

John Benjamins Publishing Company



This is a contribution from *Journal of Second Language Pronunciation* 3:1
© 2017. John Benjamins Publishing Company

This electronic file may not be altered in any way.

The author(s) of this article is/are permitted to use this PDF file to generate printed copies to be used by way of offprints, for their personal use only.

Permission is granted by the publishers to post this file on a closed server which is accessible only to members (students and faculty) of the author's/s' institute. It is not permitted to post this PDF on the internet, or to share it on sites such as Mendeley, ResearchGate, Academia.edu.

Please see our rights policy on <https://benjamins.com/content/customers/rights>

For any other use of this material prior written permission should be obtained from the publishers or through the Copyright Clearance Center (for USA: www.copyright.com).

Please contact rights@benjamins.nl or consult our website: www.benjamins.com

Using shadowing with mobile technology to improve L2 pronunciation

Jennifer A. Foote and Kim McDonough

University of Alberta / Concordia University

Shadowing has been demonstrated to improve various aspects of second language learners' pronunciation but few studies have investigated whether these changes impact untrained listeners' perceptions. In the present study, sixteen participants used iPods to practice shadowing short dialogues for eight weeks. The participants practiced at least four times per week for a minimum of 10 minutes each time, and recorded themselves while shadowing. Two tasks (a shadowing task and an extemporaneous speaking task) were administered as pre-, mid-, and post-tests, and were rated by 22 speakers of English. The shadowing task was rated for learners' ability to imitate a speech model and the extemporaneous speaking task was rated for comprehensibility, accentedness, and fluency. Interview data were also collected during the study to gauge participants' opinions of the activities. Results indicated that the participants improved significantly on all speaking measures apart from accentedness and were largely positive about the activities.

1. Introduction

Second language pronunciation has traditionally received less attention from second language acquisition (SLA) researchers than other aspects of language such as grammar and vocabulary. However, as has been noted in a recent review of studies in this area (Thomson & Derwing, 2014), the past several years have seen an explosion of interest in pronunciation. Yet there is still much that remains unknown about how to best help language learners improve their pronunciation in ways that will help them be more easily understood. As Thomson and Derwing (2014) note, there is a need for more pronunciation studies that use comprehensibility ratings as part of the criteria for determining whether or not a technique or intervention is worth using, since improvements in comprehensibility indicate that pronunciation changes are both noticeable to human listeners and helpful in making speech

easier to understand. Further, it is important that these changes can transfer from controlled practice activities to extemporaneous speech.

Comprehensibility is defined here as a subjective measure of how easy or difficult to understand a listener perceives speech to be. It is contrasted with intelligibility, which refers to objective measures of whether speech is actually understood, and accentedness which can be defined as how much a speaker's phonology differs from that of a first language (L1) speaker of that language (Munro & Derwing, 1995). Another construct often measured alongside comprehensibility and accentedness is fluency because interventions designed primarily to impact pronunciation may also provide additional benefits to speakers by improving fluency (e.g., Derwing, Munro, & Wiebe, 1998) although this is not always the case (Derwing, Munro, Foote, Waugh, & Fleming, 2014). While increased fluency is not the primary focus of most pronunciation studies and teaching methods, a pronunciation technique that also improves learners' fluency provides benefits by helping learners communicate successfully in their second language (L2).

The challenges learners face in terms of their comprehensibility will vary based both on L1 influence and learners' individual differences. One partial solution to this problem is to find activities and techniques that can give learners opportunities to notice the gaps between their own pronunciation and that of their target language without limiting learners to a specific feature. One technique that shows promise in this area is *shadowing*, a technique that has been common in pronunciation teaching manuals and classrooms for many years, but has only recently received much attention from L2 researchers. Shadowing offers learners a way to practice their pronunciation (thus potentially improving comprehensibility) without the need for explicit instruction.

Shadowing is an activity where learners imitate a presented speech stimulus "as closely and quickly as possible" (Luo, Shimomura, Minematsu, Yamauchi, & Hirose, 2008, p. 4) though the repetition can be near simultaneous or have a small delay (e.g., Goldinger, 1998; Hiramatsu, 2000; Schweda-Nicholson, 1990). While shadowing has existed for decades, its roots are not in language instruction. Shadowing has been used in cognitive psychology for testing selective attention (Bovee & Stewart, 2009) and as a language therapy to help treat stuttering (Li-Chi, 2009). In Japan, it is a popular, though somewhat controversial, technique for training simultaneous interpreters (Bovee & Stewart, 2009). In fact, searches for recent research on shadowing yield more articles on Japanese interpreter training than L2 language training; however, in Japan it has spread from interpreter training to become a popular and common classroom activity (Hiramatsu, 2000; Saito, Nagasawa, & Ishikawa, 2011). While shadowing (which can also be referred to as mirroring and tracking) may not have originated in the language classroom, it has had a place in language training for many years and is included in popular

pronunciation teaching guides and handbooks (e.g., Avery & Ehlich, 1992; Celce-Murcia, Brinton, & Goodwin, 2010). Further, for over thirty years, articles promoting shadowing for pronunciation instruction and offering suggestions on how to use it in the classroom have been published in journals and presented at conferences (e.g., Acton, 1984; Hardison & Songchaeng, 2005; Monk & Meyers, 2004; Rosse, 1999; Zielinski, Meyers, McGregor, & Reed, 2015). A search of “pronunciation shadowing” on Google[®] will reveal a multitude of instructor- and learner-oriented websites promoting shadowing for language development. With advances in mobile technology, shadowing is an activity that can easily be completed by learners outside of a classroom setting and at minimal cost, making it potentially useful for learners who do not have access to formal pronunciation instruction.

Despite the use of shadowing in many classrooms, the actual research that has been conducted on shadowing as a language-learning tool is limited. This may be partially because it is reminiscent of the much-maligned audiolingual approach to language teaching, leading detractors to argue that it is “just vocalized repetitions and only results in meaningless parrot-like practice” (Bovee & Stewart, 2009, p. 20). However, anecdotal experiences of instructors as well as research that has been done on shadowing to date, have shown the technique, when used well, is promising for pronunciation improvement. Our own interest in this subject stems from conversations with highly successful language learners who have revealed repeatedly that shadowing dialogues from TV shows or movies had been a large part of their language practice activities at home. Case studies of highly successful language learners have reported similar findings with successful learners often attributing their success in part to substantial amounts of time practicing imitating voices from speech models (Ding, 2007). Martinsen, Alvord, and Tanner (2014) investigated the role of motivation, instruction, cultural sensitivity, and time studying abroad on accentedness ratings of 102 learners of Spanish. They found that:

surprisingly, the highest rated learner in the study did not have extensive experience abroad... However, beginning as much as six months prior to studying abroad, she developed a plan to improve her pronunciation. Three to four days each week, she spent a minimum of 15 minutes listening to Spanish news-casts or other online media and imitating as closely as possible the speaker’s words and phrases (p. 74).

Of course, these individual cases of successful language learning are not sufficient evidence of the efficacy of shadowing. However, the studies that have been done indicate that shadowing is an effective technique for improving various aspects of L2 language development.

Many of the studies that have investigated the efficacy of shadowing have been primarily interested in the role of shadowing in listening comprehension. Overall,

these studies have found shadowing to be effective for this language skill (see Bovee & Stewart, 2009, and Hamada, 2014, for overviews). There has also been some interest in the effects of shadowing on general measures of speaking proficiency (e.g., Li-Chi, 2009). However, studies have also investigated the efficacy of shadowing for pronunciation improvement (e.g., Bovee & Stewart, 2009; Hsieh, Dong, & Wang, 2013; Mori, 2011; Rongna & Hayashi, 2012). All of these studies found some improvements in the speech measures they used for analysis. However, these studies have not assessed whether shadowing can lead to changes in comprehensibility, nor have they investigated whether changes can be perceived in extemporaneous speech. The speech samples used for analysis came from participants completing either a shadowing task or read aloud tasks. The measurements were most often made by computers or, in one case, a combination of computer-based acoustic analysis and one expert rater (Rongna & Hayashi, 2012). Only one of the studies used multiple human raters to gauge improvement. In that study (Bovee & Stewart, 2009), eight L1 speakers of English rated randomizations of pre- and post-test shadowing samples. However, the raters were asked to judge the speech samples “on the basis of overall quality (i.e., closeness to native-like pronunciation)” (p. 892). The other studies measured discrete features including pitch accent, intonation, final lengthening, and pronunciation of words, though one study (Hsieh et al., 2013) also included a computer analysis of “overall pronunciation” (p. 51). For this reason, all of these studies add to the evidence that shadowing can improve pronunciation; however, they do not indicate whether shadowing can help learners be better understood when speaking freely.

Due to the increase of portable technologies such as smart phones, tablets, and digital music players, shadowing is an activity that is now easy to implement as a homework activity. This is particularly true as an increasing number of apps are being produced which allow users to layer recorded tracks on top of one another. Further, for learners who are not enrolled in a pronunciation class, shadowing offers a cost-effective way of practicing pronunciation independently. This is especially important given that pronunciation does not always get much attention in language classes (Foote, Trofimovich, Collins, & Soler-Urzúa, 2016). While much of the research on computer-assisted pronunciation instruction has focused on using desktop or laptop computers, the consumer demand for pronunciation materials on mobile technology platforms has led to a wide range of commercial pronunciation and accent reduction apps appearing on the market in the past few years (Foote & Smith, 2013). However, in order for shadowing to be effective when used independently with mobile devices, learners need to see its value and choose to use it rather than seeing it as the “meaningless parrot-like practice” it has been accused of being (Bovee & Stewart, 2009, p. 20).

There is some research to draw on when looking at learners' attitudes towards shadowing. Some of the studies that have examined the impact of shadowing on pronunciation have also included surveys asking for learners' opinions of the activities. Li-Chi's (2009) study had 25 eighth grade students in Taiwan take part in shadowing activities for 15 hours over five weeks. Questionnaires and semi-structured interviews with the participants revealed that, overall, participants felt more confident about their speaking abilities after completing the shadowing project. Some participants also noted that they thought shadowing was good for studying outside of class, and that the use of recorders enabled them to find and correct their own errors. However, over a quarter of the participants reported that they found shadowing difficult and some noted that they found the activities overly repetitive and boring.

Only one study has asked learners about their opinions of shadowing activities conducted outside of class time. Bovee and Stewart (2009) had learners engage in shadowing activities as homework assignments completed using computers. They found that 67% of respondents thought their pronunciation of individual words had improved, 73% thought intonation improved, and 80% thought the activity had educational value. The participants were also given the option of writing comments about shadowing, with 64% classified as positive comments and 34% classified as negative. Of the negative comments, 30% related to technical and logistical issues, 20% were about the difficulty of the task, 20% found it too time consuming, and the remaining comments related to the conditions of the computer labs at the university where the study took place. These surveys suggest that students' experiences with shadowing have been generally, but not entirely, positive.

In sum, a number of studies have demonstrated that shadowing is a potentially useful activity for learners who want to improve their pronunciation. With the ubiquity of mobile technology, it is easy and affordable to implement shadowing in or outside of a classroom setting. However, there is little research addressing whether shadowing can lead to improvement that can be perceived by non-expert human raters. Further, if learners are able to improve through their ability to shadow enough to be detectable by human raters, will those changes also manifest in improved ratings for accentedness, comprehensibility, and fluency in extemporaneous speech? In addition, in order for shadowing to be recommended for learning outside of a classroom or lab setting, it is important to understand how learners perceive the activities.

In light of the issues and previous studies reviewed above, the research questions were as follows:

- 1) Can regular individual practice using shadowing with mobile technology in an ESL context improve ratings of advanced L2 English speakers' (a) ability to imitate a speech model and (b) comprehensibility, accentedness, and fluency in extemporaneous speech?
- 2) How do advanced L2 learners feel about doing shadowing activities with mobile technology in their own time to improve their pronunciation and other language skills in an ESL context?

2. Method

2.1 Participants

Twenty-two L2 speakers of English were recruited from an English-medium university in Montreal, Canada. Of the original 22, 16 (male = 7, female = 9) remained for the duration of the study. Their ages ranged from 18 to 38 ($M = 25.55$, $SD = 7.13$), and they came from five different L1 backgrounds (Chinese = 10, French = 3, Arabic = 1, Bengali = 1, Russian = 1). All of the participants had a high enough level of English to gain admittance to credit programs at the university. Their time in Canada ranged from one month to 60 months, with the exception of one Francophone speaker who was born and raised in Montreal but spoke English as a second language. Many of the participants were enrolled in academic ESL credit classes at the university. Some of these courses focused primarily on reading, writing, grammar, and vocabulary in academic settings. Others were enrolled in academic oral skills classes. Overall, the participants reported spending 25–60% ($M = 53.44$, $SD = 13.90$) of their time in Montreal using English as opposed to another language. Table 1 summarizes the background information of the L2 speaking participants.

Twenty-two native English listeners (male = 6, female = 16) were recruited from a Canadian university in Alberta, Canada to act as raters. Their ages ranged from 20 to 41 years of age ($M = 25.60$, $SD = 6.05$). All reported having normal hearing. Of the twenty-two speakers, six reported speaking an L2 fluently (French = 3, Chinese = 1, Japanese = 1, French and Chinese = 1). Two of the speakers had taken linguistics classes. All participants in this study were paid for their participation. This study received approval from the ethics board of the institution where the study took place, and all participants gave informed consent.

Table 1. The L2 speakers' background characteristics

Participant	L1	Age	Gender	Months in Canada	Enrolled in ESL Class ^a	Use of English (%)
1	French	28	F	11	Oral	60
2	French	34	M	whole life	Oral	25
3	Chinese	20	F	4	Writ	70
4	Chinese	19	M	4	Writ	50
5	Chinese	30	M	12	Oral	50
6	French	22	F	23	Oral	60
7	Bengali	38	M	25	Writ	70
8	Chinese	20	M	4	Both	50
9	Chinese	21	F	1	Writ	40
10	Russian	32	F	17	Oral	50
11	Arabic	38	F	60	None	30
12	Chinese	23	M	40	Writ	75
13	Chinese	19	F	4	Writ	60
14	Chinese	19	F	5	Both	50
15	Chinese	18	F	4	None	65
16	Chinese	20	M	1	Writ	50

a. *Oral* indicates that the participant was enrolled in an academic oral skills ESL class at the time of the study, while "Writ" indicates that the participant was enrolled in an academic ESL class that did not focus on oral skills.

2.2 Materials for shadowing

An iPod with a wall charger and headphones was set up for each L2 speaker in the study. The iPods were loaded with eight audio dialogues, one for each week of the study. The dialogues were taken from popular television sitcoms including *Friends*, *The Big Bang Theory*, *How I Met Your Mother*, *Raising Hope*, and *New Girl*. Each of the dialogues had only two speakers, and all were close to one minute in length. Each clip contained a small but cohesive scene within the larger context of the show (See Appendix A for all of the clips used). All of the videos were available to watch on YouTube⁸, so that participants would have the option of viewing the scenes as well as listening to them. The iPods had a free app installed on them called *Multi Track Song Recorder*¹ and e-mail accounts set up on the iPods that the participants could use for the duration of the study. This app was originally

1. A full description of the app can be found at <https://itunes.apple.com/ca/app/multi-track-song-recorder/id390599090?mt=8>

designed for musicians, but it allowed the participants to easily load dialogues from the iTunes library onto the app, and enabled learners to listen to the recordings using ear buds while simultaneously recording themselves shadowing the dialogue. This gave users the ability to listen to their recordings being played over the original speakers, or in isolation. The app allowed for easy saving and emailing of recordings.

A booklet was also created for each participant, which included all instructions and information needed to complete the study, including clear instructions on how to use the app. The scripts were included for each dialogue as well as a URL for each of the corresponding YouTube® videos. The participants received both paper and electronic versions of the booklet.

2.3 Testing instruments

2.3.1 *Language assessments*

Two different language tests were given at each testing time. One was a picture narrative task, commonly referred to as The Suitcase Story. For this task, the participants were asked to look at a series of pictures, and when they were comfortable and familiar with the story they were asked to tell the story in the pictures. This task (available in IRIS at www.iris-database.org/) was developed by Derwing, Munro, Thomson, and Rossiter (2009) and is commonly used in pronunciation research (e.g., Derwing et al., 2014; Isaacs & Trofimovich, 2012). The second task was a shadowing dialogue similar to the types of dialogues used for practice throughout the study. For the shadowing test, the participants were given an iPod with the dialogue loaded onto it and were also given a laptop with the YouTube® video loaded and ready to play if they wished to watch it. Each participant was given 10 minutes to practice the dialogue, during which time, the researcher left the room. They were instructed to only shadow one of the speakers in the dialogue. After 10 minutes, the researcher returned and the participants were audio-recorded doing the shadowing dialogue they had just practiced.

2.3.2 *Interviews*

In-person interviews were conducted at each testing point. The interviews included Likert-scale and open-ended questions. The first interview included questions about the participants' language use and views on pronunciation. The second and third interviews asked for the participants' opinions about the shadowing activities and also asked for suggestions for improvements. The third interview was of primary interest, as the participants participated in this interview after completing all of the shadowing activities. The questions used in this interview can be found in Appendix B.

2.4 Procedure

2.4.1 *L2 speakers*

The L2 participants met with the researcher individually three times (at the start and end of the study and during week six). The first sessions lasted for about an hour each, though some were longer if the participants had difficulty with the iPods or gave longer than usual responses during their interviews. At the first session, participants were given training on how to use their iPod and clear instructions about what was expected of them throughout the course of the study. They practiced shadowing and using the app to e-mail recordings of the practice attempts. They then completed their first interview and the two pretests.

Every week for eight weeks, the participants practiced shadowing using the dialogue assigned for that week. As part of joining the study, they were asked to commit to practicing at least four times per week, for at least 10 minutes each time, though more practice was allowed and encouraged. To ensure that participants practiced at least four times per week, they were asked to submit a sample recording via email each time they practiced. They were also required to email a report each week outlining how long they practiced at each session. Each new practice week started on a Monday and if no audio files had been e-mailed by the following Friday, the participants were e-mailed a reminder. They also received a reminder if their weekly reports were not submitted by Tuesday morning of the following week. In order to be invited to the second and third sessions, the participants had to keep up with sending in their audio files and reports (though small lapses such as missing one audio file in a week, were overlooked). Each Tuesday, the participants would either receive an e-mail thanking them for sending everything in, or an e-mail reminding them of what they were missing. If a technical problem arose that could not be solved via e-mail, the researcher would meet with the participant as soon as possible and either fix the problem or issue the participant a new iPod.

The participants were not given strict rules about how to practice shadowing apart from having shadowing explained and demonstrated for them and being told they must spend at least some of the practice time shadowing, and some of the time listening to their own recordings. This choice to have a less rigid practice structure was made in order to uncover how learners like to practice when they aren't being given strict instructions. This also simulates more accurately how this activity would be used outside of a research study. Along with the language assessments and interviews, numerous tests of individual differences were also administered both at the second and third sessions, but these are not reported in this paper. In the second session, the participants also completed a language background questionnaire. At each session participants were paid for their time.

2.4.2 *Native English-speaking raters*

The raters scheduled 90-minute individual appointments with the researcher. Language background questionnaires were administered. Then suitcase stories from all three sessions were played on a computer using a computer-based rating program, created in MATLAB, developed by Saito, Trofimovich, and Isaacs (2015). Following the conventions of previous studies (e.g., Derwing & Munro, 2013; Derwing et al., 2008) the first 20 seconds of the narratives were extracted for the ratings, and initial dysfluencies were removed. The MATLAB program uses a sliding scale for the ratings and the placement of the slider results in a score between 1 and 1000. The raters then listened to each speech sample, and rated it for accentedness, comprehensibility, and fluency. They were given explanations of the meaning of all three constructs. When these ratings were completed, they were asked to rate the shadowing dialogues from all three times with instructions to rate how well the L2 speakers were able to imitate the speech model. To ensure that the raters were very familiar with the speech sample, and to give them a sense of the difficulty of the task, they were trained on the iPods, and given eight minutes to practice the test dialogue before they began rating. In both The Suitcase Story rating task and the shadowing rating task, every rater heard a different randomization of the files, and the raters were not aware of which testing time each speech sample was taken from. The participants listened to each speech sample one time only before making their ratings. Intraclass correlation coefficients were used to calculate the inter-rater reliability of the raters. A high degree of agreement was found for all four measures: shadowing ($\alpha = .86$), accentedness ($\alpha = .91$), comprehensibility ($\alpha = .89$), and fluency ($\alpha = .93$).

2.5 Analysis

2.5.1 *Speech measures*

The ratings from the L2 listeners for the pre-, mid-, and post-tests were all analyzed using a one way repeated-measures ANOVA. In cases where the ANOVA was significant, post hoc tests were run.

2.5.2 *Interview data*

Responses from the interview data were analyzed based on the nature of the questions. Likert-scale questions were asked at the midpoint and again at the end, asking participants to rate on a 9-point scale, how much they enjoyed the shadowing activities and how much they thought the activities helped improve their pronunciation. Ranges and average scores were calculated for these questions. Other questions had short answers, and for these, results were analyzed based on re-

sponses. Questions with longer detailed responses were analyzed using emergent coding and illustrative quotes were extracted.

3. Results

3.1 Speech measures

The first research question asked whether the shadowing activity would improve learners' ability to imitate a speech model, and further, whether it could improve accentedness, comprehensibility, and fluency in extemporaneous speaking tasks. The means and standard deviations of the ratings of the L2 participants can be seen in Table 2.

Table 2. Means and standard deviations for the four tasks at Time 1, Time 2, and Time 3

Measure	Time 1		Time 2		Time 3	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
Shadowing	439.47	132.66	494.79	140.20	488.90	132.66
Accentedness	473.75	66.36	454.10	102.29	447.13	102.26
Comprehensibility	653.05	116.71	664.62	133.22	682.69	117.67
Fluency	505.38	122.67	510.55	123.50	548.68	127.24

Note. Scores are based on 1–1000 ratings.

The speakers showed overall improvement on all measures apart from accentedness. In order to see whether these changes were significant, one-way repeated measures ANOVAs were run on the each of the four variables with alpha for significance testing set at .0125 due to the large number of tests run (see Table 3). None of the variables violated Mauchly's test so the results assume sphericity. Overall, the participants showed significant improvement on all measures apart from accentedness, where the scores were slightly lower, albeit non-significantly.

Table 3. One-way way repeated-measures ANOVAs for the four tasks

Measure	<i>df</i>	<i>F</i>	<i>Sig.</i>	<i>partial</i> η^2
Shadowing	2	14.26	.0001*	.501
Accentedness	2	3.30	.05	.136
Comprehensibility	2	5.00	.01*	.192
Fluency	2	8.42	.0001*	.286

* $p < .0125$

Because the ANOVAs for shadowing, comprehensibility, and fluency were significant, post hoc tests were conducted with a Bonferroni adjustment. For shadowing there was a significant improvement from Time 1 to Time 2 ($p < .001$; $d = .41$) and Time 1 to Time 3 ($p < .002$; $d = .36$) but there was not a significant change from Time 2 to Time 3 ($p < 1$; $d = -.04$). Post hoc comparisons of the comprehensibility scores showed significant improvement from Time 1 to Time 3 ($p < .002$; $d = .25$) but not from Time 1 to Time 2 ($p < .934$; $d = .09$), or Time 2 to Time 3 ($p < .197$; $d = .14$). Fluency scores showed significant improvement from Time 1 to Time 3 ($p < .006$; $d = .35$) and Time 2 to Time 3 ($p < .013$; $d = .30$) but not for Time 1 to Time 2 ($p < 1$; $d = .04$).

3.2 Interviews

3.2.1 Overall opinions

In order to get an overall sense of how the participants felt about the shadowing activities, at the Time 2 and Time 3 interviews the participants were asked to rate their overall enjoyment of the activities and their perceptions about the effectiveness of these activities. These ratings were completed using 9-point Likert scales, with 1 being the lowest, and 9 being the highest rating. The results can be seen in Table 4.

Table 4. Overall opinions about shadowing activities on 9 point likert-scale

Questions	Time 2	Time 3
How much do you like the shadowing activity?	7.50 (4–9)*	7.63 (6–9)
How much do you think it is helping your pronunciation?	6.81 (4–9)	7.5 (6–9)

* figures in parentheses connote range

Similar open-ended questions were also asked to get a more nuanced sense of participants' opinions about shadowing. These responses are taken from the third interview, when participants had completed all of the required activities. When asked, "How did you find the shadowing project overall?" all but one of the comments were generally positive. The negative comment indicated that the activity could be tedious.

If I could do it with somebody for me it would be more, how do we say it, more interesting... yeah it was not bad. We had the challenge of, you know, doing this four times. Sometimes I don't feel like doing it, I say, okay, I've got to do it.

(Part 2)

There was also one participant who noted that at the start of the project she didn't believe the activities worked, but as time passed she felt they were improving

her fluency. One other indicated that although the activities were helpful, he still needed more practice. Participants gave several reasons as to why they felt positive about the shadowing activities. Some participants liked the obligatory nature of the practice activities because it pushed them to improve. This can be seen in the comment below:

I think it's useful... we have to you know, it's like a job. When you accept a task it's like a job and it force you to practice it like every week so it will keep you on the, uh, it's like accent. You can perfect your accent and sound and like the speed and you can also learn some interesting story from the dialogue so, yes, I find it's good.
(Part 3)

Others noted specific skills that they felt had improved, the relatively small amount of time required to complete the activities, and having access to authentic materials.

The participants were also asked "Do you think this is an effective technique for improving pronunciation?" All participants answered positively and some offered reasons for why they found instruction effective. One participant indicated that he found the activity more useful than traditional classroom instruction.

I mean compared to the class, the classical class I'm taking, I mean the university courses you know, always testing the rule... but it's not the right thing to do when you really want to change the student, the real way of speaking you know.
(Part 5)

Others indicated that it was useful in the absence of a speaking partner. For example, Part 10 stated, "Yes, it's um, especially it useful when there is no one who can shadow you. So the recording device would serve as a partner."

When asked about whether they believed their own pronunciation had improved, 10 of the participants gave positive responses. However, one did not believe he had improved:

I won't say a lot but I think my French background will always stays. I think the better way for me, I think to be really Anglophone Anglophone is like to leave Quebec, live at New York for one year and be really obliged to speak English and I cannot speak French.
(Part 2)

Five others gave qualified responses. For example, Participant 14 wasn't sure how much she had improved: "Uh, yes it improved, but I don't know how good I improve." Another hoped she had improved but wasn't sure, and one thought he had improved a little bit. Two participants indicated that they thought a large improvement would take longer than the eight weeks of the project.

Uh, I think so. But I want be better. Yeah sometimes still some Chinese accent so I want like you, like the you know native speaker, but it requires a long time.

(Part 4)

Yes, little bit but it's not too long time so if I go on and practice I think I will improve more.

(Part 13)

Among the participants who felt positive about their improvement, one noted that she felt she could better express her emotions though her pronunciation, and one mentioned that she felt more confident talking with native speakers.

The participants were also asked if there was anything they thought would improve the shadowing activities. The responses to this question are summarized in Table 5.

Table 5. Suggested improvements for the shadowing activity

Improvement	Frequency of comments
Technical improvements	4
Content improvements	4
Getting feedback	2
Practicing with others	2
More intensive practice	2

The most common responses related to technical improvements and content improvements. The technical comments included issues with glitches with e-mail and YouTube[®] videos, as well as suggestions for improvements to the app such as a pause function. The comments about content included requests for more varied types of content (e.g., stories as well as dialogues), for participants to have a range of materials from which to choose, for more difficult dialogues, and for dialogues that slowly increased in difficulty as the study progressed.

Finally, the participants were asked whether they would recommend the shadowing activities to friends who wanted to improve their pronunciation. All of the participants said that they would and two mentioned that they already had. One of the participants said that she had recommended it to her friends in China, and another said that she planned to continue with shadowing when she returned to China and would no longer have access to L1 English-speaking interlocutors.

One question that needs to be addressed as part of the results is the six participants who were not interviewed because they did not complete the study. There is a possibility that these participants did not complete the study because they did not like the activity or became bored with it, which could skew the interview data. However, it should be noted that of these six participants, four did not complete

even the first week. Therefore, while they may not have liked shadowing, it is also fair to say that they did not attempt it long enough for boredom to be a factor or to really get a sense of whether or not it was helping their speech. As for the other two, one dropped out of the study after week 3, and the other week 6. For these two participants, it is possible that boredom and frustration with the activity was a contributing factor to their decision to leave the study.

4. Discussion

This study investigated whether using mobile technology for individual practice with shadowing could improve advanced English learners' ability to imitate speech models and also improve their comprehensibility, accentedness, and fluency when completing an extemporaneous speaking task. Listener judgments indicated that the participants significantly improved in their ability to imitate a speech model, and also improved in terms of comprehensibility and fluency. However, accentedness did not improve. The study also investigated how language learners felt about doing shadowing activities on their own time to improve their pronunciation. Overall, the responses indicated that participants were positive about the shadowing activities, and saw them as an effective way to improve their pronunciation.

4.1 Speech ratings

As was discussed earlier, when judging the utility of pronunciation training techniques, it is important to know whether changes in pronunciation after an intervention are such that they can actually make a difference in how a learner's speech is perceived by other listeners. For this reason, human raters rather than linguistic measures were used to assess whether the participants' ability to imitate the speech model improved. The ratings showed that the changes in the participants' speech were noticeable to untrained listeners. More importantly, this improvement was also borne out in the comprehensibility ratings of the extemporaneous speaking task and the effect sizes for all of the significant results were large. Further, the improvements in the participants' fluency ratings indicate that shadowing may carry a potential additional benefit to learners in helping them improve their fluency. Combined, these results indicate that shadowing shows a great deal of promise for helping learners improve their L2 speech. The post hoc analysis revealed that while there was overall improvement on these measures, there was not always improvement from one testing time to the next. This suggests that there may be minimal thresholds for noticeable improvement and possible points at which the

benefits of shadowing decline. However to understand this fully, more research would need to be done.

The one measure that did not show improvement was accentedness. This finding reinforces the evidence demonstrating that accentedness is only partially related to comprehensibility (Munro & Derwing, 1995) and echoes similar findings from a pronunciation intervention by Derwing et al. (2014) which found that the comprehensibility ratings of their participants improved after a pronunciation course, but that accentedness scores actually worsened significantly when completing the same suitcase story task used in this study. However, in the Derwing et al. study, two different extemporaneous speaking tasks were used, and accentedness did improve on the other task. It is well documented that listeners are extremely sensitive to accent, with Flege (1984) finding that participants were able to detect a foreign accent in speech even when hearing only a part of one /t/ segment. It may be that very small shifts in pronunciation during a given utterance can lead to differences that are detectable to listeners but unimportant to having speech that is easy to understand. Given that the participants did show improvement in comprehensibility, a lack of improvement in accent is not of great concern.

4.2 Learner opinions about shadowing

This study strove to have high ecological validity by asking learners to complete the activities in their own time, and, for the most part, to practice how they liked. However, the participants were held accountable for a minimum amount of practice by their weekly reports and by the four speech samples they e-mailed the researcher each week. Further, the participants were paid for their participation, and self-selected to do the study. For learners studying on their own, the only accountability they will have is to themselves. If learners don't like shadowing, or don't believe it is effective, they are unlikely to do it for very long. Further, an instructor asking learners to spend a significant amount of time outside of class on shadowing activities would want to know that learners are likely to find these activities reasonably enjoyable and beneficial. Overall, the participants in this study were very positive about the shadowing activities. All of the participants believed that it was an effective technique for improving pronunciation, though not all were confident that they had made large amounts of improvement. While some participants mentioned that they would have found the activities more enjoyable if they were able to work with other people, it was also noted that the activities were an effective way to get speaking practice when speaking partners were not available.

It is also interesting that the Likert scale ratings for how much the participants enjoyed the activity, and for how effective they thought it was, were higher at the end of the study than they were partway through. This suggests that despite Bovee

and Stewart's (2009) claim that shadowing is seen as "meaningless parrot-like practice," learners actually tend to appreciate it more when sticking with it over a longer period of time. This finding is echoed in Ding's (2007) findings from interviews with highly successful language learners who, when talking about imitation-based activities, noted that they "had been initially forced to use these methods but gradually came to appreciate them" (p. 272). This is not to say the participants initially disliked the activities, but rather that their appreciation of the activities and their belief in their efficacy increased over time rather than waning.

4.3 Implications for instruction

Shadowing activities with mobile technology may offer a valuable tool for learners looking for ways to improve pronunciation on their own, or for instructors hoping to provide additional support to language learners who struggle with pronunciation. However, it should be noted that while the activities used in this study show promise, they should not be considered as a replacement for classroom-based pronunciation instruction. While shadowing may help learners become more comprehensible and fluent in their L2, there is not sufficient evidence to suggest that shadowing alone can help learners improve all aspects of speech that may impact comprehensibility.

If choosing to use shadowing activities with learners, it is important to give careful consideration to the models used for shadowing. While many participants liked using sitcom dialogues because of their idiomatic language, others would have preferred different content. This study had learners use the same dialogues in order to control content as a variable in the study. However, in practice it may make sense to allow learners to choose their own speech models based on their language learning goals, or to choose speech models that are appropriate for specific learner groups.

4.5 Limitations and suggestions for future research

This study had several limitations that could be addressed in future research. First, due to participant self-selection and a lack of a control group, the results must be interpreted with caution. Many of the learners were enrolled in ESL classes during the course of the study, and all had exposure to English outside of class. However, it should be noted that in a study tracking language development for L1 Mandarin and Slavic speakers of English speakers in an ESL context, Derwing, Munro, and Thomson (2008) found that for the Mandarin participants there was no statistically significant improvement in comprehensibility or fluency after two years and only a small improvement for the Slavic speakers. These participants had just

arrived in Canada at the start of the study, spent the first year of the study in full time ESL classes, and were assessed using *The Suitcase Story* that was used in the present study. These findings do not make up for a control group, but do suggest that attributing improvements to the shadowing activities is not unreasonable.

Another limitation is the use of self-report for tracking practice time. While participants were asked to email in examples of their shadowing practice each week, the amount of time learners spent practicing was emailed in weekly reports, and as such it is not possible to be certain that all of the participants practiced as much as they claimed. Another limitation is the possibility of a practice effect from using the same tasks at three different times. However, given that the tests were spaced fairly far apart, the impact of a practice effect should be limited, and a number of other similar studies have used identical elicitation tasks for pre- and post-tests (e.g., Derwing et al. 2008).

Future studies could investigate a wider range of learner variables. Language background would be an interesting variable to study further as specific segmental and suprasegmental difficulties that learners face, which are often strongly influenced by the L1, may be more or less amenable to improvement through shadowing activities. Context is another potentially interesting variable. This study took place in an ESL context, where learners have more access to other types of English input and feedback outside of class. It would be interesting to see whether the efficacy of the activities, and learners' opinions about the activities, would be different in an EFL context. Finally, the proficiency level of the participants could moderate the impact of shadowing activities.

The use of recorders with shadowing would also be a variable that could be explored in greater detail in future studies. Shadowing does not by definition require audio recordings, and in practice, using recorders with shadowing is common, but not universal. Mobile technology makes recording easy, and may help learners notice their own errors, thus facilitating change. Using learner recordings is commonly advocated as a method for pronunciation instruction (e.g., Walker, 2005). However, this study did not compare shadowing with and without recorders so it is impossible to know how much the use of recordings impacted the efficacy of shadowing. Repetition is another aspect of shadowing that warrants further investigation as shadowing does not require learners to repeatedly practice the same material, and it would be helpful to know if repetition plays a facilitating role in improvement. Finally, the amount of time learners spent practicing was only minimally controlled in this study. Future studies could examine this variable further to try to uncover how much practice is required to see improvement.

5. Conclusion

This study demonstrated that shadowing shows promise as a way to help learners improve their pronunciation and fluency. The use of mobile technology means that this technique may be a practical solution for learners who do not have access to pronunciation instruction, and instructors who are looking to help students who need extra help with pronunciation. Due to its repetitive nature, there may be a reluctance to recommend this activity to learners, regardless of its efficacy. However, interviews with participants indicate that learners enjoyed shadowing and saw it as an effective way to improve their pronunciation. The findings of this study are in line with other studies that have investigated shadowing as a technique for pronunciation improvement; this study extends these findings by demonstrating that these improvements can be detected by untrained listeners, and lead to improved comprehensibility of extemporaneous speech making this activity of potentially high utility for learners who want to communicate more effectively.

References

- Acton, W. (1984). Changing fossilized pronunciation. *TESOL Quarterly*, 18, 71–86. doi: 10.2307/3586336
- Avery, P., & Ehrlich, A. (1992). *Teaching American English Pronunciation*. Oxford: Oxford University Press.
- Bovee, N., & Stewart, J. (2009). The utility of shadowing. In A.M. Stoke (Ed.), *JALT 2008 Conference Proceedings*. Tokyo: JALT.
- Celce-Murcia, M., Brinton, D., Goodwin, J., & Griner, B. (2010). *Teaching pronunciation: A course book and reference guide* (2nd Edition). New York: Cambridge University Press.
- Derwing, T. M., & Munro, M. J. (2013). The development of L2 oral language skills in two L1 groups: A 7-year study. *Language Learning*, 63, 163–185. doi: 10.1111/lang.12000
- Derwing, T. M., Munro, M. J., & Thomson, R. I. (2008). A longitudinal study of ESL learners' fluency and comprehensibility development. *Applied Linguistics*, 29, 359–380. doi: 10.1093/applin/amm041
- Derwing, T. M., Munro, M. J., & Wiebe, G. (1998). Evidence in favor of a broad framework for pronunciation instruction. *Language Learning*, 48, 393–410. doi: 10.1111/0023-8333.00047
- Derwing, T. M., Munro, M. J., Foote, J. A., Waugh, E., & Fleming, J. (2014). Opening the window on comprehensible pronunciation after 19 years: A workplace training study. *Language Learning*, 64, 526–548. doi: 10.1111/lang.12053
- Derwing, T., Munro, M., Thomson, R., & Rossiter, M. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition*, 31, 533–557. doi: 10.1017/S0272263109990015
- Ding, Y. (2007). Text memorization and imitation: The practices of successful Chinese learners of English. *System*, 35, 271–280. doi: 10.1016/j.system.2006.12.005

- Flege, J. E. (1984). The detection of French accent by American listeners. *Journal of the Acoustical Society of America*, 76, 692–707. doi: 10.1121/1.391256
- Foote, J. A., & Smith, G. (2013, September). Is there an App for that? Paper presented at the Pronunciation in Second Language Learning and Teaching conference, Ames, Iowa.
- Foote, J. A., Trofimovich, P., Collins, L., & Soler-Urzúa, F. (2016). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal*, 44, 181–196. doi: 10.1080/09571736.2013.784345
- Goldinger, S. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological Review*, 105, 251–279. doi: 10.1037/0033-295X.105.2.251
- Hamada, H. (2014). The effectiveness of pre- and post-shadowing in improving listening comprehension skills. *The Language Teacher*, 38, 3–10.
- Hardison, D. M., & Sonchaeng, C. (2005). Theatre voice training and technology in teaching oral skills: Integrating the components of a speech event. *System*, 33, 593–608. doi: 10.1016/j.system.2005.02.001
- Hiramatsu, S. (2000). A differentiated/integrated approach to shadowing and repeating. Retrieved from: <<http://cur-ren.cjc.ac.jp/detail/58720100222082933?!=en>>
- Hsieh, K-T, Dong, D-A., & Wang, L-Y. (2013) A preliminary study of applying shadowing technique to English intonation instruction. *Taiwan Journal of Linguistics*, 11, 43–66.
- Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility. *Studies in Second Language Acquisition*, 34, 475–505. doi: 10.1017/S0272263112000150
- Li-Chi, L. (2009). *A study of using shadowing as a task in junior high school EFL program in Taiwan*. Master's Thesis. National Taiwan University of Science and Technology, Teipei, Taiwan.
- Luo, D., Shimomura, N., Minematsu, N., Yamauchi, Y., & Hirose, K. (2008). Automatic pronunciation evaluation of language learners' utterances generated through shadowing. *Interspeech 2008*, 2807–2810.
- Martinsen, R.A., Alvord, S.M., & Tanner, J. (2014). Perceived foreign accent: Extended stays abroad, level of instruction, and motivation. *Foreign Language Annals*, 47, 66–78. doi: 10.1111/flan.12076
- Monk, M., & Meyers, C. (2004). Documenting prosodic acquisition using the mirroring project. SPLIS News, 2(1). Retrieved from: <[http://www.tesol.org/read-and-publish/newsletters-other-publications/interest-section-newsletters/as-we-speak-\(splis\)/2011/11/03/splis-news-volume-2-1->](http://www.tesol.org/read-and-publish/newsletters-other-publications/interest-section-newsletters/as-we-speak-(splis)/2011/11/03/splis-news-volume-2-1->) (December 2004).
- 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.
- Munro, M. J., & Derwing, T. M. (1995). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45, 73–97. doi: 10.1111/j.1467-1770.1995.tb00963.x
- Rongna, A., & Hayahi, R. (2012). Accuracy of Japanese pitch accent rises during and after shadowing training. Proceedings from 6th International Conference on Speech Prosody, (n.p.). Retrieved from: http://www.speechprosody2012.org/uploadfiles/file/sp2012_submission_98.pdf (May 2012).
- Rosse, M. (1999). Tracking – A method for teaching prosody to ESL learners. *Prospect*, 14, 53–61.
- Saito, K., Trofimovich, P., & Isaacs, T. (2015). Using listener judgements to investigate linguistic influences on L2 comprehensibility and accentedness: A validation and generalization study. *Applied Linguistics*. Advance online publication. <<http://applied.oxfordjournals.org/>>
- Saito, Y., Nagasawa, Y., & Ishikawa, S. (2011). Effective instruction of shadowing using a movie. In A. Stewart (Ed.), *JALT 2010 Conference Proceedings* (pp. 139–148). Tokyo: JALT.

- Schweda-Nicholson, N. (1990). The role of shadowing in interpreter training. *The Interpreters' Newsletter*, 3. Retrieved from: <http://www.openstarts.units.it/dspace/bitstream/10077/2152/1/07.pdf> (August 2012).
- Thomson, R. I., & Derwing, T. M. (2014). The effectiveness of L2 pronunciation instruction: A narrative review. *Applied Linguistics*, 36, 326–344. doi: 10.1093/applin/amu076
- Walker, R. (2005). Using student-produced recordings with monolingual groups to provide effective, individualized pronunciation practice. *TESOL Quarterly*, 39, 550–558. doi: 10.2307/3588495
- Zielinski, B., Meyers, C., McGregor, A., & Reed, M. (October 2015). *In search of a teachable model of intonation*. Paper presented at the Pronunciation in Second Language Learning and Teaching conference, Dallas, Texas.

Appendix A

Note: All transcripts are available on request from the first author.

Week 1: The Big Bang Theory

Title on YouTube[®]: Why can't Sheldon lie convincingly?

URL: http://www.youtube.com/watch?v=UKBc_Y2FibM (Time: 0:23–1:30)

Lorre, C. & Prady, B. (Creators). (2007–). *The Big Bang Theory* [Television series]. Burbank, CA: Warner Bros. Television.

Week 2: How I Met Your Mother

Title on YouTube[®]: How I Met Your Mother – Drawing a Blank

URL: <http://www.youtube.com/watch?v=Qjhugz-H7Ck> (Time 0:00–1:15)

Bays, C. & Thomas C. (Creators). (2005–2014). *How I Met Your Mother* [Television series]. Los Angeles, CA: 20th Century Fox Studios.

Week 3: Friends

Title on YouTube[®]: Lazy Chandler wins over Monica

URL: <http://www.youtube.com/watch?v=MbgJ80Rtyg4> (0:00–1:01)

Crane, D. & Kauffman, M. (Creators). (1994–2004). *Friends* [Television series]. Burbank, CA: Warner Bros.

Week 4: Raising Hope

Title on YouTube[®]: Burt Runs Out Of “Burt Bucks” from “Deja Vu Man” | RAISING HOPE | FOX BROADCASTING

URL: <http://www.youtube.com/watch?v=Co7JAwjqcYs> (0:00–0:58)

Garica, G.T (Creator). (2010–2014). *Raising Hope* [Television series]. Los Angeles, CA: 20th Century Fox Studios.

Week 5: New Girl

Title on YouTube[®]: Confession from “Double Date” | NEW GIRL | FOX BROADCASTING

URL: <http://www.youtube.com/watch?v=5Uk5-ksi0Wk> (whole clip)

Meriwether, E. (Creator). (2011–). *New Girl* [Television series]. Los Angeles, CA: Fox.

Week 6: Friends

Title on YouTube[®]: Friends – The Secret Closet [Joey & Chandler]

URL: <http://www.youtube.com/watch?v=R4VCrWYxXbQ> (0:00–1:12)

Crane, D. & Kauffman, M. (Creators). (1994–2004). *Friends* [Television series]. Burbank, CA: Warner Bros.

Week 7: The Big Bang Theory

Title on YouTube[®]: The Big Bang Theory – sheldon asks penny out on a date

URL: <http://www.youtube.com/watch?v=0TSht3i7wgA> (0:07–1:13)

Lorre, C. & Prady, B. (Producer). (2007–). *The Big Bang Theory* [Television series]. Burbank, CA: Warner Bros. Television.

Week 8 Car: New Girl

Title on YouTube[®]: Jess Hits Schmidt With Her Car from “Longest Night Ever” | NEW GIRL | FOX BROADCASTING

URL: <http://www.youtube.com/watch?v=A4cBWkgX-z0> (0:00–0:53)

Meriwether, E. (Creator). (2011–). *New Girl* [Television series]. Los Angeles, CA: Fox.

Appendix B**Questions from Final Interview**

How did you find the shadowing project overall?

1. Did the way you used the dialogue and practiced change as the experiment went on?
2. Do you feel like being able to listen to your own recorded speech was helpful or was it unnecessary?
3. Do you find the dialogues difficult or easy? Is one week too much, too little, or about right for practicing?
4. What did you think of the length of the study? Do you feel like you are getting tired of the shadowing or would you like to keep going if you could?
5. Is there anything that would improve the shadowing for you?
6. Do you think this is an effective technique for improving pronunciation?
7. Do you feel your pronunciation is improving as a result of the shadowing activities?
8. Do you feel your listening skills are improving?
9. Are there any other skills you think this type of practice is helping you with?
10. Would you recommend this technique to friends?
11. Do you feel that aside from the money, you benefitted from taking part in this project?
12. Can you think of anything else that you would like to say about this project or about the shadowing?

Rate out of 9

How much you like the shadowing activity

I don't like it

I love it

1 2 3 4 5 6 7 8 9

How much do you think it is helping your pronunciation?

Not at all

A lot

1 2 3 4 5 6 7 8 9

How much is it helping your listening skills

Not at all

A lot

1 2 3 4 5 6 7 8 9

Authors' addresses

Jennifer A. Foote
English Language School
Faculty of Extension
10230 Jasper Avenue
Edmonton, AB T5J 4P6
Canada

Kim McDonough
Department of Education, S-FG 6151
1610 St. Catherine W.
Montreal, QC H3H 2S2
Canada