

Colloque des spécialistes en enseignement :
État, enjeux et identité au 21^e siècle

How to approach pronunciation training in the L2 classroom

Paul John, Ph. D. et Ludovik Béchard-Tremblay

Professeur au Département des langues modernes et de traduction
Étudiant à la maîtrise
Université du Québec à Trois-Rivières
paul.john@uqtr.ca
ludovik.bechard-tremblay@uqtr.ca

Abstract. Our aim is to help second language (L2) teachers incorporate the role of ‘accent coach’ into their professional identity by showing them an approach to pronunciation training in the classroom and beyond. Initially, we address what research tells us about L2 phonological acquisition, focusing on findings related to Quebec francophone problems with English /θ ð/ and /h/ (the so-called th- and h-sounds). These findings suggest that learners develop approximate or ‘fuzzy’ representations for L2 phonemes, which result in varying degrees of accuracy in perception and production. Next, we consider the implications for pronunciation training. Conceivably, focused pronunciation practice can improve learners’ ability to use fuzzy representations for accurate perception and production. Finally, we present an approach to pronunciation training implemented in a course for older learners at the *Université du troisième âge*. Our activities make use of Google Translate, a multipurpose and user-friendly tool that promotes autonomous and ubiquitous learning.

Introduction

The motivation for this paper comes from a contradiction between L2 learners’ achievements and L2 teachers’ practices with regard to pronunciation. On the one hand, learners typically exhibit a non-native accent, including pronunciation errors in the

production of L2 phonemes absent from the first language (L1) inventory.¹ For example, Quebec francophones (QFs) often struggle with the interdental fricatives /θ ð/ and the glottal fricative /h/. The tendency is to substitute /t d/ for /θ ð/ (*think that* → *tink dat*) and to delete /h/ (*happy holidays* → *_appy_olidays*) or even to insert /h/ hypercorrectly before vowel-initial forms (*old age* → *hold hage*). On the other hand, despite the challenge pronunciation represents for learners, L2 teachers typically avoid pronunciation activities (Foote, Trofimovich, Collins, & Soler Urzúa, 2013). Possibly this is because such a focus-on-form is considered incompatible with communicative language teaching (Isaacs, 2009), but also the pronunciation activities provided in textbooks are often inadequate (Derwing, Diepenbroek, & Foote, 2012) and ESL teachers may receive insufficient instruction in pronunciation training (Breitkreutz, Derwing, & Rossiter, 2001). Our aim is thus to help teachers to incorporate the role of ‘accent coach’ into their professional identity by presenting an approach to pronunciation training for the communicative classroom and beyond.

First, we review research on L2 phonological acquisition, focusing on our own findings related to QF perception and production of English /θ ð/ and /h/ (John, 2019). We then discuss the implications of these findings for pronunciation training. Finally, we explain how we approached pronunciation in an ESL course with older learners at the *Université du troisième âge*. Briefly, our activities make use of the text-to-speech and speech-to-text functionality of Google Translate. This highly flexible and easy-to-use (even for seniors) tool promotes autonomous and ubiquitous (anytime-anywhere) learning, constituting a form of ‘strategy development’ (Chapelle & Jamieson, 2008).

¹ This is at least the case for learners with little exposure to native speakers outside the classroom (e.g., John, 2019) and for older immigrant learners despite considerable exposure over many years (e.g., Flege, Munro & Mackay, 1995; Munro, Flege, & Mackay, 1996), so context and age of acquisition clearly condition L2 phonological achievement.

Background

The consensus in the literature is that L2 pronunciation errors stem from problems in perception that lead to inaccurate underlying representations (URs). This is certainly the view in previous work on QF perception and production of /θ ð/ (e.g., Trofimovich & John, 2011) and /h/ (e.g., John & Cardoso, 2009; Mah, Goad & Steinhauer, 2016). Under perceptual assimilation (Best, 1994), QFs mishear /θ ð/ as tokens of the L1 phonemes /t d/, recording items in the lexicon according to the perceived categories. As a result, *think that* is stored as *tink dat*, which explains why learners pronounce the words this way. Paradoxically, so-called pronunciation errors are in fact faithful realizations of URs. With regard to /h/, apparently no L1 category is sufficiently similar for assimilation to take place. Consequently, QFs fail to detect this segment in the speech input and by extension omit it from URs. Not all learners, however, mispronounce novel L2 phonemes. Learners with consistently accurate production of /θ ð/ and /h/ have presumably overcome perceptual assimilation and developed accurate URs.

The problem is that learners frequently exhibit variable realization of L2 phonemes. This was certainly the case in a recent study where 50 QF participants (34_{female}, 16_{male}; M=25.9 yrs) performed production and perception tasks targeting /θ ð/ and /h/ (John, 2019). Briefly, a reading-aloud task containing 30 phrases (e.g., *expensive theatre tickets; heart surgery*) and 20 sentences (e.g., *True love is hard to find; My brother is thirty-nine years old*) revealed that none of the QF participants showed categorically inaccurate production of /θ ð/ and /h/. This finding runs counter to the scenario whereby, under perceptual assimilation, the L2 lexicon contains inaccurate URs leading to inaccurate output. Some of the participants showed categorically accurate production of the target phonemes, a finding consistent with their having acquired the target phoneme category. The majority, however, exhibited variable production, alternating between accurate and inaccurate output such that *thank* → [θæŋk]~[tæŋk] and *hold* → [hould]~[_ould]. If URs are inaccurate, how is it that learners sometimes generate accurate [θæŋk] and [hould]? Or conversely, if URs are accurate, why do learners sometimes produce errors?

Interestingly, the variable producers exhibited a wide range of pronunciation accuracy on the reading-aloud task, with error rates evenly distributed across the range. For illustrative purposes, Figure 1 shows the data for /θ ð/: on 39 tokens of /θ ð/, while 14 QFs produced no errors, the remaining 36 showed th-substitution rates spread out between the extremes of 1 to 31 errors.

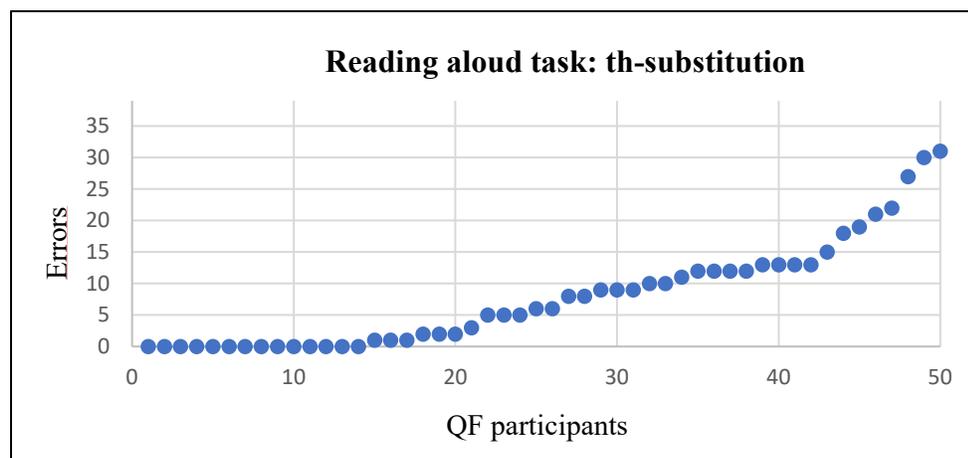


Figure 1. Error rates for /θ ð/ production (39 tokens) in the reading-aloud task

The production profile for /h/ is similar to that for /θ ð/: on 21 tokens of /h/, the 38 variable producers showed deletion rates ranging from 1 to 16, with incremental differences in accuracy between participants.

Interestingly, the same pattern of variable performance was observed in two perception tasks. One task used trials of 4 items in an oddball paradigm (e.g., *tank-tank-tank-thank*; *hold-hold-hold-old*); the other used semantically incongruous sentences based on the unexpected presence or absence of /θ ð/ or /h/ (e.g., *The soccer game ended in a thigh*; *Your car tires need more hair*) (see John, 2019, for details). On both tasks, while some QFs showed accuracy rates similar to those of 11 native speaker controls, the majority showed variable accuracy, with rates spread out across a wide range. The accuracy rates for /θ ð/ in the oddball paradigm shown in Figure 2 are representative of QF performances in both tasks and for /h/ as well.

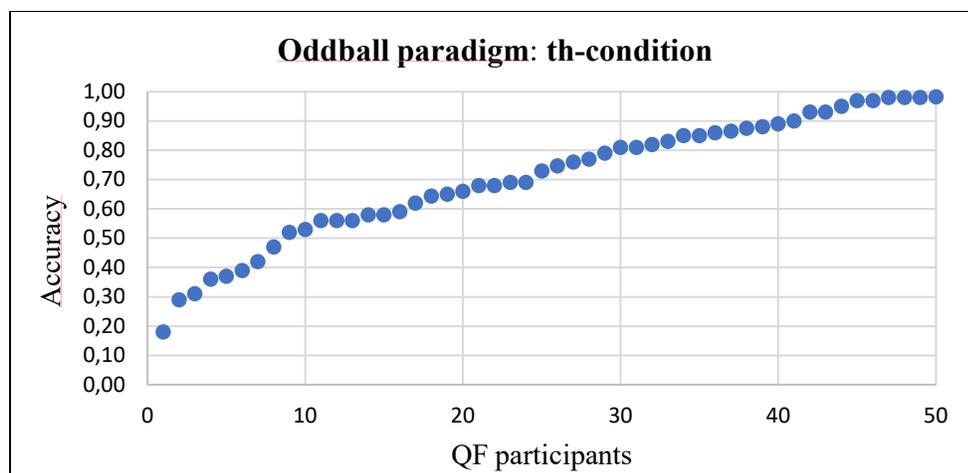


Figure 2. Accuracy rates for /θ ð/ perception in the oddball task

To summarize, most of the QF participants in John (2019) showed variable production and perception of /θ ð/ and /h/, with gradient differences in individual performances across the range of accuracy. To the extent that cross-sectional data are indicative of longitudinal development, acquisition of L2 phonemes appears to proceed incrementally. Accuracy appears to improve gradually over time, although conceivably learner development may cease at a stage short of nativelike performance.

We return now to the question of how /θ ð/ and /h/ are represented in the L2 lexicon. In our view, the observed patterns of variably accurate production/perception are consistent with so-called ‘fuzzy representation’ of L2 phonemes (e.g., Darcy, Daidone, & Kojima, 2013; see also ‘murky specification’ in John 2006). Fuzzy representation/murky specification involves tagging certain instances of /t d/ or vowel-initial forms as special. This tagging can be captured graphically via a superscript question mark such that the L2 lexicon contrasts bona fide /t d/ as in /tænjk/ ‘tank’ and /dej/ ‘day’ with special /t[?] d[?]/ as in /t[?]æŋjk/ ‘thank’ and /d[?]ej/ ‘they’. Likewise, true vowel-initial forms as in /art/ ‘art’ contrast with special vowel-initial /[?]art/ ‘heart’. Arguably, such compensatory diacritic representations develop when L2 phonemes require distinctive features that are absent from L1 specifications and hence unavailable or hard to access (Brown, 1998).

We propose that fuzzy representations differ from true phonemic representations in that they do not lead to categorically accurate production or perception; such ad hoc approximate representations only enable variably accurate perception and production. Nonetheless, with effort and practice, learners can improve their ability to use these diacritics to guide perception and production. This would explain the range and distribution of QF accuracy rates for /θ ð/ and /h/. In addition, this scenario of gradual development in accuracy suggests that learners can benefit from pronunciation training that practices using fuzzy representations for perception and production. The pronunciation activity presented next was designed to provide precisely such focused practice.

Pronunciation activity

The following pronunciation activity was developed for an ongoing ESL course with 12 QF students (10_{female}, 2_{male}; age range 59-68, mean 63.3 yrs) offered at the *Université du troisième âge*, a continuing education programme for seniors. The course itself employs a communicative approach, emphasizing oral comprehension and expression in meaning-focused activities. Most of the core activities involve listening, whether to recorded texts on the *Voice of America Learning English* site or to authentic videos (e.g., documentaries on CBC Gem), preceded by warm-up discussion and vocabulary activities. Students also answer comprehension questions as they listen and afterwards engage in group discussions where they share their ideas and personal experiences.

For one class, we designed an activity based on the *Voice of America* text targeting the idiomatic expression 'go out on a limb' (available at: <https://learningenglish.voanews.com/>). To complement this communicative activity, a follow-up pronunciation activity focuses their attention on instances of /θ ð/ in the recording. This focus-on-form is facilitated by virtue of the content having already been processed. Students listen again to an excerpt and follow along with the written text, paying particular attention to the pronunciation of words containing /θ ð/ (e.g., *thing, growth, that, another, with*) and /t d/ (e.g., *trees, take, do*) highlighted in bold. Next, using

the text-to-speech functionality of Google Translate (<https://translate.google.com>), they listen to a series of minimal pairs based on these phoneme contrasts (e.g., *thigh-tie* and *they-day*) focusing on the differences in phonetic realization. Students are invited to reflect on the characteristics of these sounds and to recognize that /t d/, used also in their L1 French, have similar properties to the equally familiar voiceless/voiced stop pairs /p b/ and /k g/. The phonemes /θ ð/ are absent from the French inventory, but are similar to the familiar voiceless/voiced fricative counterparts /f v/ and /s z/. The idea is thus to draw students' attention to the phonetic properties of familiar speech sounds in order to better attune their ears to novel sounds. Next, working with a partner, students copy one member of a /θ/-/t/ or /ð/-/d/ minimal pair into Google Translate for their partner to listen to. The partner's role is to identify which member of the minimal pair they hear. This perceptual awareness-raising activity is thus organized such that learners proceed from being sensitized to new phonemes (comparing and discriminating /θ ð/ vs /t d/) to then introspecting on the phonetic features of sounds and finally identifying whether /θ ð/ or /t d/ are heard.

A follow-up activity intended for the subsequent class involves practicing production of /θ ð/. Briefly, students are given explicit instructions on how to tackle the articulation of /θ ð/ and invited to practice producing these speech sounds in isolation. Then, students repeat the pronunciation of individual words, longer expressions and full sentences copied into Google Translate. Next, students read aloud these same words, expressions and sentences to verify whether the speech-to-text function recognizes their realization. Eventually, in future research, our intention is to test the efficacy of this pronunciation training for perception and production accuracy. We also wish to investigate the extent to which our students take to using Google Translate beyond the classroom, as a readily accessible pronunciation model for autonomous learning. Our prediction is that even older learners such as those at the *Université du troisième âge* will find Google Translate beneficial and easy to use for *ad hoc* feedback.

Conclusion

In our view, L2 learners can benefit from pronunciation training such as the approach we have presented here using the text-to-speech and speech-to-text functionalities of Google Translate. Research findings (e.g., John, 2019) suggest that learners develop fuzzy representations that lead to variable perception and production of L2 phonemes. Conceivably, pronunciation practice can help L2 learners improve their ability to use these approximate representations both to detect and realize novel phonemes. Despite the potential benefits of focusing on phonetic form, however, teachers have tended to avoid pronunciation training. Possibly such a focus-on-form is considered incompatible with communicative language teaching, but also teacher training in pronunciation and the activities provided in textbooks are frequently inadequate. As we have shown, however, pronunciation work is easily incorporated into meaning-focused activities with simple technology such as Google Translate. With this kind of an approach, communicative language teachers can readily add 'accent coach' to their professional identity.

Acknowledgements

We would like to acknowledge our co-developers of the UTA ESL course, Mariane Gazaille, Walcir Cardoso, Angelika Bacher, and Michel Monier. The UTA project was funded via an *Entente Québec-Canada* grant (*enseignement des langues secondes*) awarded to Paul John and Mariane Gazaille.

Références

- Best, C. T. (1994). The emergence of native-language phonological influences in infants: a perceptual assimilation model. In J.C. Goodman & H.C. Nusbaum (Eds.). *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167-224). MIT Press.
- Breitkreutz, J.A., Derwing, T.M., & Rossiter, M.J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal*, 19(1), 51-61.
- Brown, C. (1998). The role of the L1 grammar in the acquisition of L2 segmental structure. *Second Language Research*, 14(2), 136-193.
- Chapelle, C.A., & Jamieson, J. (2008). *Tips for Teaching with CALL: Practical Approaches to Computer-Assisted Language Learning*. Pearson-Longman.
- Darcy, I., Daidone, D., & Kojima, C. (2013) Asymmetric lexical access and fuzzy lexical representations in second language learners. *The Mental Lexicon*, 8(3), 372-420.
- Derwing, T. M., Diepenbroek, L. G., & Foote, J. A. (2012). How well do general-skills ESL textbooks address pronunciation? *TESL Canada Journal*, 30(1), 22-44.
- Foote, J.A., Trofimovich, P., Collins, L., & Soler Urzúa, F. (2013). Pronunciation teaching practices in communicative second language classes. *The Language Learning Journal*, 1-16.
- Isaacs, T. (2009). Integrating form and meaning in pronunciation instruction. *TESL Canada Journal*, 27(1), 1-12.
- John, P. (2006). *Variable h-epenthesis in the interlanguage of francophone ESL learners*. Unpublished Master's thesis. Concordia University.
- John, P. (2019). Pronunciation errors and variation in second language. In E. Rosales, R. Pinsonneault, & J. Payeras (Eds.). Proceedings of the 8th annual Meeting on Language Teaching. UQÀM.
- John, P., & Cardoso, W. (2009). Francophone ESL learners' difficulties with English /h/. In M.A. Watkins, A.S. Rauber, & B.O. Baptista (Eds.), *Recent research in second language phonetics/phonology: Perception and production* (pp. 118-140). Cambridge Scholars Publishing.

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, 1-14.

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). John Benjamins.