

VARIATION IN THE L2 FRENCH AUDIOVISUAL INPUT: YA BASIC!

Rhonda Chung, Concordia University

Walcir Cardoso, Concordia University

The struggles that second language learners experience when navigating the sociophonetic variability of speakers is often explained by the lack of exposure to varied input in the classroom because of its emphasis on teaching (usually invariable) standard varieties. In the realm of French as a second language (FSL) learning, our understanding of variational input in the classroom comes primarily from textbook studies; little empirical evidence has quantified the amount and kind of social speech markers (e.g., age) found in the FSL audiovisual curriculum. This study examines the audiovisual input of two FSL classrooms. Interviews and a questionnaire were used to elicit FSL instructors' criteria for selecting input, and their experiences with and attitude towards including variation in their lessons. Additionally, an analysis of the audiovisual input derived from one semester of each instructor was categorized by clip length and by five social markers: age, gender, race, region, and native speaker status. Results showed that the instructors held positive viewpoints towards including variation; however, the audiovisual input in both settings was invariant across multiple social markers, accounting for less than 5% of total class time. Suggestions for incorporating more varied input in the language learning curriculum will be discussed.

Cite as: Chung, R., & Cardoso, W. (2022). Variation in the L2 French audiovisual input: Ya basic!. 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.13264>

INTRODUCTION

Research recognizes that high variability in second language (L2) audiovisual input predicts robust perceptual learning able to generalize to new talkers and novel linguistic items, facilitating more accurate production (Hardison & Pennington, 2021). This improved perceptual performance can be explained by usage-based theory, which describes learning as a composite of individual learning experiences, including the central tendency of those aggregated events (i.e., the prototype; Ellis & Wulff, 2020), providing learners with multiple vantage points to process and attune to incoming stimuli.

The implications of high variability studies have direct repercussions for the L2 classroom—an environment known to favour standard or prestigious varieties, which are sociophonetically invariable (Fairclough, 2015), and therefore do not predict learning beyond that standard variety (Pisoni et al., 1994). Since high variability is based on two primary factors, namely the amount of input (“how much”) and the kinds of variability (“what kind”), this study focuses on: (a) quantifying and qualifying the audiovisual input used in two adult-level FSL classrooms, a university and government-sponsored *francisation* course, and (b) considers instructors’

motivation for selecting this audiovisual input. To contextualize and motivate the study, a review of core concepts related to L2 speech perception and audiovisual social speech markers will be examined. Our discussion will focus on the cognitive importance for using highly varied input in the language learning classroom, and pedagogical strategies to expose learners to more varied talkers both in and outside the classroom.

L2 Speech Perception

Usage-based theory explains L2 learning as sensitive to frequencies in the input (Ellis & Wulff, 2020). When applied to high variability phonetic training (HVPT), ample positive evidence of varied voices and linguistic features provide learners with more opportunities to sharpen their perception skills and supports their developing phonology of all linguistic varieties within the repertoire (Flege & Bohn, 2021), including dialectal varieties.

Dialects distinguish themselves from other dialectal varieties (including the standard) by their unique lexical, grammatical, phonetic (both segmental and prosodic), pragmatic, and multimodal features (Nycz, 2015). During interaction, individuals aurally perform at least one dialect, signalling their membership to a particular speech community (Bucholtz & Hall, 2005), meaning each talker has multi-dialectal competence in their known languages. As speakers age and move, they cross time continuums and territorial borders, which surface in their speech as diachronic and regional variations (Nycz, 2015). These idiosyncrasies of individual speech are referred to as social speech markers, which are inseparable from the talker's linguistic content (Pisoni et al., 1994).

Multi-Modal Variation in the Input: Social Speech Markers

Our understanding of social speech markers, as Eckert (2012) describes, fall under three waves of variation research. The first explains dialect learning as a phonological process that learner's self-monitor to develop towards a speech style that is more invariable (i.e., standardized), moving away from their native vernaculars. As a result, social (class, gender, age) and regional speech markers were viewed from a macro-social perspective, which enabled researchers to classify and distinguish it from the more prestigious standard. The second wave viewed these markers through a community-based, ethnographic perspective, where speech was understood to actively signal a speaker's membership to specific speech communities. The current wave focuses on how learners negotiate meaning using these social speech markers, expressed through phonological, lexical, and grammatical productions, as semiotic expressions of social identity, focusing on how certain stylistic markers are normalized while others are often stigmatized.

Speech perception, therefore, is a semiotic and multimodal process (i.e., visual, gestural, facial expressions; Hardison & Pennington, 2021). As Pitts and Gallois describe (2019), learning involves attunement to both the linguistic (lexis/grammar, accent) and non-linguistic (vocal tract size, age, race, gender) markers in the speech signal. Audiovisual speech markers are often culture-specific and processing them involves indexing them back to the various speech communities. For example, interacting with an elderly speaker entails visually processing markers related to physiology (grey hair), aurally processing suprasegmental productions related to aging (articulation patterns, vocal jitters), and linguistically processing unique speech styles. Speech markers related to gender report that female-gendered speech is often produced with a higher pitch and includes more polite forms and tag questions than male-gendered speech. Visual markers

related to race have reportedly been so strong that they override acoustic and linguistic markers whereby listeners hallucinate accents for racialized speakers from within their own speech community (Kutly, 2020).

Audiovisual Variation in the FSL Input

The reliance on invariable standards in the FSL classroom, particularly from Paris (Baker & Smith, 2010), can be understood as continuing first and second waves views of variation, which focus on the role of standardization, rather than how language is used in diverse speech communities. This pedagogical popularity of standards is explained by the fact that this is the variety most learners are tested on (Genesee, 1987). Yet FSL research demonstrates that learners who can navigate and negotiate variation in speaker style (e.g., subject doubling; *ne*-deletion) and pronunciation (i.e., accent) are viewed as socio-culturally advantaged (Oakes & Peled, 2018). The continued difficulties that learners continue to have in comprehending spoken French outside the classroom, raises questions as to “how much” (i.e., frequency) and “what kinds” (i.e., variability) of audiovisual input learners are exposed to within FSL classrooms.

Our understanding of social and regional dialectal variation in the FSL classroom, however, has come primarily from textbook studies (Duchemin, 2017), which typically do not provide audiovisual input (e.g., videos), meaning that supplementing the classroom with audiovisual material falls under the purvey of instructors. Curriculum design, and the materials that constitute it, is influenced by teachers’ beliefs and language experiences (Borg, 2018). To better understand how audiovisual input is selected by instructors, and how much and what kinds of social markers are found in the audiovisual input of the classroom, a contrastive case study of two adult-level FSL classrooms (see rationale below) was conducted to answer the following research questions:

RQ1 What kind of audiovisual input do FSL instructors select for their classrooms?

RQ2 How much and what kinds of social markers (age, gender, race, native speaker status, regional dialect) comprise the audiovisual input of these two FSL settings?

Context of study

The two institutional settings were located in Montreal, a multilingual urban city in Quebec, Canada. The university is a well-documented language setting, but the government-funded *francisation* class is comparatively understudied. To assure consistency across settings, classrooms were targeted that catered to the same population (adult speakers) of the same proficiency level (intermediate); however, access to these linguistic resources differ quite substantially.

University

Modeled after the *Grandes écoles* in France, the university's stated mission was to cultivate language skills transferable to professional environments. Tuition for FSL classes in Quebec universities are protected by a provincial tuition exemption, keeping fees low and accessible to both international and Canadian students. The intermediate-level FSL class had targeted oral and written production goals, including mid- and final-term exams required to pass the course. Classes were held weekly, running three-hours in duration for 12 consecutive weeks. The instructor was a Quebecois woman, in her mid-30s, with ten years of FSL teaching experience in multiple contexts with learners of different proficiency levels.

Francisation

The *francisation* program was created in 1968, during Quebec's Quiet Revolution, and was specifically designed for social integration into francophone society; its courses focus on developing reception and production skills (MEES, 2015). Again, classes are kept to a nominal fee and open to all citizenship statuses (citizens, international students, visa workers). Classes were held twice a week for seven weeks, running two hours in duration for seven consecutive weeks and emphasized oral production; advancement to the next level is assessed through Ministry-mandated exams. The instructor was Quebecois, in her mid-30s, with ten years of FSL teaching experience in multiple contexts with learners of different proficiency levels.

METHODS

Teacher Interviews Questionnaire

To answer the first research question, twenty semi-structured questions were developed to gauge teachers' beliefs and experiences surrounding audiovisual input use, in general, and social markers, in particular, to better understand the motivations behind their selections. Interviews were conducted remotely and responses were video recorded using an online software. The data were transcribed and organized according to the pre-established themes of the questionnaire. Responses to open-ended questions were analyzed through an iterative process of descriptive coding to inductively identify themes (Saldaña, 2021).

Audiovisual Input

To answer the second research question, instructors completed a log providing all audiovisual material used in one semester of their course. The university used a total of seven online audiovisual inputs: four news reports (one international, two national, and one provincial); and three interviews. The *francisation* course used 25 items from a variety of sources: one episode of a dramatic television series, one short film (DVD), an episode of a romantic-comedy (DVD), five music videos, five sketch comedies, and 13 dramatic short sketches.

Following the conventions of L2 speech perception studies (Derwing et al., 2004), speech input that lasted longer than 20 seconds constituted the basis of unique talker speech, meaning that any input that was under this threshold was eliminated from analysis. Since the input was sourced entirely from mass media, sociodemographic data was readily available from multiple websites to confirm the backgrounds of actors, singers, or interviewed professionals (e.g., IMDB, LinkedIn).

RESULTS

RQ1: FSL Instructors' Use of Audiovisual Input

The first research question asked instructors to describe their decision-making process and criteria for using audiovisual input in their classrooms. Two main themes emerged related to the physical parameters of an ideal audio clip (i.e., audio quality, length), and to the appropriateness of the input, which in turn, had three sub-levels related to appropriateness of: the content or subject matter; learner proficiency levels; and the social markers under investigation (see Table 1 for a summary of the findings).

In terms of parameters, both instructors stated that the primary factor for excluding input would be a noisy background or "unclear" speech. Unclear speech, for the university instructor, meant

accents from rural regions within Quebec or in Eastern Canada--varieties less likely to be encountered in the urban Montreal setting; whereas, for the *francisation* instructor, it referred to the acoustic quality of the recording. The university instructor had specific guidelines for length, genre, and where to source materials. The *francisation* instructor had no such preferences for length, but emphasized that different genres (e.g., comedy, drama) was an important feature of the curriculum.

Both instructors held similar views related to using input that was appropriate for the curriculum and for proficiency levels, but varied regarding their reasons for the inclusion of certain social markers in the audiovisual input. With regards to age, only the university instructor specified that speech derived from children would not be relevant for the adult-learning context. She also noted that generational accents, like the rolling of /r/ by older talkers of the Quebec variety, should be pointed out to learners. Although neither instructor reported preferences for markers related to gender, the university instructor did state that some learners claimed to have difficulty understanding male voices over female ones. Both instructors stated that native and proficient L2 speakers be represented in the audiovisual input; as the latter group, the university instructor explained, encourages FSL learners to notice that fluency is possible for French learners. With regards to regional dialect, again both instructors reported positive feelings towards the inclusion of regional dialects but stated a clear preference for using the Quebecois variety. The *francisation* instructor explained that European varieties were less pertinent to language learning in North America. This sentiment was echoed by the university instructor who also avoided European varieties but noted that FSL students had more experience with varieties from France, and therefore greater attention needed to be placed on the Quebecois variety.

Table 1*Themes from Instructor Interviews*

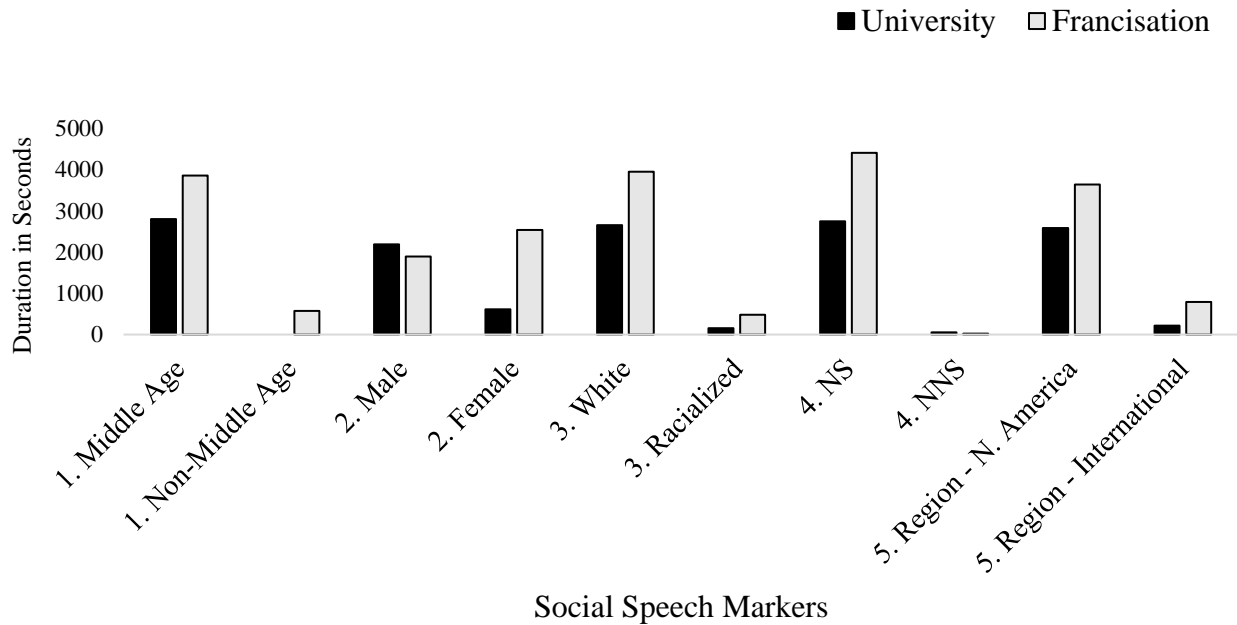
Themes	Institutional Setting	
	University	Francisation
Physical parameters	<ul style="list-style-type: none"> • Under 10 minutes • Clear and audible speech (standard accent) with no background noise • Preference for news sources 	<ul style="list-style-type: none"> • Range from one minute to over an hour • Preference for digital recordings over discs with clear acoustics (no static or background noise) • All media and genre types; preference for arts and entertainment
Appropriateness: (a) Curriculum content	<ul style="list-style-type: none"> • Be related to the theme of the lesson • Evoke learner interest, be socially relevant, include current events • Should be used in each class 	
(b) Learner proficiency level	<ul style="list-style-type: none"> • Beginners should be exposed to controlled input to ease processing (e.g., DVDs with controlled dialogue) • Intermediate and advanced learners should be exposed to natural speech from authentic sources to habituate listening 	
(c) Social markers: age, gender, race, native speaker status, regional dialect	<ul style="list-style-type: none"> • Age: children’s voices inappropriate • Gender: No preference • Race: No preference • Native speaker: Preference for native and advanced non-native speakers • Regional dialect: Other dialects important, but standard Quebec variety the priority 	<ul style="list-style-type: none"> • Age: No preference • Gender: No preference • Race: No preference • Native speaker: Preference for native and advanced non-native speakers • Regional dialect: Other dialects important, but Quebec variety the priority

Research Question 2: How Much and What Kinds of Audiovisual Input

The second research question centered on the amount and kind of social markers that comprised the audiovisual input for each setting in terms of age, gender, race, native speaker status, and regional dialect (see breakdown of markers in Figure 1). During the categorization process for age, a variable of “middle-aged” emerged, representing talkers found to be between the ages of 20 to 50 years old.

Figure 1

Social Speech Markers in Audiovisual Input



University. Of the seven audiovisual inputs used, 16 unique talkers were identified from the input, who spoke for a total of 46 minutes and 49 seconds, accounting for 2.17% of class time in the course’s 36-hour length, for an average talker speech length of 2 minutes and 56 seconds. Focusing more closely on the five markers, all sixteen talkers were middle-aged, indicating no speech from children, adolescent, or elderly talkers. In terms of gender, nine male talkers comprised 78.14% of the total audiovisual input. For race, 13 talkers presented as white, accounting for 94.62% of the input. Fifteen native French speakers accounted for 98.04% of the input. Finally, for regional dialect, 11 talkers from the Quebec region accounted for 92.20% of the input, and included five talkers from dialects in Europe, the Middle East, and North Africa.

Francisation. Of the 25 audiovisual inputs used, 66 unique talkers were identified from the audiovisual input, speaking for a total of 74 minutes, accounting for 4.40% of total class time in the course’s 28-hour duration, for an average talker speech length of one minute and seven seconds. Fifty-three middle-aged talkers accounted for 86.98% of the total audiovisual input, with adolescent (2.73%) and elderly (10.29%) comprising the rest; there was no input from children. Although the 35 male talkers outnumbered the 31 female talkers, females spoke for longer, comprising 57.32% of the input. Fifty-one talkers presented as white, comprising 89.12% of the input. Sixty-five native French speakers accounted for 99.55% of the input. The dialect from the region of Quebec represented 82.13% of the input spoken by 44 talkers, and the remaining 22 talkers came from regions in Europe (Belgium, Switzerland), Haiti, Lebanon, Morocco, Polynesia, the United States (Louisiana, New England, Maine), and West Africa (Burkino Faso and Senegal).

DISCUSSION

In this study, whether institutional setting was intended to cultivate professional skills (university) or to promote social integration (*francisation*) among their adult, intermediate-level student body, FSL instructors held positive views towards including social speech marker variation in their audiovisual curriculum. Opinions diverged, however, on: where this input should be sourced from, its genre, and the overall length of each audiovisual clip. Based on the findings, input from the *francisation* course contained four-times more individual talkers than the university course, but the duration of individual talker speech was nearly three times shorter. It is an empirical question as to whether less diverse input with a longer speech signal leads to better perceptual skills compared to more diverse input with a substantially shorter signal.

Despite instructors reporting positive feelings towards using audiovisual content with a wide range of social speech markers (age, gender, race, regional variety, native speaker status), four of the five targeted markers were invariable. This disconnect between the teachers' positive feelings towards variation (i.e., beliefs) and the actual variation found in their audiovisual input is a known issue in L2 teacher training research, and can be bridged if instructors engage in self-reflective activities to ensure that their teaching outcomes mirror their learning objectives (Borg, 2018). Because the instructors in this study had similar demographic backgrounds, future research should be conducted with teachers from varying demographics to provide a more nuanced picture of how FSL instructors design their audiovisual curriculum.

The invariable measures related to age, race, native speaker status, and regional dialect are in large part due to the input being drawn from mass media sources, which function to diffuse sociopolitical and raciolinguistic ideologies of the private industries who fund and produce them, often instructing actors to minimize their non-standard dialects as a condition of employment (Fairclough, 2015; Flores & Rosa, 2015). The presence of non-mass media, as found in the *francisation* class, partially supports this claim. Sourced from a publicly-funded government institution (the National Film Board of Canada), this audiovisual input contributed to injecting more variation for social speech markers related to race and regional dialect, featuring Francophones from regions in the Antilles, Africa, Asia, and Polynesia.

The monodialectalism of Quebecois French in both settings reflects an intentional strategy employed by instructors who both shared the belief that this variety: a) best assisted learners in acoustically acclimating to their environment, despite Montreal being one of the most multilingual cities in Canada (Galante & dela Cruz, 2021); and b) mitigated their exposure to the monodialectalism of European dialects found in previous classroom experiences. This stated preference for and use of the Quebec variety in the classroom runs contrary to multiple claims in the FSL literature that Parisian French was the preferred and de facto standard of the French-learning classroom (Baker & Smith, 2010). This dominance of the Quebecois variety in Canada can be understood the fruit of long-standing provincial policies, one of which includes keeping French courses affordable to newly arrived citizens and university students, rooted in the province's nation-building goals to protect itself from linguistic encroachment by an overwhelmingly Anglophone continent. These protective measures keep French courses affordable to newly arrived peoples and serve to codify and standardize the language by specifically eliminating, for example, regional variations (Oakes & Peled, 2018).

Based on these results, FSL instructors should specifically target audiovisual material containing social speech markers found to be under-represented in mass media input, namely: the elderly,

children, adolescents, racialized individuals, non-Quebec French varieties, and non-native speakers. Input sourced from unscripted (i.e., natural speech found outside the classroom), and non-mass media (e.g., social media, independent media) should be privileged, as this is more likely to contain diverse social speech markers that reflect the kind of speech learners are likely to encounter in their linguistic environment. Noticing activities, sourced from critical ethnographic methodologies (e.g., transcribing overheard conversations; Alim, 2010), can be assigned for outside classroom use. These listening activities all aim to raise learners' multi-modal awareness of the inherent variation of spoken French and precipitate robust perceptual learning for processing speaker variation.

Pedagogical interventions, which focus on variability enhance FSL learners' ability to decode natural speech and to recognize the social significance of variability, including the possibility of producing these same variable forms for their own social benefit (French & Beaulieu, 2020). This strategic use of variation as a pedagogical norm transforms FSL instruction into one where variation is centered, not tolerated, as a unit of study, compelling learners to expect variable input, and to develop perceptual strategies poised to decode and attune.

In sum, a lack of exposure to variable language use is increasingly positioned as socially burdensome for FSL learners (Oakes & Peled, 2018), particularly those living in multilingual urban environments, because it fails to equip learners with social adaptability skills (Schaefer & Warhol, 2020). The invariability across social speech markers in mass media does not provide enough diverse input that predicts robust perceptual learning. Therefore, when instructors use mass media—privileging the standard by default—they diffuse invariable input, which does not afford opportunities to sharpen learners' attunement skills to the variations found in speech. Ensuring there are multiple speech varieties in the audiovisual curriculum, therefore, constitutes a pedagogy that supports learners in developing perceptual attunement skills, and in expressing their identities so as to form bonds with francophone communities.

ABOUT THE AUTHORS

Rhonda Chung (rhonda.chung@mail.concordia.ca) is a doctoral candidate in Education (Applied Linguistics) at Concordia University whose research interests include ecological theories of perception, dialectal variation and plurilingual, inclusive pedagogies, and decolonial curricula.

Walcir Cardoso (walcir.cardoso@concordia.ca) is a professor at Concordia university. He conducts research on the L2 acquisition of phonology, morphosyntax and vocabulary, and the effects of computer technology on L2 learning.

REFERENCES

- Alim, H. (2010). Critical language awareness. In N. Hornberger & S. McKay (Eds.), *Sociolinguistics and language education* (pp. 205–231). Bristol: Multilingual Matters.
- Baker, W., & Smith, L. (2010). The impact of L2 dialect on learning French vowels: Native English speakers learning Québécois and European French. *The Canadian Modern Language Review*, 66(5), 711-738.

- Borg, S. (2018). Teachers' beliefs and classroom practices. In P. Garrett, J. Cots (Eds.) *The Routledge handbook of language awareness* (pp. 75–91). Routledge.
- Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, 7(4-5), 585-614.
- Derwing, T., Rossiter, M., Munro, M., & Thomson, R. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54, 655-679.
- Duchemin, M. (2017). Les représentations associées aux français nationaux, aux espaces géographiques et aux locuteurs dans les manuels de français langue étrangère et de français langue seconde: étude comparée entre la France et le Québec. *Canadian Journal of Applied Linguistics*, 20(2), 51–70.
- Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual review of Anthropology*, 41, 87-100.
- Ellis, N. & Wulff, S. (2020). Usage-Based Approaches to L2 Acquisition. In B. VanPatten, G. Keating, and S. Wulff (Eds.) *Theories in second language acquisition: An introduction* (3rd ed. pp. 63-82). Routledge.
- Fairclough, N. (2015). *Language and power* (3rd ed.). Abingdon, Oxon: Routledge, Taylor & Francis Group.
- Flege, J., & Bohn, O. (2021). The revised Speech Learning Model (SLM-r). In R. Wayland (Ed.), *Second Language Speech Learning: Theoretical and Empirical Progress* (pp. 3-83). Cambridge University Press. doi:10.1017/9781108886901.002
- Flores, N., & Rosa, J. (2015). Undoing appropriateness: Raciolinguistic ideologies and language diversity in education. *Harvard Educational Review*, 85(2), 149–171.
- French, L., & Beaulieu, S. (2020). Can beginner L2 learners handle explicit instruction about language variation? A proof-of-concept study of French negation. *Language Awareness*, 29(3-4), 272-285.
- Galante, A., & dela Cruz, J. (2021). Plurilingual and pluricultural as the new normal: an examination of language use and identity in the multilingual city of Montreal. *Journal of Multilingual and Multicultural Development*, 1-16.
- Genesee, F. (1987). *Learning through two languages; studies of immersion and bilingual education*. Newbury House.
- Hardison, D. & Pennington, M. (2021). Multimodal second-language communication: Research findings and pedagogical implications. *RELC Journal*, 52(1), 62-76.
- Kutlu, E. (2020). Now you see me, now you mishear me: Raciolinguistic accounts of speech perception in different English varieties. *Journal of Multilingual and Multicultural Development*, 1-15. DOI: 10.1080/01434632.2020.1835929

- MEES. (2015). Programme d'études, francisation: Domaine des langues - formation générale des adultes. Retrieved from http://www.education.gouv.qc.ca/fileadmin/site_web/documents/dpse/educ_adulte_action_comm/Adultes_FGA_progetudes_domainelangue_francisation_2015.pdf
- Nycz, J. (2015). Second dialect acquisition: A sociophonetic perspective. *Language and Linguistics Compass*, 9(11), 469-482.
- Oakes, L. & Peled, Y. (2018). *Normative language policy: Ethics, politics, principles*. Cambridge: Cambridge University Press.
- Pisoni, D., Lively, S. & Logan, J. (1994). Perceptual learning of nonnative speech contrasts: Implications for theories of speech perception. In J. Goodman & H. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (p. 121–166). The MIT Press.
- Pitts, M., & Gallois, C. (2019). Social markers in language and speech. *Oxford Research Encyclopedia of Psychology*. <https://doi.org/10.1093/acrefore/9780190236557.013.300>
- Saldaña, J. (2021). *The Coding Manual for Qualitative Researchers* (4th ed). Sage.
- Schaefer, V., & Warhol, T. (2020). There ain't no doubt about it: Teaching EAL learners to recognize variation and switch/shift between varieties and registers is crucial to communicative competence. *TESOL Journal*, 11(3).