

THE EFFECTS OF SELF-ASSESSMENT ACTIVITIES ON ACCURACY AND AWARENESS IN ESL PRONUNCIATION CLASSES

Kathleen Brannen, Université du Québec à Montréal

Emily Rosales, Université du Québec à Montréal

Isabelle Wouters, Université du Québec à Montréal

Paul John, Université du Québec à Trois-Rivières

In self-assessment activities, learners judge their performance in their second language (L2) with the goal of improving learning outcomes. Research has shown that these activities can have a positive impact. However, few studies have examined the effects of self-assessment on pronunciation, especially regarding segmentals. The present study focuses on some typically problematic sounds: /θ/, /ð/ and /h/. Learners tend to substitute /t, d/ or /s, z/ for /θ, ð/ (*I think that* → *tink dat/sink zat*) and speakers of certain native languages (L1s) tend to delete and, in some cases, epenthesize /h/ (*heart* → *_eart*; *apple* → *happle*). Adult learners (18F; 11M) from different L1 backgrounds and proficiency levels read short phrases aloud in the pre-/post-tests to establish progress in the pronunciation of these sounds. Test groups, with self-assessment, were compared to control groups in equivalent courses without self-assessment. Qualitative data were gathered from interviews with a subset of the test group participants. Preliminary results from quantitative data do not indicate significant differences between the test and control groups. However, interviews reveal that self-assessments improved learners' awareness of their errors and provided self-assurance in addressing them, a necessary first step to eventual improvement.

Cite as: Brannen, K., Rosales, E., Wouters, I., & John, P. (2022). The effects of self-assessment activities on accuracy and awareness in ESL pronunciation classes. In J. Levis & A. Guskaroska (eds.), *Proceedings of the 12th Pronunciation in Second Language Learning and Teaching Conference*, held June 2021 virtually at Brock University, St. Catharines, ON. <https://doi.org/10.31274/psllt.13262>

INTRODUCTION

For adults, pronunciation is perhaps the most difficult aspect of learning a second language (L2). It is notoriously resistant to change, with many adult learners fossilizing in their pronunciation progress at some point (Derwing, Munro, & Wiebe, 1997). As such, teachers are challenged to find effective methods that help students improve their pronunciation and to integrate activities that raise learners' awareness of their oral production and guide them through the learning process. One of these activities is the self-monitoring or self-assessment of one's own pronunciation.

Self-assessment is a pedagogical tool in which students judge their own comprehension and/or production in their L2 with the goal of raising awareness and improving learning outcomes. Self-assessment activities are increasingly used in formative assessment (Noonan & Duncan, 2005) and

have been shown to have a positive impact on learning (Harris & Brown, 2018; Ross, 2006). Most studies on self-assessment in language learning have dealt with grammar, but there are some that address pronunciation. Of these studies on pronunciation, many have focused on the validity of self-assessment (Foote, 2010; Maass, 2019; Trofimovich et al., 2016); fewer have addressed whether self-assessment improves pronunciation. Of the latter, investigations into the pronunciation of vowels (Lappin-Fortin & Rye, 2014; Meritan & Mroz, 2019) and suprasegmentals (Ingels, 2010) have shown that self-assessment leads to improvement, but to our knowledge, no studies have looked specifically at consonants.

The present pilot study investigates the effects of self-assessment activities on the production of three consonants that are known to be problematic for learners of L2 English: the interdental fricatives /θ, ð/ and the glottal fricative /h/. Several studies have shown that learners of L2 English tend to substitute the interdental fricative sounds /θ, ð/ with either /t, d/ (*think that* → *tink dat*) or /s, z/ (*think that* → *sink zat*); and for learners from certain L1 backgrounds, the /h/ sound is sometimes omitted (*heart* → *_eart*) and sometimes added (*apple* → *happle*). These errors occur in perception as well as in production (Brannen, 2002, 2011; Mah, Goad, & Steinhauer, 2016; Trofimovich & John, 2011; Wouters & John, 2020).

This project builds on work by Wouters and John (2020). In that study, it was shown that francophone learners of L2 English were able to detect errors in the production of /θ, ð/ and /h/ when listening to their own recordings as well as when listening to recordings of their peers. Although the authors noted that participants remarked that the perception task made them aware of their errors, the impact of this consciousness-raising on pronunciation progress was not studied.

Thus, the present study aims to determine if learners use this heightened awareness to make improvements to their production of the L2 sounds. Learners from various L1 backgrounds and proficiency levels completed weekly guided self-assessment activities and took part in a guided interview on their experience with the self-assessment activities.

Research questions

1. Do self-assessment activities lead to an improvement in the production of the interdental fricatives /θ, ð/ and in the appropriate placement of the glottal fricative /h/?
2. What do participants report concerning the self-assessment activities? Do these activities increase their awareness of their errors? Do participants feel their pronunciation improves? What is the impact on their self-confidence in speaking in their L2?

METHODS

Participants

Participants were English as a second language (ESL) students at a francophone university in Montreal, enrolled in one of three levels of pronunciation courses: Pronunciation I, II, and III, corresponding to the Common European Framework of References for Languages (CEFR) levels B1, B2, and C1. There were 29 participants, (M=11; F=18), from a variety of native languages (L1s): French (N=14), Chinese (N=4), Spanish (N=4), other (N=7). Speakers of all these L1s are known to have difficulty with English /θ, ð/ (Brannen 2002, 2011); not all these L1s show problems

with /h/, but, notably, francophones do (Mah, Goad, & Steinhauer, 2016). Of these participants, 21 were in the test groups, and for the Pronunciation I and III levels, there were corresponding control groups (N=4 for each level). (Data for the Pronunciation II control group is forthcoming.)

Procedure

Quantitative data were collected from a pre-test and post-test conducted during Week 2 of the semester and Week 15, respectively. These consisted of 10 sets of four short phrases (see Appendix A) and 25 sentences. (Only the data from the phrases are presented in this article; analysis of the sentence data is forthcoming.) Since the study was conducted during the Covid-19 restrictions, students recorded themselves remotely via the university’s online learning platforms. Throughout the semester, the test group participants completed 10 weekly self-assessments. First, the students recorded themselves reading five sentences (Pronunciation I, see Appendix B) or a short text (Pronunciation II and III, see Appendix C), then they listened to a native-speaker model reading the same text. Next, they typed their responses to questions which guided them in their self-assessments (see Appendix B). The instructor subsequently gave written feedback on their responses (for the Pronunciation II and III groups only; see Appendix D). The control groups did the pre- and post-test but did not do any part of the self-assessment activity. After the end of the semester, a subset of students from the test groups participated in interviews via Zoom (N=14). The interviews were conducted in French or English, according to the participant’s preference. Interviewees were asked how they felt about their pronunciation at the beginning of the course compared to the end of the course, with certain questions probing their impressions of their pronunciation of <th> and <h>. They were also questioned on the effect of the self-assessments on their attitudes, perceptions, and awareness.

Rating procedure

Two experienced English pronunciation instructors listened to all participants’ phrases in the pre- and post-tests and recorded the number of errors produced. A third English instructor was called upon to settle cases of disagreement. All raters were native English speakers. The rate of agreement was very high as can be seen in Table 1.

Table 1

Rates of inter-rater agreement

Inter-rater agreement			
Segment	Pre-test	Post-test	Total tokens
/h/-epenthesis	100%	99.12%	220
/h/-deletion	99.64%	99.83%	796
/θ/-substitution	95.67%	97.43%	299
/ð/-substitution	89.72%	91.82%	107

RESULTS

Quantitative results

For each group, mean error percentages were calculated for each of the four categories of segmental errors. As well, we calculated the effect size with Cohen's *d*. The relatively low number of participants in this study gives an abnormal distribution, therefore results were submitted to non-parametric tests (Wilcoxon Signed-Rank and Mann-Whitney). The Wilcoxon test was used to compare scores between the pre- and post-test for each group and to compare error type scores at pre-test and then at post-test. The Mann-Whitney test was used to compare scores between groups at pre- and post-test. Figures 1, 2, and 3 show the results from Pronunciation I, II, and III respectively. In each figure, error percentages and standard error bars are shown for each error type, and the number of target segments is indicated next to each type.

Figure 1

Percentage of errors per type for the Pronunciation I groups. Bars show standard error. Number of target tokens is shown next to the type.

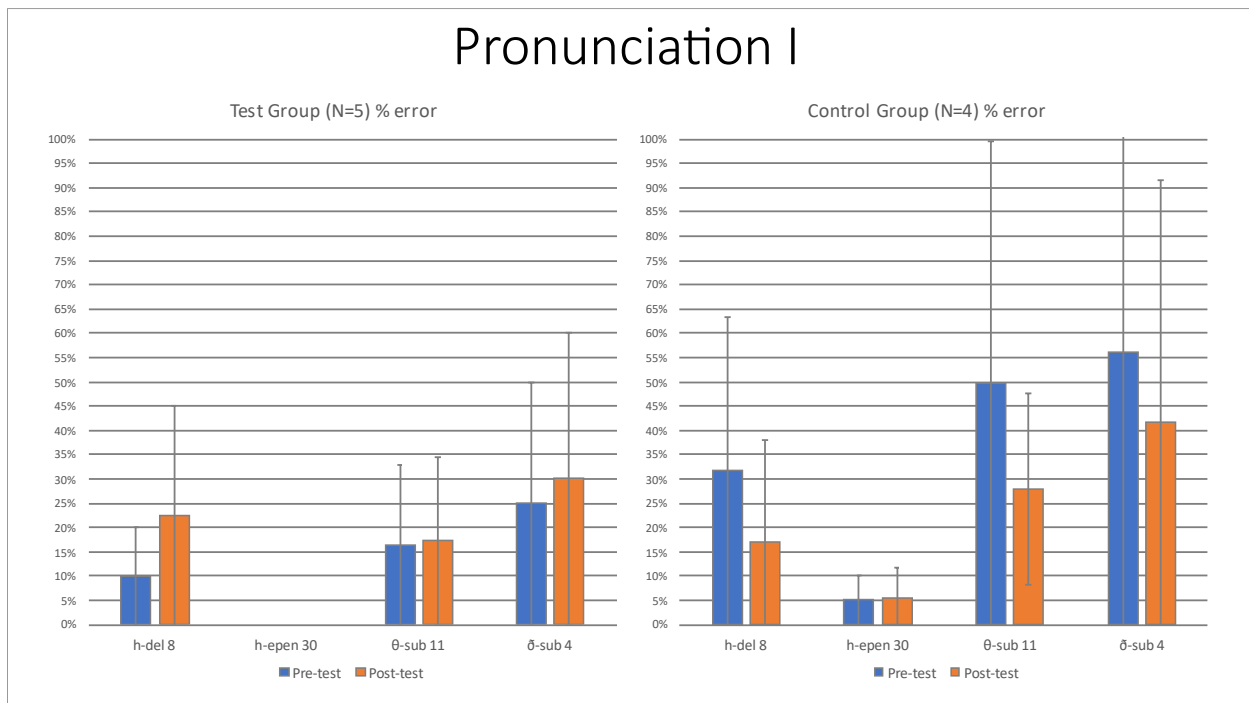


Figure 2

Percentage of errors per type for the Pronunciation II group. Bars show standard error. Number of target tokens is shown next to the type

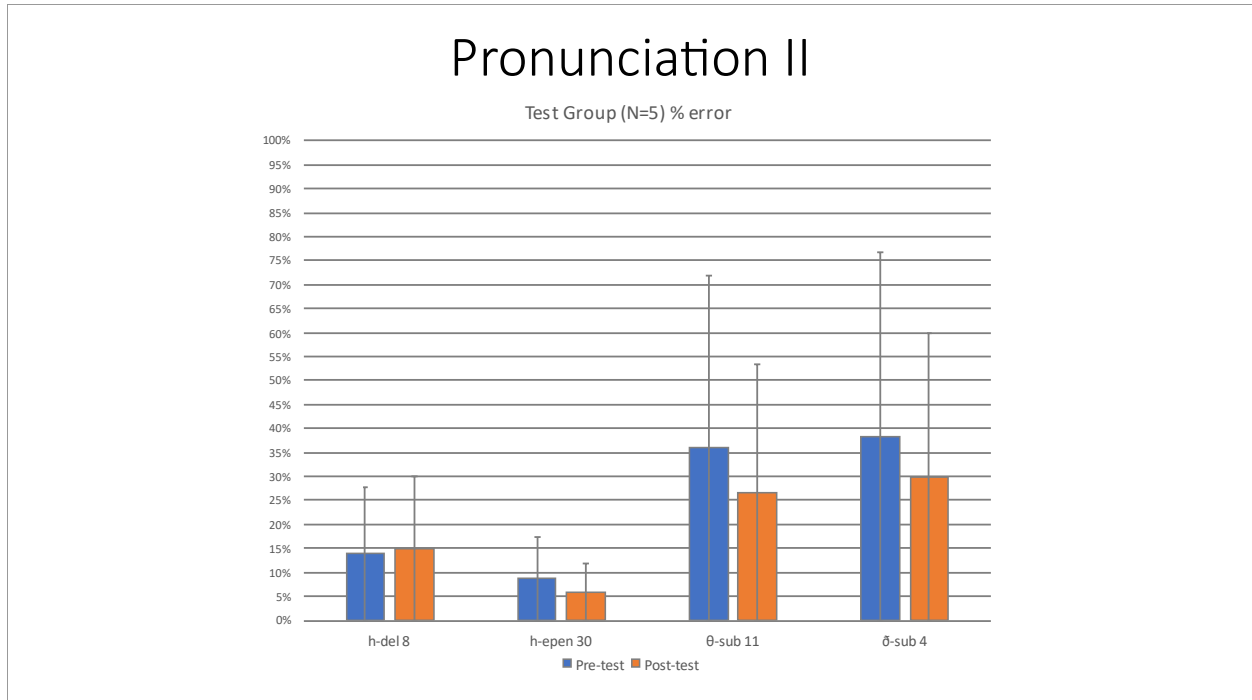
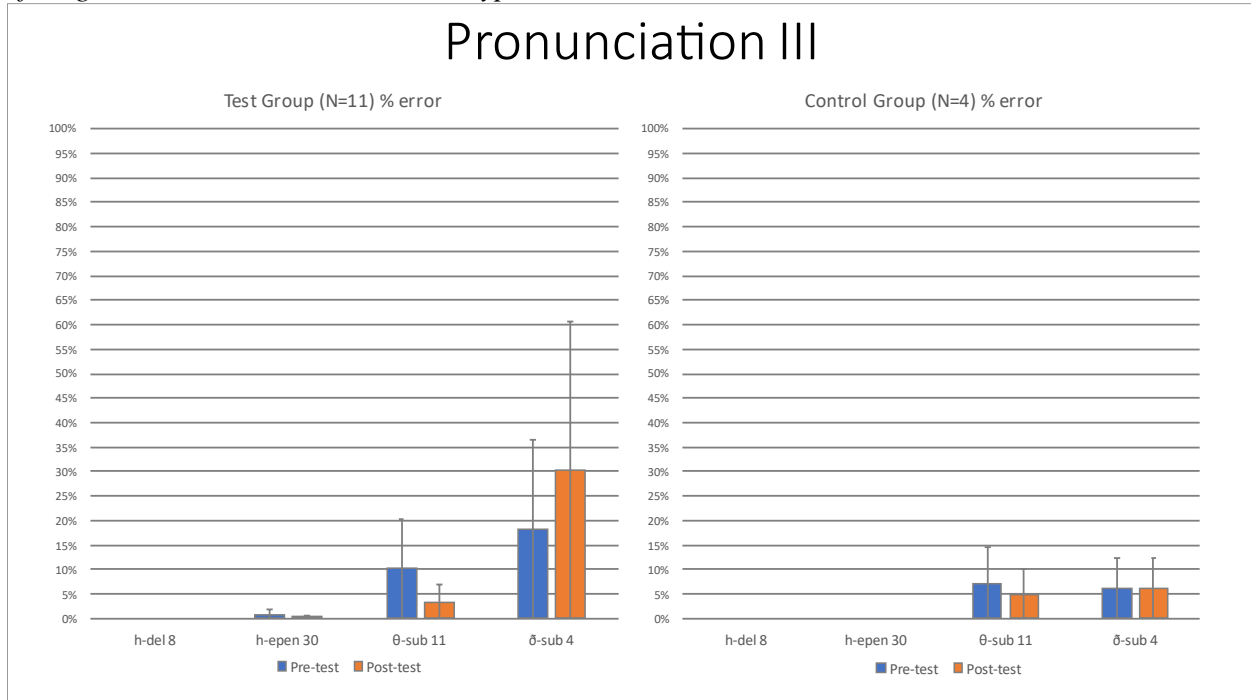


Figure 3

Percentage of errors per type for the Pronunciation III groups. Bars show standard error. Number of target tokens is shown next to the type



Analyses did not reveal any significant differences between the test and control groups nor between the pre- and post-tests within each group. The only significance found was between items when all levels were grouped together. These are shown in Table 2 (pre-test) and Table 3 (post-test). Table 2 indicates that during the pre-test for all test groups, learners made significantly more /θ/-substitution errors (15%) than /h/ errors: /h/-deletion (5%) ($z = -2.40, p = .016, d = 0.75$) and /h/-epenthesis (2%) ($z = -2.93, p = .003, d = 1.25$). There were also more /ð/-substitution errors (23%) than /h/ errors: /h/-deletion (5%) ($z = -2.49, p = .013, d = 0.89$) and /h/-epenthesis (2%) ($z = -2.88, p = .004, d = 1.24$).

Table 2*Statistical analyses, pre-test, all test groups and levels combined*

Pre-test - Test Groups (<i>n</i> = 21)				
Segment		Wilcoxon Signed-Rank Test		Cohen's <i>d</i>
<i>vs</i>		<i>z</i>	<i>p</i>	<i>d</i>
/h/-del	/h/-epen	-	-	0.33
/h/-del	/θ/-sub	-2.40	.016*	0.75
/h/-del	/ð/-sub	-2.49	.013*	0.89
/h/-epen	/θ/-sub	-2.93	.003*	1.25
/h/-epen	/ð/-sub	-2.88	.004*	1.24
/θ/-sub	/ð/-sub	-1.29	.197	0.34

*sig < .05

Table 3 indicates that on the post-test for the test groups, learners made more interdental fricative errors than /h/-epenthesis errors (2%) (/θ/-substitution: 12% [$z = -3.09$, $p = .002$, $d = 0.99$]; /ð/-substitution: 26% [$z = -3.22$, $p = .001$, $d = 1.69$]). Furthermore, there were more /ð/-substitution (23%) than /θ/-substitution (15%) errors ($z = 2.42$, $p = .016$; $d = 0.73$).

Table 3*Statistical analyses, post-test, all test groups and levels combined*

Post-test - Test Groups (<i>n</i> = 21)				
Segment		Wilcoxon Signed-Rank Test		Cohen's <i>d</i>
<i>vs</i>		<i>z</i>	<i>p</i>	<i>d</i>
/h/-del	/h/-epen	-	-	0.51
/h/-del	/θ/-sub	-1.49	.137	0.14
/h/-del	/ð/-sub	-1.49	.136	0.72
/h/-epen	/θ/-sub	-3.09	.002*	0.99
/h/-epen	/ð/-sub	-3.22	.001*	1.69
/θ/-sub	/ð/-sub	-2.42	.016*	0.73

*sig < .05

There were no significant differences in the control groups. Standard error bars in Figures 1, 2, and 3 show substantial variation between participants, especially on /θ, ð/ substitutions.

Qualitative results

In general, comments from the guided interviews were positive (see Appendix E). Learners found the self-assessment activities were useful, especially when accompanied by teacher feedback. Many students found the activities difficult at first, but with practice they were able to develop their self-assessment skills. Comments also revealed that these activities helped raise awareness of their pronunciation, and they felt able to extend this type of self-monitoring to their use of English in their daily lives. Interestingly, several participants noted that they paid more attention

to suprasegmentals than to segmentals in these activities. Ultimately, all participants felt the self-assessments helped them improve their English pronunciation.

DISCUSSION

Regarding the quantitative results, we did not observe significant differences between any of the test and control groups nor were there any significant differences between performance in the pre-test versus post-test. Thus, it seems that the self-assessment activities did not have a measurable impact on the pronunciation of /θ, ð/ and /h/ in short phrases. However, trends do indicate improvement in the post-test for the Pronunciation II test group, especially for the interdental fricatives /θ, ð/. (Recall that data from the Pronunciation II control group has not yet been analyzed.) For Pronunciation III, we see a trend towards improvement of the pronunciation of /θ/ in both the test and control groups.

Surprisingly, the Pronunciation I test group performed more poorly on the post-test than on the pre-test, whereas the corresponding control group showed a trend towards improvement. Closer analysis of the test group revealed that this was largely due to the one francophone participant in this group, who had poorer results on the post-test on all items except /h/-epenthesis. This participant substituted /ð/ and deleted /h/ twice as much on the post-test compared to the pre-test. If this data were removed from the group, data from the Pronunciation I test group would indicate trends towards improvement in post-test. This highlights the fact that there was considerable variation in the data, indicating a heterogeneous group of participants. To garner a better understanding of the effects of self-assessment on the L2 learners' pronunciation of these consonants, data from more participants are needed.

Regarding the differences between items, we have seen that, in general, participants made significantly more errors on /θ, ð/ compared to /h/. These differences are reflected in the effect sizes. This finding is consistent with results from Wouters and John (2020), who found that francophones were more skilled at perceiving their errors on /h/ than those on the interdental fricatives. Segments that are difficult to perceive are the same that cause more errors in production (Brannen, 2011). But why is /h/ easier to perceive and produce compared to /θ, ð/? It may be because errors on /h/ involve the addition or deletion of an entire segment, whereas for /θ, ð/, the errors involve substituting one sound for another.

On the post-test, we also observed that the test groups had significantly more /ð/-substitution than /θ/-substitution errors. This may also be due to perceptual saliency: voiced /ð/ is less salient than /θ/ because of its shorter duration and weak acoustic energy (Cho & Giavazzi, 2009). Also, distributional differences may contribute to the difference. Voiced /ð/ occurs in perceptually weak positions: in function words and in the middle or end of content words.

Qualitative data gathered from the interviews provided insight into participants' experiences. Students found the self-assessments to be difficult at first; but, with practice, they were able to develop their skills and subsequently found the activities useful, especially when accompanied by teacher feedback.

As noted above, several participants remarked that they paid more attention to suprasegmentals than to segmentals in these activities. This may partially explain why the quantitative data did not reveal significant improvement in the test groups versus the control groups. Perhaps learners are unable to pay attention to both segmental and suprasegmental aspects of pronunciation simultaneously while speaking. Limited attentional and processing resources, especially for less proficient learners, may cause them to focus on one or the other, and it may be that suprasegmentals prevail over segmentals. Future research should investigate this possibility by incorporating longer streams of speech and spontaneously produced utterances. Forthcoming analyses of sentence data may provide us with more insight on this question.

Overall, students felt that the self-assessment activities increased awareness of their pronunciation and boosted their confidence in speaking the L2. Furthermore, these activities are deemed to foster independent learning, enabling the student to extend self-monitoring to their use of English in their daily lives. Despite the mixed results from the quantitative data, all participants commented that they felt that their English pronunciation had improved thanks to self-assessment.

ABOUT THE AUTHORS

Kathleen Brannen is a lecturer at the Université du Québec à Montréal, where she teaches English pronunciation and speech perception to ESL learners and future ESL teachers. She is also a lecturer in the Department of Linguistics at the University of Ottawa where she teaches phonology, phonetics, bilingualism, first and second language acquisition, and psycholinguistics.

Emily Rosales has been a *maître de langues* at the Université du Québec à Montréal since 2004. She teaches primarily oral production courses (conversation, speaking and pronunciation) and specializes in the use of educational technologies to foster learner autonomy.

Isabelle Wouters is a PhD candidate in the Département de didactique des langues at the Université du Québec à Montréal. Her research interests include L2 pronunciation teaching, task-based language teaching and computer-assisted language learning.

Paul John is an associate professor in the Département des langues modernes et de traduction at the Université du Québec à Trois-Rivières. His research focuses on L2 phonological acquisition.

REFERENCES

- Brannen, K. (2002). The role of perception in differential substitution. *Canadian Journal of Linguistics*, 47(1), 1–46.
- Brannen, K. (2011). *The perception and production of interdental fricatives in second language acquisition* (Doctoral dissertation, McGill University). Retrieved from <https://www.bac-lac.gc.ca/eng/services/theses/>
- Cho, H. & Giavazzi, M. (2009). Perception of voicing in fricatives. *Proceedings of the 18th International Congress of Linguists*, 986–1006.

- Derwing, T. M., Munro, M. J. & Wiebe, G. (1997). Pronunciation instruction for fossilized learners. Can it help? *Applied Language Learning*, 8(2), 217–235.
- Foote, J. A. (2010). *Second language learners' perceptions of their own recorded speech*. Edmonton: Prairie Metropolis Centre Working Papers.
- Harris, L. R. & Brown, G. T. L. (2018). *Using Self-Assessment to Improve Student Learning*. New York: Routledge.
- Ingels, S. (2010). The effects of self-monitoring strategy use on the pronunciation of learners of English. In J. Levis & K. LeVelle (Eds.). *Proceedings of the 1st Pronunciation in Second Language Learning and Teaching Conference* (pp. 67–89). Iowa State University.
- Lappin-Fortin, K. & Rye, B. J. (2014). The use of pre-/posttest and self-assessment tools in a French pronunciation course. *Foreign Language Annals*, 47(2), 300–320.
- Mah, J., Goad, H. & Steinhauer, K. (2016). Using event-related brain potentials to assess perceptibility: The case of French speakers and English [h]. *Frontiers in Psychology*, 7, Article 1469.
- Maass, M. (2019). Teaching pronunciation through self-monitoring and reflection. *The Journal of Seigakuin University*, 31(2), 19–29.
- Meritan, C. & Mroz, A. (2019). Impact of self-reflection and awareness-raising on novice French learners' pronunciation. *Foreign Language Annals*, 52(4), 798–821.
- Noonan, B. & Duncan, C. R. (2005). Peer and self-assessment in high schools. *Practical Assessment, Research, and Evaluation*, 10(17), 1–8.
- Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment Research & Evaluation*, 11(10). Retrieved from <http://pareonline.net/getvn.asp?v=11&n=10>
- Trofimovich, P., & John, P. (2011). When three equals tree: Examining the nature of phonological entries in L2 lexicons of Quebec speakers of English. In P. Trofimovich & K. McDonough (Eds.), *Applying Priming Methods to L2 Learning, Teaching and Research: Insights from Psycholinguistics* (pp. 105-129). Amsterdam: John Benjamins.
- Trofimovich, P., Isaacs, T., Kennedy, S., Saito, K., & Crowther, D. (2016). Flawed self-assessment: Investigating self- and other-perception of second language speech. *Bilingualism: Language and Cognition*, 19(1), 122–140.
- Wouters, I., & John, P. (2020). ESL learner perception of pronunciation errors. In E. Rosales, R. Pinsonneault & J. Payeras (Eds.), *Proceedings of the 8th Meeting on Language Teaching (MeLT-RéEL-CELE)*, UQAM.
- Zoom Video Conferencing <https://zoom.us>

APPENDIX A

List of stimuli for the pre- and post-tests

1. out of order – open bar – expensive theatre tickets – push the button
2. rush around – heart surgery – easy to find – another final exam
3. ring the bell – Halloween and Easter – until nine thirty – information technology
4. behind the bushes – a sore throat – a cup of coffee – this month of August
5. on the porch – ham and eggs – three strikes – a thick sweater
6. a bad idea – a thousand dollars – pull a ligament – throw out
7. their pitcher of iced tea – from Ireland – heaven and earth – dodge a bullet
8. some hot apple pie – a large lion – my mother tongue – give to the poor
9. a broken arm – think about it – a hotel by the beach – a good judge
10. my heavy suitcase – new music video – thank you very much – a wool scarf

Note: Target segments are in red

APPENDIX B

Example of self-assessment activity for the Pronunciation I group

Instructions for Student Recording	Read the following sentences. Do your best to read as naturally as possible. <ol style="list-style-type: none">1. This is a great website to read about the weather.2. How many Canadians sat and watched all the hockey games?3. Who do you think will want to go on a hike to the top of the hill?4. I thought the group of young children was too loud.5. Anne stood on the stool to reach a cloth on the highest shelf.
Model Recording	Now listen to the model pronouncing these same sentences.
Instructions for Self-Assessment	Compare your own recording with that of the model, then answer the following questions as best you can. <ol style="list-style-type: none">1. Did you pronounce the vowels graphemes in a similar way to the model?2. Did you pay attention to the different pronunciations of the same vowel graphemes; in particular “ea”, “a”, “o”, “ou”, “oo”?3. Were there any vowel graphemes that caused you problems? If so, which ones?4. Did you pronounce the “th” sounds in the same way as the model?5. Did you pronounce the “h” correctly? Did you avoid adding “h” where it doesn’t belong?

Note: All self-assessments contained questions 4 and 5

APPENDIX C

Example of self-assessment activity for the Pronunciation II and III groups

Instructions for Student Recording	Read the following text. Do your best to read as naturally as possible.
	<p>Each winter, Canada’s capital boasts the largest naturally frozen skating rink in the world as designated by Guinness World Records: The Rideau Canal Skateway! From roughly January to late February or early March, Ontario’s only UNESCO World Heritage Site transforms into an epic urban skating rink which winds its way through downtown Ottawa. Don’t miss this unique bucket list site, especially while it celebrates its 50th anniversary in 2020!</p> <p>Nearly one million people – including visitors from around the world – experience the Rideau Canal Skateway each year. Stretching 7.8 kilometres (4.8 miles) from Ottawa’s downtown core near the National Arts Centre to spacious Dows Lake, the Skateway’s skating surface totals 165,621 square metres. That’s bigger than 105 National Hockey League rinks or more than 90 Olympic-sized hockey rinks!</p> <p>When the Skateway is open (weather permitting), it’s free and accessible 7 days a week, 24 hours a day! There’s nothing more romantic on a wintry evening than gliding with a loved one on this magical skating rink.</p> <p>Source: https://www.ottawatourism.ca/ottawa-insider/rideau-canal-skateway/</p>
Model Recording	Now listen to the model pronouncing this same text.
Instructions for Self-Assessment	After comparing your recording of <i>Rideau Canal Skateway</i> with the model, write your comments on your pronunciation here. Consider these aspects, but feel free to write other comments as well:
	<ol style="list-style-type: none">1. Did you pronounce the vowels correctly? Were there any vowels in particular that caused you problems?2. How was your pronunciation of the consonants? Did you pronounce the "th" sounds correctly? Did you pronounce the /h/ where it should be pronounced and avoid inserting /h/ where it does not belong?3. The Self-evaluations also provide you with an opportunity to ask me questions. Feel free to ask any questions you might have or to make comments.

Note: All self-assessments contained question 2.

APPENDIX D

Examples of teacher feedback on self-assessments

Pronunciation II

“To answer your question, one has vibration in your vocal cords and the other does not. The tip of your tongue should come out a little.”

“Actually, in <are happy>, we shouldn’t link because the <h> in <happy> (adj.) must be pronounced.”

Pronunciation III

“Remember that in content words that begin with <th>, this sound should be voiceless, and in most cases, when <th> is in the middle of a word, such as in <mother> or <father>, the <th> is voiced.”

“At the beginning of content words, the <h> should be pronounced, but there are a few exceptions to this rule, for example, the <h> in <hour> and <honest> is silent.”

APPENDIX E

Some participants’ comments from the guided interview

“I wouldn’t pretend I became a “th” expert, but it was more easy this time to choose between both “th”, the voiced or the voiceless. “Although”, “other” and “together” were pronounced - I believe - correctly voiced. I am more conscious now of all the rules to apply when speaking but I am still looking for how “to shape” them into a coherent whole. When I focus on the pronunciation of “th” or “h” for instance, I forget the sentence stress; when focusing on the sentence stress, I forget the aspiration of /p,t,k/ and /h/... Stress and weak forms are probably the rules I have most problems with. ”

“ [Avant le cours] Pour moi, de savoir que je ne prononçais pas les mots comme il faut ou que j’avais des difficultés, c’était comme un handicap... J’ai quand même vu de l’amélioration. J’étais même surpris de voir que ça allait quand même plus vite que je pensais. ”

Translation: “ [Before the course] For me, to know that I was not pronouncing words as I should or that I was having difficulty, it was like a handicap...Still, I saw an improvement. I was actually surprised to see that it went faster than I expected.”

“En fait, je me suis rendu compte que, souvent, même quand je parlais en anglais, c’était par habitude, ça fait que je ne portais pas attention, mais là, ce cours m’a comme obligée à prendre conscience de certaines choses. ”

Translation: “In fact, I realized that often, even when I was speaking English, it was out of habit, so I didn’t pay attention; but now, this course forced me to be aware of certain things.”