rEFLections
Vol 31, No 3, September - December 2024

Effects of Video-based Shadowing on Suprasegmental Features: EFL Learners' Pronunciation Performance and Attitudes

PHAN THI TUYET VAN*
School of Foreign Languages, Can Tho University, Vietnam
LY KIM LIEN
School of Foreign Languages, Can Tho University, Vietnam
NGUYEN DUY KHANG
School of Foreign Languages, Can Tho University, Vietnam
Corresponding author email: pttuyetvan@ctu.edu.vn

Article information
Article history:
Received: 19 Sep 2023
Accepted: 30 Aug 2024
Available online: 10 Sep 2024

Keywords:
Video-based shadowing
Pronunciation
English suprasegmental features

Abstract
The study was conducted to investigate the effects of video-based shadowing on pre-intermediate EFL learners' pronunciation performance of suprasegmental features as well as explore their attitudes toward the technique. The study employed three instruments including pre-and post-tests, a questionnaire, and semi-structured interviews, to triangulate the data. The results showed that participants made an improvement in their pronunciation performance of suprasegmental features, especially intonation. In other words, video-based shadowing has positive effects on the participants' pronunciation performance of suprasegmental features. Besides, they had positive attitudes toward video-based shadowing and outlined some benefits of using the technique. Furthermore, the participants' preference for how to practice video-based shadowing was also discovered.

INTRODUCTION

Nowadays, English has become a global language in which citizens around the world can exchange information on a daily basis. In fact, English skills play an important role in employment prospects in non-native English-speaking countries, especially speaking skills (Clement & Murugavel, 2018). Therefore, everyone needs to master English communication skills to have a chance to succeed in their career. To achieve this goal, the very first thing that every EFL/ESL learner should do is to develop their English pronunciation (Garrigues, 1999), as it is one of the most essential components of communicative competence (Morley, 1991). Gilakjani (2011) explained that if the speakers do not pronounce English correctly and clearly, their messages cannot be comprehensible to listeners, no matter how excellent the speakers are at grammar, excluding the cases of using sign language and other forms of nonverbal communication. Similarly, pronunciation causes many difficulties for learners and becomes one of the most difficult parts of language teaching and learning, in which mispronunciation can lead to misunderstandings. (Bøhn & Hansen, 2017; Pedrazzini, 2016, Rahimi & Ruzrokh, 2016).

According to Burns and Claire (2003), English pronunciation includes two main types of features: segmental and suprasegmental features (or prosodic features). The former features refer to English phonemes such as vowels (monophthongs and diphthongs) and consonants (voiced and voiceless). Meanwhile, the latter are units extending over more than one sound in an utterance, such as linking or stress. Levis (2018) pointed out that intelligibility in pronunciation (the speaker produces sound patterns that are recognizable as English) plays a central role in developing English communicative skills. Moreover, Burns and Claire (2003) suggested that to boost learners' intelligibility, it is necessary to develop their pronunciation concentration on higher levels of suprasegmental features, including stress, linking, and intonation. Nevertheless, English suprasegmental features are pretty challenging for EFL learners in Vietnam due to their limitations of knowledge about distinct phonological differences between English and their mother tongue and the lack of proper practices. According to Dang (2017), suprasegmental features are usually underestimated in many institutions in Vietnam when it comes to teaching and learning pronunciation. In this context, these features may cause trouble for both teachers and learners with difficulties in the acquisition of pronunciation. One possible solution to this problem is to use shadowing techniques with different steps of listening and imitating the target language so that learners can acquire pronunciation naturally. The concept of shadowing was first introduced by Cherry (1953), and it was originally a training technique for simultaneous interpreters. The use of shadowing in EFL classrooms was officially published in an academic paper by Tamai (1997, as cited in Hamada, 2019). At that time, shadowing as a technique for training interpreters was imported as a bottom-up technique for teaching listening skills in the context of Japan. Imitating the language input helps learners use the phonological loop to enhance their phonological coding and speech perception skills (Lambert et al., 2016). Recently, shadowing has become popular not only in other Asian countries (Hsieh et al., 2013; Omar & Umehara, 2010) but also in other continents (Foote & McDonough, 2017; Martinsen et al., 2017). Noticeably, recent research has indicated that shadowing can positively impact learners' pronunciation performance of the suprasegmental features (Hamada, 2018; Niimoto, 2022).

It can be seen that the number of related studies on shadowing in Vietnam, especially in the Mekong Delta is still relatively limited, and in particular, the main concern of those studies was not directly linked to suprasegmental features (Dang, 2020; Le et al., 2022). Thus, it is necessary to launch an investigation into video-based shadowing to find out its effects on Vietnamese EFL learners' pronunciation performance of suprasegmental features and their attitudes towards the technique.

LITERATURE REVIEW

Shadowing and video-based shadowing

Shadowing is a metaphor that connotes one's shadow following that person and copying his or her every single movement. Manseur (2015) stated that shadowing mimics the heard auditory input as soon as possible or repeats the exact words of an auditory input of the target language. Notably, Sumarsih (2017) considered shadowing an active and highly cognitive activity in which shadowers imitate the speech while simultaneously listening attentively to

the upcoming input. Shadowing was classified depending on the shadower's purpose. Murphey (2001) classified shadowing into three main types, namely complete shadowing, selective shadowing, and interactive shadowing. His way of classification is mainly based on the amount of repeated information.

In this study, video-based shadowing is synonymous with visual-auditory shadowing, which was mentioned for the first time by Nakayama (2011, as cited in Nakayama & Mori, 2012). Obviously, it is the combination of two other types of shadowing, namely visual shadowing and auditory shadowing. According to the author, auditory shadowing is a task where learners only listen to the auditory input and then repeat it immediately. In contrast, visual shadowing is an online read-aloud task where learners read the visual input (a sentence) aloud, synchronized with the auditory input's speed. In other words, for video-based shadowing, learners start with auditory shadowing first by listening to and vocalizing the auditory input (Mori, 2011; Thomson & Derwing, 2015). After that, they begin visual shadowing by reading the visual input on the screen aloud.

English pronunciation features and suprasegmental features of pronunciation

It is vital for learners to pay attention to pronunciation features to improve their pronunciation (Burns & Claire, 2003). The features of English pronunciation are illustrated in Figure 1 below.

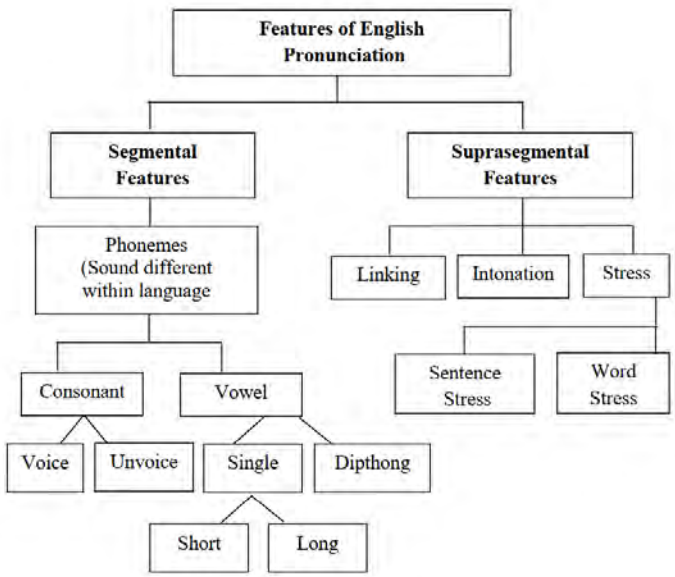

Figure 1 Features of English pronunciation (adapted from Pourhosein, 2012)

As in Figure 1, English pronunciation consists of two main features: segmental features and suprasegmental features. The former features refer to English phonemes, including vowels (monophthongs and diphthongs) and consonants (voiced consonants and voiceless consonants), while the latter features indicate wider aspects beyond the segmental features. Suprasegmental features include linking, intonation, and stress (word stress and sentence).

Regarding linking, Alameen and Levis (2015) referred to the term linking to "situations in which the ending sound of one word joins the initial sound of the next" (p. 162). They stated that linking is functional to "make two words sound like one without changes in segmental identity" (p. 162). Linking aims to avoid articulatory breaks at word boundaries (Allerton, 2000). Regarding the forms of combination for linking, Ashton and Shepherd (2012) classified linking into three main types, including consonant-to-vowel linking (C-V) occurs when the final consonant of a word connects to the initial vowel of the next word (Hieke, 1984), vowel-to-vowel linking (V-V) takes place when a word ending in a high or mid-tense vowel, then followed by a word beginning with a vowel (Celce-Murcia et al., 2010), and consonant-to-consonant linking (C-C) happens when the two identical consonants meet at word boundaries, they are pronounced as one slightly prolonged sound, such as *red dress* [redres].

About stress, this important suprasegmental feature refers to the prominence a speaker gives to certain syllables of a word or even certain parts in an utterance (Roach, 1991) and plays a significant role in speakers' and listeners' mutual intelligibility (Burns & Claire, 2003). The prominence is the combination of different factors, including increased loudness, vowel length, and changes in pitch and quality (Trujillo, 2017). Burns and Claire (2003) classified stress into word stress and sentence stress. First, word stress is a crucial factor in proper pronunciation and English communication because misplacing stress in a word often results in miscomprehension or changing the meaning of the word (Harmer, 2007). Both native speakers and L2 listeners pay special attention to the primary stress but not the weaker forms (secondary stress or unstressed syllables) to capture the word's meaning (Essberger, 2008). However, unlike other languages, word stress in English is arbitrary since stress can fall on virtually any syllable (Collins & Mees, 2006). Second, sentence stress is traditionally defined as stress on the sentence level (Gimson, 1980). The main function of sentence stress is not only highlighting semantically important words but also forming the rhythm of an utterance. However, not all word stresses are phonetically realized in an utterance. According to Kingdon (1958), content words normally receive stress on the utterance level because they deliver major semantic information and, therefore, require listeners' attention, while function words do not.

Regarding intonation, Burns and Claire (2003) described intonation as "the melody of the language" because the speaker's voice rises or falls depending on the context and meaning when making a conversation (p. 7). Furthermore, Betti and Al-Jubouri (2015) viewed intonation as one of the important suprasegmental features that refers to the pitch variation of the voice making an utterance. Similarly, intonation is considered the backbone of English pronunciation (Bailey & Nunan, 2005), because it plays a significant role in conveying the message (Low, 2015) and achieving intelligibility in communication (Burns & Claire, 2003). From a different perspective, Wells (2006) and Roach (2009) presented four main functions of intonation: attitudinal, accentual, grammatical, and discoursal. First, the attitudinal function helps the speaker express certain kinds of feelings, attitudes, and emotions, such as surprise, pleasure, anger, interest, boredom, etc. Second, the accentual function is also known as the focusing or informational function because it indicates that the placement of stress is somewhat determined by intonation (Betti & Al-Jubouri, 2015). In other words, the positions of the tonic stress can be changed based on the speaker's purpose of emphasis (Betti & Hasan, 2020). Another function of intonation is to help listeners distinguish between statements and questions. In addition,

in conversation, discourse functions signal what kind of response the speaker expects from the listener.

Three components of attitudes

Attitude plays a significant role in determining the ultimate level of success in one's learning process, especially language learning (Genc & Aydin, 2017). According to Ajzen (2001), attitude is “a summary evaluation of a psychological object” in the dimension of good versus bad, likable versus dislikeable, and the like (p. 28). Eagly and Chaiken (1993) defined attitude as a psychological tendency to evaluate objects with some degree of liking and disliking. Moreover, attitude refers to a disposition or tendency to respond with some degree of favorableness or unfavorableness to psychological objects (Fishbein & Ajzen, 1974). Eagly and Chaiken (1993) provide a multi-dimensional view of attitudes encompassing three different components, including cognitive, affective, and behavioral components, which are applied in this study to analyze the attitudinal factors.

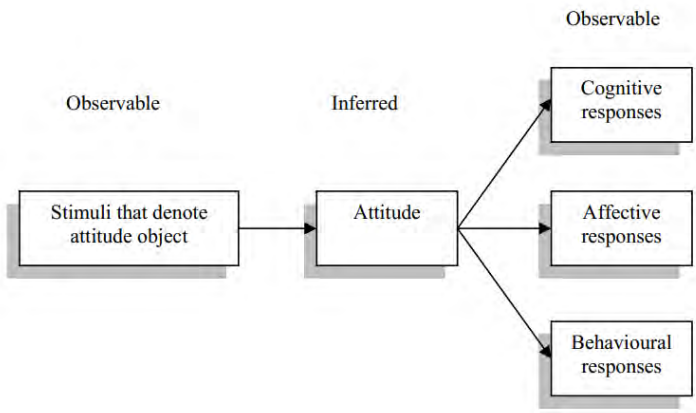

Figure 2 The multi-component model of attitudes (Eagly & Chaiken, 1993), adapted from Hassanein (2015, p. 50)

Related studies on the effects of shadowing on EFL learners' pronunciation performance of suprasegmental features

According to Nguyen and Dao (2018), learners were able to conduct the mechanism of video-based shadowing with the acts of imitation and reading on the screen to practice pronunciation. Additionally, a major difference between the two languages was that "syllables" would cause some trouble for Vietnamese learners. Le et al. (2022) conducted a study to shed light on the impact of the shadowing technique with the aid of spoken text features by Google Text-to-Speech on removing learner's flat tones as well as achieving basic English intonation in text. The results analyzed by Speech Analyzer showed that the participants benefited from the training in terms of intonation in the text. In Dang's (2020) study, the results illustrated that participants had a positive attitude toward shadowing, with about 80% of the participants agreeing that shadowing was a useful technique. Some common difficulties were also found related to using video-based shadowing, including the speaker's speaking rate, uninteresting topics, and unknown vocabulary. These difficulties were unavoidable because Vietnamese EFL

learners tend to pronounce words separately in sentences under the influence of their mother tongue (Nguyen & Ingram, 2004). The failure to link sounds together becomes a very typical problem for most Vietnamese learners in gaining normal natural speech. Additionally, the passive learning style was shaped by "the Confucian heritage culture" (Tran, 2013, p. 71).

Sugiarto et al. (2020) found that the shadowing technique had a positive and significant impact on the experimental group's pronunciation, at different levels, including segmental features (monophthongs, diphthongs, triphthongs, semi-vowels, consonants, and consonant cluster sounds) and suprasegmental features (strong and weak forms, linking, word stresses, sentence stresses, pitch, and intonation). In another empirical study by Niimoto (2022), the effect of shadowing training indicated that shadowing practice yielded statistically significant gains concerning suprasegmental productions and listening skills, while there was no significant difference in segmental productions and comprehensibility. As for learners' attitudes towards shadowing, Angel and Erika's (2018) findings showed that most participants felt more confident after using shadowing. Besides, they also felt more motivated to improve their speaking skills with shadowing. In Mıcık's (2020) study about the effects of using video-based shadowing on comprehensibility, pronunciation (individual sounds, intonation, and speech rate), and learners' attitudes towards video-based shadowing practices, the findings indicated that participants had a positive attitude towards video-shadowing. It can be implied that video-based shadowing had positive effects on the participants' pronunciation performance of suprasegmental features (Hamada, 2018; Sugiarto et al., 2020).

In general, although the shadowing technique has drawn the attention of researchers, it still appears to be a new technique that needs exploring in the context of Vietnam. Finally, it is also important to examine learners' attitudes towards video-based shadowing.

METHODOLOGY

A quasi-experimental study with a one-group pre-test and post-test design employed mixed methods to obtain both quantitative and qualitative data for examining the effects of video-based shadowing on the pronunciation of suprasegmental features from 30 EFL students in grade 10 and their attitudes towards this technique. Ten participants who agreed to be interviewed were invited for the in-depth qualitative data collection. The selected purposive participants were from the same class at a high school in the Mekong Delta, and their English proficiency was just on the threshold of the pre-intermediate level (CEFR level A2) based on their scores from the placement test. This test is yearly organized by the Department of Education and Training in each province and city across the country when enrolling 10th-grade students according to Vietnamese government regulations. The participants were recruited for the study mainly because shadowing has a particularly significant impact on lower-proficiency learners (Hamada, 2014) and they had not known or practiced video-based shadowing in any of their English classrooms before the study.

The two following questions were asked to satisfy the research aims:

1. Does video-based shadowing affect pre-intermediate EFL learners' pronunciation performance of suprasegmental features?

2. What are learners' attitudes toward video-based shadowing for improving the performance of suprasegmental features?

Regarding the tentative answers, the hypotheses were formulated as (1) video-based shadowing would have positive effects on pre-intermediate EFL learners' pronunciation performance of suprasegmental features, and (2) learners would have positive attitudes towards video-based shadowing.

The study utilized three main instruments: pre-and post-tests, a questionnaire, and semi-structured interviews.

The tests were used to examine the effects of video-based shadowing on learners' pronunciation performance of suprasegmental features, including linking, stress, and intonation. It consisted of ten items adapted from the content in the two series used as shadowing materials in this study. The test includes two main parts: a read-aloud task (seven items) and a free response task (three items) (See Appendix 1). According to Martinsen et al. (2017), the read-aloud task is one of the traditional techniques for assessing pronunciation. Regarding this, Brown and Abeywickrama (2021) described that in read-aloud tasks, learners can read either a sentence or a paragraph and then record their output for scoring. They also commented that read-aloud tasks are not only relatively easy for teachers to administer but can also be "a surprisingly strong indicator of overall oral production ability" (p. 182). For the first part, the participants were assigned to read aloud seven items which were statements and questions. This part was aimed at assessing the participants' pronunciation performance of suprasegmental features at the controlled level. The second part consists of three free-response items that are interview questions adapted from one of two selected series in which participants will listen carefully to the questions and then answer them. By designing this part, the researchers wished to re-examine whether video-based shadowing could make any difference in the participants' pronunciation performance of suprasegmental features when they did the free response task to verify the results of the study by Martinsen et al. (2017), which showed that learners showed significant improvements in their pronunciation when reading aloud but they failed to do that in the free response part.

Although the pre and post-tests had the same content, the items in each part were shuffled randomly to ensure the reliability and validity of the instrument. More importantly, these items were not shadowed by the participants during the experimental period to ensure that the results would not be affected by extraneous factors such as participants' familiarization with the content. With regard to the scoring method, the participant's performance was assessed based on the five-point rubric for assessing suprasegmental features proposed by Sugiarto et al. (2020) (See Appendix 2).

Two main kinds of materials used in this study are textbooks and videos. Regarding the textbooks, the researchers utilized the content in Mastering the American Accent (Mojsin, 2009) and

Collins Work on Your Accent (Ashton & Shepherd, 2012) to provide learners with some basic knowledge of suprasegmental features, including linking, stress (word stress and sentence stress), and intonation. The videos used for shadowing were the two series, namely Video Practical English and Video on the Street from American English File 2 (2nd edition) published by Oxford University Press. The videos were chosen based on the vocabulary, grammar, content, and, more importantly the participants' current level of proficiency, and this was verified by a foreign volunteer working at a university who had experience using shadowing for teaching pronunciation. Additionally, all the target suprasegmental features could be shadowed via the selected videos.

The questionnaire consisted of 26 five-Likert-scale statements divided into two main parts, including the participants' evaluative responses on the impact of video-based shadowing and participants' attitudes towards this technique. The first cluster of the questionnaire (8 statements) was developed based on the single-component model by Fishbein & Ajzen (1974). The 18 statements in the second part of the questionnaire were devised based on the three-component model of attitudes (Eagly & Chaiken, 1993). The values from the questionnaire were analyzed using Oxford's (1990, p.300) framework for understanding the average between 1.0 and 5.0 (Low: 1.0 – 2.4, Medium: 2.5 – 3.4, High: 3.5 – 5.0).

In terms of interviews, Schuman (1982) developed the three-interview series in which the first thing the interviewer must do in the interview is establish the “context of the participant's experience” (Seidman, 2006, p. 17) by asking them to tell as much as possible in light of the topic up to the present time. The second and third stages of the series are about the details of the participants' experience and their reflections, respectively. In the interview of the current study, the first and second items were to serve this purpose. Similarly, the rest of the items sought to investigate further information from the participants in terms of their evaluation of the effects of video-based shadowing, their attitudes, and their problems when using the technique.

The research procedure is as follows. In Week 1, the course and the test were introduced and applied to EFL learners. From Week 2 to Week 8, the treatments were implemented. In the ninth week, the post-test was delivered to learners. In Week 10, learners were asked to complete the questionnaire, and ten of them participated in the interviews.

FINDINGS

Participants' pronunciation performance of suprasegmental features before the intervention

Before the intervention, thirty participants took the pre-test. After the data was collected, a scale test was run to check the reliability of the pre-test. The result showed a high-reliability coefficient (Cronbach's Alpha = .935) (Cohen, 1992).

A Descriptive Statistics Test was run to find out the common digits of the participants' pronunciation performance of suprasegmental features before the intervention. Table 1 below presents the results:

Table 1

Performance of suprasegmental features before the intervention (Min = 1.0, Max = 5.0, adapted from Sugiarto et al., 2020)

Suprasegmental Features	N	Min	Max	Mean (Max = 5.0)	SD
Stress	30	1.00	2.60	1.62	.509
Linking	30	1.00	2.20	1.26	.345
Intonation	30	1.00	2.80	1.85	.474
Overall	30	1.00	2.40	1.58	.415

The data in Table 1 showed a low learners' overall pronunciation (M = 1.58). It presented that their ability to pronounce the three suprasegmental features was very low.

Pronunciation performance of suprasegmental features after the intervention

After the intervention, all participants took a post-test to check whether any differences were detected in the pronunciation performance of suprasegmental features. The post-test reliability coefficient was high (Cronbach's Alpha = 0.941).

Independent Samples t-tests were performed to test whether participants' ability to pronounce suprasegmental features before and after the intervention differed, as shown in Table 2 below.

Table 2

Pronunciation performance of suprasegmental features before and after the intervention

		N	Min	Max	Mean	SD	d	p
Stress	**Pre**	30	1.00	2.60	1.62	.509	1.08	.000
	Post		1.18	2.85	2.14	.482		
Linking	**Pre**	30	1.00	2.20	1.26	.345	1.16	.000
	Post		1.00	2.73	1.85	.509		
Intonation	**Pre**	30	1.00	2.80	1.85	.474	1.50	.000
	Post		1.28	3.38	2.62	.515		
Overall	**Pre**	30	1.00	2.40	1.58	.415	1.29	.000
	Post		1.15	2.98	2.20	.480		

The results indicated the differences between the mean scores before and after the intervention and the effect size, specifically stress (d_{stress} = 1.08, p_{stress} = .000), linking ($d_{linking}$ = 1.16, $p_{linking}$ = .000), and intonation ($d_{intonation}$ = 1.50, $p_{intonation}$ = .000). It means that the intervention had a large effect size (d >= 0.8 equals large effect size) on the participants' pronunciation performance of all the suprasegmental features. Interestingly, the largest effect size among the suprasegmental features was intonation ($d_{intonation}$ = 1.50, $p_{intonation}$ = .000). It implies that the intervention would bring more influences on the performance of these suprasegmental features in a bigger size of the research population and the length of the intervention. The results also showed significant differences could be observed in terms of the mean scores of the three suprasegmental features. It means that intonation was the most affected feature while stress was the least affected one.

Table 3

Participants' scores of read-aloud and free-response tasks before and after the intervention

Tasks		N	Min	Max	Mean	d	SD	p
Reading aloud	**Pre**	30	1.04	2.62	1.64	1.25	.457	.001
	Post	30	1.14	3.14	2.27		.505	
Free response	**Pre**	30	1.00	2.25	1.44	1.24	.374	.000
	Post	30	1.11	2.61	2.04		.484	

From Table 3 above, the mean scores for the read-aloud task ($M_{ReadPre}$ = 1.64, $M_{ReadPost}$ = 2.27) before and after the intervention were higher than the mean scores for the free response task ($M_{FreePre}$ = 1.44, $M_{ReadPost}$ = 2.04). Both tasks show a similarity and have a large size effect if they were applied (d > 0.8, $d_{reading\ aloud}$ = 1.25, $d_{free\ response}$ = 1.24).

Learners' attitudes towards video-based shadowing

A questionnaire was conducted after the intervention to collect quantitative data about learners' attitudes toward video-based shadowing of suprasegmental features. Semi-structured interviews were also carried out not only to gain a deeper understanding of their attitudes but also to clarify the information gathered from the questionnaire.

A scale test was run to check the reliability of the questionnaire. The result showed that the reliability coefficient was high (Cronbach's Alpha = .886).

+ Participants' evaluative response to the effect of video-based shadowing

The effects of video-based shadowing on (1) overall pronunciation, (2) suprasegmental features (stress, linking, intonation, and awareness of using suprasegmental features), and (3) evaluation of possible problems when using video-based shadowing were considered.

Table 4

Summary of the participants' evaluative responses to the effects of video-based shadowing

Items	Strongly disagree and disagree	Neutral	Agree and strongly agree
Video-based shadowing has positive effects on my English pronunciation, especially suprasegmental features.	0%	1.67%	98.33%
Video-based shadowing helps me link English sounds better.	0%	6.67%	93.33%
Video-based shadowing helps me improve my word stress when I speak English.	0%	5%	95%
Video-based shadowing helps me improve my sentence stress when I speak English.	0%	3.33%	96.67%
Video-based shadowing helps me distinguish English intonation patterns.	0%	0%	100%
Video-based shadowing helps me use English intonation patterns to express my ideas better.	0%	0%	100%

Items	Strongly disagree and disagree	Neutral	Agree and strongly agree
Video-based shadowing helps raise my awareness of using English suprasegmental features.	8.33%	16.67%	75%
I do not encounter any problems when practicing video-based shadowing.	10.33%	5%	84.67%

The results indicated that the participants highly evaluated the first four items related to pronunciation performance. It meant that video-based shadowing was generally highly evaluated (98.33%). The results from the questionnaire were also compatible with the data from the interviews. Most of the interviewees (n = 10) made positive evaluations of the effects of video-based shadowing on their pronunciation. Generally, they believed that video-based shadowing helped them pronounce more naturally, more accurately, and more fluently.

> *"It's an effective method because it helps me speak more fluently and pronounce more accurately"* (Student 01, Block 16)

> *"I think I can remember pronunciation more easily and pronounce more accurately"* (Student 02, Block 14)

Other advantages of video-based shadowing were also pointed out, such as improving accentedness, expanding vocabulary, using intonation and linking, and pronouncing faster and more confidently.

> *"…. it helps me improve my English pronunciation and know more English words"* (Student 08, Block 14)

> *"….I think I can read better and use linking and intonation naturally…."* (Student 09, Block 22)

Notably, as shown in Table 4, participants all agreed that this technique helped them distinguish and use English intonation patterns to express their ideas better (100% agreement).). In other words, from the context of this study, participants found that intonation was most positively affected by video-based shadowing compared to linking and stress. Seven participants explained,

> *"I think it is intonation. […] because when speaking English, we use different kinds of intonation in different situations. And when we use shadowing, we can imitate …"* (Student 01, Blocks 20 and 22)

> *"Intonation. […] because in the videos, the speakers have different expressions from daily exchanges to showing worries, we will improve these expressions when speaking if we can control it better."* (Student 03, Blocks 26 and 28)

> *"Intonation. […] because intonation is the most recognizable..."* (Student 04, Blocks 26 and 28)

> *"It's intonation. […]it makes our intonation sound better and more natural."* (Student 05, Blocks 28 and 30)

As for the participants' evaluative responses to the effects of video-based shadowing on linking and stress, only two students thought stress was the most affected feature. They said,

> *"It may be word stress. […] I used to make mistakes in word stress when taking an exam, but now I can do it better and more accurately."* (Student 02, Blocks 24 and 26)

> *"Ah, … maybe stress. […] because it is the easiest and most recognizable. When practicing, I can practice this one alone but when combing many features, I don't feel confident."* (Student 07, Block 22)

Besides, only one student believed that linking is the most affected feature. She admitted,

> *"I think it's linking. […] because it makes my sentence shorter and helps me speak faster."* (Student 08, Blocks 28 and 30)

In fact, linking was found to be one of the three suprasegmental features that had the least influence from video-based shadowing on seven students' pronunciation practice. This finding of the interviews is consistent with the results presented in item 2 of Table 4.

> *"…linking is affected the least. We have been accustomed to pronouncing words slowly and clearly since elementary school, so when speaking we often forget it…"* (Student 03, Block 30)

> *"… linking is in many sentences, if there are no subtitles, I can't make out what people are saying, and I don't know whether there's linking. So I think it is the least affected."* (Student 04, Blocks 34 and 36)

> *"I think it is linking. […] some words link with each other if we do not pay attention to … then we don't know whether it's linking or a word itself is pronounced that way."* (Student 06, Blocks 34 and 36)

Meanwhile, stress is the feature least affected by intervention compared to linking and intonation (Item #7, Table 4). They shared,

> *"It's stress. […] because, before the study, I learned the linking rules already, I knew them already…"* (Student 01, Blocks 24 and 26)

> *"It may be sentence stress. […] my teacher taught me when to stress and not to stress. but I often read without stress."* (Student 02, Blocks 28 and 30)

Only one student thought that the technique had some effects on both intonation and linking, but less than that on the feature of stress.

> *"[…] linking and intonation are the same. […] I can realize and do them but I don't feel as confident as when I deal with stress. […]"* (Student 07, Blocks 30 and 32)

When looking more closely at the explanation of the least affected feature, the main reason is that linking is considered the most difficult feature to master. Two students admitted that they had learned linking rules and had not paid enough attention to this feature when practicing video-based shadowing.

However, about 10.33% of participants agreed that video-based shadowing could also cause some problems. The findings indicated three main problems that learners often encountered during the intervention. First, the most common problem was that they could not hear the words correctly due to the speaker's speaking speed.

> *"I cannot catch what people say […]"* (Student 01, Block 28)

> *"…. I cannot hear the words because they are spoken a little fast…"* (Student 02, Block 32)

> *"There are many videos in which the speakers speak so fast that I cannot hear clearly or I mishear them."* (Student 04, Block 38)

To solve this problem, learners suggested either replaying the video again or slowing down the video speed.

> *"If I cannot catch…, I will replay or I will practice many times."* (Student 01, Block 32)

> *"Maybe, I will play the video again."* (Student 02, Block 34)

> *"I can adjust the speed of the video via the software"* (Student 04, Block 42)

The second problem was that the topics were not interesting enough and the technique was quite mechanical. They shared,

> *" If I like the video, I find it interesting, but if I don't like the video, I find it not natural."* (Student 05, Block 50)

> *"[…] And normally the topics are not interesting."* (Student 06, Blocks 14 and 38)

Their solution to this problem is to find other engaging videos to practice.

> *"I often go to some websites, for example, YouTube, to find videos that are suitable for me […]"* (Student 05, Block 52)

> *"[…] if the topics are not interesting I will find other topics by myself."* (Student 06, Block 42)

Finally, it is difficult for them if they do not know how to pronounce the words in the video.

> *"[…] there are some words I cannot pronounce"* (Student 01, Block 28)

> *"The way people say is different from what I know, the way I pronounce is different from them."* (Student 06, Block 38)

Their solutions are either checking the dictionary, replaying the video, or asking for help.

> *"[…] for the words I don't understand, I will check them up in the dictionary."* (Student 01, Block 32)

> *"I will replay the video to watch again. But if I cannot read, I will ask other people to show me how to do [...]"* (Student 06, Block 42)

Surprisingly, two interviewees admitted that they did not have any difficulties practicing video-based shadowing during the study.

> *"I don't have a lot of difficulties. I think everything went very well."* (Student 03, Block 32)

> *"I have no difficulties."* (Student 10, Block 56)

Furthermore, 8.33% of participants disagreed that video-based shadowing had an impact on their perception of using suprasegmental features, and 16.67% of them felt uncertain about this impact. This may be because the intervention time is quite short, and the effect on their awareness was unclear.

+ Participants' attitudes toward video-based shadowing

The second cluster of the questionnaire focuses on investigating the participants' attitudes towards video-based shadowing. A Descriptive Statistics test was run to check the minimum, the maximum, the mean scores, and the standard deviation of the responses.

Table 5

Descriptive statistics of the participants' attitudes to video-based shadowing

	N	Min	Max	Mean	Std. Deviation
Cognition	30	3.44	5.00	4.51	.473
Affection	30	3.75	5.00	4.45	.471
Behavior	30	3.50	4.67	4.22	.370

From the information presented in the Table above, the mean scores of cognition ($M_{cognition}$ = 4.51), affection ($M_{affection}$ = 4.45), and behavior ($M_{behavior}$ = 4.22) are considered a high level (mean >= 3.5) of agreement according to Oxford (1990). In detail, three smaller parts will be presented, namely the participants' **(1) cognitive components, (2) affective components, and (3) behavioral components**. The Table below summarizes the participants' cognitive components.

Table 6
Summary of the participants' cognitive components

Items	Strongly disagree and disagree	Neutral	Agree and strongly agree
Video-based shadowing is helpful for practicing English pronunciation, especially suprasegmental features.	1.1%	7.8%	91.1%
Video-based shadowing is helpful for practicing linking.	0%	0%	100%
Video-based shadowing is helpful for practicing intonation.	0%	0%	100%
Video-based shadowing is helpful for practicing word stress.	0%	10%	90%
Video-based shadowing is helpful for practicing sentence stress.	0%	3.3%	96.7%
Video-based shadowing makes linking sound practices easier.	0%	10%	90%
Video-based shadowing makes intonation practices easier.	0%	0%	100%
Video-based shadowing makes word stress practices easier.	0%	16.7%	83.3%
Video-based shadowing makes sentence stress practices easier.	0%	10%	90%

It can be seen in Table 6 that most of the participants believed video-based shadowing is helpful for practicing their English pronunciation as well as suprasegmental features. Besides, they also thought that the technique could make their practices of suprasegmental features easier. Noticeably, 100% of participants agreed on the effects of video-based shadowing on intonation, which was congruent with the aforementioned findings from the interviews.

Table 7
Summary of the participants' affective components

Items	Strongly disagree and disagree	Neutral	Agree and strongly agree
I like using video-based shadowing to practice English pronunciation.	0%	3.3%	96.7%
I feel motivated to practice my English pronunciation when using video-based shadowing.	0%	10%	90%
I feel more confident when using video-based shadowing.	0%	23.3%	76.7%
I prefer to practice video-based shadowing with my friends.	10%	3.3%	86.7%
I prefer to practice video-based shadowing with my teacher.	6.7%	23.3%	70%
I prefer to practice video-based shadowing on my own.	23.3%	13.3%	63.4%

The results from the interviews showed that most of the interviewees did not know or practice this technique before the study. Instead, there are six main ways to practice pronunciation including following teacher instructions, practicing with friends, using movies or music videos, using Google Translate to look up the dictionary, reading text aloud, and using pronunciation apps. For example:

> *"I usually practice pronunciation in class or use Google Translate. […] I repeat after my teacher or ask my friends how to pronounce."* (Student 02, Blocks 06 and 08)

> *"I learn English via apps such as Elsa or Cake on the Internet."* (Student 04, Block 05)

> *"I read some passages in class, but just read and read randomly."* (Student 09, Block 06)

Moreover, some of the participants did not use any specific method or activity to practice their pronunciation.

> *"I haven't used any methods…"* (Student 08, Block 12)

More importantly, even though it was their first experience with this technique, 96.7% of participants admitted that they liked using it to practice English pronunciation. There are many different reasons for this.

> *"… listening to native speakers is easier for us to imitate"* (Student 02, Block 38)

> *"… If there is no one to practice with, I can also practice speaking English at home on my own. ….. I can also imitate people to improve my level of pronunciation."* (Student 04, Block 46)

> *"……When we watch videos, we can see how people talk,… we hear and see the shape of their mouths."* (Student 09, Block 62)

> *"I think it is easy to understand, easy to do and I can do it myself."* (Student 10, Block 60)

However, only one student felt uncertain about the technique because of its complexity. She explained,

> *"[…] I find it suitable for me, but I'm not sure I can do it well. […] it's a little bit complicated for me […]"* (Student 06, Blocks 44 and 46)

Besides, the participants felt confident when using the technique to practice their pronunciation. They said,

> *"[…] I know how to speak...it makes me more confident."* (Student 05, Block 22)

> *"[…] I can pronounce fluently, a bit faster with more confidence."* (Student 10, Block 20)

In terms of participants' preferences for video-based shadowing, practicing with a friend was the most preferred choice (86.7%), while practicing alone was the least favorite one (63.4%) (see Table 7). The results are also supported by the interview data. Indeed, half of the interviewees liked to practice video-based shadowing with their friends. They outlined several benefits of practicing with friends, including feeling more natural, exchanging feedback more easily, receiving help and correction from others, and mutual understanding. They admitted,

> *"When practicing with friends, we can correct our mistakes easier than with our teachers […]"* (Student 02, Block 44)

> *"[…] if I practice with my friends, I find it easier to accept and learn because we are the same age and have the same views."* (Student 05, Block 60)

Additionally, three interviewees enjoyed practicing video-based shadowing with either their teachers or friends. The most reasonable explanation for their choice is that their mistakes will be easily recognized and corrected. Some explained they feel more focused and serious when someone else is with them.

> *"[…] If it's with my teacher, they can correct me, and if it's with my friends, they can recognize my pronunciation mistakes and we can give each other feedback .."* (Student 08, Block 52)

> *"I find practicing with teachers and friends will be more effective because they can give me feedback and correct mistakes to improve my speaking ability"* (Student 10, Blocks 62 and 64)

However, one student said she would rather practice video-based shadowing with her teacher than with friends or alone.

> *"…with my teachers, because there will be someone to guide me. … with my teachers, there will be something that makes me feel more serious."* (Student 07, Blocks 46 and 48)

Only one student thought he would do better if he practiced alone. His reason is that he can practice at his own pace. He shared,

> *"… when practicing alone, I can listen and adjust myself, without being influenced by my friends, I do not need to follow my friends, I have more freedom."* (Student 01, Block 44)

Furthermore, qualitative data from the interviews also yielded important findings on reasons for not choosing to practice with a teacher or alone. When practicing with teachers, students' biggest weakness is feeling shy because the generation gap leads to not being able to share opinions.

> *"[…] I feel shy when practicing with my teacher."* (Student 02, Block 44)

"[…] there will be some gap which makes us feel unnatural." (Student 09, Block 68)

Regarding the main disadvantages of practicing alone, interviewees explained that their mistakes will not be recognized and corrected and they may also feel more lonely.

"[…] I feel very lonely, then it would make me feel depressed..." (Student 05, Block 60)

"… when practicing alone, I cannot recognize my mistakes in my pronunciation." (Student 08, Block 52)

Table 8

Summary of the participants' behavioral components

Items	Strongly disagree and disagree	Neutral	Agree and strongly agree
I will continue to practice video-based shadowing to improve my English pronunciation.	0%	0%	100%
Video-based shadowing should be frequently included as an activity in English classrooms.	0%	0%	100%
I intend to recommend video-based shadowing to my friends as a new habit.	0%	10%	90%

From Table 8 above, all participants determined to improve their English pronunciation with video-based shadowing. More importantly, it is recommended that this technique be incorporated into classroom activities when learning English. Similarly, 100% of interviewees agreed with using this method in the classroom. They elaborated,

"…… because it is compatible with activities in the textbook that I am learning with." (Student 01, Block 48)

"…….I think it will improve our English pronunciation, help us to speak like native speakers." (Student 02, Block 40)

"….. because it helps us improve many aspects when speaking […] when we have daily conversations, we will speak better. And when we practice listening, we can listen better." (Student 10, Blocks 70 and 72)

Last but not least, one-tenth of participants (item#3 in Table 8) hesitate to recommend this technique to their friends. It may be due to the aforementioned problems of the technique.

DISCUSSION

This section discusses the key findings regarding (1) the effects of video-based shadowing on EFL learners' pronunciation performance of suprasegmental features and (2) their attitudes toward the technique.

Video-based shadowing had positive effects on the participants' pronunciation performance of suprasegmental features. This result was compatible with the studies of Hamada (2018) and Sugiarto et al. (2020). The shadowing technique and the video-based one both bring a positive and significant improvement in EFL learners' pronunciation of suprasegmental features, which is relevant to the study of Mıcık (2020). Furthermore, these features were found influential from the measurement of both tasks of "read aloud" and "free responses" with a difference in the general suprasegmental features of stress, linking, and intonation compared to the more detailed aspects of strong and weak forms, linking, word stresses, sentence stresses, pitch, and intonation as in Sugiarto et al. (2020).

Intonation is the suprasegmental feature that is most improved after applying the technique. This result is in line with Nguyen and Dao (2018) that the "visual and audial display of English intonation" (pp. 12-13) can help learners imitate the intonation of a target input. The visual presentation of different sound sequences and loudnesses has been overlooked by any Vietnamese using a monosyllabic language from different perspectives. That might help explain why visualizing sound patterns in terms of intonation can be improved more than other suprasegmental features.

A similar explanation can also be used for the cases of linking, as the participants achieved the lowest scores from the two measurement times. Vietnamese EFL learners tend to pronounce words in a sentence separately (Nguyen & Ingram, 2004) due to not being able to connect sounds together, which is typical in natural English speech.

The result is congruent with those from the study by Martinsen et al. (2017) about a clear improvement in the read-aloud task over the free response task. Issues related to learners' pronunciation performance and auditory control (Thomson & Derwing, 2015) tend to improve across reading-aloud tasks that are more easily achieved through pronunciation instruction. In other words, participants performed the controlled task better. That leads to agreement with Mori's (2011) suggestion that shadowing should only be combined with reading-aloud tasks instead of free-response tasks to improve the suprasegmental features of learners to the highest possible level.

Regarding participants' positive attitudes about the influence of the video-based shadowing technique on suprasegmental features, this study is consistent with previous research (Dang, 2020; Salim et al., 2020) on how principles of second language acquisition can underpin the shadowing technique versus the video-based shadowing one. It is true that learners are trained to articulate the language input at the beginning of the process of acquiring a second language. That process helps learners utilize the phonological loop to improve their phonological coding and speech perception (Lambert et al., 2016). Besides, the participants also outlined some additional advantages of using video-based shadowing such as improving accentedness, expanding their vocabulary, using intonation and linking as well as pronouncing faster and more confidently. The result is mostly in line with many other related studies (Dang, 2020; Salim et al., 2020). Furthermore, the participants encountered some common difficulties when using video-based shadowing, including the speaker's speaking rate, uninteresting topics, and unknown vocabulary. This result is supported by Dang's (2020) study.

CONCLUSION AND IMPLICATIONS

The pretest results showed that participants ' initial scores were very low regarding the impact of video-based shadowing on pre-intermediate EFL learners' ability to pronounce suprasegmental features. After the intervention, the results of the post-test showed an improvement in their pronunciation performance for suprasegmental features. In other words, it can be affirmed that the above technique positively impacted their pronunciation ability. Besides, based on the interviews and questionnaire findings, the participants also admitted that the technique helped them improve their pronunciation performance of suprasegmental features. Regarding the study's first hypothesis, it can be seen that the hypothesis was validated. In brief, the findings affirmed that video-based shadowing had positive effects on pre-intermediate EFL learners' pronunciation performance of suprasegmental features.

Regarding learners' attitudes towards video-based shadowing, the data gathered from the questionnaire and the interviews revealed that the participants had positive attitudes towards video-based shadowing. The majority of the participants considered video-based shadowing an effective technique to improve their pronunciation performance of suprasegmental features, especially intonation. Besides, the idea of integrating video-based shadowing into the English classroom was supported by all of the participants. Thus, the findings confirmed the second hypothesis of the study.

Based on the present study's findings, some pedagogical implications can be drawn for EFL teachers and learners, especially those concerned about improving their pronunciation performance of suprasegmental features.

First, the shadowing technique should be applied for EFL learners, especially those with low English proficiency and/or difficulty pronouncing suprasegmental features. As shown in the results of this study, video-based shadowing positively influenced participants' pronunciation performance. If learners practice this technique appropriately and regularly, they can achieve better pronunciation results, especially for Vietnamese EFL learners who often have difficulty pronouncing English in general, and suprasegmental features in particular due to the huge difference between the two languages.

Second, due to its benefits, video-based shadowing should be integrated into teaching pronunciation or English-speaking skills for EFL teachers. Furthermore, teachers should be aware of some possible difficulties that learners may encounter when using video-based shadowing, such as the speaker's speaking speed, uninteresting topic, and unknown vocabulary. With an awareness of these common difficulties, teachers can facilitate learners' experience of implementing video-based shadowing.

Besides, to integrate this technique into English lessons, teachers need to increase interaction between students through group or pair activities. This helps increase mutual understanding, feel more natural, exchange feedback more easily, and receive mutual help and correction. Furthermore, teachers should guide students on how to find suitable videos to perform shadowing. This helps children be more proactive and interested in this practice.

Last but not least, administrators should consider incorporating this technique into the general curriculum to help EFL students improve their pronunciation, especially suprasegmental features. This brings many benefits to learners, based on the research results that show that learners do not know how to improve their English pronunciation skills.

LIMITATIONS

Although the research objectives have been achieved, unfortunately, there are still some unavoidable limitations. Actually, research is necessary with the widespread participation of many learners to have a more multi-dimensional view of the effectiveness of the shadowing technique. In addition, the study also needs a control group to determine the difference between this technique and conventional techniques.

THE AUTHORS

Phan Thi Tuyet Van is an associate professor at the School of Foreign Languages, Can Tho University, Vietnam. She holds a PhD in Pedagogy of Social Sciences from the University of Gdansk, Poland. Her research interests include critical pedagogy, language education, literary criticism, and borderless education from multicultural perspectives. *pttuyetvan@ctu.edu.vn*

Ly Kim Lien holds an MA in Principles and Methods of English Teaching from the School of Foreign Languages, Can Tho University, Vietnam. Currently, he is a freelance English teacher for Langmaster whose head office is in Ha Noi, Vietnam. He is interested in teaching pronunciation to youngsters and is especially fond of ESP and EAP. *lklienatwork1996@gmail.com*

Nguyen Duy Khang is an associate professor at the School of Foreign Languages, Can Tho University, Vietnam. He holds a PhD in Pedagogy from the University of Gdansk, Poland. He is interested in language teaching and learning research, didactics, ICT in the classroom, classroom action research, and other practical educational theories. *ndkhang@ctu.edu.vn*

REFERENCES

Ajzen, I. (2001). Nature and operation of attitudes. *Annual Review of Psychology, 52*(1), 27–58. https://bit.ly/3C8GR8Z

Alameen, G., & Levis, J. M. (2015). Connected speech. In M. Reed & J. M. Levis (Eds,), *The handbook of English pronunciation* (pp. 157–174). John Wiley & Sons, Inc. http://dx.doi.org/10.1002/9781118346952.ch9

Allerton, D. J. (2000). Articulatory inertia vs 'systemzwang': Changes in liaison phenomena in recent British English. *English Studies, 81*(6), 574–581. http://doi.org/10.1076/enst.81.6.574.9183

Angel M. D., & Erika V. E. J. (2018). Using the shadowing technique to improve Ecuadorian English learners' speaking intelligibility. *International Journal of Current Research, 10*(12), 76770–76772. https://www.journalcra.com/sites/default/files/issue-pdf/33624_0.pdf

Ashton, H., & Shepherd, S. (2012). *Work on your accent: Clearer pronunciation for better communication*. Collins.

Bailey, K., & Nunan, D. (2005). *Practical English language teaching: Speaking*. McGraw-Hill.

Betti, M. J., & Al-Jubouri, C. F. (2015). *Approaches and methods of teaching English as a foreign language*. Nippur Publishing.

Betti, M. J., & Hasan, A. A. (2020). The Iraqi EFL learners' ability to use speech acts in MA and Ph.D. theses defense. *Education, Language and Sociology Research, 1*(2), 41–65. http://doi.org/10.22158/elsr.v1n2p41

Bøhn, H., & Hansen, T. (2017): Assessing pronunciation in an EFL context: Teachers' orientations towards nativeness and intelligibility. *Language Assessment Quarterly, 14*(1), 54–68, http://doi.org/10.1080/15434303.2016.1256407

Brown, H. D., & Abeywickrama, P. (2021). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson.

Burns, A., & Claire, S. (2003). *Clearly speaking: Pronunciation in action for teachers*. National Centre for English Language Teaching and Research.

Celce-Murcia, M., Brinton, D. M., & Goodwin, J. M. (2010). *Teaching pronunciation hardback with audio CDs (2): A course book and reference guide*. Cambridge University Press.

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://bit.ly/3JVpB8P

Clement, A., & Murugavel, T. (2018). English for the workplace: The importance of English language skills for effective performance. *The English Classroom, 20*(1), 1–15. https://bit.ly/3QEIH5b

Cohen, J. (1992). A power primer. *Psychological Bulletin, 112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155

Collins, B., & Mees, I. M. (2006). *Phonetics of English in the nineteenth century*. Routledge.

Dang, T. N. D. (2017). Suprasegmental errors, pronunciation instruction and communication. *Accent Asia, 9*(1), 98-120. https://bit.ly/3X6H9FY

Dang, T. T. (2020). *Improve speaking skill of first-year English majors at Haiphong management and technology university by shadowing method*. [Doctoral dissertation, Đại học Dân lập Hải Phòng]. https://bit.ly/3QBi7da

Eagly, A. H., & Chaiken, S. (1993). *The psychology of attitudes*. Harcourt Brace Jovanovich College Publishers.

Essberger, J. (2008). *Word stress and sentence stress: The golden key for English pronunciation*. English Club. https://www.englishclub.com/esl-articles/199810.php

Fishbein, M., & Ajzen, I. (1974). Attitudes towards objects as predictors of single and multiple behavioral criteria. *Psychological Review, 81*(1), 59–74. http://doi.org/10.1037/h0035872

Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation, 3*(1), 34–56. http://doi.org/10.1075/jslp.3.1.02foo

Garrigues, S. (1999, October 1–3). *Overcoming pronunciation problems of English teachers in Asia*. An Asianbridges Webcast. http://asianbridges.com/pac2/presentations/garrigues.html

Genc, Z. S., & Aydin, F. (2017). An analysis of learners' motivation and attitudes toward learning English language at tertiary level in Turkish EFL context. *English Language Teaching, 10*(4), 35–44. https://doi.org/10.5539/elt.v10n4p35

Gilakjani, A. P. (2011). A study on the situation of pronunciation instruction in ESL/EFL classrooms. *Journal of Studies in Education, 1*(1), 1–15. http://dx.doi.org/10.5296/jse.v1i1.924

Gimson, A. C. (1980). *An introduction to the pronunciation of English* (3rd ed.). Edward Arnold.

Hamada, Y. (2014). The effectiveness of pre- and post-shadowing in improving listening comprehension skills. *The Language Teacher, 38*(1), 3–10. https://doi.org/10.37546/jalttlt38.1-1

Hamada, Y. (2018). Shadowing for pronunciation development: Haptic-shadowing and IPA-shadowing. *Journal of Asia TEFL, 15*(1), 167–183. http://doi.org/10.18823/asiatefl.2018.15.1.11.167

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

Hassanein, E. E. A. (2015). *Inclusion, disability and culture*. Sense Publishers.

Harmer, J. (2007). *The practice of English language teaching*. Pearson Longman.

Hieke, A. E. (1984). Linking as a marker of fluent speech. *Language and Speech, 27*(4), 343–354. https://doi.org/10.1177/002383098402700405

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), 43–66. https://bit.ly/4fZAjuH

Kingdon, R. (1958). *The groundwork of English stress*. Longmans.

Lambert, C., Kormos, J., & Minn, D. (2016). Task repetition and second language speech processing. *Studies in Second Language Acquisition, 39*(1), 167–196. https://doi.org/10.1017/S0272263116000085

Le, P. N., Vu, H. M. L., & Tran, M. N. (2022). Improving EFL students' intonation in-text using shadowing technique with the implementation of Google text-to-speech. *AsiaCALL Online Journal, 13*(1), 93–121. https://bit.ly/4ejPIob

Levis, J. M. (2018). *Intelligibility, oral communication, and the teaching of pronunciation*. Cambridge University Press.

Low, E. L. (2015). The rhythmic patterning of English(es) implications for pronunciation teaching. In M. Reed & J. M. Levis (Eds.), *The handbook of English pronunciation* (pp. 125–138). John Wiley & Sons, Inc.

Manseur, R. (2015). *Exploring the role of shadowing in the development of EFL learners' speaking skill: A case study of third- year students of English at Mohamed Kheider University of Biskra*. [Master' thesis, Mohamed Kheider University Of Biskra]. Mohamed Kheider University of Biskra. https://bit.ly/3X6Tv0O

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*(4), 661–680. https://bit.ly/4ejQTE7

Mıcık, S. (2020). *Development of L2 oral proficiency through video-based shadowing practices*. [Master's thesis, Pamukkale Üniversitesi Eğitim Bilimleri Enstitüsü]. Pamukkale GCRIS Database. https://gcris.pau.edu.tr/handle/11499/38421

Mojsin, L. (2009). *Mastering the American accent*. Barron's Educational Series.

Morley, J. (1991). The pronunciation component in teaching English to speakers of other languages. *TESOL Quarterly, 25*(3), 481–520. https://doi.org/10.2307/3586981

Mori, Y. (2011). Shadowing with oral reading: Effects of combined training on the improvement of Japanese EFL learners' prosody. *Language Education & Technology, 48*, 1–22. http://dx.doi.org/10.24539/let.48.0_1

Murphey, T. (2001). Exploring conversational shadowing. *Language Teaching Research, 5*(2), 128–155. http://doi.org/10.1177/136216880100500203

Nakayama, T. (2011). Bijuarushadoingu no kouka [Effectiveness of visual shadowing]. *Rimediaruky-ouikukenkyu, 6*, 151–159.

Nakayama, T., & Mori, T. (2012). Efficacy of visual-auditory shadowing. *Eibeibunka: Studies in English Language, Literature and Culture, 42*, 55–68. https://www.jstage.jst.go.jp/article/eibeibunka/42/0/42_KJ00008046865/_pdf

Nguyen, A. T. T., & Dao, D. M. (2018). The acquisition of question intonation by Vietnamese learners of English. *Asian-Pacific Journal of Second and Foreign Language Education, 3*(1), 1–13. https://link.springer.com/article/10.1186/s40862-018-0044-4

Nguyen, T., & Ingram, J. (2004, December). *A corpus-based analysis of transfer effects and connected speech processes in Vietnamese English* [Paper presentation]. The Tenth Australian International Conference on Speech Science & Technology, Sydney, Australia. https://bit.ly/4ejHpJ1

Niimoto, S. (2022). Shadowing to develop comprehensibility for Japanese university students. *Kobe English Language Teaching, 37*(1), 38–50. https://www.jstage.jst.go.jp/article/kelt/37/1/37_38/_pdf

Omar, H. M., & Umehara, M. (2010). Using 'a shadowing technique' to improve English pronunciation deficient adult Japanese learners: An action research on expatriate Japanese adult learners. *Journal of Asia TEFL, 7*(2), 199–230. https://bit.ly/3T9qWPb

Oxford, R. (1990). *Language learning strategies: What every teacher should know.* Harper Collins Publisher.

Pedrazzini, L. (2016) The principle of 'correct pronunciation': Teaching English as a foreign language in the early twentieth century. *Language & History, 59*(1), 63–72, http://doi.org/10.1080/17597536.2016.1176292

Pourhossein, G. A. (2012). A study of factors affecting EFL learners' English pronunciation learning and the strategies for instruction. *International Journal of Humanities and Social Science, 2*(3), 119–128. https://bit.ly/3XkM5Z8

Rahimi, M., & Ruzrokh, S. (2016). The impact of teaching lingua franca core on English as a foreign language learners' intelligibility and attitudes towards English pronunciation. *Asian Englishes, 18*(2), 141–156. https://doi.org/10.1080/13488678.2016.1173466

Roach, P. (1991). *English phonetics and phonology: A practical course*. Cambridge University Press.

Roach, P. (2009). *English phonetics and phonology paperback with audio CDs (2): A practical course*. Cambridge University Press.

Salim, A., Terasne, T., & Narasima, L. (2020). Enhancing the students' pronunciation using shadowing technique at senior high school students. *Journal of Languages and Language Teaching, 8*(1), 20–28. https://doi.org/10.33394/jollt.v8i1.2212

Schuman, D. (1982). *Policy analysis, education, and everyday life: An empirical reevaluation of higher education in America*. DC Heath.

Seidman, I. (2006). *Interviewing as qualitative research: A guide for researchers in education and the social sciences.* Teachers College Press.

Sugiarto, R., Prihantoro, P., & Edy, S. (2020). The impact of shadowing technique on tertiary students' English pronunciation. *Linguists: Journal of Linguistics and Language Teaching, 6*(1), 114–125. https://doi.org/10.29300/ling.v6i1.3298

Sumarsih, S. (2017). The impact of shadowing technique on tertiary EFL learners' listening skill achievements. *International Journal of English Linguistics, 7*(5), 184–189. https://doi.org/10.5539/ijel.v7n5p184

Tamai, K. (1997). Shadowing no koka to chokai process niokeru ichizuke. [The effectiveness of shadowing and listening process]. *Current English Studies, 36*, 105–116.

Thomson, R. I., & Derwing, T. M. (2015). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics, 36*, 326–344. https://doi.org/10.1093/applin/amu076

Tran, T. T. (2013). The causes of passiveness in learning of Vietnamese students. *VNU Journal of Science: Education Research, 29*(2), 72–84. https://js.vnu.edu.vn/ER/article/view/502

Trujillo, F. (2006, June 25). *Teaching pronunciation*. E-lecture. http://www.ugr.es/~ftsaez/fonetica.html

Wells, J. C. (2006). *English intonation: An introduction*. Cambridge University Press.

Appendix 1

Pre-test
Part 1: Read aloud the following statements and questions

No.	Questions/Statements
01	I love it because it's an easy job.
02	What do you like about it?
03	I eat a lot of vegetables every day.
04	Did you have a good time on Tet holiday?
05	Do you prefer movies with sad or happy endings?
06	How many brothers and sisters do you have?
07	If you have a problem, who do you ask for advice, friends or family?

Part 2: Listen to the questions and answer

08. What kind of music do you like?
09. Where did you go for your last vacation?
10. What did you do last night?

Post-test
Part 1: Read aloud the following statements and questions

No.	Questions/Statements
01	If you have a problem, who do you ask for advice, friends or family?
02	I eat a lot of vegetables every day.
03	What do you like about it?
04	How many brothers and sisters do you have?
05	Do you prefer movies with sad or happy endings?
06	Did you have a good time on Tet holiday?
07	I love it because it's an easy job.

Part 2: Listen to the questions and answer

08. Where did you go for your last vacation?
09. What did you do last night?
10. What kind of music do you like?

Appendix 2

Suprasegmental Feature Assessment Rubrics (adapted from Sugiarto et al., 2020)

	Stress	Intonation	Linking
5	Having a native-like ability to place word and sentence stress accurately.	Having a native-like ability to use intonation patterns flexibly and correctly to express different meanings and intended functions.	Native-like use of linkings to speak smoothly.
4	Correct use of word stress and sentence stress to emphasize the meaning of the message.	Correct use of English intonation patterns to convey the intended message.	Correct use of linkings to speak smoothly.
3	Fairly correct use of word stress and sentence stress to emphasize the meaning of the message.	Fairly correct use of English intonation patterns to convey the intended message.	Fairly correct use of linkings to speak, but sometimes with non-native pauses, which does not interfere with comprehensibility.
2	Incorrect use of word stress and sentence stress to emphasize the meaning of the message, which interferes with comprehensibility.	Incorrect use of English intonation patterns to convey the intended message, which inteferes with comprehensibility.	Incorrect use of linkings to speak, with many non-native pauses, which interferes with comprehensibility.
1	Completly incorrect use of word stress and sentence stress to emphasize the meaning of the message, which causes significant problems in comprehensibility.	Completely incorrect use of English intonation patterns to convey the intended message, which causes significant problems in comprehensibility.	Completely incorrect use of linkings to speak, with many non-native pauses, which causes significant problems in comprehensibility.

Appendix 3

Questionnaire

I. Background Information:

Your name (Optional): _____

Your gender: ☐ Male ☐ Female ☐ Other

Your experience in learning English:
☐ Less than 5 years ☐ From 5 years to 8 years
☐ More than 8 years

Your frequency of using video-based shadowing after class
☐ Never ☐ Rarely ☐ Sometimes

II. Learners' self-evaluative response to video-based shadowing

Please check the number which is applicable to you. Each number refers to the following description: 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree.

No.	Statements	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1	Video-based shadowing has positive effects on my English pronunciation, especially suprasegmental features.					
2	Video-based shadowing helps me link English sounds better.					
3	Video-based shadowing helps me improve my word stress when I speak English.					
4	Video-based shadowing helps me improve my sentence stress when I speak English.					
5	Video-based shadowing helps me distinguish English intonation patterns.					
6	Video-based shadowing helps me use English intonation patterns to express my ideas better.					
7	Video-based shadowing helps raise my awareness of using English suprasegmental features.					
8	I do not encounter any problems when practicing video-based shadowing.					
	Others, please specify:					

III. Learners' attitude towards video-based shadowing

Please check the number which is applicable to you. Each number refers to the following description (1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree)

No.	Statements	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
9	Video-based shadowing is helpful for practicing English pronunciation, especially suprasegmental features.					
10	Video-based shadowing is helpful for practicing linking.					
11	Video-based shadowing is helpful for practicing intonation.					
12	Video-based shadowing is helpful for practicing word stress.					
13	Video-based shadowing is helpful for practicing sentence stress.					
14	Video-based shadowing makes linking sound practices easier.					
15	Video-based shadowing makes intonation practices easier.					
16	Video-based shadowing makes word stress practices easier.					
17	Video-based shadowing makes sentence stress practices easier.					
18	I like using video-based shadowing to practice English pronunciation.					
19	I feel motivated to practice my English pronunciation when using video-based shadowing.					
20	I feel more confident when using video-based shadowing.					
21	I prefer to practice video-based shadowing with my friends.					
22	I prefer to practice video-based shadowing with my teacher.					
23	I prefer to practice video-based shadowing on my own.					
24	I will continue to practice video-based shadowing to improve my English pronunciation.					
25	Video-based shadowing should be frequently included as an activity in English classrooms.					
26	I intend to recommend video-based shadowing to my friends as a new habit.					
	Others, please specify:					

Thank you for your valuable time!

Appendix 4

Interviews

1. Before attending this study, have you ever heard of or practiced shadowing?
2. Before the study, how did you practice your English pronunciation?
3. What is your opinion about video-based shadowing?
4. How does video-based shadowing affect your English pronunciation performance of suprasegmental features?
5. What problems did you encounter when using video-based shadowing?
6. In your opinion, is it better to practice video-based shadowing with your teacher, your friends or on your own?
7. Do you think video-based shadowing should be included in English speaking and pronunciation classes?