



The Effect of Voice Recording, Video Dubbing, and Shadowing Tasks on the Pronunciation Ability: A Study of Iranian EFL Learners

Maryam Jalilvand¹, Narjes Ghafournia*¹

¹Department of English Language, Neyshabur Branch, Islamic Azad University, Neyshabur, Iran.

*Corresponding Author's Email: narjesghafournia@yahoo.com

Received: 30-09-2022, Accepted: 06-01-2023

ABSTRACT

The present study utilized a quasi-experimental design to examine the probable significant effect of three novel techniques of video dubbing, voice recording, and shadowing on improving Iranian EFL learners' pronunciation. To meet the end, 75 Iranian advanced EFL learners, as the target sample, participated in the study. Initially, 120 language learners participated in the study. With regard to the results of the Barron's TOEFL test, 75 advanced participants were selected, as the homogenous language proficiency sample. Then, they were randomly assigned into three experimental and one control groups. Each of the experimental groups was exposed to 15 hours of treatment on one of the three techniques. The progress of learners' pronunciation was checked via a pre-test/post-test system. Descriptive as well as inferential statistics were utilized to examine the hypotheses. The findings showed that there was a significant difference between the impacts of the three techniques on improving the learners' pronunciation. The findings also demonstrated that the voice recording proved to be a less effective technique in developing learners' pronunciation than video dubbing and shadowing. Moreover, shadowing and video dubbing were equally effective in boosting learners' pronunciation proficiency. The results could have various pedagogical implications for improving language learners' pronunciation.

KEYWORDS: Pronunciation; Shadowing; Video-Dubbing; Voice Recording

INTRODUCTION

Pronunciation flourished in dominance with the emergence of direct method and audiolingualism (Larsen-freeman, 2000), but was pushed away into the margins of language learning once again with the arrival of communicative language teaching, and Krashen's Natural Approach perspectives (Richards & Rogers, 2001). This fluctuation of pronunciation throughout the history of language teaching, just as the fluctuation of other skills such as writing or listening, is the typical of traditional language teaching methods, where extreme views were adopted on different language items: one day one component of the language was in the spotlight of attention and all scholars emphasized its significance; and the other day, it was exiled to the ghettos, being considered to be trivial and unimportant by all scholars. Nonetheless, in the modern sense of applied linguistics and SLA as a science, and in the light of the more progressive methods – namely effective teaching, reflective teaching, task-based teaching, and post-method era – a more moderate stance has been opted for when it comes to teaching pronunciation (Kumaravadivelu, 2001). It is currently believed that pronunciation is one of the most significant keys to having a communicative and thorough command of a foreign language (AMPE, 2002).

Much scientific research has been carried out regarding the essential role of explicitly teaching pronunciation. The reported results reveal that explicit exercise-based instruction can be an effective method in improving language learners' segmental and suprasegmental pronunciation performance (e.g. Kissling, 2013; Reis & Hazan, 2013; Saito, 2015). This study could be deemed as significant in various aspects. First and foremost, the three techniques, which this study assumed to be effective in boosting learners' pronunciation, are all novel techniques, which seem that they have not been previously systematically used for the purpose of teaching pronunciation.



Another aspect that made this study significant is that the three strategies it introduced for the improvement of learners' pronunciation focused both on segmental and suprasegmental pronunciation features. Through once practicing these strategies, learners would gain mastery over a set of pronunciation skills.

The nature of strategies that were applied in this study was absolutely learner-centered. Not only did this mean that these strategies are in line with democratic education principles and learner-oriented teaching paradigms, but also it means that they have the optimum amount of efficiency in the learning process of the learners as these strategies are further required to do the whole task. In fact, except for an initial briefing, learners could carry on doing all the three tasks by themselves, provided that they had sufficient and adequate materials. So, these strategies could be used for self-study purposes, as well.

The first technique proposed in this study is the shadowing technique. Shadowing could be defined as listening to a recording and reading simultaneously as the learner in the recording is reading the script (Arguelles, 2010). This technique was originally developed by the prominent American linguist, Noam Chomsky (1967). He called this technique reading-while-listening (RWH), but in his version, learners were required to read the passage silently as they were listening to it. There were three differences between the shadowing applied in this study, and that of Chomsky's. First, in this study, learners were supposed to read aloud as they were listening rather than Chomsky's silent reading. Second, Chomsky originally used this technique on scripts only, but here the researcher applied them on conversations. And third, Chomsky wished to improve fluency with reading while listening, whereas the researcher investigated its impacts on learners' segmental and suprasegmental pronunciation. The researcher assumed that by listening to a conversation and repeating the lines simultaneously as they are hearing it, learners would have a chance to put their production in harmony with their reception, and hence their pronunciation would thrive via copying native speakers' pronunciation.

Hamada (2018) examined the effect of "haptic shadowing" – also referred to as, the act of "punching" stressed syllables while listening to an audio input – on the language comprehensibility as well as segmental and suprasegmental productions. 58 Japanese second-year students in near-intermediate and intermediate levels of English proficiency took part in the study, including 15 sessions of prosodic training, in which they are required to read a passage of 450-550 words focusing on segmental (i.e., specific sounds) and suprasegmental elements. Statistical reported findings demonstrated a significant improvement in the three indicated features. In another study by Yamane et al. (2019), a group of English learners were asked to mirror the beat gestures, which were produced by a video recorded speaker while performing speech shadowing, whereas another group only shadowed the audio version of the same speech. Acoustic analyses revealed that the embodied shadowing group increased its pitch range after training in comparison with the audioshadowing group.

The second technique applied in this study was voice recording. Frisby (2004) called voice recording a key instructional technique in language teaching, and Teaching and Learning with ICT Frisby asserted that L2 learners could gain immense benefits from voice recording activities such as digital learning portfolios, digital storytelling, and web-based narration. In this study, the researcher assumed that through recording their own voice, listening to it, comparing it with the models provided by the native speakers, and recoding themselves, EFL learners could improve their command of English pronunciation.

The third technique applied in this study was video dubbing. Burston (2005) believed that video dubbing is the action of recording one's voice on a silent piece of video. In this study, the researcher assumed that through dubbing their voice on the motion picture, learners would have to tune up the pace of their speech with the lip movement of the actors, tune up their voice with the moods and body language of the actors, and copy the actors' native-speaking pronunciation. Overall, this study could be considered as a pioneer study in the area of teaching pronunciation since it investigated the impact of three novel state-of-the-art strategies to teach English pronunciation that might have not been applied for this purpose earlier. Moreover, this study focuses on teaching pronunciation strategies that improve learners' pronunciation implicitly.

To accomplish the objectives of this study, the researcher chose to formulate the following research questions:

Q1: Does video dubbing have any significant effect on the pronunciation proficiency of Iranian advanced English language learners?



Q2: Does audio recording have any significant effect on the pronunciation proficiency of Iranian advanced English language learners?

Q3: Does shadowing tasks have any significant effect on the pronunciation proficiency of Iranian advanced English language learners?

REVIEW OF LITERATURE

In the recent years, two major changes have taken place in the area of teaching pronunciation. One major transformation is the introduction of technology to the world of language teaching, and the increasing interest in the utilization of this modern opportunity, and the merits that its resources bring about to enrich second/foreign language instructional settings (Murphy, 1992). No doubt, pronunciation instruction is not an exception here; rather, it is the area that has embraced technology and its use in SLA/FLA more passionately than other areas since it demands voice and audio recording, playing, and analysis devices that did not exist earlier (Hardison, 2005; Murphy, 1992).

The second paradigm shift that has occurred in the teaching of pronunciation in recent years is that approaches toward teaching L2 pronunciation have changed substantially, shifting from the accurate production of sounds to broader aspects of communicative utterances (Richards & Renandya, 2002). This is backed up by Morley (1991), who held what is taking place in the recent history of SLA is a systematic sound use of priorities and redirecting them with more central focus on the critical significance of suprasegmentals. He, also, noted that the way the suprasegmentals is utilized to communicate meaning in discourse context has come to be highlighted.

To achieve advanced pronunciation skills, various techniques have been devised, implemented, and studied. Jones and Evans (1995) explored the role of voice quality in teaching pronunciation and proved that voice quality encompassed various phonological aspects; therefore, it can provide an insightful view for further pronunciation work. In their study, they applied a particular technique, in which voice quality was utilized in communicative practice to provide the participants with the opportunity to identify the meaningful aspects of suprasegmental pronunciation to check how they match the overall pattern of connected speech. Thompson (2000) explored current teaching models for intonation of questions and offered an alternative approach to teach the falling intonation in *wh*- questions, which gave priority to the speaker's communicative intentions. Believing that non-native speakers (NNSs), who are using English for international communication now outnumber their native speakers, and this matter has serious implications for ELT pedagogy, Jenkins (2002) put an attempt to empirically establish phonological norms and classroom pronunciation models for English as an International Language (EIL), in which intelligibility for non-native speakers rather than native speaker receivers was the primarily essential motivation by drawing data from non-native speakers interaction are provided to exemplify the kinds of empirical evidence, necessary to enable us to make informative claims about phonological intelligibility in English.

Derwing and Rossiter (2003), on the other hand, investigated the perceptions of 100 adult English-as-a-second-language learners from a variety of first language backgrounds concerning their major difficulties in pronunciation as well as the employed strategies in case of breakdown in communication. They discovered that the major identified pronunciation problems were segmental. Although their most commonly used strategies were paraphrasing, self-repetition, writing/spelling, and volume adjustment when they had not been understood. In another study, Herron, Hanley, and Cole (1995), held that improving the feedback quality of a computer-based system for pronunciation training needs much detailed and precise knowledge about the nature of actual mispronunciation in a language learners' speech. To provide this kind of information, they developed the necessary components for the automatic pronunciation error localization and correcting errors methods.

Despite all the above-mentioned and the very many other studies that have taken place to develop excellent techniques in teaching pronunciation, there are still many facets of teaching pronunciation that lurk in the dark, and are in need of further research and a broader scope of techniques to creep into the light. This study attempts to utilize three different SLA/FLA techniques as a means to teach pronunciation, and draw a comparison between them.



The results of academic systematic research in the area of second/foreign language pronunciation instruction, as discussed above, put great emphasis on segmental and suprasegmental phonological features of the language (Hardison, 2005). More increasingly, teachers of our time are motivated to pay attention to the pedagogical needs of language learners to develop and sequence the most appropriate language teaching materials in any area of language teaching to reach the most desirable results (e.g. Celce-Murcia, et al., 1996; Morley, 1991). This means that pronunciation should not be taught in isolation and as an independent component of the language, but it has to be incorporated in discourse, meaning, and communication. In this spirit, this study aimed at taking teaching pronunciation beyond the level of teaching vowels and consonants' articulation or teaching sentence stress in isolation. Yet, it intended to involve pronunciation with meaning, and swap explicit teaching of pronunciation with imitation of native-like pronunciation. Better put, in this study, figuring out the segmental and suprasegmental pronunciation of the material was a burden that is put on the shoulder of the learners. Moreover, three different techniques of practicing pronunciation are presented to the learners, which have rarely been applied to improve pronunciation, and whose effects have never been compared and contrasted earlier.

METHODOLOGY

PARTICIPANTS

The target sample of this study was made up of 75 advanced level EFL learners, all studying in Shokuh English Institute, located in Mashhad-Iran. The participants were homogenized both in terms of general English mastery.

The initial participants comprised of 120 language learners, who had to answer a general proficiency test of TOEFL. After scoring the papers, 75 participants were considered as the target sample, based on their scores on the test. The participants were adults – above 19 years of age – and their age distribution could be found in Table 1.

Table 1.
Age Distribution of the Sample

Age Range	Population
Between 19 and 25	24
Between 25 and 30	29
Between 30 and 40	18
Over 40	4

The learners in the sample were both male and female, and they had a wide range of occupations, varying from stay-home mothers to medical doctors. The participants, also, possessed various levels of education, whose distributions are illustrated in Table 2.

Table 2.
Education Distribution of the Sample

Education	Population
Doctorate	5
Master's	20
Bachelor's	37
Associate's	6
Undergraduate Diploma	7

In order to increase the level of generalizability of this study, the researcher chose not to extend any types of delimitations to the sample in terms of individual differences (e.g. learning styles, personality variables, socio-economic status, etc.)



INSTRUMENTS

Two instruments were utilized to collect the data in the pre-test and post-test phases of the study. The first instrument was the FCE test, taken from Quintana (2003). This test has four sections, each section covers one of the key skills of listening, reading, speaking, and writing, and each section shapes up 25 percent of the test scores. In this study, the researcher utilized the FCE test as an instrument to make sure the entire sample had almost equal knowledge of general English. Geranpayeh (2004, p. 22) noted that the composite reliability for the complete FCE test has been consistently measured about 0.92 since 2000. This means that the correlation between all parts of the test is high, reflecting a reliable examination.

The second instrument, applied in this study, was the *Pronunciation Assessment Scale*, which was developed by Yates, Zielinski, and Pryor (2008), in the Australian University of Macquarie, and the University of Melbourne. As the researcher had to measure the pronunciation proficiency of the learners in the pre-test and the post-test, and the learners were in the advanced level, the researchers applied the assessment scale. In order to maintain the psychometrics of this rubric, the researcher decided to pilot this scale on 20 advanced learners, who had similar characteristics with the sample. This scale contains 10 major criteria, which are commonly applied in the assessment of pronunciation, and the raters were supposed to assess each scale based on a 5-point Likert scale, which ranged from 1 (poor) to 5 (excellent).

It should be added that prior to the application of this scale, the researcher checked its psychometric features in a pilot study on 38 learners, who had features that closely resemble those of the sample. In the pilot study, the internal consistency of this scale, measured by the Cronbach's Alpha analysis, was found to be 0.639.

MATERIALS

The materials used in the course of the treatments of this study were 36 short video clips (each between 20 to 50 seconds), which were extracted from the American series, produced in the recent years. These series were Person of Interest, The Good Wife, The Mentalist, 90210, Dexter, Southland, Grey's Anatomy, How I Met Your Mother, Scandal, and Magic City, and the episodes from which the clips were extracted were all produced and released in 2010 or afterward.

The researchers personally extracted the clips and chose the parts where conversations were fluent, and comprehensible. Besides, to enable the learners to easily understand the clips and reproduce them orally in the limited time of the class, the researchers made sure that all the conversations centered on routine life topics and did not contain hard and scholarly words. It should also be mentioned that the researchers chose the extracts only from the parts of the series where native-born American actors were talking. They did this to make sure the learners were exposed to standard American accent all throughout the treatment. The researchers used the same material in the treatment of all the groups, and the amount of the material that all the groups worked on were exactly similar, as well. Hence, it could be concluded that the only differences among the treatments were the types of instruction and practice that the learners received.

DATA COLLECTION PROCEDURE

Initially, the researchers selected 120 adult advanced learners, studying English in Shokuh Language Institute in Mashhad. For homogenizing the participants in terms of language proficiency, Barron's TOEFL Paper Test (version 15) was utilized. After scoring the papers, the learners with scores within one standard deviation from the mean were selected as the homogenous learners. This shrank the sample to 75 learners. Then, the sample was divided into three 20-member treatment groups and one 15-member control group, and the pronunciation pre-test was administered to them. In order to promote reliability in scoring, each learner was scored via Pronunciation Assessment Scale (Yates, Zielinski, & Pryor, 2008) by two different raters. After checking the inter-rater reliability, the researchers made sure of the inter-group homogeneity within the three groups both in terms of general English and pronunciation, and then, the treatment began.

The first treatment group practiced shadowing. In this group, the learners watched each clip several times in order to make sure they comprehended all the words and phrases. Then, they would use the subtitle of the clips as the script, and read out loud the script as they were watching the clip and hearing the actors pronouncing the sentences. Learners would repeat each clip at least 20 times and shadowed the actors/actresses as they were listening to them. The researcher pointed out to the learners that they needed to match the



pronunciation of the words and the intonation of the sentences with what they were hearing in the clip. In other words, learners listened to the sentences the actors were uttering and read them at the same time, trying to put their mouth in harmony with their ears.

The second treatment group practiced voice recording. Learners of this group were first asked to listen to each clip several times to make sure they understood everything in the conversations. Then, they used the script of the clip as their text, read through it, and recorded their voice. These learners were encouraged to listen to the pronunciation of the words in the clip very carefully, attend to the stress and intonation of the sentences, and copy the actors while recording their voice. They were also encouraged to go back to the clip and listen to words or sentences they fail to pronounce while recording their voice. Each learner was encouraged to attempt at least 10 times to record and rerecord themselves until they make sure they have recorded their best performance.

The third group practiced video dubbing. In this group, learners were asked to watch the clips a few times so that they make sure they fully perceived all the phrases and statements. Then, they used the subtitles as the script, and with the help of Sony Vegas Pro Software, they recorded their voice on the clip. In fact, they had to act as voice actors/dub artists, and dub their voices on the clip. They were supposed to keep the exact same pace of speech as actors, since they were supposed to make the dubbing look realistic. Overall, all the learners in the three groups worked on the same material, and the nature of their practices was the mechanical listening and repeating the activities. The learners in the shadowing group repeated the lines simultaneously as they listened, whereas the learners of the other two groups repeated with delay. Learners in the voice recording group recorded the lines at their own pace and without noticing the gestures of the actors, whereas the learners of the video dubbing group were pushed to tune up their speed with the actors and had all their gestures and body language at their disposal while repeating the lines. It should also be added that to prevent from the chaos of the voices of 20 learners, the treatment classes were held at the large computer lab of the institute, where the learners sat with ample distance from each other and used computers and headphones to do their practices.

The whole treatment was offered in 10 sessions of 90 minutes, which makes the total period of the treatment in this study to be 15 hours. The classes met twice a week, so the entire treatment took 5 weeks. Table 3 contains the summary report of the treatment that was offered to the three groups.

Table 3.

Overview of the Material Covered in the Treatments

Session	Session Period	No. of Extracts	Extracts' Length	Series
1	90 minutes	3	40 to 50 seconds	Person of Interest
2	90 minutes	3	40 to 50 seconds	The Good Wife
3	90 minutes	3	40 to 50 seconds	The Mentalist
4	90 minutes	4	25 to 35 seconds	90210
5	90 minutes	3	40 to 50 seconds	Scandal
6	90 minutes	4	25 to 35 seconds	Dexter
7	90 minutes	3	40 to 50 seconds	Southland
8	90 minutes	4	25 to 35 seconds	Grey's Anatomy
9	90 minutes	5	20 to 30 seconds	How I Met Your Mother
10	90 minutes	4	25 to 35 seconds	Magic City

One week after the treatment was over, the researchers administered the post-test. The pronunciation post-test took 10 minutes, and involved two sections: reading a passage/conversation text book of *New American Streamline Connections* (Hartley & Whiney, 2003) for 3 minutes, and speaking on topics of general interest for 7 minutes. In order to promote reliability in scoring, each learner was scored via Pronunciation Assessment Scale (Yates, Zielinski, & Pryor, 2008) by two different raters. After checking the inter-rater reliability of the scores concerning the views of the two raters, the researchers applied inferential statistics to compare the results of the three groups and answer the research questions.



RESEARCH DESIGN AND RESEARCH VARIABLES

This study was a quasi-experimental study, with three treatment groups and a control groups. Since this study only enjoys one dependent variable, it is a univariate study, and the impact of the treatments were measured and compared with a pre-test post-test design. The variables in study are as follows:

Independent Variable(s): video dubbing, shadowing, and voice recording

Dependent Variable(s): the progress of learners' pronunciation mastery

RESULTS

In order to promote reliability in scoring, each learner was scored via Pronunciation Assessment Scale (Yates, Zielinski, & Pryor, 2008) by two different raters. After checking the inter-rater reliability of the scores by the two raters, the researcher applied inferential statistics to compare the results of the three groups and answer the research question.

Table 4.
Descriptive Statistics for the FCE Test

	N	Mean	Std. Deviation	Variance
FCE Test	71	61.9859	12.75762	162.757

As Table 5 suggests, the mean of the homogenized sample was 60.71, and the standard deviation was 8.1.

Table 5.
Descriptive Statistics of FCE Test after Homogenization

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		
						Statistic	Std. Error	Ratio
FCE Test	75	49.00	75.00	60.71	8.10	.479	.309	1.55

After making sure that the learners of the sample had almost the same level of English mastery, the researcher divided them into four groups, one control group and three experimental groups on a random basis. Table 6 illustrates that the means of the three groups were 61.2, 61.6, and 58.35 while the mean for control group was 57.4.

Table 6.
Descriptive Statistics of FCE Test within Groups

	N	Mean	Std. Deviation	Skewness		
				Statistic	Std. Error	Ratio
Voice Recording	20	61.20	8.02365	.952	.512	1.86
Shadowing	20	62.60	7.89670	-.102	.512	-0.20
Video Dubbing	20	58.35	8.20959	.821	.512	1.60
Control Group	15	57.4	5.36960	.695	.512	1.20

The Skewness ratio of the three treatment groups were 1.86, -0.2, and 1.6 and Skewness for control group was 1.20. This suggested that the data distribution was normal. Hence, in order to check the inter-rater homogeneity between the four groups, the researcher administered the Analysis of the Variance. As it could be observed in Table 7, ANOVA (df between groups: 3, df within groups: 71, F: 1.45, and sig: 0.243 > 0.05) did not report any statistically significant differences between the FCE scores of the three groups. Hence, it could be concluded that the learners in the four groups enjoyed almost equal levels of English.



Table 7.
ANOVA Results

	Sum of Squares	Df	Mean Square	F	Sig.
Between Groups	187.633	3	93.817	1.450	.243
Within Groups	3688.550	71	64.711		

After the administration of the treatments to the three treatment groups, a pronunciation post-test was administered to the sample. Table 8 signifies that the statistical mean for the first rater was 24.75, and for the second rater, it was 25.05. The standard deviations were 6.18 and 5.95, respectively.

Table 8.
Descriptive Statistics of Pronunciation Post-Test within Raters

	N	Mean	Std. Deviation	Skewness		
				Statistic	Std. Error	Ratio
Rater 1	60	24.75	6.18534	.276	.309	0.89
Rater 2	60	25.05	5.95584	.197	.309	0.64

Since the distribution of the scores of both raters were not markedly Skewed, the researcher applied the parametric correlation of Pearson in order to check the inter-rater reliability. As Table 9 suggests, the p value for the correlation was $0.000 < 0.05$, and the correlation coefficient was 0.934. This proved that the two raters' scores enjoyed an ample degree of inter-rater reliability.

Table 9.
Correlations within Raters on Post-Test

		Rater 1	Rater 2
Rater 1	Pearson Correlation	1	.934
	Sig. (2-tailed)		.000
	N	75	75
Rater 2	Pearson Correlation	.934	1
	Sig. (2-tailed)	.000	
	N	75	75

Since the two raters' scores enjoyed inter-rater reliability, the researcher used their average to determine participants' scores in the post-test pronunciation test. As depicted in Table 10, the shadowing group, with the mean of 28.75, staged the best performance in the post-test, followed by video dubbing group, with the average of 25.85. The voice recording group, with the average of 20.1, achieved the poorest scores. The Analysis of the Variance was administered in order to determine whether or not the differences between the means of the three groups in the post-test were significant.

Table 10.
Descriptive Statistics of Pronunciation Posttest within Groups

	N	Mean	Std. Deviation	Skewness		
				Statistic	Std. Error	Ratio
Voice Recording	20	20.10	3.09754	-.110	.512	-0.36
Shadowing	20	28.75	4.86799	-.345	.512	-1.12
Video Dubbing	20	25.85	6.04609	.054	.512	0.11
Control group	15	19.21	3.7656	.265	.512	.365

As depicted in Table 11, the p value of the ANOVA was $0.000 < 0.05$ (df between groups: 2, df within groups: 57, F: 16.65), hence, the Analysis of the Variance did report a meaningful difference between the scores of the three groups. The null hypothesis of this study was rejected, and it was confirmed via data analysis that a statistically significant difference did exist between the impact of shadowing, voice recording, and video dubbing on the improvement of learners' pronunciation progress.



Table 11.

ANOVA on Pronunciation Post-Test

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	775.300	3	387.650	16.650	.000
Within Groups	1327.100	71	23.282		
Total	2102.400	74			

In order to pinpoint the place of the mean differences that ANOVA reported, the Scheffe test was administered. As it could be depicted in Table 12, the difference between the means of shadowing and video dubbing groups was not significant (sig: 0.174 > 0.05). However, the differences between the means of voice recording and shadowing (0.000 < 0.05) and video dubbing (0.002 < 0.05) were significant.

Hence, it could be argued that the voice recording, with the average of 20.1, was less effective in the development of learners' pronunciation than the other two methods. Nonetheless, although the mean of shadowing group was 28.75 and higher than the mean of video dubbing group (25.85), shadowing did not have a stronger impact in improving learners' pronunciation than video dubbing since Scheffe test did not report any significant differences between the two means.

Table 12.

Scheffe Test on Post-Test Scores

(I) Grouping	(J) Grouping	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
Voice Recording	Shadowing	-8.65	1.52	.000	-12.4852	-4.8148
	Video Dubbing	-5.75	1.56	.002	-9.5852	-1.9148
	Control group	-.89	1.32	.36	-11.2563	-3.3366
Shadowing	Voice Recording	8.65	1.52	.000	4.8148	12.4852
	Video Dubbing	2.90	1.52	.174	-.9352	6.7352
	Control group	-9.54	1.46	.032	6.6852	9.3652
Video Dubbing	Voice Recording	5.75	1.52	.002	1.9148	9.5852
	Shadowing	-2.90	1.52	.174	-6.7352	.9352
	Control group	-6.64	1.36	.046	5.3625	8.6663
Control group	Shadowing	9.54	1.46	.032	-9.3652	-6.6852
	Video Dubbing	6.64	1.36	.046	-8.6663	-5.3625
	Voice recording	.89	1.32	.36	-3.3366	-11.2563

Overall, based on the data analysis of this study, voice recording proved to be a less effective technique in developing learners' pronunciation than video dubbing and shadowing. Moreover, the data analysis of this study proved that shadowing and video dubbing were both equally effective in boosting learners' pronunciation proficiency.

DISCUSSION AND CONCLUSION

This study proved that although shadowing and video dubbing are more effective means of teaching pronunciation than voice recording, they are equally effective in improving learners' pronunciation proficiency. Burston (2005) proposed that to gain native-like mastery in pronunciation, the mechanical repetition of sounds is not sufficient. Phonetic accuracy, stress placement, intonation, rhythm, timing as well as paralinguistic voice features (e.g., surprise, sadness, joy, impatience, frustration, etc.) need to be fully practiced on. It could be argued that voice recording did not provide learners with a chance to do so. Although learners listened to the actors acting out the conversation several times, they would almost lose the mood while they were recording themselves. However, when they were shadowing, they would simultaneously go with the flow of the conversations that they were hearing and they hence practiced on stress placement, intonation, and paralinguistic voice features. The video dubbing group, as well, would synchronize their timing with the actors since they were asked to make dubbing look real, and the mood and the body language of the actors guided them in abiding



by the paralinguistic features exactly as they were required. That is the reason why learners in the shadowing and video dubbing groups improved more noticeably than those in the voice recording group.

Studies such as de Bot (1983), Hardison (2005), and Weltens & de Bot (1984) highlighted that a visual display of a pitch contour is an essentially effective tool for teaching L2 learners to produce more native-like prosody. Similarly, learners who practiced dubbing on videos experienced stronger degrees of progress compared to those who practiced with voice merely. Hence, it could be argued that visual aids, both in forms of pitch contour and video context – the settings of the scenes, the clothing of the actors, the body language and facial mimics – could promote learners' pronunciation skills.

Brooke (2003) found that working with videos strikingly increased learners' motivation to do the task. Burston (2005) backed this up by stating since accurate lip synchronization is involved in video dubbing; it requires students to pay particular attention to timing, gestures and facial expressions. This makes it much easier for learners to put themselves into the persona of the characters whose voices they are dubbing, and this feeling of identification with characters increase their motivation on the one hand, and increases their chance of giving more natural native-like delivery on the other hand. The findings of this study are in line with Pennington's (1996), who emphasized that phonological differences can signal differences at several different levels of the language, for instance, differences in lexical meaning, grammatical meaning, and utterance meaning, and raising noticing on how pronunciation participates in the meaning increases chances of pronunciation progress.

The findings are also in line with Tannen (1989), who proposed that shadowing, as an interpersonal involvement strategy, is a natural methodology of learning, since humans intuitively are built to imitate. This could be backed up not only by the results of this study, but also with that of Acton (1984), who used shadowing to defossilise the pronunciation mistakes of foreigners, having lived in the United States for a long time.

Nonetheless, the findings are different from the findings of Firth (1992), who proposed that pronunciation is the skill that develops gradually in its nature. Shadowing and video dubbing, in this study, nonetheless, proved to improve learners' pronunciation in a short term.

Moreover, the findings of this study do not approve Muller's (1980) findings, which indicate videos take a fairly long time to cast their full impact of language learners' English skills. In only 5 weeks' time, video dubbing in this study managed to noticeably promote learners' pronunciation skills in English language. This study could also propose a solution to Ho (2003). He proposed that the problem with communicative teaching methodology today is lack of chances for speaking practice in authentic environments, and that is why learners fail to develop proper commands of production abilities, particularly pronunciation. Through shadowing and video dubbing, learners would have a chance to practice their pronunciation in authentic environments of the series and movies, as the video provides them with the setting, mood, and motivation to carry on producing authentic material in authentic context. This is also in line with the findings of Crouse and Noll (1980), who discovered creating English dubbed videos as classroom projects would greatly impact learners' fluency.

PEDAGOGICAL IMPLICATIONS

The findings of this study can suggest some implications, useful for teaching pronunciation. One suggested technique is the application of video dubbing technique as well as the shadowing technique in improving learners' pronunciation. Through encouraging language learners to apply these techniques in a variety of listening and speaking interactive activities, language teachers can get better results. Moreover, English teachers at private language institutes could also utilize this technique to improve their learners' pronunciation.

Another implication of this study could be for those who do not have time to take classes or do not have access to English classes, but intend to self-study for learning English. Since both shadowing and video dubbing are techniques that do not need a teacher to administer, and could easily be used while self-studying, these learners could make the best out of them to master English segmental and suprasegmental pronunciation.

In another aspect, teachers could utilize video dubbing and shadowing as homework for their learners of different levels in order to have them practice on their pronunciation at home. Traditionally, voice recording was the technique that teachers applied in order to set assignments for their learners, but this study proved that video dubbing and shadowing were more effective assignments.



With a brisk glance at the material available in the market, it could easily be perceived that not too many books and videos have been developed as the standard material for shadowing and video dubbing. It would be a great idea for material developers to develop some audio-visual material for learners of different levels with appropriate speech pace and level of difficulty so that all learners with all levels of proficiency could enjoy these instructional techniques.

Another area of language teaching that video dubbing and shadowing could be useful is in the training of anchormen and television hosts in English language. These techniques could be widely used on a variety of material for all the programs whose objective is to train non-natives of English language to be anchormen or showmen who could speak English without accent. Moreover, programs that aim at training simultaneous interpreters could utilize video dubbing and shadowing to assist the simultaneous interpreters they develop in becoming more proficient in their pronunciation. And finally, these techniques have to be included in teacher education programs. While student teachers are being trained to become EFL teachers, they should learn video dubbing and shadowing as techniques that noticeably improve learners' pronunciation so they could use it as they teach pronunciation or prescribe pronunciation homework later on when they are teachers in service.

REFERENCES

- AMPE Fact Sheets. (2002). *Fact Sheets: What is Pronunciation?* Available at:
<http://www.nceltr.mq.edu.au/pdamep>
- Arguelles, A. (2010). *Guide to Autodidactic Foreign Language Study*. Retrieved on February 24th 2014 from http://www.foreignlanguageexpertise.com/foreign_language_study.html
- Brooke, S. (2003). Video production in the foreign language classroom: Some practical ideas. *The Internet TESL Journal*, 9(10). Retrieved 3 March 2008 from <http://iteslj.org/Techniques/Brooke-Video.html>.
- Burston, J. (2005). Video Dubbing Projects in the Foreign Language Curriculum. *CALICO Journal*, 23 (1), 79-92.
- Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (1996). *Teaching Pronunciation*. Cambridge: Cambridge University Press.
- Chomsky, N. (1967). *Appendix to Lenneberg, E. H. Biological Foundations of Language*. New York: Wiley, 1967.
- Crouse, G. B., & Noll, A. (1980). Using video tapes to teach foreign languages. *Foreign Language Annuals* 13(5), 391-93.
- De Bot, K. (1983). Visual feedback of intonation 1: Effectiveness and induced practice behavior. *Language and Speech*, 26(4), 331-350.
- Derwing, T. M., & Rossiter, M. J. (2003). The effects of pronunciation instruction on the accuracy, fluency, and complexity of L2 accented speech. *Applied Language Learning*, 13, 1-17.
- Firth, S. (1992). Pronunciation syllabus design: A question of focus. In P. Avery & S. Ehrlich (Eds.), *Teaching American English Pronunciation* (pp. 173-183). Oxford: Oxford University Press.
- Frisby, R. (2004). *Sound recording for language teaching*. Middlesex University Language Centre.
- Goodwin, J. M. (2004). Imitative conversations. *Demonstration presented at the international TESOL Conference*, Miami, FL.
- Hamada, Y. (2018). Shadowing for Pronunciation Development: Haptic-Shadowing and IPA- Shadowing. *The Journal of Asia TEFL*, 15(1), 167-183.
- Hardison, M. D. (2005). Contextualized Computer-based L2 Prosody Training: Evaluating the Effects of Discourse Context and Video Input. *CALICO Journal*, 22(2), 175-190.
- Hartley, B., & Viney, p. (1985). *American Streamline Connections*. Oxford University Press: Oxford.
- Herron, C., Hanley, J., & Cole, S. (1995). A comparison study of two advance organizers for introducing beginning foreign language students to video. *Modern Language Journal*, 79(3), 387-394.
- Ho, Y. K. (2003). Audiotaped dialogue journals: An alternative form of speaking practice. *ELT Journal*, 7(3), 269-277.
- Jenkins, J. (2002). A sociolinguistically based, empirically researched pronunciation syllabus for English as an International Language. *Applied Linguistics*. 23(1), 83-103.
- Jones, R. H., & Evans, S. (1995). *Teaching pronunciation through voice quality*. Retrieved online on September 17, 2011 From <http://eltj.oxfordjournals.org/content/abstract/html>



- Kissling, E. M. (2013). Teaching pronunciation: is explicit phonetics instruction beneficial for FL learners? *The Modern Language Journal*, 97, 720-744. [10]
- Orion, G. F. (2012). Pronouncing American English: Sounds, stress, and intonation. Boston: Heinle.
- Kumaravadivelu, B. (2001). Toward a Postmethod Pedagogy. *TESOL Quarterly*, (35)4, 537-560. doi:10.2307/3588427,
- Larsen-Freeman, D. (2000). Grammar: Rules and reasons working together. *ESL/EFL Magazine*, 4, 10-12.
- Morley, J. (1991). The Pronunciation Component in Teaching English to Speakers of Other Languages, *TESOL Quarterly*, 25, 3, 481-520.
- Mueller, G. (1980). Visual Contextual Clues and Listening Comprehension: An experiment. *Modern Language Journal*, 64, 335-40
- Murphy, J. M. (1992). From Caterpillar to Butterfly: Using Modern Technology within the Current Paradigm of ESL Instruction. *TESL Canada Journal/Revue TESL Du Canada*, 9(2).
- Pennington, M. C. (1996). *Phonology in English language teaching*. Addison Wesley Longman: Harlow.
- Quintana, J. (2003). *PET Practice Tests*. Oxford University Press: Oxford.
- Reis, J., & Hazan, V. (2011). Speechant: a vowel notation system to teach English pronunciation. *ELT Journal*, 66, 156-165.
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge University Press.
- Richards, J.C., & Rodgers, T. (2001). *Approaches and methods in language Teaching*. Cambridge: Cambridge University Press.
- Saito, K. (2014). Experienced teachers' perspectives on priorities for improved intelligible pronunciation: The case of Japanese learners of English. *International Journal of Applied Linguistics*, 24, 250-277. [13]
- Saito, K. (2015). Experience effects on the development of late second language learners' oral proficiency. *Language Learning*, 65, 1-32.
- Tanen, D. (1989). *Talking voices, repetition, dialog and imagery in conversational discourse*. Cambridge: OUP.
- Teaching and Learning with ICT. (2011). Retrieved on January 22nd 2014 from <http://Education.qld.gov.au/smartclasses>
- Weltens, B., & De Bot, K. (1984). Visual feedback of intonation II: Feedback delay and quality of feedback. *Language and Speech*, 27, 79-88.
- Yamane, N., Shinya, M., Teaman, B., Ogawa, M., & Akahoshi, S. (2019). Mirroring beat gestures: Effects on EFL learners. *Proceedings of the 19th International Congress of Phonetic Sciences* (pp. 3523- 3527). Canberra, Australia: Australasian Speech Science and Technology Association Inc.
- Yates, L., Zielinski, B. & Pryor, E. (2008). The assessment of pronunciation and the new IELTS Pronunciation scale. *IELTS Research Reports*, 12.