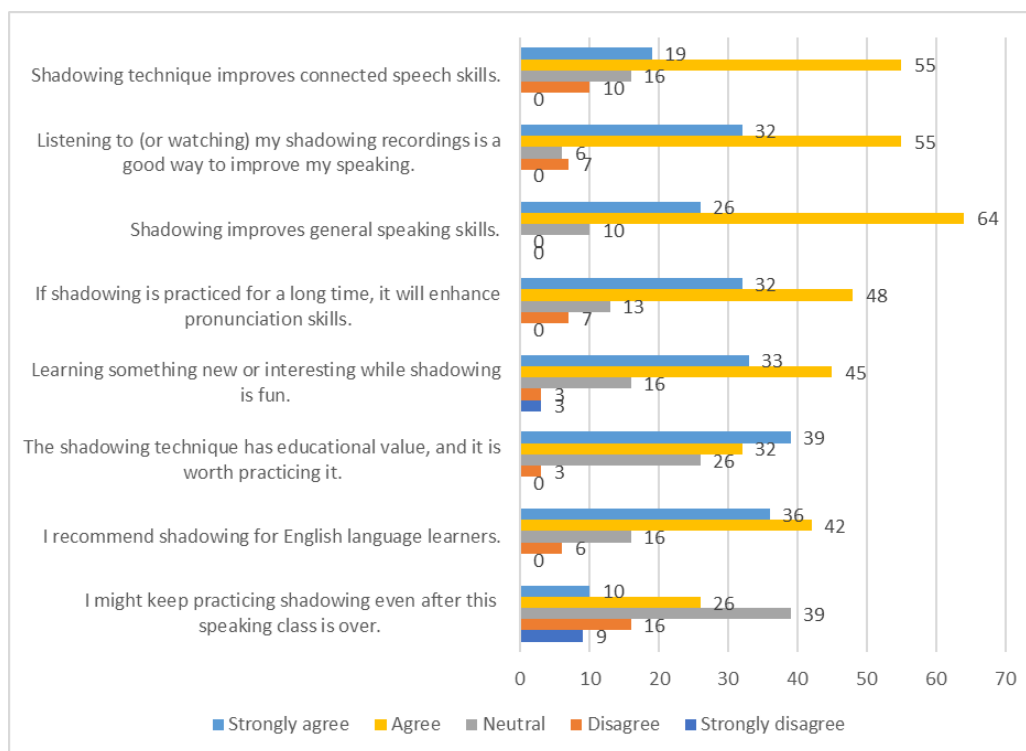


Figure 6. The evaluations of the potential of shadowing (%)

Correspondingly, 74% agreed (19% strongly) that shadowing has the potential to improve connected speech skills. Overall, 87% of the participants agreed that listening to or watching their shadowing recordings is beneficial for enriching speaking skills. In the same way, 90% agreed (26% strongly) that shadowing improves general speaking skills. The majority (80%) thought (32% strongly) that shadowing had the potential to enhance pronunciation if performed for a long time. Learning something new or interesting was regarded as an added advantage of the technique, as reflected in 78% of the confirming responses. In general, 71% believed the shadowing technique has pedagogical value, while 26% were undecided. The majority (78%) showed a willingness to recommend the shadowing technique to other L2 learners. Despite these positive responses, only 36% agreed that they might continue shadowing, 39% opted for the neutral option and 27% showed disagreement.

The two open-ended questions in the survey were related to the positive and negative aspects of the shadowing technique. According to content analysis, 29% of the participants stated that shadowing was efficient in terms of improving pronunciation. Similarly, 31% believed that shadowing boosted their general oral proficiency. Overall, 29% of the responses pointed out that shadowing is also useful for enhancing listening skills. Furthermore, 10 participants reported that they improved both listening and speaking skills and four participants also argued that they could comprehend native speakers of English better. On the other hand, considering the disadvantages, 33% of the participants stated that shadowing has no negative aspects. While 40% indicated that practicing shadowing took a lot of time, 17% of the participants explained that shadowing

was boring because of the difficulty of keeping the utterances in memory and recording the video, and 10% stated they had difficulty finding a quiet place to record themselves.

Results of the Interview

Within the scope of the second research question, we gathered qualitative data related to shadowing through a semi-structured interview and analyzed it through content analysis. The most frequently uttered phrases in the interviews were counted and coded. Eleven keywords emerged and were categorized into four main themes: overall improvement of L2 learning through video-based shadowing (31%), the effects of shadowing on oral skills (28%), attitudes toward shadowing (20%), and shadowing experiences and procedures (21%).

Overall improvement in second language skills through video-based shadowing was the most frequently mentioned theme (See Table 4).

Table 4. Emerging themes from the interviews

<i>Theme</i>	<i>Frequency (f)</i>
<i>Overall improvement in L2 learning</i>	185
<i>Effects of shadowing on oral L2 skills</i>	
Pronunciation	52
Suprasegmental features	50
Speaking	45
Fluency	16
	163
<i>Attitudes towards shadowing</i>	
Fun, easy, motivating	65
Difficult, exhausting, burden, boring, time-consuming	52
	117
<i>Shadowing experiences and procedures</i>	
Self-monitoring, correction of errors, comparison, awareness	48
Feedback, instructor	31
Memorization	23
Repeating, recording again	21
	123
TOTAL	588

Words and phrases such as ‘develop, affect, improve, useful, helpful, and beneficial’ ($f = 185$) were frequently used to refer to the benefits of shadowing. Participants mentioned general improvement in the second language, second language use, and listening. Some illustrative quotations were:

P10: We had very little speaking and listening practice in high school. Shadowing developed especially listening and speaking. I learned new language structures.

P2: After I practiced shadowing for some time, I think it benefited me. I think it has a lot of positive aspects.

P8: We learned both new words and how to pronounce them.

The second theme was the positive effects of shadowing on oral skills. Within this theme, 52 of the responses referred to the effects of shadowing on enhancing pronunciation and decreasing articulation mistakes in individual words. While 50 of the replies under the oral skills theme revealed that shadowing was favorable for improving specific speech features such as intonation, rhythm, connected speech, and word stress, 16 replies suggested that participants viewed shadowing as a means of practicing speaking. The remaining 16 responses pointed to benefits in ‘fluency’ after practicing shadowing. To illustrate, Participant 10 reported feeling more confident and motivated in speaking skills in addition to improving suprasegmental features:

P10: It positively affected fluent speaking and gave me more confidence and motivation to speak. I participated in the lesson less in the first term. At first, I was shy, but I felt more relaxed in the second term. Especially when talking about an important thing and emphasizing an important thing, it affected my intonation. I noticed that native speakers raised their voices at some points. I started to do the same. I also started to reflect connected speech features into my daily English speech.

Participant 7 was less certain about the improvement in fluency but concurred that her intonation, stress, and connected speech features were improved:

P7: We wouldn’t know if our fluency improved without talking to a foreigner. I just completed a speaking exam, and I still hesitated many times. I was already anxious about speaking, but on the whole, I would predict it would improve my fluency. I guess I am not at that point exactly. However, shadowing improved my stress and intonation immensely. I paid attention to them in my presentation. I was aware of when to raise and lower my voice. For example, I start by saying “however” (*raises her voice*) and I continue speaking in a lower tone. I don’t know if I use connected speech in my regular speech, but I try to do it in read-aloud practice. For example, I try to connect “of’s. My awareness increased because I did not know what it was before shadowing.

The third common theme in the interviews was related to attitudes towards shadowing practices. Two contrasting trends were notable as indicated in the keywords: “fun, easy, motivating, good” versus “difficult, burden, boring, exhausting and time consuming”. The content analysis implied that even though the participants enjoyed shadowing ($f = 65$), they also found the procedure hard ($f = 52$). Some representative interview responses were:

P8: Actually, it is enjoyable on the whole, but when our mode is low, we may not want to do it. I would go on doing it when I’m alone because it’s enjoyable.

Similarly, Participant 1 complained that the frequency of the assignments was high; however, he wrongly recalled that the assignments were weekly, although they were biweekly. Despite the somewhat negative stance towards the frequency of shadowing, he continued to practice it on his own.

P1: The only problem was that since it was a weekly assignment, I viewed it as a kind of homework load at the beginning. I think the frequency was a bit high. But later, although it is no longer an assignment, I do it on my own sometimes.

The last common theme emerging from the interviews comprised statements related to the shadowing experience and procedures. This theme was classified into four subcategories: self-monitoring, feedback, rerecording, and memorization. Within the self-monitoring category, the statements indicated that the key process in the shadowing procedure was the recurrent self-comparison of participants' speaking performance with the model speech. As participants recounted, this process helped them become aware of their pronunciation errors and the suprasegmental features of English:

P3: I think I pronounce a word correctly. When I relisten to the voice recording, I notice my error and try to pronounce it correctly. It was useful to listen to your own voice.

Another participant mentioned feeling better after self-monitoring:

P5: I watched the two videos [the original video and the shadowing] and when I saw that the two were similar, I felt self-confident.

Although self-monitoring was an important emerging theme, the importance of external factors, such as instructor feedback, was also emphasized:

P2: Feedback was important for pronunciation. The instructor showed words, like "if you pronounced a word in this way, it would be good". I noticed the words I mispronounced and corrected them. These words also appeared in the courses a few times.

P5: I didn't pay attention to my recordings initially, but after the instructor's feedback, I improved them.

On the other hand, 23 statements concerned efforts to memorize all the subtitles of the model speech at the beginning of the task to be flawless in the final recording, which was challenging to accomplish. Although the participants were instructed to avoid memorization, some participants mentioned doing this in their initial practices:

P1: At first, everybody was confused. Including me, everyone tried to memorize the subtitles. Actually, we were not supposed to do it. We could sometimes look at the lines. The aim is to pay attention to intonation and speech. We understood this later, which took some time. The actual aim of shadowing is to follow someone like their shadow. We were able to grasp this. At first, we memorized; actually, not memorized, we just read. Later, in the feedback the instructor told us not to do direct reading; she told us to practice the intonation and rhythm. Then, everybody started to read together with the video. That was much better.

Nevertheless, 21 of the statements involved complaints about the initial hardship of recording their final shadowing performance, which decreased gradually.

P3: At first, I had to watch the video 7-8 times, but later I could easily handle it, paying attention to the [speaker's] pronunciation 2-3 times.

P5: The hardest part was recording again and again.

Some participants also complained about the difficulty of recording their voices in a noisy environment:

P6: The greatest challenge was the environment and noises because I live in a dorm. I had to wait for my roommates to be quiet and I had little time to be alone, so sometimes I had to go outside. The internet connection in the dorm was also slow. Sending it [the recording] was a problem. I only experienced technical problems. I had no other problems.

Regarding the procedural suggestions, some participants talked about the recording format and the materials. Participant 5 mentioned that working on pre-selected video materials was more convenient.

P5: My video selections were also good, but I'd prefer the instructors' selections. It is also more practical that way.

Some students mentioned that they preferred recording their voice in an audio file rather than in a video file.

P9: In the first term, we video-recorded ourselves. In the second term, audio recording in place of video recording made me more comfortable. Apart from that, a higher frequency [of shadowing practice] would be better, but then we have other courses, so it's not that possible.

P10: Redoing the shadowing after getting feedback would be more effective, but maybe it would be boring. I'm not sure. You learn your mistakes.

At the end of the interview, the participants were also asked whether they would recommend shadowing to language learners. Some representative interview responses were as follows:

P1: It's worth doing it. I'll use it when I become a teacher and recommend it to others. It would be beneficial for young learners. Everyone has a tool that records voice in their home. It could start at secondary school. It's not a difficult task at all. I wish I had done it earlier. It boosts self-confidence and helps you speak more.

P10: It doesn't take a lot of time. You can listen to it 5-6 times. You can record your voice 5-6 times and it wouldn't take even an hour. In a country like Türkiye where English is not spoken frequently, it is an effective technique for learning a language with the correct pronunciation.

Discussion

This study aimed to examine the recorded and perceived impact of the shadowing technique on comprehensibility, pronunciation of individual sounds, intonation, and speech rate of L2 learners and their views about the shadowing practices. Four raters' evaluations of the participants' pre- and post-shadowing voice recordings indicated that the experimental group had slightly higher scores after the intervention in comprehensibility, intonation, and speech rate. However, the speech ratings did not indicate statistically significant differences between the experimental and control groups in terms of comprehensibility, pronunciation, intonation, and speech rate. These results were not in line with past studies, such as that of Foote & McDonough (2017), which recorded improved comprehensibility scores following shadowing in L2 English speakers at a Canadian university. The results also contrasted with previous research in terms of improvement of pronunciation and intonation after shadowing in contexts where the target language is not spoken (Bovee & Stewart, 2009; Hsieh et al., 2013; Martinsen et al., 2017; Mori, 2011) and in contexts where the target language is spoken (Mishima & Cheng, 2017). The discrepancy in the results might be related to differences in the rating procedures and the number of raters. Unlike the past studies, the raters in this study assessed four speech features at the same time by listening to the recording once, which might have made it more difficult to detect the nuanced changes. As we did not have native speakers within our immediate context, the rating had to be completed online in two sessions. Increasing the number of raters and pursuing the rating in a physical environment with more sessions or using different rating methods would have been more effective and consistent. As Saito and Plonsky (2019) explain, computer-assisted acoustic analyses of pronunciation generally yield more consistent ratings than human raters and the "lack of statistical significance does not necessarily diminish the effects of pronunciation teaching, it may rather indicate individual differences in human perception of spontaneous speech" (p. 695).

Individual differences in L2 learners might also have obscured any potential effects of shadowing. Linguistic aptitude, musical aptitude, imitation ability, and L2 learning motivation have been discussed as factors that influence L2 pronunciation learning (See Nagle, 2022; Suzukida, 2021 for comprehensive reviews). Differences in auditory processing have also been found to affect L2 phonological proficiency (Saito et al., 2022). The measurement of such individual differences was not possible within the scope of our study and should be considered in future investigations.

Apart from individual differences in raters and participants, a comprehensive meta-analysis of pronunciation training studies over 25 years has indicated that focusing on the measurement of specific aspects of L2 pronunciation, such as minimal pairs, the pronunciation of specific segments, etc. is more effective and easier to quantify improvement than that of global pronunciation or comprehensibility (Saito & Plonsky, 2019). The measurements in the present study focused on the rating of global segmental and suprasegmental features, the progress of which can be harder to track.

Despite the lack of quantitative improvement in L2 learners' speech features, the attitudes of the experimental group towards the shadowing technique were positive according to the results of the survey and the semi-structured interview. In the current

study, most of the participants found shadowing helpful for developing the speech features of pronunciation, intonation, word stress, connected speech, and rhythm. Additionally, as supported in the interview responses, the group believed it was worth practicing shadowing and recommended the activity to other L2 learners. These results were in accordance with previous survey studies (Bovee & Stewart, 2009; Sumiyoshi & Svetanant, 2017) documenting the perceived contributions of shadowing in the pronunciation of individual sounds and intonation. As in these studies, most participants were convinced of the pedagogical value of the shadowing technique. The majority of the participants indicated that it was also fun to learn new and interesting facts through shadowing. Moreover, the majority found shadowing to be useful for learning new vocabulary and chunks. Although listening comprehension was not the main focus of the present study, the participants' perspectives of shadowing as a contributor to listening comprehension development were in line with that of past research (Bovee & Stewart, 2009; Sumiyoshi & Svetanant, 2017; Teeter, 2017). Some participants also commented that their self-confidence in speaking skills increased, in line with the results of Teeter (2017).

Along with the overall positive comments, there were complaints that the shadowing technique is time-consuming, similar to the comments in Bovee and Stewart, (2009). This could be one reason why the activity received slightly lower overall ratings than in some previous studies (Foote & McDonough, 2017). Another reason might be that carrying out shadowing tasks was a course requirement for the experimental group in the present study, while in Foote and McDonough (2017), the group was paid for their participation in the study, which might have increased their motivation to perform the shadowing tasks.

Also, unlike North American contexts, the students in our context had a high course load, which might have challenged their time management. Based on some participants' interview responses, their past L2 learning habits and experiences have led them to test-oriented, short-term study patterns. The lack of skills-oriented language learning in K12 education in Türkiye might have led the learners to initially perceive shadowing or similar pronunciation training activities as challenging and time-consuming. In our context, we have also observed that the participants were not used to doing skills-based L2 practice as homework. Coupled with this, in the freshmen year, the learners did not seem to have enough self-autonomy and time management skills to cope with the assignment load of 10 compulsory courses. Although two weeks were given for students to ease task completion, the program's overall workload might have led to anxiety, especially at the initial stages of the shadowing practices, as some participants mentioned. In such a context, we believe giving students feedback about their shadowing practices was crucial in terms of motivating them to continue practicing, as the survey and interview responses and the previous literature confirmed (Sumiyoshi & Svetanant, 2017). However, it should be noted that listening to a large number of student recordings and giving individual feedback was challenging on the part of the instructor. We believe practicing shadowing in smaller classes more frequently and with more feedback would be effective in contexts where students have less course load.

The participants also reported technical difficulties in voice recording and homework file submission due to the slow Internet connection. To alleviate this concern,

we asked participants to send sound files instead of videos in the second term, which made the participants more relaxed in terms of recording anxiety as well.

The participants' reported practice times varied and diminished gradually as they mastered the procedure. In the beginning phases, students tried to memorize the model speaker's lines, which gradually transformed into a more natural shadowing performance. However, even after long-term practice, the participants reported pausing the video to listen to the model speech and found the model's speech rate somewhat fast. If possible, using mobile applications that automatically send reminders, record voice, give feedback, and track the pauses and timing of practice would be more practical and effective as was done in (Foote & McDonough, 2017). In our case, such applications were not free of charge and therefore we could not adopt them for monitoring students' practice time. We expect that more varied and free shadowing applications will be available in the future, following the improvements in artificial intelligence.

Despite the technical and procedural challenges, all participants found shadowing efficient in improving L2 speech features. Although some participants reported they initially found shadowing difficult to practice, time-consuming, and somewhat boring, they thought it was worth practicing as they got used to the procedure. This accords with the findings of Foote and McDonough (2017), which showed that the participants' early thoughts about the effectiveness of shadowing differed till the end of their study.

In the light of these findings, the pedagogical and research-related implications can be summarized as:

1. As the participants' statements indicate, shadowing practices contribute to increasing L2 learners' awareness of segmental and suprasegmental features, listening comprehension, and vocabulary. Apart from linguistic benefits, participants report improvement in self-monitoring their pronunciation errors, and increased self-confidence, which was also reported in Teeter (2017).
2. Measurement of global speech features (e.g., comprehensibility, general pronunciation, etc.) after shadowing practices might not yield statistically significant rates of improvement in experimental studies. As Saito and Plonsky (2019) suggest, focusing on the measurement of more specific features, such as certain minimal pairs, or problematic sounds for the specific L1 group may yield more noticeable improvements in speech features.
3. Supporting shadowing with more explicit instruction on specific aspects of pronunciation and intonation could be of value, especially with less autonomous learners and in foreign language learning environments.
4. Homework-based shadowing and pronunciation practices might be challenging for learners in written test-oriented educational settings, rather than performance-based settings. More external support and feedback should be provided to such learners. As the participants report, regular feedback is crucial in terms of maintaining learner motivation in pronunciation training. For crowded classes, using practical mobile applications or learning management systems, which automatically track practice times and give feedback, would be highly beneficial for learners (Foote & McDonough, 2017). In addition, consistency and distributed practice over a long

time is important for students to familiarize themselves with the procedure (Martinsen et al., 2017).

5. If automated practice systems are not available, presenting learners with a variety of videos to choose from, instead of only asking them to select videos to shadow seems to be timesaving, as the participant responses indicated.
6. Asking learners to record their voices in audio files, rather than video files makes them feel more comfortable, and decreases file submission problems due to slow Internet connection.

Conclusion

Video-based shadowing practices have increased L2 learners' awareness of pronunciation, intonation, and fluency although this has not resulted in quantifiable changes in their speech features, as rated by human raters after the intervention. Participants reported benefiting from comparing their shadowing performances with a model speaker and feeling more self-confident in oral skills and listening comprehension. As for the negative aspects, participants reported that shadowing was time-consuming and technically challenging, especially in the early stages.

The present study had some limitations. One issue was the number of participants, which was limited. In the beginning, there were 60 first-year ELT students in the experimental group in intact classes, but only 32 of them could complete all of the eleven shadowing tasks since assignment completion rates are in general low in the specific school context. Hence, it is difficult to generalize the results of this study to a larger population from different backgrounds. To get more externally valid results, more participants are required. Additionally, the shadowing tasks were carried out at home. So, it was difficult to control location effects and to track practice times and durations more accurately.

The lack of a validated survey to measure attitudes to shadowing was another limitation. Within the scope of the present research, it was not possible to validate the attitude survey without pre-administering it to 300-400 university students who have practiced shadowing for a long time. Using an unvalidated survey also complicated the statistical appropriacy of using more generalizable inferential statistics on the items. The lack of native speaker raters within our immediate context was another limitation. Using a higher number of raters or computer-assisted acoustic analyses should be considered in future studies. The study of individual differences in shadowing training is another topic that should be investigated further.

References

- Allan, D. (2004). *Oxford quick placement test*. Oxford University Press.
- Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829–839. <https://doi.org/10.1038/nrn1201>
- Baddeley, A., Gathercole, S., & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological Review*, 105(1), 158–173. <https://doi.org/10.1037/0033-295X.105.1.158>

- Bovee, N., & Stewart, J. (2009). The utility of shadowing. *JALT 2008 Conference Proceedings*. Tokyo: JALT, 888–900.
- Cherry, E. C. (1953). Some experiments on the recognition of speech, with one and with two ears. *The Journal of the Acoustical Society of America*, 25(5), 975–979. <https://doi.org/10.1121/1.1907229>
- Cicchetti, D. V. (2001). Methodological commentary the precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology*, 23(5), 695–700. <https://doi.org/10.1076/jcen.23.5.695.1249>
- Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.
- Dauer, R. M. (1993). *Accurate English: A complete course in pronunciation*. Prentice-Hall.
- Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, 3(1), 34–56. <https://doi.org/10.1075/jslp.3.1.02foo>
- Hamada, Y. (2016a). Shadowing: Who benefits and how? Uncovering a booming EFL teaching technique for listening comprehension. *Language Teaching Research*, 20(1), 35–52. <https://doi.org/10.1177/1362168815597504>
- Hamada, Y. (2016b). Wait! Is it really shadowing? *The Language Teacher*, 40(1), 14. <https://doi.org/10.37546/JALTTLT40.1-3>
- Hamada, Y. (2017). *Teaching EFL learners shadowing for listening: Developing learners' bottom-up skills*. Routledge Taylor & Francis Group.
- Hamada, Y. (2019). Shadowing: What is it? How to use it. Where will it go? *RELC Journal*, 50(3), 386–393. <https://doi.org/10.1177/0033688218771380>
- Hsieh, K.-T., Dong, D.-H., & Wang, L.-Y. (2013). A preliminary study of applying shadowing technique to English intonation instruction. *Taiwan Journal of Linguistics*, 11(2), 43–65.
- Kadota, S. (2007). *Shadowing to ondoku no kaagaku [Science of shadowing, oral reading, and English acquisition]*. CosmoPier Publishing.
- Kadota, S. (2021). *Shadowing as a practice in second language acquisition: Connecting inputs and outputs*. Routledge.
- Kadota, S., Kawasaki, M., Shiki, O., Hase, N., Nakano, Y., Noro, T., Nakanishi, H., & Kazai, K. (2015). The effect of shadowing on the subvocal rehearsal in L2 reading: A behavioral experiment for Japanese EFL learners. *Language Education and Technology*, 52, 163–177.
- Koo, T. K., & Li, M. Y. (2016). A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Lin, L. (2009). *A study of using "shadowing" as a task in junior high EFL program in Taiwan* [Unpublished master's thesis]. National Taiwan University of Science and Technology, Taipei.
- Martinsen, R., Montgomery, C., & Willardson, V. (2017). The effectiveness of video-based shadowing and tracking pronunciation exercises for foreign language learners. *Foreign Language Annals*, 50. <https://doi.org/10.1111/flan.12306>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage Publications.
- Mishima, M., & Cheng, L. (2017). The impact of a computer-mediated shadowing activity on ESL speaking skill development: A pilot study. *L2 Journal*, 9(1). <https://doi.org/10.5070/L29132493>
- Mori, Y. (2011). *Shadowing with oral reading: Effects of combined training on the improvement of Japanese EFL learners*. https://doi.org/10.24539/let.48.0_1
- Nagle, C. L. (2022). Individual difference factors for second language pronunciation. In S. Li, P. Hiver, & M. Papi, *The Routledge handbook of second language acquisition and individual differences* (1st ed., pp. 269–281). Routledge. <https://doi.org/10.4324/9781003270546-23>
- Rongna, A., & Hayashi, R. (2012). Accuracy of Japanese pitch accent rises during and after shadowing training. *Proc. Speech Prosody 2012*, 214–217.

- Saito, K., & Plonsky, L. (2019). Effects of second language pronunciation teaching revisited: A proposed measurement framework and meta-analysis. *Language Learning*, 69(3), 652–708. <https://doi.org/10.1111/lang.12345>
- Saito, K., Sun, H., Kachlicka, M., Alayo, J. R. C., Nakata, T., & Tierney, A. (2022). Domain-general auditory processing explains multiple dimensions of L2 acquisition in adulthood. *Studies in Second Language Acquisition*, 44(1), 57–86. <https://doi.org/10.1017/S0272263120000467>
- Saito, K., Trofimovich, P., & Isaacs, T. (2015). Using listener judgments to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*. <https://doi.org/10.1093/applin/amv047>
- Saito, Y., Nagasawa, Y., & Ishikawa, S. (2011). Effective instruction of shadowing using a movie. *JALT2010 Conference Proceedings*, 139–148.
- Sumiyoshi, H. (2019). The effect of shadowing: Exploring the speed variety of model audio and sound recognition ability in the Japanese as a foreign language context. *Electronic Journal of Foreign Language Teaching*, 16(1), 5–21.
- Sumiyoshi, H., & Svetanant, C. (2017). Motivation and attitude towards shadowing: Learners perspectives in Japanese as a foreign language. *Asian-Pacific Journal of Second and Foreign Language Education*, 2(1), 16. <https://doi.org/10.1186/s40862-017-0039-6>
- Suzukida, Y. (2021). The contribution of individual differences to L2 pronunciation learning: Insights from research and pedagogical implications. *RELC Journal*, 52(1), 48–61. <https://doi.org/10.1177/0033688220987655>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th ed.). Pearson.
- Takeuchi, H., Maruyama, T., Taki, Y., Motoki, K., Jeong, H., Kotozaki, Y., Shinada, T., Nakagawa, S., Nouchi, R., Iizuka, K., Yokoyama, R., Yamamoto, Y., Hanawa, S., Araki, T., Sakaki, K., Sasaki, Y., Magistro, D., & Kawashima, R. (2021). Effects of training of shadowing and reading aloud of second language on working memory and neural systems. *Brain Imaging and Behavior*, 15(3), 1253–1269. <https://doi.org/10.1007/s11682-020-00324-4>
- Teeter, J. (2017). Improving motivation to learn English in Japan with a self-study shadowing application. *Languages*, 2(4), 19. <https://doi.org/10.3390/languages2040019>

Appendix

Shadowing Instructions

1. Choose a 2-minute part of a video from TED talks, etc. Choose a speaker with natural pace, not too fast, not too slow OR watch the assigned video without subtitles (with headphones if possible).
e.g. <https://www.youtube.com/watch?v=RcGyVTAoXEU> (The first two minutes)
2. Watch the video without subtitles (with headphones if possible).
3. Shadow (=speaking at the same time with the speaker) the sentences at least twice.
4. Watch the video with subtitles/interactive script.
5. Practice saying the sentences by shadowing the speaker (=speaking at the same time with the speaker) while following the subtitles/script.
6. Practice shadowing the speaker paying attention to prosody (e.g. rising intonation, falling intonation, etc.), word stress, etc.
7. Review and re-record your voice until satisfied with the quality. (Recording duration: approximately 2 minutes).
8. From time to time, you can look at the subtitles or the transcript while recording yourself (in order to remember some words), BUT do not directly read the script.
9. Note down how long you practiced (in minutes), how many times you recorded the speech.
10. Submit “your” final shadowing recording.