Rewilding second language learning: Non-formal learning through songs

Ross Sundberg

A Thesis

In

the Department

of

Education

Presented in Partial Fulfillment of the Requirements

For the Degree of

Doctor of Philosophy (Education) at

Concordia University

Montréal, Québec, Canada

July 2024

CONCORDIA UNIVERSITY

SCHOOL OF GRADUATE STUDIES

This is to certify that the thesis prepared

By: Ross Sundberg

Entitled: Rewilding second language learning: Non-formal learning through songs

and submitted in partial fulfillment of the requirements for the degree of

Doctor Of Philosophy

complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the final Examining Committee:

_____ Chair
Dr. Denis Liakin

_____ External Examiner
Dr. Jérémie Séror

_____ Arm's Length Examiner
Dr. Fanny Macé

_____ Examiner
Dr. Saul Carliner

_____ Examiner
Dr. Pavel Trofimovich

_____ Thesis Supervisor
Dr. Walcir Cardoso

Approved by _____
Dr. TBC, Graduate Program Director

August 23, 2024 _____
Date of Defense Dr. Pascale Sicotte, Dean of Arts and Science

Abstract

Rewilding second language learning: Non-formal learning through songs

Ross Sundberg, Ph.D.

Concordia University, 2024

The goal of this manuscript-based dissertation is to explore French second language development through songs in the non-formal learning space to understand learners' progression towards successful integration into informal environments. To explore the self-regulatory process required to increase autonomy and the degree to which music can aid learners in this pursuit, I examine new technologies and learners' behaviours towards the use of music in second language contexts. This is done by following Cardoso's (2022) chronological framework for examining new technologies in second language contexts because when learners engage with L2 music, they do so through technology. After providing an introduction and background on autonomy and the non-formal learning space (Chapter 1), the framework is explored through three manuscripts to investigate different aspects of language learning through songs and the use of the *Bande à Part* app (which was used as a tool to understand this approach).

The first manuscript (Manuscript A) investigates theoretical affordances provided by music as a form of L2 phonological exposure (e.g., acoustic comparisons between singing vs speaking) and form of content (e.g., type of vocabulary, rhyme schemes, and repetition). This manuscript highlights affordances provided by overlapping cognitive processes shared between language and music processing. For example, music encourages increased focus on form, which could help with pronunciation and exposure to drawn out vowels, leading to better formation of vowel categories. In terms of content, music is repetitive and motivating—two factors that are

difficult to incorporate into the classroom. Empirical work on the impact of music on language development is interpreted with this in mind, supporting the theoretical affordances that were highlighted. Gaps are highlighted for future studies with the subsequent two chapters shedding light on some of these gaps.

The second manuscript (Manuscript B) assesses learners' perceptions of suitability and acceptability of learning French with the *Bande à Part* music app (development and release documented in Sundberg & Cardoso, 2019). Results demonstrate that learners believe the app helps improve their French and that it is relatively intuitive to use, although more instructions and more songs should be included. These perceptions are gathered because they are predictors of whether or not learners will use a tool, critical to self-directed contexts. The feedback gathered is also used to make iterative improvements to the app. The manuscript discusses principles and takeaways for future material developers who are interested in student-centred tools.

The third manuscript (Manuscript C) investigates learners' awareness of language, the language learning process, and ability to self-regulate their own pronunciation development. This is accomplished through think-aloud exercises where participants listen to and then comment on recordings of themselves imitating sung and spoken audio recordings. This research helps to illuminate where scaffolding is needed and the degree to which learners can learn on their own.

The main findings emphasize the value of self-directed pedagogical tools, particularly songs, in enhancing L2 development. Educators and developers can use these insights to create more effective materials for language learners and support beyond the classroom learning endeavours.

Acknowledgements

I first want to thank my supervisor Dr. Walcir Cardoso. Few are fortunate to have a relationship like the one we have developed. Dr. Cardoso is a true mentor. He has met me where I am at every step of the way and his support has extended well beyond the classroom. We also share a number of common interests, including music! As I move on in my career, I feel inspired to adopt the relational pedagogical and research outlook he has upheld. When common ground is established, critical feedback can flourish and inspire growth. Dr. Cardoso has been able to simultaneously support my interests while pushing me to do my best work. The zone of proximal development, which I discuss in my research, is a space where you achieve more than you can do on your own through the help of others and Dr. Cardoso is a master at scaffolding this space for his students. I cannot say enough about how important it is to undertake challenges, such as a doctorate degree, under the right supervision.

I also want to thank my Dr. Pavel Trofimovich and Dr. Saul Carliner for taking part in my committee and doctoral journey. The time I spent doing reading courses with each of them was especially important to the development of my ideas and methods. They have helped me refine my work while also encouraging me along this journey. I also thank Dr. Jérémie Séror and Dr. Fanny Macé for their part in my examination committee and Dr. Denis Liakin for chairing the defense. I am grateful for the time each of you dedicated to reading and engaging with my research.

I want to thank my colleagues for journeying alongside me. In particular, Dr. Mike Barcomb and Dr. Clint Hendry played crucial roles reading my work and providing feedback on my ideas. Mike, you were there for me on short notice for questions and bouncing ideas around. You could relate to the process and struggles I experienced in ways that I am unsure I could have

handled on my own. We were in the trenches together. Your expertise also pushed me to ensure my work was the best it could be. Clint, from travelling to conferences and sharing meals during work breaks to discussing strategies around dissertation writing while caring for an infant at home, your support and inspiration has been essential. I'll forever be grateful for the support from each of you.

I would like to thank my parents, Becky and Perry Sundberg, for the years of support they have poured into my life. I am fortunate to have parents that crucially supported the development of a mindset that enabled me to take on this challenge. It is because of them that I feel adaptable and persevere when life gets hard. It is because of them that I know I can grow.

Finally, this degree would have been impossible without my wife, Jacquelyn, and my son, Kasper. Jacquelyn cared for Kasper as I spent weekends writing. She engaged with my ideas and trialled research instruments. She supported my mental and physical health when I struggled. She understands what drives me and gives me space to be myself, a freedom that allows for so much growth. She inspires new ways of seeing the world that expand my life. I love you, Jacquelyn, and thank you for everything. My son, Kasper, has taught (and is still teaching) me to slow down. He knows what is important in life and lives in the moment. He constantly reminds me that there is much more to life than intellectual pursuit. He restores balance and provides so much purpose and meaning. I will forever enjoy watching you develop and engage with the world.

Jacquelyn is due with our second son any day now. I look forward to meeting you, son, and your imminent arrival has given me a personal deadline. I am wrapping up this dissertation and will have more time to spend with you.

Contribution of Authors

This manuscript-based dissertation consists of five chapters. The first chapter is an introduction that sets up the framework and rationale for the three main manuscripts, which make up chapters two through four. Chapter five concludes the research program. Manuscript B was presented at an international conference (EUROCALL) in 2019 and was published *in The Computer Assisted Language Instruction Consortium* (CALICO) in 2022. The remaining two manuscripts are being prepared for journal submission.

CRediT author statement (Manuscript A) – Ross Sundberg: conceptualization; writing – original draft; writing – review & editing; funding acquisition. Walcir Cardoso: conceptualization (supporting); writing – review & editing (supporting); supervision; funding acquisition (supporting).

CRediT author statement (Manuscript B) – Ross Sundberg: conceptualization; methodology; formal analysis; investigation; resources; data curation; writing – original draft; writing – review & editing; visualization; project administration; funding acquisition. Walcir Cardoso: conceptualization (supporting); methodology (supporting); formal analysis (supporting); investigation (supporting); writing – review & editing (supporting); visualization (supporting); supervision; funding acquisition (supporting).

CRediT author statement (Manuscript C) – Ross Sundberg: conceptualization; methodology; software; formal analysis; investigation; resources; data curation; writing – original draft; writing – review & editing; visualization; project administration; funding acquisition. Walcir Cardoso: conceptualization (supporting); methodology (supporting); formal analysis (supporting); investigation (supporting); writing – review & editing (supporting); visualization (supporting); supervision; funding acquisition (supporting).

Table of Contents

List of Figures

List of Tables

Chapter 1: General Introduction

Rewilding Language Education in Non-Formal Environments

When examining out-of-school factors that affect target language learning, Lindgren and Muñoz (2013) found that young Europeans interact with their target language through songs more than any other singular activity (e.g., movies, social media). Case studies of self-directed learners, by Chik and Ho (2017) and Murray (2004), for example, also demonstrate the desire that learners have to interact and learn through media and pop culture in their target language, with music playing a central role. This natural (informal) learning environment, where language learners intentionally or incidentally develop language skills through activities such as listening to songs sung in their target language, has been referred to as "learning in the wild" by some researchers because learners are initiating activities in their target language outside of an institutional setting (Little & Thorne, 2017; Sundqvist, 2019). The wild is where learners ultimately want to use and practice their target language; one way that students can regulate their learning process naturally in this environment is through music. *Rewilding*, on the other hand, emphasizes "designing supportive conditions for goal-directed interaction in spaces outside of classrooms" (Thorne et al., 2021).

Learner-initiated activities in settings beyond the classroom, however, require a high capacity to self-regulate (Darasawang & Reinders, 2010; Murray, 2004; Reinders, 2020) and variable, often unpredictable levels of target language proficiency. Further, not all target language exposure is as effective as it could be. For example, movies with subtitles may have a larger effect on foreign language listening and reading development than music, despite learners having more exposure to music on average (Lindgren & Muñoz, 2013). However, this begs further investigation into the benefits from listening to music, especially with regard to

pronunciation (including prosody), and vocabulary development—skills that were not analysed in the Lindgren and Muñoz (2013) study. Additionally, rewilding pedagogies can help learners adapt to learning from music in informal settings by progressively restructuring the experience, turning it into a deliberate language learning interaction. In essence, rewilding is the process of helping learners progress from formal learning environments to informal ones by either removing structure from, or introducing "structured unpredictability" into the learners' environment (Little & Thorne, 2017). This can be done by introducing "intentionally underspecified tasks" that require student-initiated exploration (Thorne et al., 2021) within the learner's zone of proximal development (ZPD; Little & Thorne, 2017)—ZPD refers to the learning space beyond what learners can do on their own but not beyond what they can do with assistance (Ortega, 2009).

Although rewilding pedagogies can be present in formal learning environments, the locus of study for this dissertation is in a space that I consider to be between formal and informal learning environments, known as "non-formal." Stickler and Emke (2011) characterize the non-formal learning space as being learner-initiated and not involving credentialing (like informal environments) but with materials that have been designed for learning (like formal environments). One benefit of the non-formal learning space is that it can increase learners' control over initiating learning activities, compared to formal environments, because of the lack of generic assessment required. In credentialed programs, learners are required to conform to specific competencies for the purposes of gatekeeping. In turn, non-formal learning through rewilding pedagogies has the potential to foster autonomy and, thus, liberate learners to take control of their learning in the environments of their choosing, while also maintaining some structure, or "performance support" (see Carliner, 2013), designed to target a learner's ZPD. A

simple and traditional example of this would be students who voluntarily choose to work through a series of graded readers to progress their reading ability. By analysing, categorizing, enhancing, and presenting songs in a likewise structured and non-formal way for learners, it is possible that they could increase the efficacy and enjoyment of language learning through songs, which is an activity so many are already choosing to take part in on their own. To provide this type of performance support, a music app, called *Bande à Part* (*BàP*), was created prior to this investigation to help learners of French learn non-formally. This app is referenced and used throughout the dissertation to varying degrees. The development of this app is discussed in Sundberg and Cardoso (2019).

The goal of this manuscript-based dissertation is to explore second or foreign language (L2) development, in this case French, through songs in the non-formal learning space in order to better understand the progression towards successful interaction in more informal and, thus, *wild* environments. It responds to Thorne et al.'s (2021) renewed call to pursue ecological alignment in education. To do so, three manuscripts are included to begin this investigation.

The first manuscript details a conceptual framework for understanding the affordances of listening to music in one's target language. This is primarily from a cognitive perspective and is a cross-disciplinary exercise aimed at formally introducing L2 researchers to the literature on the relationship between language and music more generally.

The second manuscript explores learners' perceptions of *BàP* because these perceptions influence the uptake and continued use of new tools and materials—an essential aspect of initiating one's learning in the non-formal space.

The third manuscript explores what learners can learn on their own when imitating French lyrics either sung or spoken. The think-aloud exercises target the reflective processes that

learners go through, helping us to understand the differences in learners' self-regulatory potential when using music and spoken resources when learning on their own. The purpose is for instructional designers and teachers to better understand where structure can be introduced to support learners.

Along with introducing each of the manuscripts, the aim of this chapter is to provide a background to the concepts and theories that motivate the dissertation as a cohesive project. I open with an overview of sociocultural theory, which underpins the research on rewilding and self-regulation. I then provide background on learner autonomy and non-formal environments, proposing the non-formal space as an ideal space to develop one's autonomy through learner-initiated activities with semi-structured materials to scaffold learners towards environments with natural language interactions.

Theoretical Framework

This project is interpreted through a sociocultural theory (SCT) of learning (Vygotsky, 1978). As learners interact with their environment, they use symbols (language) as tools to mediate between reality and oneself (Ortega, 2009). Overtime, one moves from object- to other- to self-regulation, which means that one develops increasing liberty and capacity to act autonomously. In the context of this dissertation, that means there is a goal of progressing towards interacting in more informal environments. By exploring the relationships between the individual and the individual's learning context (Ortega, 2009), we can better understand their ZPD, where a learning potential emerges as a result of activities that have been undertaken (Wells, 1999). As one gains confidence within the ZPD, their self-regulatory capacity increases, and one can act with more autonomy.

Perspectives on Learner Autonomy

In their book, *Autonomy in Language* Education, Raya and Vieira (2020) give tribute to Henri Holec, Yves Châlon, and Paulo Freire for the work they have done on liberating pedagogies that empower learners (p. 2). In Raya and Vieira's opening remarks, they contend that Holec's (1981) definition of learner autonomy, "the ability to take charge of one's own learning" (p. 3), is still relevant today. According to Holec, this definition of autonomy requires that learners are responsible for:

1) Determining the objectives [of learning];
2) Defining the contents and progressions;
3) Selecting methods and techniques to be used;
4) Monitoring the procedure of acquisition (rhythm, time, place, etc.);
5) Evaluat[ing] what has been acquired

Despite Holec's influence today (e.g., Benson, 2007, 2013; Murray, 2020), Raya and Vieira (2020) highlight that empowering learners to act autonomously is not yet mainstream pedagogy, and that autonomy, as a concept, is absent from the frameworks of large institutions, such as The Common European Framework of Reference for Languages (p. 4). This is not necessarily because of a lack of research on the topic, but more likely because it is challenging to hierarchical systems. Empowerment undermines conformity and conformity is important to assessment-oriented practices (Châlon, 1970; Holec, 1981; Raya & Vieira, 2020). Relatedly, research into language teacher education programs has also found almost no content related to developing learner autonomy (Reinders & Balcikanli, 2011). As such, the pursuit of researching and implementing empowering pedagogies that are self-directed is still underway.

Keeping in mind the emancipatory potential of encouraging learner autonomy, the end of the 20th century and start of the 21st century has seen a critique of autonomy as an enlightenment ideal (e.g., Benson, 2007; van Lier, 1996; Mackenzie & Stoljar, 2000). These postmodern critiques have acknowledged that independence and sovereignty over one's life has consequences that impinge on others, and we need to recognize some form of interdependence. If autonomy allows individuals to self-determine, thus liberating us from dependence on existing hierarchies and systems, then relational autonomy (Benson, 2007), as one might call the result of postmodern critiques, adds a relational component through a more macroscopic attention shift to our interdependent nature in this biosphere. My actions affect your freedoms, and vice versa. In fact, although Holec's vision for autonomy has been positioned as an extension of the body of works produced by both Châlon and Freire, these prior contributions also integrate well with these postmodern critiques of autonomous learner ideals since Châlon (1970) wrote about negotiated pedagogies and Freire worked on power dynamics—both implying cooperative engagement. Ultimately, and I think Holec might agree too, his vision after all was one where learners had more influence; consequently, when learners self-organize, complex eco-social systems emerge (Murray, 2020). In my research, I view autonomy through this relational lens that considers our social environments.

Learning Beyond the Classroom

In a separate vein of research, critiques of industrial models of education and the rise of lifelong learning have led to research on informal learning. Informal environments have several names, including learning in the wild, extramural English, or even digital wilds when highlighting the common digital context. Definitions of these terms are not often consistent across the literature, but definitions will be provided here in ways that best delimit the categories

that pertain to this research. Generally, these umbrella terms cover all learning, intended or not, that happens outside of formal educational contexts (like a classroom) and through activities initiated by the learner (Carliner, 2013; Sauro & Zourou, 2019; Sundqvist & Sylvén, 2016). However, while learning in the wild has been explicitly defined to only include environments that do not exist for the primary goal of language learning (Sauro & Zourou, 2019), the examples of extramural English education provided by Sundqvist and Sylvén (2016), such as playing online games, watching TV, listening to music, etc., only implicitly indicate that they do not include working through materials that are expressly designed for language learning. When engaging with past research, it is important to consider characteristics of the learning environment like whether or not the learning is deliberate or not since it is not always indicated. Future research should be intentional about clarifying these shortcomings.

In the interest of integrating with the wider literature, it is worth looking beyond language learning contexts. Drawing on the work of Eraut (2000), Carliner (2013) similarly describes formal, non-formal, and informal learning environments and processes, but elucidates five important "levers" that can be determined to some extent by learners when involved in task initiation. According to the author, learners may have more or less control over (adapted from Carliner, 2013, p. 5, with operative questions in italics): 1) process (*who controls and assess the learning process?)*; 2) location (*is the place where learning occurred intended for learning?*); 3) purpose (*is learning a primary or secondary goal of the activity?*); 4) content (*is the content abstract, such as concepts; or practical, related to an everyday skill?*); and 5) consciousness (*when it happens are people aware that learning occurred?*). Manipulating these levers allows for "intentionally underspecifying tasks" during the transition from formal to informal environments while still providing some guidance, as discussed earlier (Thorne et al., 2021).

Applying this framework to the current context, the expectation is that learners progress from what Eraut (2000) describes as the initial pattern recognition of their target skill to more complex metacognitive understandings, as a result of managing everything involved with their own learning strategies. For example, a learner wants to learn French pronouns but realizes they also need to plan how to do so.

If the informal learning environment has a more naturalistic bent, describing the learning that happens when engaging with a target language in situations that were not designed for learning, then Stickler and Emke's (2011) definition of non-formal learning is especially useful for describing learner-initiated activities that *have* been designed (structured) with a goal of improving target language skills but are not provided by an institution and likely do not lead to a credential. Examples of non-formal learning spaces include language cafés (intended for learning but unstructured, usually guided by a community of learners; see Murray, 2017 for an example), or in digital contexts where online materials are structured to scaffold learners but with enough freedom to allow them to initiate activities themselves. In workplace learning, "performance supports" are provided to help workers carry out tasks (Carliner, 2013). Because the non-formal space is the focus of my research, it is helpful to draw on the idea of performance supports from the workplace learning literature where self-directed learning is common. The non-formal space as described here, not only acts as a bridge between formal and informal environments, but also a powerful space for encouraging learner autonomy, capable of scaffolding learners towards more wild, naturalistic, and unstructured settings. In sum, I view the non-formal space as an emancipatory rewilding space.

Surely, the potential for learning in the wild today is extraordinary because of the amount of language content available online. Further, for many busy adults, a few hours a week in class

does not provide enough time to learn a language well, anyway (Collins & Muñoz, 2016). Supplementing exposure outside of the classroom is necessary. However, not all learners are ready to jump into the wild and learn exclusively from unstructured content that was never designed for learners (Winne, 2001). There are benefits to learners choosing their own tools and materials according to their proficiency levels and interests and simultaneously developing their capacity to manage more natural environments. Let us now bring together the autonomous learning and learning in the wild traditions so we can situate our understanding within an ecology of language learning.

A Rewilding Language Learning Ecology

Both learning in the wild, which has always highlighted the learning environment, and sociocultural definitions of autonomy, which explore autonomy within an eco-social system, have converged in their incorporation of Vygotskian theory and postmodern critiques. This has been acknowledged by researchers like Little and Thorne (2017), who have been following these developments for several decades and argue that both traditions have a lot to offer. I too argue that we need to embrace how they overlap in order to progress the narrative. In the next sections, I propose that current frameworks for learning beyond the classroom (e.g., those by Benson, 2011 and Reinders, 2017, 2018, 2020), as well as our frameworks for self-regulation, can help us integrate our research and understanding of learning in the wild and learner autonomy, as well as managing the learning progression towards more natural empowering learning environments.

The Formal/Informal Web of Learning

Nobody takes classes forever. Learners eventually move beyond the classroom and continue developing their language on their own (in the wild; as supported by Reinders, 2020). Traditionally, this progression has meant that teachers can serve as facilitators in developing

their students' strategies, metacognitive and other, in addition to providing direct interaction with the target language. In the past, an institution acted as the primary point of access. In most cases, the learning trajectory was more linear than it is today. It is increasingly recognized that most learning happens outside of the classroom, whether one takes part in formal learning (for the expressed purpose of a credential), in addition to informal (not involving a credential), or not (e.g., Carliner, 2013; Reinders, 2020). One can imagine that learners' experiences today would involve several *nodes* of experience (to borrow a term from Reinders, 2017) within a network of different nodes. Some nodes are more formal than others and consist of varying degrees of learner autonomy—hopefully according to the learners' needs. Technology has brought access to a variety of learning trajectories; however, the need to develop an ability to self-regulate in order to manage one's autonomy remains.

An Ecology of Learning Beyond the Classroom

What does an ecology of learning beyond the classroom (LBC) that encourages learner autonomy look like? Reinders (2020) describes a model for developing, delivering, and investigating language learning ecologies for learning beyond the classroom that he and Benson have been researching for the past ten to fifteen years. Reinders' use of the term suggests that LBC is informal and learner-initiated but offers a bridge between formal and informal settings. The extreme ends of the spectrum progress from the structured, teacher-directed, and assessment-oriented classroom, to the unstructured, self-directed wild, where one uses their target language to accomplish whatever everyday tasks they choose. To be more specific, Reinders (2017) views in-class and beyond class as interrelated, and that an ecology for learning should include efforts to: (1) encourage LCB, (2) prepare for LBC, (3) support LBC, and (4) involve LBC.

It could be argued that, with the right task design, any environment could move a learner along this axis, but that the first two stages would require more structure than the last two because, initially, learners are less able to self-regulate. Stickler and Emke (2011) also point out that even "non-formal learning presupposes a high level of autonomy and reflection on the part of the learner". Citing Darasawang and Reinders (2010), the authors state that "autonomy is not an innate trait of the language learner but needs support or training to develop" (p. 154). The understanding that increasing one's capacity to act with autonomy is better achieved with support or training is also shared by others in the field (e.g., Andrade & Bunker, 2009; Murray, 2004; Reinders & Balçikanli, 2011).

The next axis, from Reinders and Benson (2017), consist of four dimensions to characterize LBC: location, formality, pedagogy, and control. Location defines the space, such as learning with a teacher and pupils in a classroom or learning from an online course where one progresses at one's own pace. Formality refers to whether the competencies are credentialed. Pedagogy refers to the nature of instruction or non-instruction. Control refers to whether tasks are learner-initiated or other-initiated. Imagine now that task progression can be designed around these dimensions for the purpose of encouraging, preparing, supporting, or involving LBC. This progression involves pulling structure away from a formal classroom environment by manipulating the four dimensions of LBC. Or in Little and Thorne's (2017) rewilding terms, introducing *structured unpredictability*. This dissertation is founded upon the premise that providing students access to a music application with a collection of songs graded according to vocabulary level (similar to *Bande à Part*) allows students to explore independently (i.e., less structure and more control) within their ZPD while still providing supports to the learning process.

Finally, Reinders (2020, p. 70) suggests that each of the stages of the self-directed learning process be covered (which orbit around one's individual reflections, motivations, and interactions). The stages include: (1) identifying learning needs, (2) setting goals, (3) planning learning, (4) selecting resources, (5) selecting strategies, (6) practice, (7) monitoring progress, and (8) assessment and revision. Taken together, we have three axes that can be used to move learners towards more autonomy: the ecology of learning, which encourages the progression towards informal interactions beyond the classroom; the four dimensions characterizing LBC, which is concerned with the control and pedagogy, etc.; and the stages of self-directed learning, which I refer to as self-regulated learning and, consequently, has to do with a learner's ability to manage this progression.

Self-Regulation

If the goal is to empower language learners such that they can navigate less structured environments on their own, the next question is how? Over the past two decades, Second Language Acquisition (SLA) researchers have been narrowing in on the construct of self-regulation as a way to understand how learner autonomy develops (e.g., Andrade & Bunker, 2009, Gao & Hu, 2020). If autonomy is a spectrum (people can act with more or less control over their learning), then one's capacity to self-regulate determines how capable one is moving along the spectrum (Andrade & Bunker, 2009). For example, the better learners can evaluate their needs, set goals, plan, and monitor their progress, the more effectively they can take control of their learning experience and, thus, act with autonomy. However, this differentiation is not yet consistent in the literature. Gao and Ho (2020) also bring up the fact that, occasionally, researchers use autonomy and self-regulation interchangeably, while others differentiate self-regulation from self-regulated learning, so it is worthwhile at this point to take a short detour to

cover some definitions and the historical development of these related constructs in order to negotiate this space.

Though the term *self-regulation* has been in use since the 1970s (e.g., Mlott, Marcotte, & Lira, 1976), and models of self-regulation emerged in the late 1990s (e.g., Zimmerman, 1998), self-regulation is only more recently receiving attention in SLA. After decades of research on language learning strategies that struggled for definitional consistency between "observable behaviours and inner mental operations, or both" (Tseng, Dörnyei, & Schmitt, 2006, p. 80), Tseng et al. (2006) proposed replacing learner strategies with the concept of self-regulation, which could better integrate with the broader field of Educational Psychology. In recounting the evolution of this term and synthesis of constructs involved (learning strategies, motivation and behaviour, cognition, and metacognition), Gao and Ho (2020) suggest that it is still a work in progress.

Framework for Self-Regulation

Andrade and Bunker (2009) propose a framework and definition that is broad enough to include most of the related constructs to date and, though is intended for distance learning contexts, is easily adapted to Benson and Reinders' (2011) LBC context. More specifically, they do so in a way that links SLA research in this area to the broader research on autonomy, such as with Zimmerman (1994). Andrade and Bunker include four psychological components, namely cognition, metacognition, motivation, and behaviour (as is recommended by Raya & Vieira, 2020, p. 33-35), each of which inhabit six dimensions: motive, method, time, physical environment, social environment, and performance (from Zimmerman, 1994). As such, language learning strategies, goals, context, etc. are all factored in. A learner's objective is to *influence* the psychological components and gain more control over the six dimensions. For example, if a

learner improves how they plan their study (a metacognitive component), they will have to exert control over several dimensions including method, time, and environment. The *structure* and *dialogue* included in learning materials and instruction scaffold this process—that is, they are part of the *influence* and can contribute to developing a learner's ability to control the four psychological dimensions. Overtime, learners move from higher scaffolding needs and lower autonomy to lower needs and higher autonomy.

Using the six dimensions attributed to autonomy (i.e., motive, method, time, physical environment, social environment, and performance) is important because if we want to understand learners' relationships to *when* they study, for example, you need to ask questions about their beliefs about how long they should study (cognitive), how they plan their study (metacognitive), when or how long they actually study (behaviour), and why they do so (motivation). Learners that improve their ability to self-regulate demonstrate "a movement from object- or other-regulation to self-regulation, which is a clear sign of the internalization of the resources that were once necessarily outside of the individual and that have subsequently become resources for the individual to act under conditions of greater autonomy" (Little & Thorne, 2017, p. 17). This progression towards self-regulation and autonomy is a key aspect of Andrade and Bunker's (2009) model, which is employed in chapter 4 of this dissertation to analyze learners' autonomous development in the context of L2 learning with music.

To summarize, the umbrella under which this research is situated is the non-formal space in which, from a rewilding perspective, learners are enabled to progressively initiate more of their learning experiences while improving their ability to self-regulate in order to then operate autonomously in informal contexts.

This Dissertation

Considering the context outlined above, in this dissertation, I explore how self-directed pedagogical tools fit into learners' target language development with a focus on the use of songs beyond the traditional classroom. I recommend the non-formal space as an environment to employ rewilding pedagogies where learners initiate their own use of structured materials in order to progress to wilder environments with less structure. I investigate aspects of this development through songs and the *Bande à Part* (*BàP*) app, focusing on an extensive exploration of the affordances provided by listening to songs in a target language (Manuscript A), learner perceptions of the self-directed *BàP* French music app (Manuscript B), and self-regulatory processes through an analysis of learner reflections when participants attempt to improve their pronunciation after imitating sung and spoken recordings (Manuscript C).

Songs were chosen as the form of content to investigate because the average person already listens to over two and a half hours of music per day (IFPI, 2021; and almost four hours per day in the USA) and music has been reported to be the most common source of target language exposure outside of the classroom among English foreign language learners—more common than films, video games, reading, and other forms of internet use (Lindgren & Muñoz, 2013; see also Pavia et al., 2018), although social media use might surpass music soon. Despite this exposure, the use of songs to aid second language acquisition remains an under researched area of study (Engh, 2013a; Ludke, 2018; Werner, 2018).

As will be outlined below, the overall scope of this dissertation follows Cardoso's (2022) chronological framework for examining new technologies in second language contexts: 1) Development (explored in a separate study reported in Sundberg & Cardoso, 2019, not included

in this dissertation), 2) Exploration, 3) Assessing suitability, and 4) Assessing pedagogical effectiveness.

The first stage involved the development of the music app, *Bande à Part*. A discussion of its theoretical and pedagogical rationale can be found in Sundberg and Cardoso (2019). An important note to take away from this development is that it followed a design-based framework which requires iterative upgrades based on feedback (Manuscript B describes the first step in gathering feedback from *Bande à Part* users).

The second stage, an exploration of pedagogical affordances is provided in Chapter two of this dissertation (Manuscript A). However, rather than limiting the discussion to *Bande à Part*, I provide an in-depth review and analysis of the affordances of learning through songs and introduce a theoretical lens for understanding how research on incorporating songs into target language development can progress in the field of second language acquisition.

In Chapter three (Manuscript B), I present a study that evaluates learners' perceptions of the self-directed application, *Bande à Part*, designed to encourage the learning of French through songs. This app provides structure for learners, but is highly flexible and non-linear, providing a good example of a non-formal learning tool. This study assesses the suitability of this tool (stage three in Cardoso's, 2022, chronological framework) through gathering learners' perceptions of the tool. These perceptions are used to understand whether *Bande à Part* is likely to be adopted by learners, an important evaluative step if learners are initiating the activities that they take part in. This chapter is a pilot study that allows for early testing, a key component of design-based research programs (Reeves et al., 2005), and establishes whether learners would voluntarily continue to use the app on their own time—a requisite for learner autonomy and future studies on pedagogical effectiveness. The Technology Acceptance Model (Davis, 1989; Venkatesh &

Davis, 2000) is adopted to gather learners' perceptions of the app, a robust model that is used extensively across many disciplines but not yet widespread in CALL or SLA literature.

Chapter four (Manuscript C) is a first attempt at assessing pedagogical effectiveness. It is a study that explores learners' capacities to self-regulate their pronunciation development while listening to sung lyrics compared to spoken lyrics. While the *Bande à Part app* is not used in this study due to the need for constraints, lyrics from the app are still used. This is an important contribution considering the promise of language learning through songs and the lack of research on this topic (see chapter two; also supported by Engh, 2013a; Ludke, 2018; Werner, 2020), the amount of music the average adult listens to (reports of over two and a half hours per day according to IFPI, 2021), and the need to advance our understanding of self-regulation (Gao & Hu, 2020).

This same chapter investigates self-regulation as an umbrella construct consisting of cognition, metacognition, behaviour, and motivation through learners' reflections on their attempts to improve their own pronunciation after listening to sung and spoken content. Since researchers have adopted the more comprehensive term "self-regulation", there have also been calls to further study the "inner lives of learners" (Pavlenko, 2013). The intention here is to research learners' abilities to self-regulate without structured scaffolding in order to offer recommendations about what kind of scaffolding might facilitate learner autonomy for learners at different proficiency levels. This study uses a series of tasks and think-aloud protocols to explore learners' awareness and how they use that awareness to attempt to improve their pronunciation after listening to and imitating, either a singing voice, a spoken voice, or recordings of their own voice. By comparing these three conditions, we are better able to understand learners' strategies that guide pronunciation development from sung and spoken

content. This provides a baseline for understanding which features may or may not be needed in self-directed music apps (like *BàP*) to scaffold learners (e.g., textual enhancements for less salient features like liaison or stress in target French). In addition, this study highlights the potential of adding a karaoke-like feature to *BàP* that allows learners to hear their own voice versus learning from purely listening to the singing in songs.

Finally, Chapter five (concluding remarks), provides a reflection on chapters one through four bringing together what has been learned in order to provide suggestions for future research. Here, I describe how chapters three and four were selected as important contributions because they capture learner volition and strategies learners use to improve their pronunciation (one aspect of self-regulation that is not well understood with regards to learning from songs). Assessing suitability and pedagogical effectiveness are two essential facets of researching autonomous learning tools; in addition, this analysis represents the first steps of an iterative design-based research program. Further analysis of learning gains made from using *Bande à Part* are underway and constitute the next steps to be taken in this research program, some of which is described here. A more comprehensive evaluation of the effectiveness of tools like this are useful, but as Long (2015) suggests, it is difficult to holistically evaluate multifaceted materials, tools, and curricula; in addition, a piecemeal approach such as described here is more conducive to making iterative improvements. Moreover, we already know that several of the features included in the app are effective as demonstrated by other studies (e.g., including first/target language subtitles, grading content for vocabulary levels), which is why they have been included in the design of the app in the first place.

Summary: Goal of the Dissertation and Component Studies

The overarching goal is to explore the non-formal learning space through a rewilding lens and using songs as a medium from which to explore this space. This includes a thorough discussion of affordances provided by using songs in L2 learning (Chapter 2), whether learners adopt and use structured materials such as *BàP* (Chapter 3), as well as how they internalize and learn from what they hear in songs (Chapter 4). This dissertation uses the following questions to guide the research:

1. Manuscript A (Chapter 2): What are the theoretical affordances of language learning through songs? What does the empirical work on using songs for teaching and learning reveal about the potential benefits and limitations of language learning through songs? What are the future directions for using songs to develop language skills?
2. Manuscript B (Chapter 3): Are learners willing to use the self-directed *Bande à Part* app to help them learn French?
3. Manuscript C (Chapter 4): What aspects of language and language learning processes do learners notice when imitating sung vs. spoken input?

Table 1 provides a summary of the goals and contributions of each manuscript.

Table 1

Summary of Component Studies

Manuscript	Goals	Contribution	Participants
A	◊ Provide theoretical lens for understanding language and music research ◊ Review pedagogical affordances of listening to L2 music	◊ Bringing language and music research into L2 research sphere	◊ N/A
B*	◊ Gather learner perceptions on *BàP*, a non-formal music application for learning French	◊ Demonstrate importance of iterative design ◊ Add to limited literature on Technology Acceptance Model in CALL and SLA	◊ Adult Learners of French
C	◊ To explore learners' strategies and capacities to self-regulate pronunciation development from exposure to spoken and sung context	◊ Add to limited research on benefits of listening to music ◊ Important qualitative exploration of self-regulating pronunciation processes	◊ Adult Learners of French

* This peer-reviewed article has been published in the *CALICO Journal*: Sundberg and Cardoso (2022). I am the first and main author in this publication

Chapter 2: Using Songs in Second Language Learning, Teaching, and Research

Musical skill, or musicality—the term often used in the literature, is comprised of several fundamental skills, such as tonality, sense of rhythm, and various kinesthetic and aesthetic abilities (Shuter-Dyson, 1999). These skills are employed to comprehend and produce music (henceforth *music* will refer to the non-linguistic, instrumental elements of music, while *songs* will refer to music accompanied by lyrics—a distinction that is important when looking at past research). Language, too, relies on comparable skills to produce and decode prosodic features and, since they are both dependent upon the auditory system and respective connected neural networks, it should be no surprise to the phonologist that research on music, the brain, and language acquisition intersect. Musical abilities are indeed correlated with second language (L2) pronunciation abilities and development (Milovanov, 2009; Milovanov et al., 2008; Nardo & Reiterer, 2009; Sarsarabi & Sosouri, 2016; Slevc & Miyake, 2006); although implications for using songs for L2 acquisition extend beyond phonological development (Gilleece, 2006; Kraus & Chandrasekaran, 2010; Nardo & Reiterer, 2009). However, for most second language acquisition researchers, knowledge of these overlapping processes is rudimentary; there is recognition that some neural pathways are shared, justifying the incorporation of songs as L2 input, but a deeper understanding of the research originating outside of our field would lead to far reaching implications that could push our field in new directions.

Recent literature has recognized this gap. Engh (2013a) has argued for more work evaluating the use of music and song in L2 learning environments. Though he claims that incorporating music into L2 learning is theoretically and empirically justified, specific affordances are not well understood. This view is also reaffirmed by Werner (2020).

As a form of input for learning, listening to songs presents a tremendous opportunity.

After all, music is a major part of every culture and as Werner and Tegge (2020) argue, pop culture is not only a "vehicle [for] various other instructional content but… a valuable point of departure for a range of language-related competencies and topics." In fact, the average person listens to at least two and a half hours of music per day (almost four hours per day in the USA; IFPI, 2021). Converting this time spent listening to music, which is already a part of many routines, into learning time, does not require a change in priorities or an addition of study time to an already busy schedule. Especially at a time when technology has increased opportunities for self-regulated, out-of-class learning (e.g., Sundqvist, 2019), and research has increasingly directed focus on how to efficiently carry pedagogical methods through emerging mediums (Chapelle, 2003).

In this paper, I review and summarize theoretical considerations for using songs in second language (L2) learning. These considerations are an attempt to translate literature from the fields of Music Education and recent developments in Cognitive Science to highlight similarities and differences between musical and linguistic processing with reference to SLA theory. This cross-disciplinary exercise is needed to bring this knowledge into the field of SLA. As such, comparing the similarities and differences in neural processing serves as a lens for understanding the state of current research on incorporating songs into L2 learning environments, as outlined in the Theoretical Underpinnings section below. Next, I review relevant empirical literature through this lens, discussing the interface between linguistic and musical skill development, transfer effects from musical to linguistic abilities, its application to L2 learning, and, finally, research on the use of songs in L2 settings. Lastly, I discuss future directions of music-enhanced pedagogy with a focus on the affordances offered by researched-based approaches that use songs to extend the classroom via technology.

Theoretical Underpinnings

What is the difference between music and language? To understand what songs offer L2 learners, it is important to recognise the similarities and differences in the cognitive and motor processes that are used when interacting with music and language. First, I review musical and linguistic *processing* considerations by exploring vocal acoustic differences between singing and speaking. Then, I compare structural differences in song and speech as artefacts, which I will refer to as *content*. I relate these areas of research back to L2 learning theories and, later, the differences between language and music will underpin the interpretation of empirical work that uses songs in L2 learning, teaching, and research. This parallel arrangement of discussing research on *processing* (including both input and output) followed by a section on *content* (in song and speech) will continue throughout the paper.

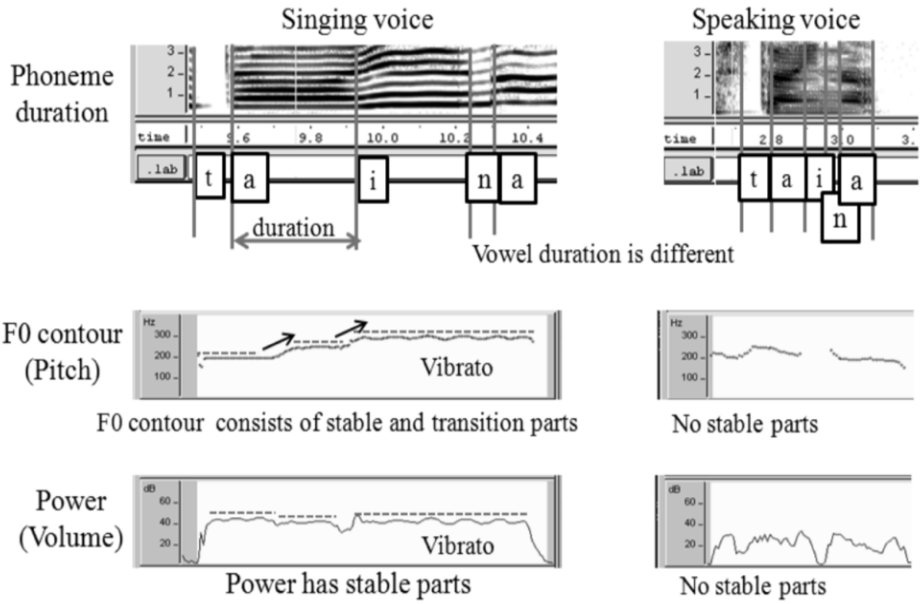
Processing: Singing versus Speaking

Acoustic Differences—pitch, duration, volume

The most commonly cited acoustic differences between singing and speaking voices are differences in the fundamental frequency (F0) contours, phoneme duration, and volume (Aso et al., 2010). Every vowel consists of an F0 (a.k.a. pitch) and several formants that resonate with this F0. Figure 1 below compares singing and speaking voices, demonstrating that the F0 pitch contour of a singing voice is more stable and higher than a speaking voice, and may modulate (Aso et al., 2010). Furthermore, sung vowels are lengthened and amplitude is synchronized with the changes to the F0. In contrast, spectral changes in speech occur more quickly (Patel, 2011).

Figure 1

Differences Between Singing Voice and Speaking Voice (From Aso et al., 2010).



For the purposes of teaching singing skills to music students, Nix (2015) echoes these acoustic distinctions between speaking and singing and adds relevant differences important to singers' goals that, interestingly, affect the benefits and limitations of using music for L2 learning. For instance, intelligibility is a primary goal when speaking, whereas with singing, intelligibility might be "sacrificed for aesthetic" purposes by manipulating phonetic properties (Nix, 2015). Nix also notes that speaking occurs at a lower intensity and lower frequency and utilizes a narrower range of intensities and frequencies whereas singing covers a greater range of intensities and frequencies and is typically louder and higher in pitch. Finally, he argues speaking has more co-articulation between consonants and vowels.

These differences have implications for L2 learning and are relevant not just for oral production activities using songs but also for listening practice. For example, the longer vowel length (and consequent fewer syllables/second processing) in sung lyrics provides a longer time

frame for listeners to register vowel quality and for speakers to practice vowels that are not present in their L1 inventory.

In addition to opportunity to focus on vowel quality, The OPERA hypothesis (Patel, 2011), which is discussed in the next section, suggests that music increases one's focus on acoustic elements, such as pitch. This suggests that if there is an additional focus on acoustic features when listening to songs, L2 suprasegmental development could be improved with respect to the three distinctions provided by Aso et al. (2010)—F0, vowel duration, and volume—and, indeed, research on music reviewed later in this paper supports this hypothesis (e.g., Kraus & Chandrasekaran, 2010). To further illustrate the importance of discerning pitch, duration, and volume, note that stress perception relies on these three elements. For instance, consider English, a language in which discriminating words like "bad" from "bat" requires not only distinguishing the quality of the final consonant (which might be glottalized, flapped, or not released), but also on distinguishing between shorter and longer vowel lengths (the "a" being longer in "bad" than "bat"). It is not only music training that improves ability to encode F0, since bilinguals continually develop their ability to encode F0 (Krizman et al., 2014), so the combination of music training and exposure to songs in one's L2 dually reinforce this development. As drawbacks, it would be unlikely to expect sentence-level exemplars of pitch to be able to be learned from songs since pitch is prescribed in music according to the composition. Finally, if intelligibility is sacrificed in a song, problems making out meaning can be expected.

Moving from acoustic features (F0, vowel duration, and volume as captured on a spectrogram) to representation and temporal processing of sound in the brain leads us to research on prosodic features of music and speech. Levitin (2006) outlines seven distinct features of musical prosody: pitch, rhythm, melody, harmony, tempo, metre, and loudness. All of which are

comparable to prosodic features in speech, though recently space and timbre have been suggested as additional elements (Jordanous, 2016).

In L2 research, prosodic elements have been tied to L2 speech ratings in several areas, including accent, comprehensibility, and fluency (Trofimovich & Isaacs, 2012; Trofimovich, Kennedy, & Blanchett, 2017). Neuroimaging research has also revealed how these elements are processed in the brain and, because converting sound to accurate prosodic representations with meaning in a target L2 are contingent upon these processes, an overview of this research is provided next.

Implications from Cognitive Science

One well supported model for music perception is offered by Koelsch (2011) and is useful for introducing the research on music and language processing. Koelsch likens his model to Friederici's (2002) model for language perception. Together, he refers to these processes as the "music-language-continuum". In crude terms, sound enters the auditory system for feature extraction (timbre, intensity, etc.), representations are formed through a gestalt process, and then a syntactic process hierarchically assembles the sequence of events. Numerous other processes are involved such as working memory, long-term memory, and the motor system; however, research on representations and syntactic processing have been the focal point for the last two decades.

The modularity and overlap of mechanisms in the brain has long been debated (Peretz & Coltheart, 2003), but since Patel's (1998; 2003) work in neuroimaging, an interesting paradox has emerged in the study of language and music. That is, that although individuals suffering from aphasia can retain musical abilities (suggesting separate areas for music and language mechanisms; summarized in Peretz & Coltheart, 2003), neuroimaging studies still show a great

deal of overlap in the processing of both music and language (Besson et al, 2007; Maess et al., 2001; Patel 1998). Patel (2003) has proposed the *shared syntactic integration resource hypothesis* to explain this paradox.

This hypothesis essentially states that representation in music and language is separate from its processing, but that many of the above-mentioned prosodic features run on the same neural networks for both music and language (Patel, 2003, 2013). This sounds intuitive, since music and language both rely on the auditory system and this hypothesis has been gaining support (e.g., Brown, Martinez, Parsons, 2006; Fedorenko et al., 2009). Recent research has even extended this further to suggest that this network deals with sequencing all temporal elements (tone in music, words in language, and even the order of operations for goal planning; Asano & Boeckx, 2015; Fitch & Martins, 2014). This hierarchical sequencing, termed a syntactic process (not to be confused with the narrower sense of syntax in linguistics), is a shared network between music and language but linked separately to their respective semantic associations and sensory-motor operations.

To be clear, this does not mean that tonality and syntax in language and music are necessarily the same. A review by Zatorre and Gandour (2008) shows that bias in hemispheric processing is dependent upon context (task or paralinguistic signaling), tone in language (e.g., Mandarin Chinese vs. English at the syllable level), and training. For example, even with bilateral involvement, there is a left-side bias if acoustic cues elicit a phonemic category or occur at a rate similar to speech patterns, and a right-side bias if pitch precision or tonal direction is the focus (Zatorre & Gandour, 2008). However, since musical and linguistic tones both consist of temporal sequences governed by internal relationships between elements, the same networks may be used and likely connect across hemispheres (Fitch & Martin, 2014; Patel, 2013). Further

exploration of these hypotheses might help explain why musical training is often reported to correlate with speech encoding (e.g., Besson et al., 2011) and L2 pronunciation ability (e.g., Sarsarabi & Sosouri, 2016). I argue that this research on representational and syntactic processing should be considered by L2 researchers interested in exploiting musical processing mechanisms for L2 suprasegmental development.

One proposed avenue for L2 researchers is through testing Patel's OPERA hypothesis, which tries to predict which conditions are necessary for musical training to positively impact neural speech encoding (Patel, 2011, p. 1-2). The conditions are as follows (Patel, 2011, p. 6):

1) Overlap: there is overlap in the brain networks that process acoustic features used in both speech and music,
2) Precision: music places higher demands on these networks than does speech, in terms of the precision of processing,
3) Emotion: the musical activities that engage this network elicit strong positive emotion,
4) Repetition: the musical activities that engage this network are frequently repeated,
5) Attention: the musical activities that engage this network are associated with focused attention.

Patel (2011) argues that this hypothesis is not to pit musical training against linguistic training, but rather to understand the systems involved in both. In addition, he mentions that music may be an attractive means of providing "home-based" repetitive training tasks (p. 11), a point I also raise in this paper.

One example of support found for the temporal aspect of Patel's *shared syntactic integration resource hypothesis* is by Jungers et al. (2016), who demonstrated that speech rate

can be primed by either music or language exposure. To test this, two experiments were carried out: 1) within-domain, where participants listened to fast and slow spoken sentences followed by speaking tasks; and 2) cross-domain, where participants listened to fast and slow melodies followed by speaking tasks. Regardless of whether participants were listening to spoken sentences or melodies, their subsequent speech rate was affected by the previous stimuli, meaning fast sentences and fast melodies primed faster speech and vice versa. In addition to supporting the shared processing hypothesis, this research implies that music could be used to progressively increase L2 learners' speech rate by starting with songs with slower lyrics and increasing to faster ones. It is also worth highlighting that the above research has primarily focused on *instrumental* music processing. Using songs in L2 settings, however, incorporates within-domain language processing and cross-domain music processing (Patel, 2013). Researchers should ensure to account for or attribute effects appropriately as well as consider that synergy or interference between domains is possible.

Beyond the similarities and differences in processing language and music, it is important to review differences in content. In the next section, I explore what music and song have to offer learners in terms of content in ways that differ from regular speech. Three avenues of comparative research I suggest for inclusion are: repetition, motivation, and exposure to prosodic cues.

Content: Song Versus Speech

Repetition and motivation

In songs, lyrics and sounds are arranged to form patterns that repeat unlike with most speech. These repeating patterns lead to anticipation on the part of the singer or listener (Salimpoor et al., 2015). Without a degree of repetition and/or predictability, music would be a

random collection of sounds and anticipation would be impossible. Evidence has suggested that the interplay between the anticipation of a sound and the surprise of an unexpected sound is what makes music enjoyable (Ekeus et al., 2012; Salimpoor et al., 2011). Our ability to anticipate might explain why pop songs are considered so accessible (and quickly grow tiresome?), while jazz takes more investment. This also helps to explain why we enjoy genres most familiar to us (Salimpoor et al., 2013). Note here that songs are not only comprised of repetitive internal elements (e.g., melodies, choruses), but also that it is a form of input to which we listen repetitively. Repetition has been extensively researched in L2 learning with noteworthy implications for memory and recall (Ellis & Collins, 2009; Ellis et al., 2015), skill acquisition theory (Dekeyser, 2007), and task design (Bui & Skehan, 2016; Gatbonton & Segalotwitz, 2005) to name a few areas.

The use of motivating material is also unmistakably important; for instance, Gatbonton and Segalotwitz (2005) have noted that repetitive tasks are often avoided precisely because they are not interesting to learners. Using songs is a natural way of combining repetitive and enjoyable target language content (recall that emotion, repetition, and attention were part of the OPERA hypothesis discussed above and are predicted to positively affect speech encoding).

Prosody in Music and Language

Arranging prosodic elements such as certain tones within a given rhythmic structure is what produces patterns in both music and linguistics. These too have been comparatively investigated (Patel & Daniele, 2003). For example, English rhythm is considered stress-timed, usually meaning the duration between stressed syllables is held constant, while rhythm in French is syllable-timed, meaning the duration between vowel onsets is held constant. In reality, empirical research indicates it might be more accurate to say that vocalic duration varies more

with stress-timed languages than syllable-timed languages (Grabe & Low, 2002; Patel & Daniele, 2003, p. 36). Nonetheless, as Patel and Daniele (2003) investigated prosody in instrumental music, they found that it was significantly influenced by the composer's first language. This insight could mean that listening to L2 songs might increase exposure to L2 rhythm. To my knowledge, no research has been carried out on teaching rhythmic structure from music composed in the target language. To carry out this research, one could compare a group that listens to instrumental music composed by someone who speaks their target language to another that does not. Pre- and post-tests could indicate if the experimental groups improved their rhythm in their target language significantly more than the control.

Summary of Findings

There are several key theoretical takeaways critical to the empirical research that is discussed in the next section. First, song and speech have acoustic and structural differences. The benefits and limitations of using songs to aid L2 learning lie in these differences. The section above, on *processing*, offered a cross-sectional look at prosodic processing as captured by neuroimaging research, while the section on *content* reviewed the structural differences between messages delivered through these two modes. This forms the lens for investigating the question: What do songs have to offer an L2 learner?

Empirical Literature

To answer this question, this chapter reviews empirical research that has explored the interface between musical and linguistic skill development. This review incorporates the above lens by analysing affordances from *processing* and *content* perspectives. I begin by looking at musicality and L1 correlates, including transfer effects (musical skills that transfer to L1 skills), sensorimotor training effects on L1 development, development of musicality, and effects on

memory. I then review research that speaks to musicality and L2 correlates, including studies that deconstruct musicality for the purpose of understanding the effects of musicality on L2 development. Next, I review studies looking at the effects of musicality (one's sensitivity to, knowledge of, or talent for music) on pitch cues, and memory on L2 pronunciation. Following this focus on music and language *processing*, I review literature on how songs contribute to vocabulary and grammar development because of the unique way in which songs present *content*. Finally, I discuss future directions, including suggestions for research and practical advice for incorporating songs into language learning contexts and for developing pedagogical materials. Recall that the main goal of this chapter is to uncover the role and potential of using music in L2 education and to set the scene for the rest of the dissertation.

The Interface between Musical and Linguistic Skill Development

To explore the interface between musical and linguistic skills and the development of those skills, two key areas must be tackled: assessing transfer effects between domains (music and language correlates) and disentangling individual attributes that comprise individuals' musicality. This second point allows researchers to operationalize abilities relevant to L2 research. It also contributes to an understanding of what songs can offer learners because researching songs directly confounds linguistic and musical variables. Investigating how one develops their musicality is much like investigating how one learns a language and, as such, the complexity is beyond the scope of this paper. The extent to which one's musicality develops as a result of listening to songs is likely specific to the listening skills involved; in addition, this ability would depend on several individual internal factors, such as the attention one pays to practicing discriminating among the prosodic features present. This is a skill probably aided by instruction and a potential that future research can explore.

Processing: Musicality and L1 correlates

Transfer effects

Musical abilities have shown many transfer effects to language skills. Besson et al. (2011) demonstrated that music training aids the "encoding" of and "sensitivity" to acoustic features that could help with speech. Kraus and Chandrasekaran (2010) showed that musicians have more accurate representations of pitch contours when processing a foreign language and that musicians have faster responses to changes from consonants to vowels (p. 601). Musical training interventions have shown an impact in a number of other areas including reading skills (Moreno et al., 2009), speech perception (Liu et al., 2015; Slater et al., 2015), segmentation (François et al., 2012), phonological awareness (Degé & Schwarzer, 2011), verbal memory (Ho et al., 2003), executive control (in as little as 20 days; Janus et al., 2016), and language remediation (Schön & Tillmann, 2015). Kraus and Chandrasekaran (2010) make the analogy to physical fitness, reasoning that musical training improves "auditory fitness" (p. 599). More specifically, it seems that listening and producing music fine-tunes attention to sound. Much of this research has been carried out in light of the work in cognitive science discussed above. L2 researchers should continue to explore the associations listed above. To note, several of these studies are carried out with children, so it is crucial for L2 researchers working with adults to take these findings into consideration.

Sensorimotor training

Sensorimotor training involves coordinating physical movement alongside aural input; this is commonly developed during the learning of an instrument. Considering effects from sensorimotor training in L2 language development may not always be convenient but is worth discussing because of the potential benefits or limitations stemming from using (or not using!)

our bodies in addition to our voices and ears. Research has indeed shown that there are benefits to musical training that come from incorporating the sensorimotor system alongside the auditory system. For example, Lappe et al. (2008) used a two-week piano lesson treatment with non-musicians to measure changes in brain plasticity and found evidence for a connection between the sensorimotor and auditory systems. In addition, piano practice produced more changes in the brain than auditory training alone. It may be that some of the sensorimotor reinforcement is difficult for non-musicians to take advantage of.

Development

It is well established that musicality can be developed. However, the first longitudinal study to demonstrate the development of musical ability and subsequent transfer to linguistic ability in children was Kraus et al. (2014), who found that syllable differentiation improved at the neurological level after a two-year community music program. They also noted that changes likely require longer than a year, since testing at the one-year mark did not reveal a significant difference. Cross-sectional correlations, such as found in the work above, suggested that this was an important avenue for research. More of this type of work should be done and an opportunity to pioneer this type of research in the field of second language acquisition is waiting. Is listening to songs with L2 lyrics over a long timeframe enough to encourage this development?

Effect on memory

Verbal memory has also been shown to improve with musical training while visual memory does not (Ho et al., 2003). The main implication for L2 learning is with regards to reinforcing vocabulary and prototypical exemplars of grammatical sequences present in songs. This comes not only from the repetition in the songs, but possibility from a musical reinforcement (tone and rhythm might trigger the memory of lexical items and syntactic

patterns). A secondary implication could come from improved working memory function since working memory is more involved while processing music (Besson et al., 2011). Work in L2 acquisition has been done in the area of lyrical recall and is discussed near the end of the next section.

Processing: Musicality and L2 Correlates

Deconstructing musicality into its subcomponents

Nardo and Reiterer (2009) exemplify how to deconstruct musicality in the field of L2 acquisition. They first break down musicality into that which is innate talent and that which is learned ability. They explain that talent is normally distributed among the population and that everyone has some "degree of potential"; most have an "average aptitude" (p. 215). Next, they identify several abilities that can be measured and improved through practice (according to Shuter-Dyson, 1999) including: tonality (relative pitch perception), rhythmic abilities (metre, tempo), kinesthetic abilities (including how this effects auditory perception as shown by Mainwaring, 1933), aesthetic abilities, and creative abilities. Nardo and Reiterer (2009) then measure which aspects of musicality best correlate with L2 language measures. In this study, they found that rhythm and pitch discrimination have the most influence, and the language measures that were most impacted were phonetic talent (e.g., the participants' ability to perceive and produce features) and grammatical sensitivity (as measured by the Modern Language Aptitude Test, which requires the identification of words that play similar roles across contexts; e.g., compare, "Mary" in "MARY is happy" to the underlined words in "From the look on your face, I can tell that you must have had a bad day"). Interestingly, rhythm and pitch discrimination both belong to the syntactic processing realm and not representational realm. In fact, all musicality measures (except the score related to participants' self-reported ability and interest in

dancing) showed correlations to language measures. These musical abilities should be further researched with respect to L2 acquisition because they are skills that can be improved (not innate) and their impact on L2 acquisition is promising.

Effect of musicality on L2 pronunciation lessons

In a pretest-posttest experiment, Sarsarabi and Sorouri (2016) measured the effect of 15 consecutive pronunciation sessions on Persian learners' pronunciation of English. Using a multiple intelligences framework, they controlled for musical intelligence and found that, although all participants improved their L2 pronunciation, participants with high musical intelligence performed better on assessments and improved more overall. This corroborates transfer effects related to music and L2 pronunciation, similar to what was previously found in L1 research. For similar findings on music training and L2 pronunciation, see Milovanov et al., (2007), Pastuszek-Lipinska (2008), and Slevc and Miyake (2006).

Zheng et al. (2022) compared music aptitude to domain-general auditory processing ability. They found that although music aptitude was associated with pitch and formant discrimination, domain-general auditory processing, defined as ability to discriminate pitch, formants, duration, amplitude, was a stronger predictor of L2 pronunciation ability. Because this was a cross-sectional study, it is not clear how auditory processing develops during the L2 learning process, but it is important to recognize that it plays a large role and is separate from but connected to music aptitude. This finding supports Patel's (2003) *shared syntactic integration resource hypothesis* because it suggests either effects specific to musical developmental processes transfer to domain-general processes, or, more likely, that these are in fact the same syntactic processes. Either way, they are connected.

Pitch cues and segmentation

Rather than correlating language outcomes to participants' musical abilities, Schön et al. (2008) evaluated learners' ability to segment and learn words according to the musical (prosodic) cues present in the input. To understand a continuous stream of words, one must first segment a sequence of syllables and words. Drawing on the findings from Patel's work on music processing discussed above, which suggest that the processing of hierarchical sequences in music and language uses a shared network, Schön and colleagues investigated foreign language segmentation from a statistical learning perspective. This research replicates research from Saffran et al. (1996a; 1996b) but using sung sequences in addition to spoken ones. In a series of three experiments with three different groups of participants, they analysed non-word learning of monotone spoken sequences (group one), sung sequences with matching musical and linguistic boundaries (group two), and sung sequences with mismatched musical and linguistic boundaries (group three). The matching boundaries condition meant that each syllable was always sung on the same pitch (e.g., the word "py.mi.so" was always sung as B5 E6 F6), while the mismatch condition meant that syllables did not always have the same pitch. The results showed that group two, who received matching musical and linguistic input, outperformed both other groups, and group three, who received mismatching musical and linguistic input, outperformed group one. This research supports previous work highlighting how pitch cues aid segmentation (similar to tones in tonal languages, or intonation in non-tonal languages in the case of group 3?) and that this holds true when learning a second language.

At this point, it is useful to recall the acoustic differences between music and speech mentioned by Aso et al. (2010)—F0, phoneme duration, and volume. The only manipulation in the Schön et al. (2008) study was pitch—phoneme duration and volume were held constant

across experiments. Pitch alone seems be enough to improve segmentation, at least with a non-real language, and it could be hypothesized that other prosodic cues might provide further assistance.

To elucidate these findings, consider Patel et al. (2010), who found that, in a quiet environment, Mandarin speakers had no trouble with intelligibility of Mandarin Chinese sentences when pitch information was removed from the input. However, as background noise was added, intelligibility depreciated. This is surprising considering that Mandarin Chinese is a tonal language. Based on these findings, I suggest that depending on the context, language in question, and experience with said language, prosodic cues may be either more or less redundant. As the context becomes more difficult (no longer in a lab but in a classroom or, in the case of listening to music, on a metro, subway, or bus) and familiarity with a language decreases, prosodic cues become more important, and redundancy is more crucial to intelligibility. It also seems to be the case that processing these cues can be greatly bolstered by experience and training with music.

Memory

Research on memory as it relates to songs and lyrical recall is unclear. Table 2 provides a brief overview of five relevant studies (two tested with L1s and three with L2s), followed by an interpretation and hypothesis. Keep in mind the theoretical lens offered above, which emphasizes that learning outcomes will stem from differences in *processing* between singing and speaking and differences in *content* between song and speech. Suggestions for future L2 learning research on memory are made based on this lens. For example, some of the outcomes reported in the table can be attributed to the differences between song and speech (such as complexity) and between singing and speaking (such as presentation rate).

Table 2

Text Recall through Songs and Speaking activities

Study	Lang. Status	Description & General Findings
Kilgour, Jackson, & Cuddy (2000)	L1	● ran experiments on the recall of sung and spoken lyrics in participants L1 ● group with sung lyrics outperformed the group with spoken lyrics when tested on their recall of the lyrics ● when presentation rate was controlled (songs were spoken and sung at the same rate)—no difference was found ● slower presentation rates showed better recall than faster ones ● musicians outperformed non-musicians
Racette & Peretz (2007)	L1	● investigated lyrical recall in participants L1 after either singing or speaking ● found no difference ● argued cognitive demand is higher while learning vocabulary and musical forms together
Li & Brand (2009)	L2	● three classes (all songs, half songs, no songs) ● learning content through a variety of activities, all delivered by the same teacher ● music class learned most, followed by half music, then no music ● it seems the content might have been unnatural for the no music group since all classes used the same materials (e.g., vocabulary lists would have been based on songs)
Salcedo (2010)	L2	● tested students listening to songs and spoken recordings of song lyrics. ● found that the group listening to songs outperformed the group with spoken recordings for 2/3 songs in the short-term, ● no long-term effect. ● testing and learning modalities were not completely consistent; students were tested through a cloze test (written) after listening to and reading lyrics in class. ● the listening to songs group also reported more mental rehearsal in a follow-up questionnaire, which could an added benefit to using songs.
Ludke et al. (2014)	L2	● addressed the limitations of the study by Racette and Peretz (2007) stating that the songs were too complex ● made changes to five variables (longer learning period, matched speaking and singing rates, higher variety of testing measures, less content to learn, and used unfamiliar language with translations rather than L1) ● found that the singing group outperformed other groups on two out of five tests (production test and delayed-recall conversation task) ● recognition, vocabulary, and English recall tests showed no differences.

The results from these studies suggest that cognitive demand on memory might be higher when one is learning both words and melodies of a song and more exposure is required to produce an effect. If the learning period is extended, there may well be an effect during the delayed-posttest. It is also possible that, in doing so, lyrical recall might be more robust as musical associations reinforce one's memory.

The theoretical lens in this paper can help shed light on these studies with regard to speech presentation rate. If a song is fast (or sped up), left-side hemispheric processing is likely to dominate and less reinforcement from right-side processing can take place; the brain focuses on meaning (Patel, 2014). However, if a song is slower, fine-grained focus from right-side processing might aid processing of prosodic cues alongside the processing of meaning (Patel, 2014). In any case, it is likely that the slower rate often presented in songs is beneficial to learning. Further research is needed to clarify these issues.

Content: Using Songs in L2 Learning

An overlap in cognitive processing is not the only reason to use songs in L2 learning, as songs also offer learners natural repetition and motivating exposure to the target language. Below, this is explored through reviewing what corpus work has revealed in terms of listeners exposure to vocabulary and grammatical forms, as well as studies on using songs in the L2 classroom.

Vocabulary

Tegge (2017) analysed the vocabulary of pop songs on the US billboard charts and found that the vocabulary required to understand English pop songs was lower than other written texts and similar to spoken corpora. Specifically, three-thousand-word families were required to understand 95% of pop songs and 6000 were needed for 98% coverage (95-98% vocabulary

comprehension is recommended for understanding written texts; Nation, 2001). Murphey (1992) referred to pop music as “musical motherese”, suggesting that the simple and repetitive content in songs scaffolds students in the way a parent repeats simple phrases to a child. Tegge’s work suggests Murphey’s proposition could be true depending on the songs selected; however, on average, these songs would require a solid lexical base to be understood. Preselecting songs according to vocabulary is certainly an option in order to match learners’ proficiency levels (Sundberg & Cardoso, 2019; Tegge, 2017).

Grammar

Sundberg and Cardoso (2016) tagged grammatical features in select French songs to demonstrate how tags can help curriculum designers, teachers, and students choose songs that align with their goals. For example, in their analysis, the song “Si” by Zaz and the song “Tout le monde en même temps” by Louis-Jean Cormier were both tagged for having higher than average occurrences of the conditional tense (si + verb in the conditional form), meaning these songs could be chosen to provide learners with natural exposure to these forms. Importantly, these analyses are not unique to songs, since any content can be analysed and curated according to learners’ proficiency levels. Outside of this article, grammatical analysis of songs is uncommon and it is generally unclear how grammar in song might differ from speech.

Songs in the L2 classroom

Research on the effects of using songs in the classroom is scarce (Engh, 2013a; Ludke et al., 2014). Engh’s (2013a) review concluded that more work needs to be done evaluating the use of songs in the classroom. Currently, teachers do use music in the classroom, but that the decision is not always grounded in research (Engh, 2013b) and teachers are not always comfortable knowing how to incorporate it (Rose, 2016). A nontrivial reason often cited for

using music is as way of delivering *fun* L2 content (Engh, 2013b). Accordingly, much of the classroom research done so far has been on motivation and demonstrates that music improves learners' motivation (Beasley & Chuang, 2008; Kara & Aksel, 2013; Mora, 2000; Rulkholm, 2011). There are also studies showing that L2 music has positive effects on pronunciation (Wilcox, 1995), grammar (Oulton, 2010), and writing fluency (Alisaari & Heikkola, 2016), though more work needs to be done in these areas.

An exciting step in the direction of supporting teachers has been led by TESOL, who released a book that includes 101 research-based ways of incorporating music into the classroom (Arnold & Herrick, 2017). The book is divided up according to skills the activities are designed to improve, including reading, writing, listening, speaking, grammar, vocabulary, cultural exploration, and integrated skills.

Overall, there is very little research on the *content* presented in songs in relation to speech. Research on music *processing* is promising but the next step of connecting the potential benefits of engaging with L2 music to concrete teaching and learning practices is in its infancy. The following section explores some of these directions.

Future Directions

Processing and L2 Research

With the review above in mind, it is possible that music and songs could shed light on issues in L2 acquisition theory. For instance, French L1 speakers have been reported to struggle with contrastive stress, leading researchers go as far as questioning if French speakers are capable of learning to perceive it (Dupoux et al., 2008). French has predetermined lexical stress and a syllable-timed rhythm, whereas English has variable lexical stress and a stress-timed rhythm, so the cue can carry more meaning in English. An interesting line of research would be

to investigate whether the additional focus on pitch required by musical training or listening to songs would influence French speakers' acquisition of these features in English, as music can shift attention from semantic representation to syntactic processes (i.e., meaning to form; Patel, 2011). It would be unreasonable to expect language learners to learn a musical instrument simultaneously to a language, but investigating how to take advantage of one's musicality or whether listening or singing activities improve musicality or are capable to shift one's focus attention are worthwhile avenues to pursue.

Another area avenue yet to be researched is treating singing along to music as a form of imitation or shadowing (see Chapter 4). L2 learning research has shown benefits from imitation and shadowing activities on both production (Foote & McDonough, 2017) and comprehension (Adank, Hagoort, & Bekkering, 2010); additionally, songs are a natural way of encouraging learners to repeat other speakers. Further, they likely encourage internal imitation even when listeners do not sing along out loud, which is an interesting consideration in light of Gambi and Pickering's (2013) Simulation Theory, which posits that speech production mechanisms can be simulated internally during listening.

Incorporating Content: Learning Beyond the Classroom

Besides research on L2 processing, it is important to consider how L2 learners can best use songs to aid learning. As mentioned in the introduction, people listen to an average of two and a half hours of music per day (IFPI, 2021). This indicates that the greatest opportunity for further integration of songs into L2 settings will likely be outside of the traditional classroom. However, that does not mean it must be limited to incidental use. Following Collins and Muñoz's (2016) recent call for research on extending the classroom in ways that allow for classrooms to be used more efficiently—since the number of hours of class time is often

insufficient—further exploration of the ways music can address this gap is needed. Of course, this is not a new idea. Murphey (1990) suggested music as an enjoyable way of practicing L2 material at home. If so, this raises the importance of determining how best to scaffold learners working on their own.

The next three sections discuss three key considerations for future research inquiring into the effects of music on language development beyond the classroom. Because engaging with music is usually an autonomous activity, the first recommendation is to build upon a framework for self-regulation. The second recommendation is to use an instructional design method for developing structured materials for autonomous learning. The third recommendation is to account for affordances provided by technology since technological devices are used to access music.

Self-Regulated Learning and Autonomy

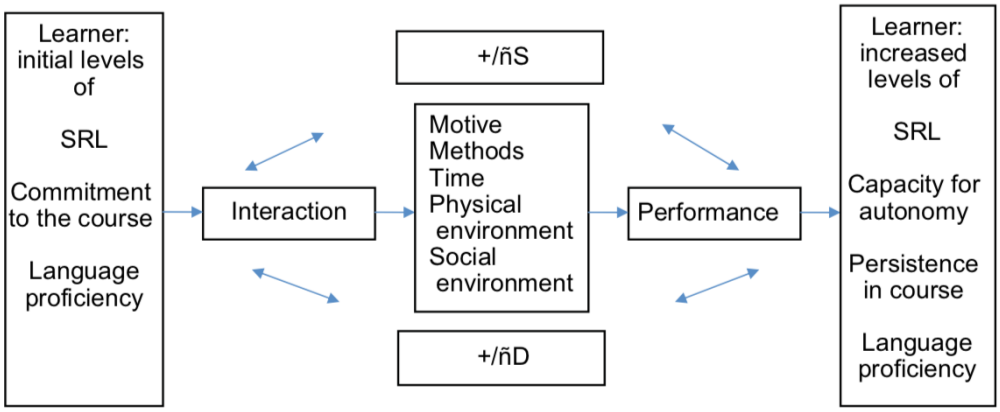
By assuming most interaction with L2 music happens outside of the classroom, teachers are faced with determining the extent to which they: 1) determine the content, and 2) need to help learners develop their autonomy. This is partly due to the enormous amount of material online, making it hard for learners to choose appropriate material themselves, and the fact that most students expect teacher involvement (Lai et al., 2016). Lai et al. (2016) recommend that teachers show students how they find appropriate material themselves and explain why they chose it rather than just presenting the content itself. Doing so is one way of bridging the classroom with learners' informal learning habits that falls in-line with Reinders (2017) Learning Beyond the Classroom (LBC) framework that recommends "encouraging LBC, developing LBC, supporting LBC, and involving LBC" (for details, see Chapter 1).

To better understand how this type of instruction helps with self-regulation and ultimately

autonomy, we can look to Andrade and Bunker's (2009) proposed model for self-regulated distance language learning (Figure 2). First, Andrade and Bunker point out six dimensions to self-regulated learning that are based on four psychological components: cognition, metacognition, motivation, and behaviour. The six dimensions are motive, method, time, physical environment, social environment, and performance (these dimensions were originally proposed by Zimmerman, 1994, p. 54). For reference, these dimensions answer the questions of why, how, when, where, with, whom, and what, with regards to self-regulation. The goal is that learners develop the ability to control these dimensions themselves by influencing each psychological component. For example, metacognitive action, like planning, can affect method of study or amount of study time. The more scaffolding required from the course and teacher ("structure" and "dialogue"), the lower one's autonomy. As learners' ability to self-regulate increases, their capacity for autonomy increases and less structure and dialogue are needed to support the learner. I highlight this model as it serves as a framework for exploring the use of self-directed pedagogical tools beyond the classroom. (See Andrade & Bunker, 2009 for specific examples of how to help learners develop control over these dimensions).

Figure 2

Model of self-regulated distance language learning; S=structure, D=dialogue, SRL=self-regulated learning (from Andrade & Bunker, 2009, p. 54)



Task-Based Language Teaching and Instructional Systems Design

In terms of providing structure, Instructional Systems Design approaches, which are presently used across numerous fields, offer a particularly appropriate roadmap for bridging out-of-class work (flipped classrooms, blended learning, and distance learning settings) and traditional course designs. Though several frameworks have potential, I specifically recommend the ADDIE framework (analysis, design, development, implementation, and evaluation; as outlined by Carliner, 2015) because its parallel structure is suitable to Task-Based Language Teaching curricula development (see Long, 2014 for comparison). Future research using ADDIE to design pedagogical materials incorporating music to learn outside of the classroom is a compelling avenue to pursue.

Affordances of Computer-Based Technology

Future research on music will likely involve the use of technology, particularly mobile technology, so opportunities to take advantage of that should be encouraged. Technology is often used to deliver online homework in Massive Open Online Courses (MOOCs) and distance

learning settings to provide curated content and interaction. With songs, technology can be used to control speed, vocabulary exposure (through corpus analysis), and to highlight less salient or redundant forms as well as host and deliver materials across environments (Sundberg & Cardoso, 2019). The benefits of textually enhancing L2 language and features are well documented (for a few examples see Chapelle, 2003; Yanguas, 2009; 2012; Yoshii, 2006); in addition, these features might even address some of the limitations encountered in songs (e.g., problems with intelligibility). Indeed, most music apps now incorporate access to lyrics automatically now. A number of mobile applications have also been released to encourage learners to study on their own, including apps developed by Turnbull et al. (2017) and Sundberg and Cardoso (2019). Pairing this type of out-of-class content with well-structured curricula can encourage learners to initiate their own non-formal learning opportunities (see Chapter 1 for definition) and progress their skills autonomously beyond the classroom.

Conclusion

Music and language share similar prosodic attributes. In order to understand how songs might aid L2 learners, I have reviewed the similarities and differences in how music and song is *processed* compared to language, as well as differences in the *content* provided by songs compared to speech. For example, F0, phoneme duration, and volume are the main acoustic differences between singing and speaking, providing a slightly different stimulus. A concrete way to think about this is to consider vowel lengthening in songs and how this might be an opportunity to focus on vowel quality. Next, language and music likely share a neural syntactic network (Patel, 2003) but do not share access to representational forms. L2 research on the use of songs should take this into consideration, such as using songs in form-focused activities for prosodic development or speech rate priming. In terms of *content*, songs offer motivating

material with natural repetition, which is important to task design (Gatbonton & Segalowitz, 2005).

Empirical evidence shows that these differences between music and speech do lead to transfer effects with musicians often outperforming non-musicians on linguistic measures. L2 research shows similar potential, such as demonstrating how pitch might aid the segmentation of word boundaries (Schön et al., 2008) and offering more redundant cues that can help in noisy environments. However, it is clear that more research needs to be done with songs and L2 learning. From a pedagogical perspective, songs can easily be incorporated into activities that take place outside of the traditional classroom, providing more hours of practice with target language content in an engaging and learner-centred manner.

Chapter 3: The Feasibility of Using Bande à Part to Aid French Language Learners

The *Bande à Part* application (app; Sundberg & Cardoso, 2019) was developed as a pedagogical tool to explore how French learners can increase their exposure to French (and consequently promote vocabulary, grammar, and pronunciation practice) outside of the classroom. For background, the app uses computer-assisted language learning principles to combine pedagogical objectives with technological affordances. It allows users to select songs according to vocabulary difficulty (a metric based on frequency bands), prominent grammatical features (e.g., verbal inflection), music genre, gender of singer, and country of origin (Figure 3). After choosing a song, users can then interact with the song in ways to help them notice linguistic features of French (i.e., they can change speed, repeat phrases, turn on a translation—for details about the app and its development, see Sundberg & Cardoso, 2019). The next steps of this project include testing the app with users, revising it according to feedback, and measuring users' learning gains in order to contribute to research on the use of songs in L2 learning settings. Specific research targets are phonological processing and pronunciation instruction, instructional design, and self-regulation.

Figure 3

Bande à Part Song Selection

Search:

Song	Artist	Country Of Origin	Gender of Singer	Genre	Vocabulary Difficulty	Grammar
Est-que tu le sais?	Les Chats Sauvages	France	Male	Retro	Easy	1st/2nd person pronouns, Verbs in 1st/2nd person, Clitic pronouns
Je T'aime	Françoise Hardy	France	Female	Retro	Easy	
Je Le Savais	Julie Doiron	Canada	Female	Indie	Easy	Adverbs, 1st/2nd person pronouns, Verbs in 1st/2nd person
Si Tu Es Un Homme	Alizee	France	Female	Pop	Medium	1st/2nd person pronouns

Context and Rationale: Extending the Classroom with Songs

Songs were chosen as content for the app to address the lack of research on the pedagogical use of songs in second language (L2) settings (Engh, 2013a; Werner, 2020) and to address the recent call for research examining how learners can effectively increase their hours of target language practice in ways that relieve classrooms of the pressure of providing learners with all their target language exposure (Collins & Muñoz, 2016).

With respect to the evidence available on incorporating songs into L2 settings, theoretical and empirical work suggests songs could have several benefits. For instance, listening to songs offers learners enjoyable target language content with naturally occurring repetition. These benefits address a concern raised by Gatbonton and Segalowtiz (2005), who highlighted the lack of repetition in many task-based learning settings. Additionally, research into music and language processing has demonstrated that musical processing shifts one's attention to prosodic elements—e.g., pitch, volume, and phoneme duration (Patel, 2011)—common components of stress and rhythm (Celce-Murcia et al, 2010). This shift in attention is hypothesized to be partly due to a slower presentation rate in music, allowing for more focus on form, and partly to an overlap in the syntactic networks that sequence features of language and music (Patel, 2003; 2011).

The hypothesized overlap in neural networks impacts language learning because some features of music and language processing are shared and, thus, musical processing may affect linguistic processing and vice versa. In L1 and L2 research, musical training and pitch discrimination have been tied to learners' facility to process language in several areas including speech segmentation (e.g., in L1 research, François et al., 2012; in L2 research, Schön et al., 2008) and the encoding of pitch (Besson et al., 2011; Kraus & Chandrasekaran, 2010). L2

research has begun to highlight how songs can contribute to L2 development (e.g., vocabulary, Ludke et al., 2014; pronunciation, Wilcox, 1995); however, since songs involve both musical and linguistic processing, it is not always clear which domain is affecting the learning gains and the extent to which music can aid learning. Future research investigating the connections between music and second language development has the opportunity to deepen this insight and aid L2 learners through isolating these variables. The *Bande à Part* app was developed with the goal of pedagogically implementing these connections between music and L2 development, under the assumption that the appropriate use of technology can improve out-of-class learning outcomes (e.g., Barcomb & Cardoso, 2020).

As such, *Bande à Part* was developed as a database that is easily accessible anytime, anywhere. In addition, it uses music as a pedagogical target. Music makes for ideal content because, on average, 15-18-year-olds already listen to three hours of music per day (Rideout, Foehr, & Roberts, 2010), so turning this period of time into a learning opportunity does not require learners to schedule more studying time into the day.

To confirm whether the proposed app meets the acceptance levels required to be pedagogically appropriate and usable, this pilot study reports findings from evaluating users' perceptions of the app, following insights from SLA research but within the auspices of the framework adopted, the TAM. Testing perceptions of new material is in keeping with Task-Based Language Learning practices (Long, 2014), a design-based research paradigm (Reeves, Herrington, & Oliver, 2005), and research into students' acceptance and integration of new technology in the learning environment (Davis, 1989). For those unfamiliar with design-based research, it is primarily characterized by a long-term iterative design process that involves the integration of "design principles with technological affordances" (Reeves et al., 2005, p. 105). In

this study, the quantitative and qualitative feedback were collected from a survey about the app and its features in order to evaluate its perceived usefulness, ease of use, and enjoyment, which are the constructs adopted to predict users' intentions to use the app according to our adapted version of the main technology acceptance model (Davis, 1989).

Background

Technology Acceptance Model

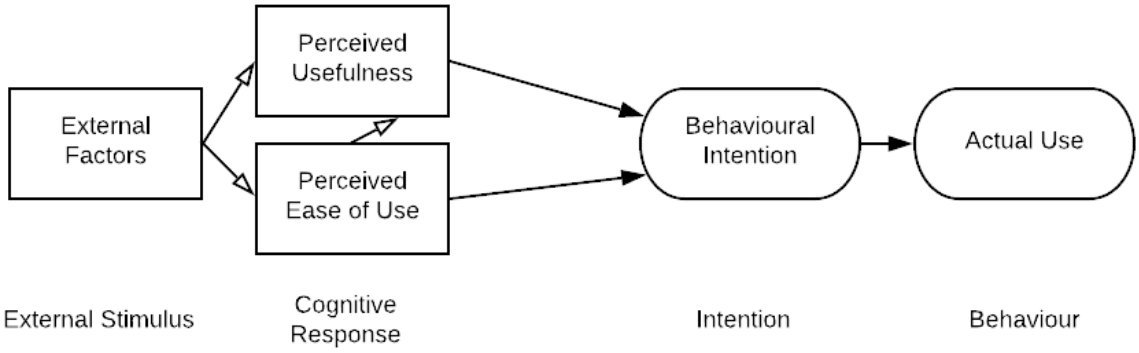
The technology acceptance model (TAM) has been adopted as a framework in order to explore the feasibility of using *Bande à Part* as a French learning tool. The TAM proposes perceived ease of use and perceived usefulness as the primary constructs to predict the integration of new technology into the lives of potential users. The model has been shown to be robust across organizational and e-learning settings (e.g., King & He, 2006; Farahat, 2012; Venkatesh & Davis, 2000) and Hsu and Chang (2013) have suggested that future mobile-assisted language learning studies adopt the TAM in order to evaluate learners' perceptions of L2 learning tools. Hsu and Chang's suggestion comes from the fact that the model is relatively new to computer-assisted language learning contexts (examples of L2 studies that use the TAM include Hsu, 2015, 2016; Hsu & Chang, 2013; Lin, 2014, Yih & Nah, 2009; and Tsai, 2015). Ultimately, insight from this model should lead to an understanding of: 1) whether learners intend to use the app, and 2) how the perceived usefulness and perceived ease of use of the app affects learners' intentions to use the app to learn French. Both are factors that will influence future studies and iterations of the app.

Davis (1989) proposed the TAM (Figure 4) to investigate the acceptance of new technology in organizations. He based the model on the theory of reasoned action (TRA) which, simply stated, argues that most behaviours rationally extend from attitudes and intentions (Davis,

1989; inspiration was also taken from the theory of planned behaviour, which elaborates on the TRA). Therefore, Davis (1989) argued that the actual use of a tool is based on previous use and one's intentions that derive from that use.

Figure 4

Technology Acceptance Model. Adapted from Davis and Venkatesh (1996).



The model has been tested across several contexts and its two main constructs, perceived usefulness and perceived ease of use, are consistently reliable predictors of users' intentions (Lee et al., 2003). The model with these two constructs will be referred to as the "main TAM"—since so many adaptations co-exist. Reviews of the TAM have found that the model explains about 40% of the variance between behavioural intention and actual usage (Venkatesh & Davis, 2000). To demonstrate that the model is applicable to L2 learning contexts, Hsu (2016) validated the strength of the relationship between users' behavioral intentions and actual usage (β=.79, $p < .001$) in a study using an ASR-based computer-assisted pronunciation training program for learners of English as a foreign language. These reliable and transferable results have led to the TAM becoming the most commonly used model for assessing users' perceptions of technology (Venkatesh, 2000). Compared to other models, such as the Technology Readiness Index (Parasuraman, 2000), the constructs are easily operationalized (Porter & Donthu, 2006).

As the TAM framework has expanded to new contexts, it has also been elaborated on. Venkatesh and Davis (1996) remark that to design usable systems, determinants of perceived usefulness and perceived ease of use must be investigated. These determinants will be discussed briefly as they relate to the current study.

Determinants of Perceived Usefulness

Venkatesh and Davis (2000) proposed the TAM2, which examined determinants of perceived usefulness in voluntary and mandatory settings across four organizations. The five main determinants investigated were: (1) subjective norm, defined as how people in the user's social circle think the user should use the system; (2) image, defined as how system use is perceived to affect social status; (3) job relevance, defined as perceived relevance to helping one achieve goals or complete tasks; (4) output quality, defined as how well the system is perceived to carry out expected tasks; and (5) result demonstrability, defined as how easy users perceive linking their gains to the system.

According to Venkatesh and Davis (2000), subjective norm only has impact in settings where the use of a system is mandatory. Since the current study evaluates voluntary individual usage, the social factors, subjective norm and image, are not assessed. Also, the duration of the pilot does not likely allow users enough time to link gains to the system, which could limit perceived usefulness. The survey questions in the current study (see forthcoming discussion) inquire about output quality, relevance, and the degree to which users believe songs can help them learn a language.

Determinants of Perceived Ease of Use

Venkatesh and Davis (1996) examined determinants of perceived ease of use and found that it is primarily determined by self-efficacy with computers, followed by more objective

measures of usability. In this study, the survey questions ask about the ease of using specific features (e.g., location of translations on screen), if the program is intuitive, and whether users have used apps for learning in the past.

Additional Constructs and Determinants

In research that uses the TAM, additional variables are often added according to context. Marangunić and Granić (2015) note that the TAM mainly focuses on extrinsic factors because these factors specifically focus on the system. However, intrinsic factors can inform researchers about the target population. Two studies that include the intrinsic factor, perceived enjoyment, are worth considering because they are similar to the current study. When researching the acceptance of an online learning system at a University in China et al. (2008) found enjoyment to be an intrinsic factor that affected behavioural intention to a higher degree than perceived ease of use, but lower than perceived usefulness. In a mobile learning setting using an extended TAM framework, Huang et al. (2007) found that perceived enjoyment moderated attitude (the antecedent to behavioral intention in their model) and perceived ease of use. Taking this into consideration, the survey in the current study asks users questions about whether they liked the music in the program and whether it had enough content. Perceived Enjoyment was added as a construct predicted to influence users' intentions to use the app (see forthcoming Figure 5 for an illustration).

In addition to perceived enjoyment, specific L2 learning determinants of perceived usefulness and perceived ease of use are worth considering. Motivation has been identified as a major predictor of success in L2 research (e.g., Dörnyei, 2009) and has been identified as a possible determinant of perceived ease of use (Hwang & Yi, 2002). Hwang and Yi (2002) conceptualized motivation as a combination of enjoyment, learning goal orientation, and self-

efficacy. They found learning goal orientations to influence self-efficacy, a prerequisite to perceived ease of use. To capture motivation in the current study, the survey asks users about why they are learning French (conceptualized similarly to learning goal orientation in Hwang and Yi's study). Two hypotheses were entertained to explore how motivation might affect the two main TAM constructs:

H_1. Intrinsically motivated users will find the app *more useful* than extrinsically motivated ones.

H_2. Intrinsically motivated users will find the app *easier* to use than extrinsically motivated ones.

Last, in a recent literature review on the TAM, Marangunić and Granić (2015) found that it was common for researchers to test whether participants' demographic information moderated users' intentions to use a new system (e.g., Porter & Donthu, 2006). Because *Bande à Part* was intended for high-beginner to low-intermediate users, we examined whether one's experience with French predicts one's interest in using the app. For instance, if beginner learners find the app less useful, it might be because the content is beyond their abilities. This could help verify if the app is indeed targeted to these users. Finally, age was included to test whether it is associated with perceived ease of use. Two hypotheses were formed to explore these demographic factors.

H_3. There is a negative relationship between age and perceived ease of use.

H_4. Users with less than one year and greater than 10 years of experience in French will not find the app useful.

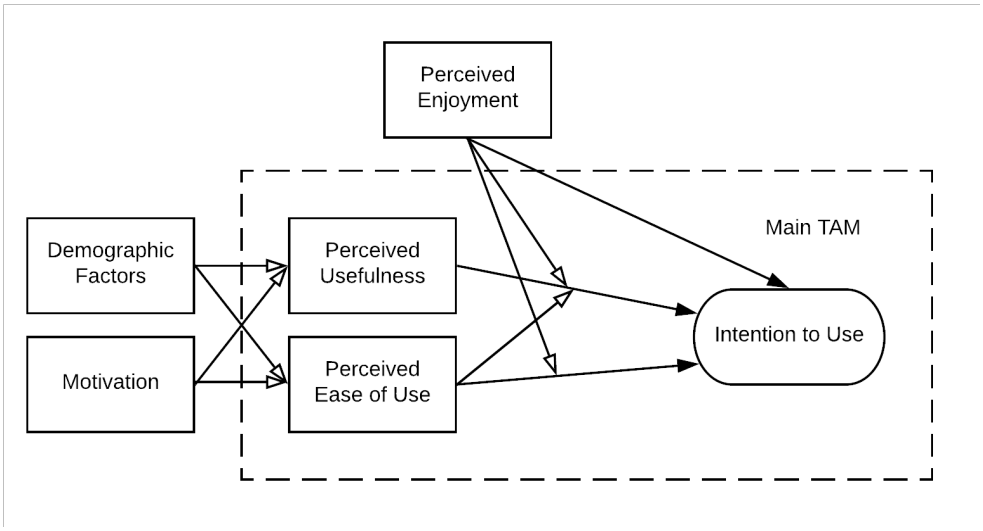
Current Study

The current study assesses the *Bande à Part* app through lens of the TAM framework; it includes additional demographic factors (age, and years of experience learning French) and

motivation as determinants of the two main predictors. Perceived enjoyment is added as a main predictor alongside perceived usefulness and perceived ease of use (see Figure 5 for an illustration of the adapted version of TAM utilized in this study). Actual use was not measured and "behavioural intention" has been termed "intention to use" in this study.

Figure 5

Proposed Technology Acceptance Model for Current Study



The following research questions were developed in order to collect users' perceptions of the *Bande à Part* app:

RQ1. Do users want to continue using the app?

RQ2. What do participants think about the app's features?

RQ3. Do the hypothesized constructs (perceived usefulness, perceived ease of use, and perceived enjoyment) predict if users want to continue using the app?

RQ4. Which users want to use the app (based on age, years of experience, and motivation)?

Method

Participants

Thirteen participants took part in the study: nine female and three male. The participants come from a variety of backgrounds (Canada and abroad) but they all live in Canada. Twelve currently live in Montréal and one lives in Alberta. Voluntary response sampling was used so as not to limit any particular group and only include participants who wanted to use the app voluntarily. Participants with diverse experiences and proficiencies were sought in order to collect a wide range of feedback. This decision does not necessarily mean that the tool is targeted towards all demographics, but that the study can help evaluate which learners are most interested in using the app.

Participants were recruited via social media and include a wide range of experience with French (0 to 10+ years). For some, French is a second language and for others it is a third or fourth. The average age of the participants was 33.75 years-old (range 25-44). Participants were not paid for their involvement.

Materials

The primary tool was the *Bande à Part* webapp, which was accessed by participants via a computer browser or a mobile device (iOS or android). A 39-question survey was distributed online. It was used to collect background information on the users and included a 6-point Likert scale (Disagree=1; Agree=6) to collect users' perceptions of the app. There were 18 questions covering the three constructs (11 for perceived usefulness, four for perceived ease of use, and three for enjoyment), 12 questions about the participants' experience (e.g., "I think this program helped my pronunciation improve", "The program is intuitive"), and three open-ended questions to collect comments about the app (i.e., "What would you add to the program?", "What would

you delete from the program?", and "Other comments?".). Six additional questions about participants use of the app were included to help explain possible findings (e.g., asking about the device used, iOS, android, desktop, laptop).

App Description

The *Bande à Part* app can be accessed via a computer or mobile device. Once a song is selected for playback, users are taken to a screen where they can interact with the song as they listen (see the app's interface in Figure 6). Users can change the speed of the music, add English subtitles, see linguistic features highlighted such as liaison (i.e., the resyllabification of "t" and "ou" [tu] in *tout oublié* 'all forgotten'), and navigate forwards or backwards through the song at the phrase level. This navigation allows users to repeat phrases over and over until they have parsed and understood them. In general, features of the app are intended to encourage learners to notice features of French by increasing the saliency of L2 forms and scaffold learners through repetition. See Sundberg and Cardoso (2019) for details about the app and its development.

Figure 6

Screenshot of the Bande à Part Application

Procedure

A recruitment message was posted on social media in a number of groups. Interested participants were emailed instructions to access the app. No minimum number of hours of use with the app were required by participants in order to qualify, since natural levels of motivation are required. After about three weeks, participants were provided access to the online survey and asked to fill it out.

Prior to data collection, the survey was piloted, pretested and discussed among known colleagues. The same instructions were provided to these colleagues as the participants in this study.

Data Analysis

As recommended by Ruel et al. (2015, p. 116), coding, preliminary analysis, and tables of results are included even though the pilot has a small sample size and, consequently, it is underpowered. As such, the interpretation of the results from the statistical tests is limited and primarily serve to test the process and design through to the end stages. In addition to reporting descriptives and frequencies, the study uses a logistic regression to construct a model to determine which variables and to what degree each of the variables contribute to predicting users' interest in continuing to use the app. The hypotheses were tested with a series of chi-squared and correlation tests to examine associations between variables. To determine the threshold for an acceptable sample size, an odds ratio of 5.28 (from Gagnon et al., 2012) and an R^2 of .55 (from Venkatesh & Davis, 2000) were plugged into the G*Power software. It was found that a minimum sample size of 70 is needed to establish a power of 0.8. Additionally, as a rule of thumb, at least 10 responses per predictor is recommended for logistic regressions (van

Belle, 2011). Planning the recruitment of participants for the future study will take this into consideration.

Coding

First, the survey questions were coded according to the following independent variable constructs: perceived usefulness, perceived ease of use, and perceived enjoyment. Eleven Likert-scale items were coded under the construct perceived usefulness, four items under the perceived ease of use, and three under perceived enjoyment. Motivation was coded as a binary variable: intrinsic if the participant included at least one intrinsic reason for learning French, or extrinsic if no intrinsic reasons were listed (meaning the intrinsic category includes participants that provided examples of both types of motivation). For example, learning French for pleasure or travel was coded as intrinsic, whereas professional and immigration purposes were coded as extrinsic. The variables were coded this way because hypotheses 1 and 2 test whether intrinsic motivation affects a user's intention to use the app. The dependent variable intention to use was measured by a single item in the survey: "I would like to continue using this app."

Results

Table 3 shows the frequencies for the outcome variable for the main portion of the TAM, captured by the question, "I would like to continue using this app". The mean is 4.77 on a possible 6-point Likert-scale (n = 13, SD = 1.536). The participants' reported frequency of use is in Table 4, with most users using the app a few times or a few times per week.

Table 3

Descriptive statistics for the participants' intention to continue using the app (n=13)

	Intention to Use	N
Agree	1	1
↑	2	0
	3	2
	4	0
	5	5
↓		
Disagree	6	5

Table 4

Descriptive statistics for actual reported frequency of use (n=13)

Reported Frequency of Use	N
A few times	7
Once per week	1
A few times per week	5
Once per day	0
More than once per day	0

As seen in Table 3, the full spectrum of the Likert-scale (1-6) was not used; this means the contingency table would be left with zeros and, consequently, a multinomial logistic regression cannot be run. Instead, the outcome variable was dichotomized (Likert-scale item responses 5 and 6 were recoded as "agree" while responses 1-4 were recoded as "disagree"). This results in 10 of 13 participants indicating that they agree with the statement "I would like to continue using the app".

Internal Reliability of Constructs

In order to evaluate whether the survey questions appropriately captured the constructs in the TAM, Cronbach's α was run to determine internal reliability of each of the aggregated constructs. For perceived usefulness, Cronbach's α was .893, suggesting the questions accurately capture the constructs; however, there were five inter-item correlations higher than .7, suggesting that these questions are not independent. Three questions were about perceptions on learning vocabulary through the application, so it is unsurprising they are highly correlated. After reducing the questions about vocabulary down to one item, the new Cronbach's α was .838. For the purpose of the study, the other correlated items were left in since it was decided there are not enough participants to justify making further changes to the construct. A future study will take this into consideration.

For perceived ease of use, Cronbach's α was .642, indicating the questions that represent the construct should be more robust. There were no inter-item correlations above .7 but deleting one question would increase Cronbach's α to .737, which is acceptable (Kline, 1999). For perceived enjoyment, Cronbach's α was .648. There were no inter-item correlations higher than .7 and the deletion of one question increases Cronbach's α to .700, also acceptable. If similar results occur in the future study, the items will be deleted to improve Cronbach's α. Overall, these results suggest that these constructs will be reliable.

Descriptive Statistics

The descriptive statistics for the main predictors (Table 5) show the following means for perceived usefulness (M = 4.27), perceived ease of use (M = 3.88), and perceived enjoyment (M = 3.95). Descriptives and frequencies for the determinants are illustrated in Tables 5, 6, and 7.

Table 5

Descriptive statistics for perceived ease of use, perceived usefulness, perceived enjoyment, and age

Constructs	M	SD	95% CI
Perceived Usefulness	4.27	.92	[3.71, 4.82]
Perceived Ease of Use	3.88	1.2	[3.16, 4.61]
Perceived Enjoyment	3.95	.93	[3.39, 4.51]
Age	35.54	8.62	[30.33, 40.75]

Note. N=13; Likert-scale items: 1=Disagree, 6=Agree

Table 6

Descriptive statistics for type of motivation and years of French experience (n=13)

Motivation	N
Intrinsic	3
Extrinsic	4
Both	6

Table 7

Descriptive statistics for years of French experience (n=13)

Years of French	N
Less than 1 year	3
1-2 years	1
2-4 years	3
5-10 years	4
10+ years	2

Collinearity diagnostics indicate that there is no multicollinearity between predictors (VIF < 10, tolerance > .1, and variance proportions do not indicate dependency between variables). The three independent variables were found to be linearly related to the logit of the dependent variable according to the Box-Tidwell (1962) procedure (a Bonferroni correction with all seven variables in the model resulted in a new α of .00714). The casewise diagnostics showed no standardized residuals > ±2.5 standard deviations.

Previous research has demonstrated that perceived usefulness is the primary predictor (Venkatesh & Davis, 2000), so a hierarchical binomial logistic regression was performed (n=13) to test whether perceived usefulness, perceived ease of use, and perceived enjoyment would predict interest in the application (Table 9). Block one included perceived usefulness, block two included perceived ease of use, and block three included perceived enjoyment as well as the interaction variables, perceived usefulness*perceived ease of use, perceived usefulness*perceived enjoyment, and perceived ease of use*perceived enjoyment. The regression model was statistically significant when including the perceived usefulness variable, $\chi^2(1) = 8.69$, $p = .003$, but the model fit deteriorated as additional predictors were added (i.e., including perceived ease of use was still significant but there was no improvement in the R^2). Thus, neither perceived ease of use nor perceived enjoyment indicated whether users were more likely to be interested in continuing to use the app. The Hosmer and Lemeshow goodness of fit test was not significant ($p > .05$), meaning the model is not a poor fit. The model explained 73.8% (Nagelkerke R^2) of the variance in participants' interest in continuing to use the app. The model also correctly classified 84.6% of the cases, compared to the base model which correctly classified 76.9% of cases (Table 8).

The classification table summarized in Table 8 shows that the percentage accuracy in classification is 84.6%, while the sensitivity and specificity are 90% and 66.7%, respectively. The positive predicted value is 90% and the negative predicted value is 66.7%.

Table 8

Classification table for predicted vs. observed cases for intention to use

		Predicted		
		Intention to Use		Percentage Correct
Observed		No	Yes	
Intention to Use	No	2	1	66.7
	Yes	1	9	90.0
Overall Percentage				84.6

Note. *Cut value is .500*

The area under the ROC curve, which plots sensitivity vs 1-specificity, represents the ability for the model to discriminate between participants who would like to continue using the app and those who do not. The area was calculated to be .967 (95% CI, .871 to 1.00), which is outstanding according to Hosmer et al. (2013).

Although, the model was significant, the Wald test was insignificant, which means the coefficient is not trustworthy and odds ratio does not indicate the true the likelihood of increasing intention to use according to perceived usefulness scores. See Table 9 below for a summary of the results.

Table 9

Model information for predicting intention to use.

	b (SE)	95% CI for Odds Ratio Lower	Odds Ratio	Upper
Perceived Usefulness	6.110 (4.284)	.102	450.492	199
Constant	-23.444 (16.899)			

Note. R^2 = *.488 (Cox & Snell), .738 (Nagelkerke). Model* $\chi^2(1)=8.690$, * $p < .05$, ** $p < .01$.

To test whether the hypothesized determinants (motivation, age, and French experience) affect the main predictors in the model (H_1, H_2, H_3, and H_4), a series of chi-squared tests were run. Because the study has a small sample size and categories in the contingency tables often have fewer than five values, the variables were dichotomized, and Fisher's exact test was used on categorical variables. The results were as follows: motivation and perceived usefulness ($p > .05$), motivation and perceived ease of use ($p > .05$), and years of experience and perceived usefulness ($p > .05$). A Pearson correlation was run on age and perceived ease of use ($p > .05$). No variable appeared to be significantly associated with another, meaning for H_1, H_2, H_3, and H_4, the null hypotheses, fails to be rejected.

Qualitative Data

The qualitative data were collected from three open-ended survey questions and coded according to the constructs in the TAM. The first question was, "What would you add to the program?" The second was, "What would you delete from the program?" The last question asked for additional comments. Most of the comments collected were about perceived ease of use closely followed by perceived usefulness. A summary of the number of comments about each construct is provided in Table 10.

Table 10

Number of comments collected about perceived usefulness, perceived ease of use, perceived enjoyment, and intention to use.

Construct	# of comments
Perceived Usefulness	15
Perceived Ease of Use	16
Perceived Enjoyment	6
Intention to use	2

Comments were coded as referring to usefulness if participants suggested adding or changing features related to the language learning process, to ease of use if they referred to the ease of using the app (including misunderstandings about how features are used because misunderstandings suggest a lack of intuitiveness), to enjoyment if they implied enjoying the app, and to one's intention to use the app whether they implied or explicitly stated their intentions. Of note, two participants recommended adding more songs; informally (outside of this research project), the inclusion of more songs has also been requested. These comments have not been coded, but they were recorded for future iterations of the project. Examples of abridged comments (Table 11-14) include:

Table 11

Theme 1: Perceived usefulness

"It would also be nice if there was a document explaining some structure. For instance, in Touts les mêmes (sic) a lyric is << y en a marre >> what are these pronouns replacing?"
"A way to save vocabulary words or expressions; for example, you click on them and it creates a list for you"
"…and it had everything I needed"

Table 12

Theme 2: Perceived ease of use

"return to menu button" [this user wanted a button to leave the current song and return to the song selection page. There is a "return to menu" button, but it is not straightforward enough for this user].
"…it would be nice if there was an app that did not require constant internet access. If this feature existed, I would use it more frequently."
"In Android, it does not work as an app…"
"I wasn't able to download the app but I know that if that having it in my phone would've allowed me to use it more often."
"I was unable to use it on my iPhone (4). :(…"
"Nothing really. I liked that it was easy to use…"
"As I mentioned, in iOS is difficult or hard to enjoy."

Table 13

Theme 3: Perceived Enjoyment

"Great app, great initiative! I would love to have it!"
"Jean Leloup!"
"I really like it."

Table 14

Theme 4: Intention to Use

"I hope this app will be available in the future (open to the public)."
"I would have used the app more if it wasn't for finals!"

Discussion

The results relating to the main TAM suggest agreement between quantitative and qualitative data. Below they are discussed as they relate to the research questions and hypotheses. As mentioned in the methods section, interpreting results from the statistical tests should be limited due to the small sample size (n=13) and to the fact that there is a high

probability of a type II error. However, possible improvements to the app can be inferred after having analysed the collected data.

In response to RQ_1, "do users want to continue using the app?", the quantitative data suggests that 10 of the 13 participants would clearly like to continue using the app. The two survey comments related to users' intention to use the app also support this (e.g., "I hope this app will be available in the future..."; see theme 4 in the results section).

In response to RQ_2, "What do participants think about the app's features?", the results suggest that the users generally find the app useful, not overly difficult, and enjoyable. The mean scores for perceived usefulness were highest (4.27/6), while the perceived ease of use was slightly lower (3.88/6), and perceived enjoyment was in the middle (3.96/6). Although the regression suggests usefulness is the main predictor of users' behaviour intention (consistent with other TAM research; Venkatesh & Davis, 2000), the qualitative data suggest that perceived ease of use is still an important factor. These comments also point to a few things that can be improved. First, several participants indicated having trouble accessing the mobile version of the app. This needs to be addressed before conducting any subsequent studies by ensuring that the instructions are clear and we follow-up with participants early on to ensure there are no problems accessing the app. Second, no participants suggested deleting features, but almost all participants suggested adding features (e.g., highlighting more grammar features, increasing the number of songs, incorporating a one-click option to toggle translations). These features will be considered as long as they are in line with the learning theories and practices used during development (see Sundberg & Cardoso, 2019 for details about the development of *Bande à Part*).

In response to RQ_3, "Do the hypothesized constructs predict if users want to continue using the app?", the regression model that fit best only included the perceived usefulness

construct. Not much can be said at this time about the model, since the study was underpowered due to a small sample size. Determining whether the proposed TAM model appropriately predicts which factors influence whether users would like to continue using the app will have to wait for more participants. However, responding to this question in the future will inform the developers of the degree to which usefulness, ease of use, and enjoyment contribute to users' intentions. For instance, if usefulness continues to be the primary predictor and users suggest improving usefulness, then support can be provided to demonstrate to users how the app is able to aid learning. Measures of actual usefulness will have to wait until experimental studies using the app assess usefulness appropriately. Likewise, if ease of use significantly predicts users' intentions, more comments from users could help in addressing problematic areas, such as those identified in this pilot (e.g., the design of a more intuitive interface, adding online instructions to demonstrate how the app works, fixing bugs with device compatibility)

In response to RQ_4, "Which users want to use the app?", the results are still unknown. As all of the null hypotheses failed to be rejected, no associations between determinants (years of experience with French, type of motivation, or age) were found with the main predictors. This is not an uncommon finding in TAM research and, from our review, explains why only two constructs, perceived usefulness and perceived ease of use, continue to form the main TAM model. If this changes in the larger study, it will be easier to probe whether or not the app appropriately appeals to high-beginner to high-intermediate learners, whether intrinsic motivation is important to continuing to use the app, or whether it appeals to learners of certain ages.

In hindsight, the survey questions should be improved to capture these determinants. Years of experience could be aggregated with a question asking about users' reported

proficiency in French, as years of experience is likely not comprehensive enough to capture learners' experience in French. The categorical nature of the years of experience question is also unnecessary. The cut-off points were arbitrary (1-2 years, 2-4 years, etc.) and more precision could be included if participants entered their exact number of years of experience.

In addition to coding type of motivation, questions about degree of motivation should be included. These Likert scale questions could be borrowed from Dörnyei's L2 motivational self-system (e.g., "Studying [French] is important to me because it is necessary for promotion in the future"; Dörnyei & Taguchi, 2010). A question could also be added about self-efficacy with these types of systems. Originally, the question asking about previous app use was meant to assess whether previous experience with learning apps increased perceived ease of use, but this question does not necessarily imply self-efficacy. Upon further research, it seems clear that asking about one's beliefs and abilities is important (Venkatesh & Davis, 1996).

Not as many participants signed up for the pilot through social media as expected. More participants outside of Montréal would be ideal so that foreign and second language contexts could be explored. In the future, recruiting participants through classrooms might provide access to a larger sample. This perception survey could also be paired with studies measuring learning gains with larger groups since the pilot has demonstrated that the app works and participants are interested in using the app to a degree that carrying out experimental studies would likely be successful.

Conclusion

A major step in this project was piloting the *Bande à Part* app with users for the first time. Before recruiting more participants, it was necessary to ensure the app worked on a variety of devices and that it is a tool that learners would want to use. As the participants were not paid

to participate in this study, we believe it adequately captures the intrinsically motivated population that are likely to use this app in natural settings, and we trust the feedback we received despite the low number of participants. The results reported here showed that there is interest in using *Bande à Part* and some improvements can be made to improve its ease of use. However, during the pilot, the only significant predictor of intention to use was perceived usefulness and none of the determinants (motivation, age, or experience with French) seemed to impact the main TAM. With more participants this can be determined with more certainty.

As it stands, the app could be used to support classroom learning, or used separately as an autonomous tool that learners could use on their own time. No technical help or training was provided to the participants, so it is likely that learners would be left to self-regulate their own learning process. The next step is incorporating some improvements to the app and then collecting more feedback on *Bande à Part* in a larger study along with testing to see what users learn through the app. The first iteration of the app includes working on a solution to allow users to add songs to the database and automating the addition of lyrics and translations. Hyperlinks to websites with explanations of features will also be added, as this is a relatively easy addition to implement. Lastly, having user interface and user design experts provide input on the navigation and layout of the app will help to make the app more intuitive and user friendly. The results from following studies will also inform future iterations of the app and contribute to two main areas of research: 1) the use of songs in second language learning, and 2) applications of the TAM in second language settings.

Chapter 4: Just Listen to Yourself: Self-Regulating French Pronunciation Development through Spoken and Sung Input

Learners ultimately want to use their target language in natural environments—for a job, romance, to speak to family, to integrate into a culture, or any number of reasons beyond the classroom. Designing materials to help learners transition to learning and using their target language in natural environments requires an understanding of their capabilities and limitations, as well as awareness of how this process changes overtime. Songs are one common form of input that learners access outside of formal environments; however, it is not well understood what can be learned from listening to songs sung in a target language and how songs might be a part of this self-regulatory process (Engh, 2013a; Werner, 2020).

Using think-alouds to reveal learners' awareness as they engage in activities is one way to explore self-regulatory processes, but much of the research on language learning through music that does exist focuses on vocabulary recall with English as the target language (Werner, 2020). Given the amount of music consumed by the average person, it would be especially helpful to have a better understanding of other aspects of language development that result from engaging with music, such as improvements to and awareness of pronunciation features in languages other than English.

Reports from the International Federation of the Phonographic Industry (IFPI) put music consumption in 2021 at over 2.5 hours per day worldwide (almost 4 hours per day in the USA; IFPI, 2021). In L2 research, Lindgren and Muñoz (2013) reported that music is the most common form of language input beyond the classroom, although it is possible that this research is out of date considering how social media has developed over the last decade. Nonetheless, this exposure represents a learning opportunity that is not well exploited or understood.

As such, the goal of this study is to examine how learners self-regulate their French pronunciation development from songs for beyond-the-classroom settings. The study investigates L2 self-monitoring strategies during oral imitation tasks by analysing reflections gathered from subsequent think-aloud exercises. The paper is exploratory in nature and designed to inquire about what learners notice about their own pronunciation after listening to recordings of themselves 1) imitating sung lyrics, 2) imitating spoken lyrics, and 3) reading song lyrics. These imitations and reading tasks are followed up with think-aloud activities to shed light on learners' strategies and reflections, which speak to Chapelle's (2003) and, more recently, Cardoso's (2022) interest in further research into how learners behave and negotiate for meaning in computer-mediated contexts.

From the perspective of sociocultural theory (SCT), how learners mediate their environment primarily involves understanding the "inner life of learners" (Pavlenko, 2013)—another area that needs more research (Reinders, 2020). That I am aware, the only conditions that have previously been compared are listening to another speaker and listening to a recording of oneself, and for the purpose of comparing *vocabulary* recall rather than investigating learners' pronunciation development or abilities to self-monitor. This is also without an imitation task like in the present study (Forrin & MacLeod, 2018). Research on language and music suggests that representational and syntactic processes are neurologically separate (Patel, 2003; to be explained in the *Learning Through Songs* section below), and exploration into phonological processing (syntactic) is warranted.

With regards to elicited imitation to evaluate language proficiency, there have been several important contributions (Erlam, 2006; Hamayan et al., 1977, Jessop et al., 2007; see Burger & Chretien, 2001 for examples exploring pronunciation through imitation) but none of

these include music. By including a condition of music, evaluating pronunciation development through elicited imitation, and using think-aloud strategies to explore learners' thought processes, we can take a snapshot of how learners interact and learn from music compared to other conditions that are better understood. The outcome of this study will inform research into scaffolding learners towards learning beyond the classroom through their music endeavours. The results will also speak to bridging formal and informal learning environments (for example, teachers assigning music related homework) by identifying what learners are and are not able to do on their own and the processes and behaviours that mediate these interactions.

This study is part of a larger project exploring ways of scaffolding learners target language development through songs in order to increase their capacities to take control of their learning process. To do so, an app, called *Bande à Part*, was developed to scaffold learners' interactions with lyrics that match their proficiency levels (detailed in Sundberg & Cardoso, 2019), and learners' perceptions of the app and the use of songs in language learning was investigated in a pilot study in order to determine whether participants would be willing and able to use the app for language learning, and to identify potential improvements to its features (Chapter 3). In this study, we have pulled lyrics from several songs in the app to gain understanding of how learners try to improve their pronunciation from listening to, and imitating, the songs. Having learners reflect on their recordings from these exercises prompts them to provide commentary on what they notice in their own pronunciation (i.e., when imitating sung and spoken content). The degree to which they think they can act on their intentions to make improvements is then explored through comparing learner awareness and behaviour across contexts. In the end, the findings from this study will guide future directions for tools and apps like *Bande à Part*.

Literature Review

Theoretical Framework

Listening to music is generally an informal activity where learning may be intentional or incidental. In formal environments, there have been many attempts from teachers to incorporate music into the language classroom (see Sposet, 2008 for an overview on second language acquisition and music in the classroom). However, there has been less research performed in *informal* environments, and therefore, unlike in formal environments, the benefits are less clear. The larger project within which this study is situated investigates learners' language development when interacting with music in the non-formal space, which is characterized by the use of structured materials that are designed for the purpose of learning (e.g., through the *Bande à Part* app; compare also to how a song might be used in a traditional classroom to highlight specific grammar or vocabulary items) but not as part of a credentialed process (Stickler & Emke, 2011). As such, learners must initiate their own activities in this non-formal space.

Because the goal is for learners to self-regulate in order to "carr[y] out an activity independently" (Ortega, 2009, p. 220), a sociocultural theory (SCT) perspective has been adopted to understand the development and interactions within the environment. In SCT, learners move from object-, to other-, to self-regulation as they interact with their environment (Ortega, 2009). The difference between what learners can do on their own and what they can do with assistance is the zone of proximal development (ZPD) that emerges and offers a potential for learning if scaffolding is provided by others and/or tools (Ortega, 2009, p. 224). The basis for this development depends on noticing and negotiating for meaning (Lyster, 2007). Exploring learners' behaviours and abilities when practicing a target language independently can help us identify where feedback and support is needed.

Learning Beyond the Classroom

Over the past 10 years, Benson and Reinders (e.g., 2011; Reinders, 2020; Reinders & Benson 2017) have contributed to a language learning ecology that bridges formal and informal learning—an informal environment being one where the materials and/or environment is not explicitly designed for learning. They have termed this framework learning beyond the classroom (LBC), which has the goal of enabling learners to initiate their own activities in informal environments (Reinders, 2020). We can imagine that today, the learning trajectory for L2 learners is more web-like than linear. Learners interact with many of what Reinders and Benson (2017) refer to as *nodes*, all requiring different degrees of self-regulation. The formal classroom might be one node, listening to music beyond the classroom might be another, social media interactions another. In the past, when there was more limited access to the target language, linear language development within a classroom was better able to address the needs of every student. Today, needs vary more widely. For instance, a student may have high reading ability from online exposure to the target language and poor oral production. From the SCT perspective, learners' relationships to these nodes would be loci of interaction that mediate not just linguistic development, but identity and cultural development too (Ortega, 2009).

To facilitate designing for this instructional space, Reinders (2020) provides three main axes that make up LBC. The first axis highlights a progression from formal to informal environments by encouraging LCB, preparing for LBC, supporting LBC, and involving LBC. In the larger project (Manuscript B), this progression is facilitated in a non-formal context. In terms of this study, however, the learning environment is slightly more formal because the tasks are more akin to homework assigned by a teacher and, consequently, it is not learner-initiated.

The second axis consists of four dimensions to characterize LBC: location (describes physical or digital space), formality (involving credentialing or not), pedagogy (structured instruction or not), and control (self or other). With regards to this project, learners are in a digital space, interacting with structured material that does not involve credentialing, while control would involve a mix of self-regulation and other-regulation (i.e., if learners engage the tools in the *Bande à Part* app). For this study, however, learners must entirely self-regulate because they did not use the app; instead, they followed a sequence of tasks available in the proposed activities.

The third axis covers the stages of the self-directed learning process, which orbit around one's individual reflections, motivations, and interactions. These stages are: (1) identifying learning needs, (2) setting goals, (3) planning learning, (4) selecting resources, (5) selecting strategies, (6) practice, (7) monitoring progress, and (8) assessment and revision.

These three axes are intended to increase learners' capacities to self-regulate in informal environments. This study strives to understand how structured language learning through songs can fit into this above ecology and progression, which will help us to better identify learner needs and observe how learners select strategies and practice in this environment. In the next section, I discuss affordances of using music to aid pronunciation development in order to better select strategies that incorporate music in target language content development.

Learning Through Songs

As mentioned above, most people listen to a lot of music, and it is one of the ways new learners use their target language. Research also shows that pronunciation is correlated with musicality (Milovanov, 2009; Milovanov et al., 2008; Nardo & Reiterer, 2009; Sarsarabi &

Sosouri, 2016; Slevc & Miyake, 2006), but to frame how pronunciation might be improved from listening to songs, we need to start with research outside of SLA.

Research on language and music can help us to identify two important areas: 1) similarities and differences in syntactic and representational processes, and 2) similarities and differences in the content of song and speech. First, when listening to music, there are notable acoustic differences in pitch (specifically F0), phoneme duration, and volume. Furthermore, while listening to music, one's attention often shifts from meaning to form (Patel, 2011). Together, these characteristics provide an opportunity to work on discrimination of pitch (important in segmenting words, Schön et al., 2008) or identification of vowels, for example. One potential drawback is that aesthetics can take priority over intelligibility (Gilman et al., 2010). A good example of this phenomenon is when the rapper Eminem famously argued that several words rhyme with "orange" if you make them (Genius, 2016).

Additionally, brain imaging research suggests that prosodic elements of language and music are processed on the same neural networks and that music tends to shift the listeners focus from meaning to form (Besson et al., 2007; Maess et al., 2001; Patel, 2003). This suggests that music could be particularly useful for learning prosodic features of a target language (e.g., rhythm, lexical stress patterns, pitch); in addition, there is evidence that music can prime speech rate (e.g., faster songs prime faster speech), which could be useful for fluency development (Jungers et al., 2016).

With regards to point two, similarities and differences in the content of song and speech, the content in songs provides target language exposure in a motivating way for two reasons. First, people enjoy listening to music for reasons besides language learning, making it an engaging and natural context for learning. Second, music offers a natural form of repetition that

occurs both within the song itself (e.g., via key phrases that are repeated in the chorus), and through the listener's tendency to repeatedly listen to songs they like. This repetition aligns with the principles of designing effective second language tasks, as argued by Gatbonton and Segalowitz (2005). It is not easy to develop repetitive materials that are engaging to learners, so this is an area to explore further. Music naturally fits non-formal and informal environments as well because of the ease in which learners can initiate their activities. It is also culturally significant and could play into Dörnyei's (2009) L2 motivational self-system if musical engagement is important to learners' L2 identities.

Imitation

Elicited imitation is often used to assess learners' competencies (as discussed by Bley-Vroman & Chaudron, 2013) in a reconstructive manner to correct L2 learners' grammar (Jessop et al., 2007), or as a memory aid (e.g., Forrin & MacLeod, 2018). Imitation is used in these ways because it involves both the processing system and the memory system (Bley-Vromen & Chaudron, 2013, p. 247). To place this into a broader understanding of acoustic and phonological processing beyond second language research, I will refer to these systems as the *syntactic* and *representational* systems (Patel, 2003), as mentioned above. This broader framework helps to frame the pedagogical use of imitation in the current study.

To illustrate, Forrin and MacLeod (2018) found that listening to recordings of oneself was beneficial compared to listening to recordings of someone else when it came to memory recall (primarily focusing on representational system), so it would be interesting to shed light on these comparisons with regards to the syntactic processing system and whether this is implicitly learned or there is explicit awareness—the latter being the focus of the current study.

More indirectly, Tost (2013) suggests receiving feedback from peers while reading aloud as a helpful way of improving one's pronunciation. If students can provide peer feedback, there is reason to believe they could provide feedback on their own speech. Think-aloud activities in the current study were designed to explore what kind of feedback learners are able to provide after listening to their own recordings.

Lastly, shadowing (a form of mimicry where learners mimic the speech of a target speaker as close to concurrently as possible) has been shown to help learners improve their target language comprehensibility and fluency (Foote & McDonough, 2017). A mix of past research on imitation, shadowing, and peer feedback motivate the design of the current study. More recently, Hamada and Suzuki (2024) have classified shadowing practices in ways that help delineate phonological and meaning-focused techniques (e.g., prosody shadowing vs. content shadowing, respectively). These techniques can guide recommendations for students according to their needs and help us interpret literature in the field.

Phonological Awareness

Previous research on phonological awareness can provide insight into the type of phonological phenomena one might expect learners to notice and the pronunciation features that likely develop or change alongside proficiency. In 2014, Kennedy et al. explored the relationship between learners' phonological development and awareness. This study took place as part of a 15-week listening and speaking course and found that exposure to a variety of pronunciation tasks likely helps build awareness. It also highlighted potential differences in learners' strategies. For example, their findings indicated that learners who focused on extracting meaning (e.g., they referred to the content in their activity, as opposed to referring to items to be memorized or rules) performed better with *Enchaînement* (i.e., across-word resyllabification; e.g., *il arrive /i.la.riv/,*

meaning “he arrives”), while those who focused on applying rules did better with *Liaison* (the resyllabification of a normally silent word-final consonant followed by a vowel-initial word; compare the pronunciation of the silent /t/ in *petit* /pe.ti/ “small” with *petit ami* /pe.ti.ta.mi/ “boyfriend”).

In another study, Kennedy and Blanchet (2014) investigated the French learners’ awareness of connected speech processes. They found that learners who focused on extracting meaning from speech improved their awareness of patterns more than those who focused on rehearsing knowledge about connected speech patterns. This reinforces the findings of their previous study (Kennedy et al., 2014) in that awareness is contingent on meaning, but explicit knowledge of rules can provide some aid.

A study by Inceoglu (2021) investigated connected speech processes and awareness during a 12-week long course focused on French pronunciation. Learners recorded read-aloud tasks and wrote reflective journals where they commented on their *enchaînements* and *liaisons* throughout the term. Their accuracy improved from about 60% to 85%; however, the degree to which the improvements are attributed to the explicit instruction versus the reflection is unknown, but the reflections help to explain the developmental processes and awareness of the learners. The results showed that learners’ awareness improved in relation to their abilities, but there was no direct link between specific instances in journals and recordings. The author suggested that reflections could also help instructors identify features to target with explicit pronunciation instruction, which is one reasons why reflections are included in the current study.

To understand pronunciation strategies of French learners, Woore (2010) utilised interview discussions that followed a read-aloud activity. These discussions suggested several prominent strategies. The most common was relying on declarative knowledge of symbol-to-

sound mapping in French (e.g., “oi” in “témoin” makes a /wa/ sound), supporting a previous finding by Erler (2003). This knowledge might have resulted from explicit instruction or implicit learning, and it was pointed out that participants in this study were not always correct. Another strategy used was analogy, which was used by three-quarters of participants. Analogy is when learners use knowledge of a known symbol/sound mapping and apply it to another word that has a similar pattern (e.g., “marteaux” has a silent x because “Bordeaux” does). Discussions afterwards demonstrated that learners would often vocalise the words to see if they “sounded right”, as learners had trouble judging whether they were correct. Further investigating learners’ strategies in relatively similar contexts is one of the goals of this paper.

In the current study, I predict that French L2 learners are more likely to notice segments that are different from those found in their first languages, such as /y/ or /ø/ (as in *tu* /ty/ “you”, and *ceux* /sø/ “those”, respectively), particularly if they have had explicit instruction about those phonemes. This hypothesis is based on the Speech Learning Model (Flege, 1995; for an updated version of this model, see Flege & Bohn, 2021) and my own experience as a practitioner teaching pronunciation and phonology to students and future teachers, as well as my own French language learning journey. I also predict that participants will comment on connected speech phenomena because *liaison* is often explicitly taught in French classes, as demonstrated in the above literature.

Self-Regulation

In addition to negotiating traditional eco-social spaces involved in language learning (from classrooms to coffeeshops), autonomy and self-regulation have been associated with out-of-class learning contexts that, in more recent years, are highly digital. To understand self-regulation in the current study, I follow the footsteps of Tseng et al. (2006), who propose

capturing learning strategies, motivation and behaviour, cognition, and metacognition under the umbrella of self-regulation to bring the construct in-line with Educational Psychology. This undoubtedly makes it difficult to study self-regulation as a single construct; instead, the concept of self-regulation is used to explore several other constructs that all feed into learners' capacities to control their learning processes. The current study uses adapted questions from the Tseng et al. study, which originally focused on vocabulary, to better understand learners' beliefs about self-regulatory processes with regards to French pronunciation (see Appendix A for a sample of the questionnaire that is used in this study).

Andrade and Bunker's (2009) model of self-regulation, which also considers self-regulation to be an umbrella construct, is used to provide themes in the current study because of the way in captures one's ability to self-regulate as a dynamic progression that determines the extent to which one can act autonomously. The following four psychological components make up the themes explored: cognition, metacognition, behaviour, and motivation. It is understood that learners' level of control over these components governs the degree and type of scaffolding needed to become more autonomous in their learning journey. Because the ability to self-regulate changes over time, it is likely that students at different proficiency levels, with different levels of awareness, will differ in their ability to improve their own pronunciation on their own.

Finally, in their review on self-regulation, Gao and Hu (2020) listed four areas for future directions: 1) How do students use strategies to self-regulate language learning? 2) How do students' strategies to self-regulate language learning develop? 3) What mediates students' adoption of strategies for self-regulated language learning? 4) How can we help students negotiate with contextual conditions in developing appropriate strategies for self-regulated

language learning? Although these questions do not explicitly guide the discussion, commentary on these questions is integrated throughout the discussion.

The Present Study

This study investigates how learners self-regulate their French pronunciation development when engaging with songs beyond the classroom. Using a sociocultural theory framework, it explores what learners notice about their own pronunciation after listening to recordings of themselves. Previous research suggests that music may aid pronunciation due to acoustic differences that draw attention to form, overlapping neural processing of prosodic elements in language and music, and engaging, repetitive content. Imitation tasks tap into syntactic and representational processing systems, presenting an opportunity to explore these systems. Learners' phonological awareness, use of strategies like analogy, and self-regulatory capacities mediated by proficiency likely impact what they notice; exploring these facets contribute to the research on self-regulation. The study aims to shed light on how learners can be scaffolded to learn independently through music by understanding learner strategies in terms of four psychological themes recommended by Andrade and Bunker (2009): cognition, metacognition, behaviour, and motivation. To gain insight into these self-regulatory constructs, the following research question guided this study:

1. What do learners notice in recordings of themselves when:

    a) imitating sung lyrics,

    b) imitating spoken lyrics, and

    c) reading lyrics?

Method

Participants

Twenty French learners were recruited from universities and private language institutions in Quebec (Canada), and online platforms such as Facebook. All learners were over 18 years old and spanned all proficiency levels in French due to the exploratory nature of the study. Proficiency levels were self-reported (self-reported measures positively correlate with objective measures of proficiency; for the rationale, see Kenji & D'Andrea, 1992 and Luc & Bialystok, 2013). To evaluate proficiency, participants were asked to rate their own speaking ability, listening, reading, and writing French using a Likert scale from 1 (very poor) to 7 (very good). Because this study targets speaking and listening skills, the average of those two skills were taken and can be seen in the Table 17 below. Participants also reported on the other languages they speak, and the total number of languages was tallied below. Average languages spoken: 4.05 and median was 4. Average French proficiency: 4.075/7 and median was 4, which would be a high intermediate level. It can be noted that this group of participants was highly multilingual, and the majority are living in Montreal.

Table 15

Participant Demographic Data

Participant ID	Age	City	Gender	First Language	Languages spoken (#)	French proficiency
1	42	Montreal	Male	Spanish	3	3
2	19	Brampton	Female	English	3	6.5
3	38	Montreal	Female	English	2	4
4	40	Montreal	Female	Spanish	3	4
5	32	Montreal	Female	English	3	3
6	43	Longueuil	Female	Ukrainian	5	6
7	53	Dorval	Female	English	2	4
8	30	Toronto	Female	Mandarin Chinese	4	2
9	35	Montreal	Male	Spanish	3	5.5

10	24	Montreal	Female	English	4	6
11	35	Saltspring Island	Female	French	6	7
12	32	Montreal	Female	Farsi	2	3.5
13	29	Hyderabad, India	Female	English	4	2.5
14	49	Saint Constant	Male	Portuguese	4	1
15	29	Montreal	Female	Portuguese	4	1.5
16	24	Warsaw	Male	Polish	10	4
17	30	Wrocław	Male	Polish	7	3.5
18	39	Montreal	Male	Portuguese	4	5.5
19	57	Montreal	Male	English	4	5.5
20	38	Dallas	Female	Spanish	4	3.5

Materials

This study was conducted using LimeSurvey. It included a background questionnaire (see sample questions in Appendix B) followed by a series of tasks and a self-regulation survey, which are described below. The tasks included reading or imitating lyrics from French songs, taken from the *Bande à Part* app (see sample lyrics in Appendix C), and providing think-aloud reflections. These songs are aimed at high-beginner to high-intermediate French learners as described in Sundberg and Cardoso (2019). 103 phrases from these songs were also included based on having a variety of phonological phenomena, including *liaison* (discussed earlier), the vowel /y/ (e.g., *tu* “you”), “h aspiré” (found whenever /h/ does not undergo liaison; e.g., *le havre*, pronounced [lə.avʁ], not [lavʁ]), “h muet” (when orthographic “h” is not pronounced; e.g., *l’heure*), and variable /s/ and /p/ production (e.g., “plus” or “trop”, which are variably pronounced or deleted). These phrases had a variety of lengths, averaging five words and seven syllables per phrase, which should not generally strain working memory. Recalling longer phrases for imitation might depend on how participants chunk them for subsequent recall (Ellis, 2003). The shortest phrase was “Alors dites-moi” (3 words and 4 syllables) and the longest phrase was "Et ce cri trop interne ne peut vraiment pas cesser" (9 words, 13 syllables as sung).

There were two other phrases with three words and four to five syllables and eight words with nine words and ten to twelve syllables. Spoken recordings were provided by a female L1 Québec French speaker who is bilingual in French and English since early childhood. The sung phrases were taken from singers with varying accents. Think-alouds were provided at the end of every page. There were between four and sixteen phrases per page depending on how many phrases were included for each song.

Lastly, a questionnaire adapted from Tseng et al. (2006) was included (see Appendix A) to capture learners' self-regulatory processes and beliefs. As mentioned in the literature review, Tseng et al. (2006) explored self-regulatory constructs related to vocabulary acquisition. Their survey consists of 20 Likert-scale questions, which were adapted to a pronunciation context, the focus of this study (e.g., "When I feel stressed about [~~vocabulary~~ pronunciation] learning, I know how to reduce this stress"). Tseng et al.'s survey (p. 85-86) has been validated for the following categories: Commitment Control ("which helps to preserve or increase the learners' original goal commitment"), Metacognitive control ("which involves the monitoring and controlling of concentration, and the curtailing of any unnecessary procrastination"), Satiation Control ("which helps to eliminate boredom and to add extra attraction or interest to the task"), Emotion Control ("which concerns the management of disruptive emotional states or moods, and the generation of emotions that will be conducive to implementing one's intentions"), and Environmental Control ("which helps to eliminate negative environ- mental influences and to exploit positive environmental influences by making the environment an ally in the pursuit of a difficult goal"). Henceforth, the adapted questionnaire is referred to as the "self-regulation survey."

Procedure

The study employed a one-shot, pretest-posttest design to capture changes before and after the intervention. As a pretest, all participants read and recorded four audio passages in French consisting of short excerpts from songs collected from the *Band à Part* repertoire (see sample questions in Appendix D). Learners were not informed that these phrases were lyrics from songs. This was called the "Read Block." After reading and recording an audio sample of each passage, participants listened to their recording and re-recorded with the intention of improving their pronunciation (note that this block does not include a French speaker for comparison). After listening to this second recording, they recorded a think-aloud reflecting on what they observed and changed in their pronunciation, and whether they were effective in achieving the changes they intended to.

Next, participants imitated L1 French speakers singing and speaking 103 short passages, split into a "Sing Block" and a "Speak Block." Half of the participants performed the Sing Block followed by the Speak Block, while the other half performed them in the reverse order to avoid testing effects (counterbalancing) and to shed light on the difference between the two conditions (singing vs speaking input). The phrases were split into chunks of two to 16 phrases per page (all phrases on a page belonged to the same song). Following a similar process to the Read Block, participants listened to their recordings followed by the original sung or spoken content again and then re-recorded, again attempting to improve their recordings. They then listened to all the final recordings on the current page and recorded a think-aloud response, once again discussing their observations, changes, and effectiveness. This series of blocks differentiate the role of learning from mimicking spoken and sung content from each other as well as what can be learned from listening to one's own voice without a comparison to French speakers. The hope is

to identify whether what learners notice in sung versus spoken content differs and to identify what learners notice in their own recordings.

Lastly, as a post-test, participants performed the Read Block again but with 50% of the content being new. In total, the survey and tasks took approximately one hour and a half to complete. For a visual summary of the design, including block order and activities, see Tables 15 and 16. Prior to data collection, the surveys and questionnaires were piloted, pretested and discussed among a small group of volunteers.

Table 16

Study Design: Block Order

	Group 1 (n=x)	**Group 2** (n=x)
Pretest	Read Block	Read Block
Block1	Sing Block x10	Speak Block x 10
Block 2	Speak Block x 10	Sing Block x 10
Posttest	Read Block	Read Block

Table 17

Study Design: Block activities

Block Activities	**Read Block**	**Sing Block**	**Speak Block**
1. Listen	Read & record phrase	Listen to **sung** recording	Listen to **spoken** recording
2. Record 1	N/A	Imitate (record voice)	
3. Listen	Listen to one's own recording and rate accuracy		
4. Listen 2	N/A	Listen to **sung** recording again	Listen to **spoken** recording again
5. Record 2	Rerecord imitation		
6. Listen 3	Listen to one's own recording and rate accuracy again		
Repeat steps 1-6 for each phrase on the page			
Think-aloud			

Data Analysis

The think-aloud task was transcribed and coded according to emerging topics, following Saldaña's (2021) descriptive first cycle coding approach using the Dedoose software program. Dedoose is a web-based qualitative and mixed methods data analysis software to organize, code, analyze, and visualize research data. Based on the framework in Andrade and Bunker (2009) and following a top-down elaborative second cycle coding approach, these codes were then put into the four following themes: cognition, metacognition, behaviour, and motivation. These themes answer the questions of *what, how, when,* and *why* learners do what they do. Codes and themes were reviewed and discussed with three colleagues for consistency, as recommended by Saldaña (2021). The singing and speaking contexts were then compared to explore notable differences in awareness in each modality. Of note, for the pre/post task reading exercises, participants exclusively pointed out that the reading tasks were easier and they understood them better. Because there were so few pre/post task reading exercises and the comments from them were homogenous, they have been left out of the analysis and discussion. To determine if there were any improvements from pretest to posttest, a future study where external raters rate the speech samples will be necessary.

The results of the self-regulation survey (adapted from Tseng et al., 2006), used to capture learners' beliefs about learning pronunciation, were descriptively analysed and triangulated with the above themes and results. Because the survey focused on learner strategies and did not dive into cognitive aspects of self-regulation, comments from the survey only relate to metacognition, behaviour, and motivation in the discussion.

Results

Codes and Themes

To understand what learners notice about their own pronunciation, think-aloud comments were coded and put into themes. This section first presents the coded think-aloud responses of participants. The thematic analysis that follows aligns with the framework of self-regulatory processes, encompassing cognition, metacognition, behavior, and motivation. Using a descriptive process, codes were first assigned to label the type of responses received from participants. They were generated based on answering the research question: "What do learners notice in sung input, spoken input, and their own pronunciation?". The codes generated were as follows:

1) Features: when participants commented on specific language features (e.g., liaison, intonation, specific phonemes);
2) Length: when comments related to the length of the phrase they were imitating (e.g., long or short, longer phrases may constrain memory);
3) Music/Speech relationship: when comments related to whether music or speech specifically affected their imitations;
4) Pronunciation: general pronunciation comments;
5) Understanding: when comments related to whether they understood the audio and whether understanding or meaning affected their ability to imitate;
6) Reasoning: when participants commented on the reason why they did something or performed a certain way;
7) Metacognitive Strategies: when participants commented on their strategy (e.g., "closing my eyes, focusing on how I'd say it in conversation");

8) Behavioural Strategies: when strategy was related to behaviour changes during the study or in the future (e.g., mentioning that they will do these types of exercises in the future on their own).
9) Metacognitive beliefs about learning: when commenting on why they believed they performed a certain way in their imitations;
10) Motivational beliefs about learning: when comments about performance influenced their motivation or enjoyment and how that impacts learning.
11) Enjoyment: when commenting on level of enjoyment.

From here, some codes were categorized into subcodes falling under main codes to group similar codes and then put into the following four themes as described by Andrade and Bunker (2009) and illustrated in blue in Figure 7. Codes and sub-codes can be seen in magenta and orange, respectively. The number on the left refers to the number of codes or themes in each category.

Theme 1: Cognition, consisting of two codes, focus on form and focus on meaning. Focus on form was comprised of the first four subcodes, “features”, “length”, “music/speech relationship”, and “pronunciation”. Focus on meaning was comprised of the subcode “understanding.”

Theme 2: Metacognition, made up of the codes, “reasoning”, “strategies”, and “metacognitive beliefs about learning”.

Theme 3: Behaviour, the “behavioural strategies” were placed in this theme.

Theme 4: Motivation, consisting of “motivational beliefs about learning” and “enjoyment”.

The codes relating to strategies and beliefs about learning began as individual codes that were later split to suit the themes better. The total number of codes are shown below in Figure 7. The majority of comments related to whether or not the participant understood what they were saying and relatively few comments were related to behaviour or motivation.

Figure 7

Total number of codes, categories, and themes (from dedoose)

The next step was to break out the comments resulting from imitating sung lyrics from spoken lyrics. The purpose of this is to compare what participants noticed in each modality. See Figure 8 for the comparison. A few responses were not included in this comparison because they were generalized responses where participants were not responding to specific recordings and thus, the responses did not apply to a given modality. Strikingly, the responses were remarkably similar in theme. One of the unsurprising differences was that there were more comments on *music/speech relationships* while reflecting on imitations of sung lyrics. Reflections on

imitations of sung lyrics also focused slightly more often on *understanding* while more reflections on imitations of spoken lyrics focused slightly more often on *pronuncation*, but these are minor differences. There does not seem to be proportional differences between learners' focus on metacognitive, cognitive, behavioural, and motivational processes across contexts.

Figure 8

Codes according to modality (singing or speaking)

Below is a collection of samples falling into each of the four main themes observed in the analysis: cognition, metacognition, behaviour, and motivation. They are separated according to whether the comments came after a singing or speaking imitation.

Table 18

Cognitive Processes in Pronunciation Learning

Focus on Form
Features – Singing
“I think I’m parodying the cadence that I’m hearing a little bit.”
“I’m noticing that I have a pretty natural liaison in the first example, ‘tout au long’.”
“There was rhyming involved, which made it easier.”
“There is a lot of liaison.”
Features – Speaking
“I'm happy with my liaison, such as when I had to say, where was it? ‘Tout au long’.”
“I'm noticing the emphasis is in the wrong spot sometimes.”
Length – Singing
“The clips were shorter so maybe it was a little easier.”
“Sentences that were long and that didn’t have vocabulary that I could understand, I really had trouble with.”
Length – Speaking
“These sentences were longer, and it was very... I don't know, it was hard to follow them all.”
“it’s getting more difficult because the sentences are getting longer and harder to remember.”
Music/speech relationship – Singing
“I couldn't hear what they were saying because the sound of the music was competing.”
“I found it much harder [to hear the lyrics] with the music.”
“I mean, rap is just easier to understand.”
Music/speech relationship - Speaking
“The sentences are harder to repeat than the songs, maybe.”
“The spoken recording are again much easier, even easier than the rap I would say.”
Pronunciation – Singing
“Here I could reproduce the sounds well, even if the words didn't make sense.”
“I try to repeat more rhythm than actual words.”
“Literally I can just repeat the rhythm, but nothing more.”
“I'm noticing more of a sort of throat-like sound as I'm listening to these recordings.”
“it feels like the quality of my words are changing a little bit, like they're becoming much shorter and more kind of staccato.”
“I notice that sometimes I'm like adding things that don't really belong.”
“Exaggerated ends of words, and that again made it easier to copy.”
“I was focusing my attention on individual sounds.”

Pronunciation – Speaking

“I'm pronouncing it with an accent of the lady and not the accent that I use regularly.”
“Because I feel like I'm mocking if I use a Quebec accent.”
“I didn't really realize how poorly I was pronouncing it.”
“I'm noticing my R's much more than I usually do”
“I do pronounce the P differently than in English.”
“I could hear more of the vowel differences”

Focus on Meaning

Understanding – Singing
“I had absolutely no idea what was being said most of the time except for the last one”
“I think it's hard when I don't understand what they are saying.”

Understanding – Speaking
"Being uncertain about what I'm saying makes fluency challenging."
“I have difficulty pronouncing things without fully understanding the meaning.”
“because I wasn't 100% sure of what they were saying, so I was just kind of mumbling along.”

Table 19

Metacognitive Processes in Pronunciation Learning

Reasoning – Singing
“because your memory is attached to like a character in your head it helps you recall the context and the meaning”
“And kind of having those two reference points, [original audio and my recording], allowed me to make the second recording”
Reasoning – Speaking
“I feel like I was able to use that first attempt to improve my second attempt after cross-checking with the original recording.”
“I'm a Spanish speaker sometimes we don't have those sounds.”

Strategy – Singing
“I tried to practice before recording the sounds.”
“I divided the sentence like in 2 or 3 parts and I tried to pay attention to each part.”

Strategy – Speaking
"I took a third take, finding it helpful for my self-study. It feels easier with more confidence."
"I'm leaning into using the technology now and finding it very helpful for my self-study."
"When uncertain, it can be difficult to repeat with gusto; leaning into using technology now."

Metacognitive beliefs about learning – Singing
“The ones I can understand, I can improve.”
“if I just didn't know the words then it was hard to improve.”
“I think that you need a mental representation of what you are trying to say.”

Metacognitive beliefs about learning – Speaking

“I was not able to improve on what I did not know.”
“I think if I'd had a third shot I might have got there eventually.”

Table 20

Behavior

Behavioral Strategies – Singing

“I feel like I'm making kind of symbols just for those sounds to write it down to remember the sequence of what I'm saying.”
“I tried to practice before recording the sounds.”

Behavioral Strategies – Speaking
"Took a third take… after two attempts that I still had a much better take in me.”
“I'm going to continue to do something like this for my own self-study and recommend it to my students because it's great.”
“try and correct the pronunciation by making it more natural, closing my eyes, focusing on how I would say it if I was in conversation.”
"I'm using the opportunity to hear myself to make improvements; it's great for my self-study."

Table 21

Motivational Aspects of Pronunciation Learning

Motivational beliefs about learning – Singing
“This recording stuff is making me self-conscious about my pronunciation. But it's kind of helping.”
“the exercise is getting so tiring that I can't even reproduce these very short clips.”

Motivational beliefs about learning – Speaking
“I feel like I had more confidence for sure, and I think the result was better.”
“it feels like I'm starting to teach myself and it's a little uncomfortable”

Enjoyment – Singing "That one was really hard and I didn't enjoy doing it." "I was so happy to know what I was saying now."
Enjoyment – Speaking "I'm really appreciating and enjoying the opportunity here because it's occurring to me that this is something I have not done with my French" "I'm going to continue to do something like this for my own self-study and recommend it to my students because it's great."

Self-regulation Survey

To complement the data collected from the think-aloud responses, participants were asked questions related to self-regulation. These questions draw on Tseng et al. (2006), which proposed the following five facets making up self-regulation: Commitment Control (CC), Metacognative Control (MC), Satiation Control (SC), Emotion Control (EmC), and Environmental Control (EnC). Each of the questions in this survey were coded to one of these facets. Table 22 shows descriptive statistics for the responses to the self-regulation survey. These results are ordered according to the mean score from lowest to highest, where satiation control (SC4) received the lowest average score, meaning participants viewed their ability to regulate their mood in response to boredom lowest, and emotional control (EMC3) received the highest score, meaning participants felt like giving up when stressed. As there are negatively worded questions in the survey (i.e., SC1 and EmC3), they were inverted before creating Table 23 comprised of the composite variables.

Table 22

Responses to self-regulation survey (adapted from Tseng et al., 2006)

	Minimum	Maximum	Mean	Std. Deviation	Variance
SC4: When feeling bored with learning pronunciation, I know how to regulate my mood in order to invigorate the learning process.	2	6	3.25	1.552	2.408
EmC1: When I feel stressed about pronunciation learning, I know how to reduce this stress.	1	6	3.45	1.276	1.629
MC3: When it comes to learning pronunciation, I have my special techniques to prevent procrastination.	2	6	3.45	1.356	1.839
MC4: When it comes to learning pronunciation, I think my methods of controlling procrastination are effective.	2	5	3.50	1.395	1.947
SC3: During the process of learning pronunciation, I am confident that I can overcome any sense of boredom.	1	6	3.55	1.820	3.313
SC2: During the process of learning pronunciation, I feel satisfied with the ways I eliminate boredom.	1	6	3.55	1.669	2.787
MC2: When learning pronunciation, I think my methods of controlling my concentration are effective.	1	6	3.65	1.565	2.450
MC1: When learning pronunciation, I have special techniques to keep my concentration focused.	2	6	3.70	1.342	1.800
EmC4: When I feel stressed about my pronunciation learning, I cope with this problem immediately.	1	5	3.75	1.446	2.092

CC2: When learning pronunciation, I believe I can achieve my goals more quickly than expected.	1	6	3.90	1.334	1.779
EnC2: When learning pronunciation, I know how to arrange the environment to make learning more efficient.	2	6	3.90	1.165	1.358
EmC2: I feel satisfied with the methods I use to reduce the stress of pronunciation learning.	2	6	3.95	1.276	1.629
CC3: When learning pronunciation, I persist until I reach the goals that I make for myself.	1	6	4.10	1.619	2.621
CC1: When learning pronunciation, I have special techniques to achieve my learning goals.	2	6	4.15	1.182	1.397
EnC1: When I am studying pronunciation and the learning environment becomes unsuitable, I try to sort out the problem.	2	6	4.40	0.995	0.989
CC4: I believe I can overcome all the difficulties related to achieving my pronunciation learning goals.	1	6	4.40	1.603	2.568
SC1: Once the novelty of learning pronunciation is gone, I easily become impatient with it.	2.00	7.00	4.5500	1.60509	2.576
EnC3: When learning pronunciation, I am aware that the learning environment matters.	2	6	4.55	1.432	2.050
EnC4: When I study pronunciation, I look for a good learning environment.	2	6	5.05	0.999	0.997
EmC3: When I feel stressed about pronunciation learning, I simply want to give up.	2.00	7.00	5.1000	1.68273	2.832

Table 23 shows the composite variable scores that were calculated by averaging the scores for each question belonging to one of the five facets. We can see here that environmental (EnC) and commitment control (CC) scored the highest with mean scores of 4.475 and 4.1375, respectively, and metacognitive control (MC) scored the lowest with a mean score of 3.575. Satiation control (SC) had the greatest variance among learners (1.710). An average of all five composite variables is included to represent an overall self-regulation score (called SelfReg), which had a mean score of 3.9950 out of 7 and a variance of 0.573 and standard deviation of 0.75688.

Table 23

Min-Max Composite Variables

Composite Variables	Minimum	Maximum	Mean	Std. Deviation	Variance
MC	2.00	5.50	3.5750	1.07941	1.165
SC	1.75	6.25	3.7250	1.30762	1.710
EmC	2.25	5.75	4.0625	0.97291	0.947
CC	2.50	6.00	4.1375	0.92294	0.852
EnC	2.50	6.00	4.4750	0.88815	0.789
SelfReg	2.95	5.90	3.9950	0.75688	0.573

Table 24 includes the composite scores for each of the participants in the study. Interesting to note, the highest proficiency French speaker did not rate her ability to self-regulate higher than average and the lowest proficiency speakers did not rate themselves lower. The prediction that proficiency might affect self-regulation is not apparent in this data set.

Table 24

Raw Data with composite variable (average for each category)

Part.	Age	City	Gender	L1	Langs spoken	French prof	CC	MC	SC	EmC	EnC	Total SelfReg
1	42	Montreal	Male	Spanish	3	3	4	2	4	5	5	4
2	19	Brampton	Female	English	3	6.5	5	5	3	5	6	5
3	38	Montreal	Female	English	2	4	3	3	3	3	4	3
4	40	Montreal	Female	Spanish	3	4	4	2	2	5	5	4
5	32	Montreal	Female	English	3	3	5	5	5	4	5	4
6	43	Longueil	Female	Ukrainian	5	6	5	5	6	4	4	5
7	53	Dorval	Female	English	2	4	3	3	2	3	4	3
8	30	Toronto	Female	Mandarin	4	2	3	4	4	4	5	4
9	35	Montreal	Male	Spanish	3	5.5	5	2	2	2	4	3
10	24	Montreal	Female	English	4	6	6	6	6	6	6	6
11	35	Saltspring	Female	French	6	7	5	4	5	5	4	4
12	32	Montreal	Female	Farsi	2	3.5	5	3	2	3	6	4
13	29	Hyderabad	Female	English	4	2.5	5	3	5	5	3	4
14	49	St Constant	Male	Portuguese	4	1	4	5	5	5	5	5
15	29	Montreal	Female	Portuguese	4	1.5	3	3	2	4	3	3
16	24	Warsaw	Male	Polish	10	4	3	3	5	3	4	4
17	30	Wrocław	Male	Polish	7	3.5	4	4	3	4	4	4
18	39	Montreal	Male	Portuguese	4	5.5	4	5	4	4	5	4
19	57	Montreal	Male	English	4	5.5	5	3	4	5	5	4
20	38	Dallas	Female	Spanish	4	3.5	4	5	5	4	5	4

Discussion

This study examined what learners notice in recordings of themselves when imitating sung and spoken lyrics. In the analysis, we consider the four psychological components that make up the themes for self-regulation (from Andrade & Bunker, 2009) explored in this study: cognition, metacognition, behaviour, and motivation. Each of these are discussed as a whole, but we also explore differences that emerged in the singing and speaking contexts.

Cognition

Cognitive aspects of self-regulation relate to *what* learners self-regulate, as opposed to *how*, *when*, and *why* they do so. Tallies from the think-aloud transcriptions show that cognitive

aspects of language were the most common comments from participants. Most often, these were general comments relating to meaning and to overall ability to pronounce a phrase accurately. Participants frequently reported challenges in comprehending the lyrics of the songs, indicating their belief in the importance of language comprehension, vocabulary recognition, and listening skills in the pronunciation learning process. The recognition that comprehension is important in the learning process aligns with the studies from Kennedy and Blanchet (2014) and Kennedy et al. (2014), which suggest that meaning-form relationships are important to language awareness. However, it would be interesting to investigate how often learners default to their L1 pronunciation habits when they understand a word or phrase. For example, it is conceivable that an English speaker learning French would imitate the word "rue" with an /ɹ/ sound rather than the French /ʁ/ if they already know the phrase and struggle with the uvular French /ʁ/. If hearing the sound as part of a phrase that one does not understand (and does not have the opportunity to read/see it), it is possible the learner will approximate the sound in their articulators because they have not mapped the target uvular French onto the English /ɹ/.

Another reason there may have been a strong focus on meaning was that participants were exposed to chopped up phrases that were not necessarily contextualized well (e.g., "N'y va pas par quatre chemins"). This could make it harder for participants compared to having the opportunity to listen to an entire song or conversation *in context*.

It is not entirely clear whether there were substantial qualitative differences between the comments coded as "understanding" between the singing and speaking contexts. There were more comments related to understanding in the singing group and upon evaluation, it is probable that understanding sung lyrics was a little more difficult than the spoken phrases. This behaviour is not entirely surprising considering Patel's (2003) hypothesis that music shifts attention to form

rather than meaning. Two examples illustrate this point, “I try to repeat more rhythm than actual words,” and “it feels like the quality of my words are changing a little bit, like they're becoming much shorter and more kind of staccato.” There is also more acoustic information to process, which is hypothesized to interfere with focus on meaning, but potentially a real opportunity to help learners notice acoustic and suprasegmental qualities of language (Patel, 2013). In addition, repetition, which is prevalent in music, may compensate for the challenge in understanding the meaning of the lyrics because repeated exposure allows for an opportunity to become familiar with both form and meaning.

The influence of regional accents on pronunciation was a recurring theme, emphasizing the role of phonological processing and accent recognition in language learning. If the singer or speaker’s accent did not match their own, participants occasionally commented on how they struggled to imitate the accent from the audio recording. These comments appear to be similar across singing and speaking contexts. Exposure to a variety of accents is likely positive over the long term, as this exposure provides a form of high variability pronunciation training, which has been shown to improve L2 pronunciation skills (e.g., Melnik & Peperkamp, 2021).

When it comes to specific pronunciation features, participants most often commented on liaison, vowels, intonation, and the length of phrases. Awareness of liaison and vowels were predicted to be common comments from participants (see literature review and insights from Kennedy et al., 2014). Comments appear to be related to declarative knowledge of sound mapping more than analogy across contexts, as discussed by Inceoglu (2021), and several participants wished they had access to text to improve pronunciation accuracy.

As far as the length of phrases goes, varying numbers of syllables were included in the exercises to see how that affected learners’ abilities and think-aloud responses. It is hard to

remember what was said if one's working memory is taxed, as is apparent from the think-aloud results (e.g., "I found it was a lot more repetition and shorter sentences, so it was easier to remember and repeat"). As for the other features, if learners are paying most attention to liaison, vowels, and intonation, then certain aspects of pronunciation, such rhythm (syllable timing) and consonants (/ʁ/ or "H" muet/aspiré, for example), are areas teachers and material developers should focus on for teaching and material development. "H" muet/aspiré was not explicitly commented on by participants despite contrasting exemplars being included in the sung and spoken activities. A follow up study that explores whether participants made unconscious adjustments to their speech could inform whether implicit learning happens here. If not learned implicitly, features that often go unnoticed such as "H" muet/aspiré might benefit most from explicit instruction.

There was a considerable amount of variability in the comments related to music/speech relationships; understandably, most were generated in the singing context. In the context of the literature discussed earlier, it seems that, depending on the song and the student, listening to music can be either beneficial (e.g., promote focus on form) or a hinderance (e.g., interfere with understanding the lyrics). One's understanding and ease of imitating sung lyrics seemed to depend on the context, genre, speed, familiarity, competing instrumental sounds, melody, and rhyming (supported by Ludke et al., 2014). Although unexpectedly several participants commented on rap being easier to understand despite the faster cadence. It would be more beneficial to engage students in exercises where one repeatedly listens to songs in their entirety (as opposed to snippets like in this study) and reads along with lyrics to help with contextualization. Practice with different speeds can also prime later speech patterns (Jungers et al., 2016). It is reasonable to expect that this is how learners would incorporate music into their

target language development outside of formal learning environments. If students are motivated by listening to L2 songs, able to choose songs with appropriate vocabulary, and provided supports (access to lyrics or translations), there are likely additional benefits to learning through music, such as increased repetition leading to easier memorization and imitation (e.g., “a lot more repetition… so it was easier to remember and repeat”, as one of the participants stated).

Participants indeed commented on remembering words due to rhyming (e.g., “I'm noticing a lot of rhyming sounds; ‘plus’ /ply/, ‘tu’ /ty/”). Awareness was higher in the singing context; however, because all passages were taken from lyrics, the speaking context still generated comments related to rhyming as well. Still, the participants generally found imitating sung lyrics more difficult. Similar activities should probably be initiated by students who want to learn this way, and instructors and material developers can contribute by helping learners understand what they might implicitly learn from songs. Even if one does not understand a word, one is still exposed to cadence, vowel quality, liaison, and other phonological and extra-linguistic patterns. Research on implicit learning from L2 music exposure and coaching learners to see value in this exposure is warranted.

Metacognition

Now let us consider learner awareness from a metacognitive perspective. Participants did acknowledge metacognitive processes and recognized areas for improvement that can be captured by Reinders’s (2020) stages of self-directed learning, which include reflections and interactions around (1) identifying learning needs, (2) setting goals, (3) planning learning, (4) selecting resources, (5) selecting strategies, (6) practice, (7) monitoring progress, and (8) assessment and revision. As seen here, metacognition has to do with recognizing when learners describe *how* they will regulate their learning. For examples, consider the following two

statements: “I took a third take on that second to last one there because I had the feeling after two attempts that I still had a much better take in me,” and, “I feel like I was able to use that first attempt to improve my second attempt after cross-checking with the original recording.” Participants also demonstrated metacognitive awareness of using technology for self-monitoring and improvement. Similar to the findings of Foote and McDonough (2017), a few students mentioned that they would do activities like this on their own in the future because they had not realized the benefits of imitating or listening to one’s own recordings before (e.g., “I'm going to continue to do something like this for my own self-study and recommend it to my students because it's great”).

Several participants also commented on tiring of the activity and losing attention, e.g., “I think I'm getting tired of repeating the same exercise and I tend to forget the last part of the sentence.” Since these tasks took students 1-1.5 hours to complete, the task was probably too lengthy to expect participants to maintain focus and to repeat similar activities. Appropriate planning of learning activities would improve this issue (Tseng et al., 2006). No differences across singing and speaking contexts were evident.

Of note, the self-regulation survey used in this study (adapted from Tseng et al., 2006) suggests that learners would benefit from improving metacognitive skills as these were the lowest rated. Helping students with planning how they study would be valuable. When thinking about bridging classroom and independent study, I recommend teachers spend some class time on metacognitive aspects of learning such that learners can better spend time developing their language skills on their own, such as recommended by Reinders (2020) and Botero et al. (2019). Metacognitive awareness is a major part of self-regulating dimensions of learning beyond the classroom, including method, time, and environment (Andrade & Bunker, 2009).

Behaviour

The low number of comments made by participants could reflect their limited awareness of behaviour with regard to pronunciation development, but it is more likely the nature of the exercise that did not provoke them to comment on it. One-shot studies do not ask learners to consider the question of *when* to study in order to accomplish objectives over a given timeframe. Of the comments gathered, however, codes demonstrate that participants noticed that iterative learning involved adjusting behavioural pronunciation strategies based on feedback and self-assessment. Some participants also commented on adjusting future behaviours based on what they experienced during this study, e.g., “I'm going to continue to do something like this for my own self-study and recommend it to my students because it's great.” Again, no major differences in awareness were recognized across singing and speaking contexts.

Lastly, challenges with longer phrases pointed to behavioral aspects of refining practice strategies for specific language elements. For example, one participant remarked, “I divided the sentence like in 2 or 3 parts and I tried to pay attention to each part” and another commented, “I had to practice many times before I recorded.” This highlights the importance of learners engaging with materials in ways that can be readily processed, whether that means in small chunks or with repetition (Ellis, 2003). Research suggests that repetition is much more enjoyable with music than with spoken content (Engh, 2013a); it is probable that if learners are given autonomy to select songs and manage time, additional benefits of music exposure would be recognized.

The self-regulation survey asked about environmental control and commitment control, which are somewhat related to behaviour. In these two areas, learners reported high confidence,

suggesting that there may not be any apparent pedagogical suggestions to be made. Studies evaluating learners' behaviour in learner-initiated contexts would offer more insight here.

Motivation

Finally, considering motivation (*why* learners study the way they do), responses from this study indicate that confidence and fluency were linked to motivation, indicating intrinsic motivation and positive reinforcement through language proficiency. The following three statements illustrate this: "if I rate myself low on one exercise, then on the second I'm less motivated to perform well;" "That one was really hard, and I didn't enjoy doing it;" and "I was so happy to know what I was saying now."

Enjoyment similarly contributed to motivation. These connections have been highlighted many times before in literature (e.g., Dörnyei & Ushioda, 2009; Hsu, 2015; Kara & Aksel, 2013). For these participants, understanding the lyrics was a major contributor to enjoyment and motivation. Thus, working to develop level-appropriate materials is important for this reason alone (Reinders, 2020). Perhaps as a surprise, reflections from the singing context were not overwhelmingly different to the speaking context. The fact that the singing context was not obviously more enjoyable than the speaking one might be due to the fact that, during these activities, participants had little control over duration of study, songs choice, and when and which lyrics to repeat, etc.

The self-regulation survey results revealed that several of the lower rated questions related to boredom or concentration, suggesting that participants have a lower ability to regulate attention and motivation compared to other self-regulatory skills. Musical content can be an excellent way to address this, as demonstrated in the study carried out by Sundberg and Cardoso, 2022 (Chapter 3). However, it is important to ensure that learners have access to level-

appropriate materials so that they can comprehend content and feel confident in their ability to learn from it.

Conclusion

The findings of this study shed light on the interplay between cognitive processes, metacognitive awareness, behavior, and motivation in the context of pronunciation learning from sung and spoken input. The implications drawn from this study contribute to the limited research on music and second language learning as well as reflective practices that contribute to self-regulatory processes when learning a second language. A major highlight was related to participants' extensive focus on cognitive aspects, including both meaning and form, which underscores the importance of contextualized language learning. This appears to be the case in both singing and speaking exercises. In fact, differences in language awareness between singing and speaking contexts are minimal across all four themes in this study. However, an important difference is that understanding content can be more challenging with sung input, so finding level-appropriate songs should be a goal of learners and teachers wishing to incorporate music into their language learning pursuit; fortunately, repetition might mitigate this challenge over time. A potential strength of this shift away from meaning is the opportunity to focus on formal elements of language that may not always be salient in spoken contexts.

Specific pronunciation features, such as liaison, vowels, and intonation, emerged as focal points for learners in both singing and spoken contexts. These may be because these are more salient features or because they are traditionally highlighted in educational environments. The degree to which these points should continue to be major curricula items depends on the degree to which learners can notice, assimilate, and practice improving these features on their own. However, the variability in comments related to music/speech relationships indicates that the

effectiveness of using music as a learning tool depends on various factors, such as context, genre, and familiarity. The influence of regional accents on pronunciation highlights the role of phonological processing and accent recognition in language acquisition. Therefore, how music is integrated into learning likely requires careful consideration.

Participants did demonstrate some metacognitive awareness through reflections on their ability to use technology for self-monitoring and improvement. This suggests that autonomous learning through technology-enhanced music can facilitate the development of self-regulatory skills. Furthermore, participants' behavioral comments revealed their adaptive learning strategies, such as adjusting pronunciation strategies based on feedback. However, when looking at how learners rated their metacognitive strengths, they are less confident in their ability to self-monitor. Thus, explicit metacognitive support helping learners to understand how to plan and approaches to learning would be helpful and facilitate the development of cognitive aspects, particularly when encouraging learners to incorporate music into learning beyond the classroom environments with the express purpose of improving their target language.

Lastly, the findings related to motivation underscore the importance of developing level-appropriate materials to sustain learners' motivation and engagement for the purposes of positive reinforcement, enjoyment, and intrinsic motivation. The effect on learning outcomes is the next step to be examined to answer the question of whether learners' perceptions about the changes they are undertaking are effective. Further research is warranted to explore the implicit learning potential of L2 music exposure and to develop strategies to enhance learners' understanding and appreciation of this exposure.

Chapter 5: Concluding Remarks

Prologue

There is a story that I often read my son called, “Mr. Tiger Goes Wild” by Peter Brown. In the story, animals live upright on their hind legs, in formal dress, and in an urban environment. Mr. Tiger feels like he is missing out on something and unable to fully express himself. The pressure to conform is immense, but he eventually breaks free. He runs four-legged and without formal attire. He spends time in nature and roars loudly. A crude analogy but it makes me think about what language learners want out of their educational experiences. They want to move beyond formal environments and express themselves naturally. As I read the literature on rewilding in second language education settings, I began to think about how to bridge formal and informal environments. On the one hand, there is much to be gained in a controlled setting where we engage in activities that target specific gaps in our knowledge and ability (practicing a grammatical tense or a vowel sound that does not exist in your L1). On the other hand, we can be limited in the degree to which we direct our learning. The concept of rewilding offers a suitable bridge because it requires us to think carefully about how learners expect to use their language when they leave our classrooms and how to help them to get there. Rewilding pedagogies embrace the tension between and affordances of structured and unstructured learning.

This echoes my own experience, especially with regards to the interconnected nodes mentioned in Chapter 1 and described by Reinders (2017). When I began learning French in Quebec, I already spoke high-intermediate Spanish. I was heavily motivated to learn through online materials and use a variety of resources. I took some classes but what I wanted from those classes was practice speaking, not a presentation of grammar, which I could find online, as

needed. I listened to French music, read French children's books, worked my way through Duolingo, watched YouTube videos, hired tutors for brief periods, and used a long list of other digital and non-digital materials. This process, however, was often frustrating. I would sit at my computer listening to a song, trying to scroll backwards in the song to repeat lyrics while flipping between browser tabs with the song's lyrics and google translate.

What I could have used the most help with is finding materials that matched my level of proficiency and finding tools that were efficient at providing feedback, even quick access to definitions or translations of vocabulary. I could have used more structured materials that also allowed for self-direction. For the most part, on my own I was able to work up to about a B2 level but found it frustrating when I was placed into classes where the focus was on writing because my declarative knowledge of French was reasonably high. Much of the material that I used to get to a B2 level was difficult and not designed for learners.

The goal of this dissertation was to explore how self-directed pedagogical tools, specifically the use of songs beyond the traditional classroom, fit into learners' target language development. Chapter 1 identified the non-formal learning space as way to situate understanding the self-regulatory processes required to encourage learners to use structured materials autonomously such that objectives, content, methods, monitoring, and evaluation (Holec, 1981) are in the learner's control. While non-formal learning describes the learning environment of interest in this dissertation, rewilding describes a method for understanding the interactions and progression between environments. I believe this novel framework is helpful for bridging theories of autonomy, self-regulation, and ecologies of language learning and adds to the literature on rewilding introduced by Little and Thorne (2017) and Thorne et al. (2021).

Songs are chosen as the primary content for investigation due to their prevalence in learners' daily lives and as a common source of target language exposure (Lindgren & Muñoz, 2013). I was also motivated because of my own experience and the experience of those I interacted with online and in similar situations.

Through the rest of this chapter, I provide a comprehensive summary and conclusion to the dissertation by synthesizing the key findings and contributions from the previous chapters. Following the introduction in chapter one, which lays out the framework for this research, the manuscripts in chapters two through four investigate: 1) the theoretical affordances provided by listening to songs in a target language; 2) learner perceptions of the self-directed *Bande à Part* (*BàP*) French music app, created to address some of the learner needs identified in Chapter 1; and 3) learners' self-regulatory processes and awareness, via an examination of the interplay between cognitive processes, metacognitive awareness, behavior, and motivation in pronunciation learning resulting from listening to and commenting on recordings of their own imitations of sung and spoken input. As a refresher, the dissertation followed Cardoso's (2022) chronological framework for conducting CALL research, continuing from where the development stage (stage 1) concluded in Sundberg and Cardoso (2019). The stages were as follows: stage 2, exploration (Manuscript A), stage 3, assessing suitability (Manuscript B), and stage 4, assessing pedagogical effectiveness (Manuscript C). The following questions were used to guide the research:

1. Manuscript A (Chapter 2): What are the theoretical affordances of language learning through songs? What does the empirical work on using songs for teaching and learning reveal about the potential benefits and limitations of language learning

through songs? What are the future directions for using songs to develop language skills?

2. Manuscript B (Chapter 3): Are learners willing to use the self-directed *Bande à Part* app to help them learn French?
3. Manuscript C (Chapter 4): What aspects of language and language learning processes do learners notice when imitating sung vs. spoken input?

Summary of Findings

Chapter 2 discussed the similarities and differences between music and language processing and content. Identifying these similarities and differences allows us to underline potential affordances of music, which, in turn, forms the theoretical lens through which I investigated the empirical research available on the use of songs for teaching and learning. The theoretical underpinnings explored included: 1) how acoustic differences between singing and speaking (pitch, duration, volume) might contribute to affordances of learning through music via the enhancement of prosodic features of language, 2) how the *shared syntactic integration resource hypothesis* (Patel, 2003) for music and language encourages focus to shift back-and-forth between form and meaning, and 3) how exposure to the motivating and repetitive content provided in songs benefits second language learners in ways that can be limited in formal environments (Gatbonton & Segalowitz, 2005).

With these considerations in mind, the empirical research reviewed suggests several benefits to language learning through music, a few of which are highlighted here with their connections to the theories above: 1) musical training can improve pronunciation development (e.g., Milovanov, 2009), and that this development is likely due to improving domain-general auditory processing rather than increased musical aptitude itself (Zheng et al., 2022). This

conclusion lends support for the *shared syntactic integration resource hypothesis* because it demonstrates that musical development impacts general domain processing to a degree that general domain processing is more predictive of pronunciation ability than musical aptitude alone. 2) Songs can aid in segmentation and word learning in a new language (Schön et al, 2008). This appears to be the case because pitch cues are important factors in identifying word boundaries, highlighting the potential that increased salience in pitch, volume, and duration may have on improving one's implicit ability to discriminate these boundaries. 3) The type of content provided in songs is repetitive (Engh, 2013a), motivating (Engh, 2013b; Kara & Aksel, 2013), and memorable (Murphey, 1990; 1992). However, classroom learning material often struggles to capture these important aspects of language learning (Gatbonton & Segalowitz, 2005); thus, the inclusion of music in the language acquisition process is well supported. Despite support for incorporating music in the L2 curriculum, further research is necessary to explore the connections between L2 language development and L2 exposure to music. In addition, advocacy is needed to promote the integration of music across various learning contexts (Engh, 2013a; Werner, 2018).

Some avenues for future research were identified including: 1) exploring how musical training might help learners acquire difficult L2 features, 2) investigating singing as a form of imitation or shadowing in language learning, 3) developing strategies for using songs in self-regulated learning outside of the classroom, 4) applying instructional design principles to create song-based learning materials, and 5) leveraging technology to enhance song-based language learning. Chapters 3 and 4 explore and comment on several of these directions.

Chapter 3 presented a pilot study on participants' use of the *Bande à Part* (*BàP*) app. *BàP* is a music app, developed according to CALL theories (Sundberg & Cardoso, 2019), that allows

students to listen to French songs, repeat phrases, access lyrics and translations, and see highlighted phenomena like liaison on screen. The TAM model (Davis & Venkatesh, 1996) was used to explore this learner-centric music app to understand whether the iterative development process of this tool led to a product that learners found useful and easy to use, the two most important characteristics that predict continued user engagement with a tool. The findings suggested that learners do enjoy learning in this format, and they believe that they can learn through *Bande à Part*. While the app was found to be user-friendly, there were still opportunities for further enhancements to optimize its ease of use. Importantly, this app provides learners with some degree of control over all five levers described by Carliner (2013; namely: process, location, purpose, content, and consciousness), which are crucial for enabling self-direction in non-formal and informal environments. A number of improvements were identified for improving the tool such as adding more music and providing instructions on how to use the tool. Investigating learning outcomes will be an important step in the future directions of this research.

Chapter 4 explored the self-regulatory processes involved in pronunciation learning listening to imitations of sung and spoken input. Insight was generated from think-aloud exercises that inquired into learners' own awareness of their pronunciation and their perceived ability to improve on their own through recordings. These responses were coded and put into themes according to Andrade and Bunker's (2009) self-regulation framework that includes cognition, metacognition, behaviour, and motivation. Findings emphasize the importance of contextualized language learning and level-appropriate input. It was recognized that learners do believe they can improve through imitation and listening to their own recordings. Studies that investigate learners' perceptions are important because if learners do not believe they can learn in a given modality, they very likely will not (e.g. Botero et al., 2019; Sundqvist & Sylvén,

2016). Connected to this, it was found that learners' beliefs about improving their pronunciation were heavily influenced by whether they understood what they were listening to. This finding emphasizes why ensuring learners can access level-appropriate material is so crucial. Target language content in the form of music can be challenging to understand but listeners are more likely to enjoy repetitive exercises when they are delivered through music (Engh, 2013a; Sundberg & Cardoso, 2022). Incorporating music into the learning process can be an ideal way to push learners beyond their current abilities, provided that they actively participate in initiating the task and, ideally, have the opportunity to select level-appropriate content with intelligible lyrics.

These results tell us a lot about what these participants can and cannot do on their own. By understanding learners' limitations, we can determine the specific points where scaffolding may be necessary to support their effective use of independent learning materials. For example, participants in this study only vaguely commented on rhythm (the reoccurring pattern of stressed and unstressed syllables in speech; Celce-Murcia et al., 2010). Does this mean they are less likely to improve rhythm on their own without scaffolding, or is it simply a lack of metalanguage to articulate differences in what they hear? Participants did, however, often comment on vowel quality. Accordingly, it would be interesting to examine the degree to which the same participants could improve their vowel production on their own through independent exercises like the ones used in the experiment. Additionally, one could investigate the degree to which imitation and access to recordings serve as sufficient feedback for self-regulated pronunciation improvement. Taken together, these findings can inform approaches to incorporating music into language learning environments and supporting learners' self-regulation skills.

Discussion

The findings from these chapters highlight the potential of songs as a powerful medium for L2 learning in non-formal environments. By combining theoretical insights with empirical evidence, the research underscores the benefits of incorporating music into the language learning process. Song-based learning tools, like the *Bande à Part* app, offer important opportunities to develop effective self-regulatory strategies that encourage pronunciation development. These takeaways are valuable for educators and researchers seeking to understand affordances of music in language learning and approaches to support autonomous, self-regulated learners.

Integration of Findings

Theoretical and Methodological Implications

This dissertation reinforces the importance of Sociocultural Theory in understanding language learning processes and its applicability to non-formal environments, particularly considering learners' current reliance on a wide array of environments and materials to improve their skills. The concept of *rewilding* offers a novel lens for understanding the interaction and progression from formal to informal environments with the added benefits of bridging existing theories on autonomy and self-regulation.

The *shared syntactic integration resource hypothesis,* stemming from work outside of the field of second language acquisition, is indeed an important consideration for those in pronunciation research and teaching because it suggests that the cognitive resources used for processing language may overlap with those used for processing music. This has implications for understanding how music knowledge and exposure might influence language learning. Likewise, the Technology Acceptance Model (TAM), which is gaining support in CALL research, proves to be a valuable framework for evaluating learner perceptions of new learning tools because it

provides a systematic approach to understanding the factors that influence users' acceptance and adoption of technology. This allows researchers and material developers to gain insights into how learners view and engage with technologies that they plan to adopt for pedagogical purposes. This dissertation contributes to the field of second language acquisition by synthesizing theories (the *shared syntactic resource integration hypothesis* and TAM) validated in adjacent fields to bolster our own work without reinventing models and frameworks that already exist.

Lastly, these chapters respond to the demand for more theoretical research on self-regulation in language learning (Gao & Hu, 2020), particularly in environments where learners have more control over the learning process. This research highlights the benefits of using think-aloud protocols to investigate self-regulatory processes in pronunciation learning, contributing to a better understanding of what Pavlenko (2013) refers to as the "inner lives of learners."

Pedagogical Implications

Educators and developers can use these findings to inform the design and implementation of self-directed language learning tools, particularly those involving music. For example, incorporating songs into classroom activities can provide additional practice opportunities and enhance learners' motivation. Doing so might also serve as a demonstration of what learners can do on their own and encourage them to initiate their own similar learning activities to increase autonomy.

The findings from each chapter contribute to an understanding of how self-directed pedagogical tools can enhance L2 development using music as the target language form of input. They underscore the importance of learner-initiation of activities that can be scaffolded, particularly in non-formal environments. They highlight the need for level-appropriate content,

contextualized learning, adaptive strategies, and motivation in language learning by probing into the reflections, motivations, and interactions of learners as suggested by Reinders (2020). The chapters collectively suggest that songs can be a valuable tool for language learners, offering motivating material with naturally-occuring repetition and providing additional practice outside the classroom. However, to maximize the benefits of using songs beyond the classroom, some learners might benefit from improving their metacognitive strategies to further enable their ability to set objectives, define contents and methods, and monitor and evaluate their progress effectively (according to Holec's, 1981 requirements for autonomy). Controlling these learner responsibilities, in turn, makes for an empowered, self-directed learner.

In sum, the theoretical, methodological, and pedagogical implications described in these chapters outline a framework that I believe is conducive to learning from music: using technologies, such as *Bande à Part,* to provide structure to the learning process, thus allowing for level-appropriate, learner-initiated activity in a non-formal environment. These technologies, which can support self-regulated learning outside of the classroom must be evaluated for their perceived usefulness and ease of use to ensure that learners will in fact choose to use them on an ongoing basis. The non-formal environment exists along a spectrum that allows for the addition or removal of supports (e.g., highlighting *liaisons* onscreen), which can be adjusted according to Carliner's (2013) levers, to encourage the progressive development of self-regulatory skills (see chapter 1 for an outline of these levers). These supports can reduce autonomy in the short-term with the intention of raising awareness in the long-term and, ultimately, paving the way for more autonomy in the future.

Limitations and Future Research Directions

While the research provides valuable insights into the use of music as a way of encouraging self-direction in the language learning process, there are limitations and next steps to be considered. First, the small sample sizes used limit the generalizability of the findings. Future research should continue to explore the use of songs in L2 learning and investigate ways to enhance learners' understanding and appreciation of this exposure. Second, future studies can build upon the findings of this dissertation to further explore the effectiveness of self-directed learning tools. Studies exploring learning outcomes over longer durations are needed. Personally, I look forward to analysing pronunciation development through these tools to better understand learning outcomes.

Conclusion

In conclusion, this dissertation highlights the potential for self-directed pedagogical tools, particularly those using songs, to promote L2 development. The findings suggest that songs offer unique affordances for language learning, including prosodic development and engagement outside of the classroom. By understanding the cognitive, metacognitive, behavioral, and motivational aspects of language learning, educators and developers can create more effective and engaging tools for language learners that allow learners to determine their own path and receive feedback at a time when they are ready to assimilate it.

The contribution of this dissertation to the fields of second language acquisition and computer-assisted language learning is accomplished through 1) exploring the non-formal learning space through a rewilding lens, 2) using songs as a motivating form of content that offers unique affordances, and 3) providing insights into how self-directed pedagogical tools can support language development. It is hoped that this research will inspire further exploration and

innovation in this area, ultimately benefiting language learners around the world, just as music and our collective digital spaces have profoundly enriched my own language learning journey.

References

Adank, P., Hagoort, P., & Bekkering, H. (2010). Imitation improves language comprehension. *Psychological Science*, *21*(12), 1903-1909.

Alisaari, J., & Heikkola, L. M. (2016). Increasing fluency in L2 writing with singing. *Studies in Second Language Learning and Teaching*, *6*(2), 271-292.

Aljaafreh, A., & Lantolf, J. P. (1994). Negative feedback as regulation and second language learning in the zone of proximal development. *The Modern Language Journal*, *78*(4), 465-483.

Andrade, M. S., & Bunker, E. L. (2009). A model for self-regulated distance language learning. *Distance education*, *30*(1), 47-61.

Arnold, J. L., & Herrick, E. (2017). *New ways in teaching with music*. TESOL Press.

Aso, S., Saitou, T., Goto, M., Itoyama, K., Takahashi, T., Komatani, K., Ogata, T., & Okuno, H. G. (2010, September 6-10). Speakbysinging: Converting singing voices to speaking voices while retaining voice timbre. *Proceedings of the 13th International Conference on Digital Audio Effects* (DAFx-10), Graz, Austria.

Barcomb, M., & Cardoso, W. (2020). Rock or Lock? Gamifying an online course management system for pronunciation instruction: Focus on English/r/and/l. *CALICO Journal*, *37*(2), 127-147.

Beasley, R. E., & Chuang, Y. (2008). Web-Based Music Study: The Effects of Listening Repetition, Song Likeability, and Song Understandability on EFL Learning Perceptions and Outcomes. *TESL-EJ*, *12*(2), n2.

Benson, P. (2007). Autonomy in language teaching and learning. *Language teaching*, *40*(1), 21-40.

Benson, P. (2011). Language learning and teaching beyond the classroom: An introduction to the field. In P. Benson & H. Reinders (Eds.), *Beyond the language classroom* (pp. 7-16). Springer.

Benson, P. (2013). Learner autonomy. *TESOL Quarterly*, *47*(4), 839-843.

Benson, P., & Reinders, H. (2011). *Beyond the language classroom.* Springer.

Besson, M., Chobert, J., & Marie, C. (2011). Transfer of training between music and speech: common processing, attention, and memory. *Frontiers in psychology*, *2*, 94.

Besson, M., Schön, D., Moreno, S., Santos, A., & Magne, C. (2007). Influence of musical expertise and musical training on pitch processing in music and language. *Restorative neurology and neuroscience*, *25*(3-4), 399-410.

Bley-Vroman, R., & Chaudron, C. (2013). Elicited imitation as a measure of second-language competence. In E. Tarone, S. Gass, & A. Cohen (Eds.), *Research methodology in second-language acquisition* (pp. 269-310). Routledge.

Botero, G., Restrepo, M., Zhu, C., & Questier, F. (2019). Complementing in-class language learning with voluntary out-of-class MALL. Does training in self-regulation and scaffolding make a difference? *Computer Assisted Language Learning, 34(8)*, 1013–1039. https://doi.org/10.1080/09588221.2019.1650780

Box, G. E., & Tidwell, P. W. (1962). Transformation of the independent variables. *Technometrics*, *4*(4), 531-550.

Brown, S., Martinez, M. J., & Parsons, L. M. (2006). Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *European journal of neuroscience*, *23*(10), 2791-2803.

Bui, G., & Skehan, P. (2018). Complexity, fluency and accuracy. *TESOL encyclopedia of English language teaching*, 1-7.

Burger, S., & Chrétien, M. (2001). The development of oral production in content-based second language courses at the University of Ottawa. *Canadian Modern Language Review*, *58*(1), 84-102.

Cardoso, W. (2022). Technology for speaking development. In T. M. Derwing & R. Thompson (Eds.), *The Routledge Handbook of Second Language Acquisition and Speaking* (pp. 299-313). Routledge.

Carliner, S. (2015). *Training design basics*. Association for Talent Development.

Carliner, S. (2013). *Informal Learning Basics.* Association for Talent Development.

Chalon, Y. (1970). Pour une pédagogie sauvage. *Melanges pedagogiques*, *1*, 1-7.

Chapelle, C. (2003). English language learning and technology. John Benjamins.

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

Chik, A., & Ho, J. (2017). Learn a language for free: Recreational learning among adults. *System*, *69*, 162-171.

Collins, L., & Muñoz, C. (2016). The foreign language classroom: Current perspectives and future considerations. *The Modern Language Journal*, *100*(S1), 133-147.

Darasawang, P., & Reinders, H. (2010). Encouraging autonomy with an online language support system. *CALL-EJ Online*, *11*(2), 11-12.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly*, 319-340.

DeKeyser, R. (2010). Practice for second language learning: Don't throw out the baby with the bathwater. *International Journal of English Studies*, *10*(1), 155-165.

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.

Dörnyei, Z., & Ushioda, E. (2009). *Motivation, language identity and the L2 self* (Vol. 36). Multilingual Matters Bristol.

Dupoux, E., Sebastián-Gallés, N., Navarrete, E., & Peperkamp, S. (2008). Persistent stress 'deafness': The case of French learners of Spanish. *Cognition*, *106*(2), 682-706.

Ekeus, H., Abdallah, S., Plumbley, M., & McOwan, P. (2012). The melody triangle: exploring pattern and predictability in music. Eighth Artificial Intelligence and Interactive Digital Entertainment Conference.

Ellis, N. C. (2003). Constructions, Chunking, and Connectionism: The Emergence of Second Language Structure. In C.J. Doughty & M.H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 63-103). Blackwell. https://doi.org/10.1002/9780470756492.ch4

Ellis, N., & Collins, L. (2009). Input and second language acquisition: The roles of frequency, form, and function introduction to the special issue. *The Modern Language Journal*, *93*(3), 329-335.

Ellis, N. C., Römer, U., & O'Donnell, M. (2015). Second language constructions: Usage-based acquisition and transfer too. In J. Schwieter (Ed.), *The Cambridge Handbook of Biligual Processing*. (pp. 235-254). Cambridge University Press.

Engh, D. (2013a). Why use music in English language learning? A survey of the literature. *English Language Teaching*, *6*(2), 113-127.

Engh, D. (2013b). Effective use of music in language-learning: A needs analysis. *Humanising Language Teaching*, *15*(5), 1-9.

Erlam, R. (2006). Elicited imitation as a measure of L2 implicit knowledge: An empirical validation study. *Applied Linguistics*, *27*(3), 464-491.

Farahat, T. (2012). Applying the technology acceptance model to online learning in the Egyptian universities. *Procedia-Social and Behavioral Sciences*, *64*, 95-104.

Fedorenko, E., Patel, A., Casasanto, D., Winawer, J., & Gibson, E. (2009). Structural integration in language and music: evidence for a shared system. *Memory & cognition*, *37*(1), 1-9.

Fitch, W. T., & Martins, M. D. (2014). Hierarchical processing in music, language, and action: Lashley revisited. *Annals of the New York Academy of Sciences*, *1316*(1), 87-104.

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233-277). York Press.

Flege, J. E., & Bohn, O.-S. (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.

Foote, J. A., & McDonough, K. (2017). Using shadowing with mobile technology to improve L2 pronunciation. *Journal of Second Language Pronunciation*, *3*(1), 34-56.

Forrin, N. D., & MacLeod, C. M. (2018). This time it's personal: the memory benefit of hearing oneself. *Memory*, *26*(4), 574-579. https://doi.org/10.1080/09658211.2017.1383434

François, C., Chobert, J., Besson, M., & Schön, D. (2013). Music training for the development of speech segmentation. *Cerebral Cortex*, *23*(9), 2038-2043.

Friederici, A. D. (2002). Towards a neural basis of auditory sentence processing. *Trends in cognitive sciences*, *6*(2), 78-84.

Gagnon, M. P., Orruño, E., Asua, J., Abdeljelil, A. B., & Emparanza, J. (2012). Using a modified technology acceptance model to evaluate healthcare professionals' adoption of a new telemonitoring system. *Telemedicine and e-Health*, *18*(1), 54-59.

Gambi, C., & Pickering, M. J. (2013). Prediction and imitation in speech. *Frontiers in psychology*, *4*, 340.

Gao, X., & Hu, J. (2020). From language learning strategy research to a sociocultural understanding of self-regulated learning. In M. J. Raya & F. Vieira (Eds.), *Autonomy in language education: Theory, research, and practice* (pp. 63-73). Routledge.

Gatbonton, E., & Segalowitz, N. (2005). Rethinking communicative language teaching: A focus on access to fluency. *Canadian Modern Language Review*, *61*(3), 325-353.

Genius. (2016, December 30, 2016). *Eminem Proves There Are Plenty Of Words That Rhyme With 'Orange'* [Video]. YouTube. https://www.youtube.com/watch?v=lPcR5RVXHMg

Gilleece, L. F. (2006). *An empirical investigation of the association between musical aptitude and foreign language aptitude* [Doctoral dissertation, Trinity College Dublin]. Centre for Language and Communication Studies. http://hdl.handle.net/2262/77993

Gilman, M., Nix, J., & Hapner, E. (2010). The speech pathologist, the singing teacher, and the singnig voice specialist: Where's the line? *Journal of Singing*, *67*(2), 171-178.

Grabe, E., & Low, E. L. (2008). Durational variability in speech and the rhythm class hypothesis. In *Laboratory phonology 7* (pp. 515-546). De Gruyter Mouton.

Hamada, Y., & Suzuki, Y. (2024). Situating shadowing in the framework of deliberate practice: A guide to using 16 techniques. *RELC Journal, 55*(1), 219-227.

Hamayan, E., Saegert, J., & Larudee, P. (1977). Elicited imitation in second language learners. *Language and Speech*, *20*(1), 86-97.

Ho, Y.-C., Cheung, M.-C., & Chan, A. S. (2003). Music training improves verbal but not visual memory: cross-sectional and longitudinal explorations in children. *Neuropsychology*, *17*(3), 439.

Holec, H. (1981). *1981: Autonomy in Foreign Language Learning.* Pergamon Press.

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). *Applied logistic regression* (Vol. 398). John Wiley & Sons.

Hsu, C.-K. (2015). Learning motivation and adaptive video caption filtering for EFL learners using handheld devices. *ReCALL*, *27*(1), 84-103.

Hsu, H.-h., & Chang, Y.-y. (2013). Extended TAM model: Impacts of convenience on acceptance and use of Moodle. *US-China Education Review*, *3*(4), 211-218.

Hsu, L. (2016). An empirical examination of EFL learners' perceptual learning styles and acceptance of ASR-based computer-assisted pronunciation training. *Computer Assisted Language Learning*, *29*(5), 881-900.

Huang, J. H., Lin, Y. R., & Chuang, S. T. (2007). Elucidating user behavior of mobile learning: A perspective of the extended technology acceptance model. *The electronic library*.

Hwang, Y., & Yi, M. (2002). Predicting the use of web-based information systems: intrinsic motivation and self-efficacy. *AMCIS 2002 Proceedings*, 149.

IFPI. (2021). *IFPI Releases Engaging with Music 2021.* https://www.ifpi.org/ifpi-releases-engaging-with-music-2021/

Inceoglu, S. (2021). Exploring the relationship between explicit instruction, pronunciation awareness, and the development of L2 French connected speech processes. *Language Awareness, 30*(4), 336–354. https://doi.org/10.1080/09658416.2021.1881527

Jakobson, L. S., Cuddy, L. L., & Kilgour, A. R. (2003). Time tagging: A key to musicians' superior memory. *Music Perception*, *20*(3), 307-313.

Janus, M., Lee, Y., Moreno, S., & Bialystok, E. (2016). Effects of short-term music and second-language training on executive control. *Journal of experimental child psychology*, *144*, 84-97.

Jessop, L., Suzuki, W., & Tomita, Y. (2007). Elicited imitation in second language acquisition research. *Canadian Modern Language Review*, *64*(1), 215-238.

Jorgensen, E. R. (2008). *The art of teaching music*. Indiana University Press.

Jungers, M. K., Hupp, J. M., & Dickerson, S. D. (2016). Language priming by music and speech: Evidence of a shared processing mechanism. *Music Perception: An Interdisciplinary Journal*, *34*(1), 33-39.

Kara, Z. E., & Aksel, A. S. (2013). The effectiveness of music in grammar teaching on the motivation and success of the students at preparatory school at Uludağ University. *Procedia-Social and Behavioral Sciences*, *106*, 2739-2745.

Kenji, H., & d'Andrea, D. (1992). Some properties of bilingual maintenance and loss in Mexican background high-school students. *Applied Linguistics, 13*(1), 72-99.

Kennedy, S., & Blanchet, J. (2014). Language awareness and perception of connected speech in a second language. *Language Awareness*, *23*(1-2), 92-106.

Kennedy, S., Blanchet, J., & Trofimovich, P. (2014). Learner pronunciation, awareness, and instruction in French as a second language. *Foreign Language Annals*, *47*(1), 79-96.

Kilgour, A. R., Jakobson, L. S., & Cuddy, L. L. (2000). Music training and rate of presentation as mediators of text and song recall. *Memory & cognition*, *28*(5), 700-710.

King, W. R., & He, J. (2006). A meta-analysis of the technology acceptance model. *Information & management*, *43*(6), 740-755.

Koelsch, S. (2011). Toward a neural basis of music perception–a review and updated model. *Frontiers in psychology*, *2*, 110.

Kraus, N., Slater, J., Thompson, E. C., Hornickel, J., Strait, D. L., Nicol, T., & White-Schwoch, T. (2014). Music enrichment programs improve the neural encoding of speech in at-risk children. *Journal of Neuroscience*, *34*(36), 11913-11918.

Krizman, J., Slater, J., Skoe, E., Marian, V., & Kraus, N. (2015). Neural processing of speech in children is influenced by extent of bilingual experience. *Neuroscience letters*, *585*, 48-53.

Lai, C., Yeung, Y., & Hu, J. (2016). University student and teacher perceptions of teacher roles in promoting autonomous language learning with technology outside the classroom. *Computer Assisted Language Learning*, *29*(4), 703-723.

Lappe, C., Herholz, S. C., Trainor, L. J., & Pantev, C. (2008). Cortical plasticity induced by short-term unimodal and multimodal musical training. *Journal of Neuroscience*, *28*(39), 9632-9639.

Lee, Y., Kozar, K. A., & Larsen, K. R. (2003). The technology acceptance model: Past, present, and future. *Communications of the Association for information systems*, *12*(1), 50.

Levey, S., Levey, T., & Fligor, B. J. (2011). Noise exposure estimates of urban MP3 player users. *Journal of Speech, Language, and Hearing Research, 54*(1)*,* 263-277. https://doi.org/10.1044/1092-4388(2010/09-0283

Levitin, D. J. (2006). *This is your brain on music: The science of a human obsession*. Penguin.

Li, X., & Brand, M. (2009). Effectiveness of music on vocabulary acquisition, language usage, and meaning for mainland Chinese ESL learners. *Contributions to music education*, 73-84.

Lin, C.-c. (2014). Learning English reading in a mobile-assisted extensive reading program. *Computers & Education*, *78*, 48-59.

Lindgren, E., & Muñoz, C. (2013). The influence of exposure, parents, and linguistic distance on young European learners' foreign language comprehension. *International Journal of Multilingualism*, *10*(1), 105-129.

Little, D., & Thorne, S. L. (2017). From learner autonomy to rewilding: A discussion. *Learner autonomy and web*, *2*, 12-35.

Liu, F., Jiang, C., Wang, B., Xu, Y., & Patel, A. D. (2015). A music perception disorder (congenital amusia) influences speech comprehension. *Neuropsychologia*, *66*, 111-118.

Long, M. (2014). *Second language acquisition and task-based language teaching*. John Wiley & Sons.

Luk, G., & Bialystok, E. (2013). Bilingualism is not a categorical variable: Interaction between language proficiency and usage. *Journal of Cognitive Psychology, 25*(5), 605-621.

Ludke, K. M. (2018). Singing and arts activities in support of foreign language learning: An exploratory study. *Innovation in Language Learning and Teaching*, *12*(4), 371-386.

Ludke, K. M., Ferreira, F., & Overy, K. (2014). Singing can facilitate foreign language learning. *Memory & cognition, 42*, 41-52.

Mackenzie, C., & Stoljar, N. (2000). Introduction: Autonomy refigured. In C. Mackenzie & N. Stoljar (Eds.), *Relational autonomy: Feminist perspectives on autonomy, agency, and the social self* (pp. 3-31). Oxford University Press.

Maess, B., Koelsch, S., Gunter, T. C., & Friederici, A. D. (2001). Musical syntax is processed in Broca's area: an MEG study. *Nature neuroscience*, *4*(5), 540-545.

Mainwaring, J. (1933). Kinaesthetic factors in the recall of musical experience. *British journal of psychology*, *23*(3), 284.

Marangunić, N., & Granić, A. (2015). Technology acceptance model: a literature review from 1986 to 2013. *Universal access in the information society*, *14*(1), 81-95.

Melnik, G. A., & Peperkamp, S. (2021). High-Variability Phonetic Training enhances second language lexical processing: evidence from online training of French learners of English. *Bilingualism: Language and Cognition, 24*(3):497-506. doi:10.1017/S1366728920000644

Milovanov, R. (2009). Musical aptitude and foreign language learning skills-neural and behavioral evidence about their connections. *Proceedings of the 7th Triennial Conference of European Society for the Cognitive Sciences of Music (*ESCOM 2009). Jyväskylä, Finland.

Milovanov, R., Huotilainen, M., Välimäki, V., Esquef, P. A., & Tervaniemi, M. (2008). Musical aptitude and second language pronunciation skills in school-aged children: Neural and behavioral evidence. *Brain Research*, *1194*, 81-89.

Milovanov, R., Tervaniemi, M., Takio, F., & Hämäläinen, H. (2007). Modification of dichotic listening (DL) performance by musico-linguistic abilities and age. *Brain Research*, *1156*, 168-173.

Mlott, S. R., Marcotte, D. B., & Lira, F. T. (1976). The efficacy of programmed instruction in the training of paraprofessionals. *Journal of clinical psychology*, *32*(2), 419-424.

Mora, C. F. (2000). Foreign language acquisition and melody singing. *ELT journal*, *54*(2), 146-152.

Moreno, S. (2009). Can music influence language and cognition? *Contemporary Music Review*, *28*(3), 329-345.

Murphey, T. (1990). The song stuck in my head phenomenon: a melodic Din in the LAD? *System*, *18*(1), 53-64.

Murphey, T. (1992). The discourse of pop songs. *TESOL Quarterly*, *26*(4), 770-774.

Murray, G. (2004). Two stories of self-directed language learning. In: Reinders, H., Anderson, H., Hobbs, M. & Jones-Parry, J. (Eds.) *Supporting independent learning in the 21st century. Proceedings of the inaugural conference of the Independent Learning Association, Melbourne September 13-14, 2003,* (p.112-120). Independent Learning Association Oceania.

Murray, G. (2020). Learner Autonomy and Holec's Model: A Complexity Perspective. In M. J. Raya & F. Vieira (Eds.), *Autonomy in Language Education* (pp. 89-102). Routledge.

Murray, G. (2017). *Social spaces for language learning: Stories from the L-café.* Springer.

Nardo, D., & Reiterer, S. M. (2009). *Musicality and phonetic language aptitude*. De Gruyter Mouton.

I.S.P. Nation. (2001). *Learning Vocabulary in Another Language.* Cambridge University Press.

Nix, J. (2015). Singing vs. Speaking. [PDF]. Retrieved from https://music.utsa.edu/pdfs/61_SpeakingvsSinging.pdf

Ortega, L. (2009). *Understanding Second Language Acquisition*. Hodder Education.

Oulton, C. S. (2010). *Focus on form through singing in a first grade Yugtun immersion classroom*. University of Alaska Fairbanks.

Parasuraman, A. (2000). Technology Readiness Index (TRI) a multiple-item scale to measure readiness to embrace new technologies. *Journal of service research*, *2*(4), 307-320.

Pastuszek-Lipinska, B. (2008). Influence of music education on second language acquisition. *Journal of the Acoustical Society of America*, *123*(5), 3737.

Patel, A. D. (1998). Syntactic processing in language and music: different cognitive operations, similar neural resources? *Music Perception*, *16*(1), 27-42.

Patel, A. D. (2003). Language, music, syntax and the brain. *Nature neuroscience*, *6*(7), 674-681.

Patel, A. D. (2008). Talk of the Tone. *Nature*, *453*(7196), 726-727.

Patel, A. D. (2011). Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Frontiers in Psychology*, *2*, 142. https://doi.org/10.3389/fpsyg.2011.00142

Patel, A. D. (2013). Sharing and nonsharing of brain resources for language and music. In M. A. Arbib (Ed.), *Language, Music, and the Brain: A Mysterious Relationship* (pp. 329-355). MIT Press.

Patel, A. D., & Daniele, J. R. (2003). An empirical comparison of rhythm in language and music. *Cognition*, *87*(1), B35-B45.

Patel, A. D., Xu, Y., & Wang, B. (2010). The role of F0 variation in the intelligibility of Mandarin sentences. In *Proceedings of Speech Prosody 2010-Fifth International Conference*. May 11-14, 2010. Chicago, USA.

Pavia, N., Webb, S., & Faez, F. (2019). Incidental vocabulary learning through listening to songs. *Studies in Second Language Acquisition*, *41*(4), 745-768.

Pavlenko, A. (2013). The affective turn in SLA: From 'affective factors' to 'language desire' and 'commodification of affect'. In D. Gabryś-Barker & J. Bielska (Eds.), The Affective

Dimension in Second Language Acquisition (pp. 1-28). Multilingual Matters. https://doi.org/10.21832/9781847699695-001

Peretz, I., & Coltheart, M. (2003). Modularity of music processing. *Nature neuroscience*, *6*(7), 688-691.

Porter, C. E., & Donthu, N. (2006). Using the technology acceptance model to explain how attitudes determine Internet usage: The role of perceived access barriers and demographics. *Journal of business research*, *59*(9), 999-1007.

Racette, A., & Peretz, I. (2007). Learning lyrics: to sing or not to sing? *Memory & cognition*, *35*(2), 242-253.

Raya, M. J., & Vieira, F. (2020). *Autonomy in Language Education: Theory, Research and Practice*. Routledge.

Reeves, T. C., Herrington, J., & Oliver, R. (2005). Design research: A socially responsible approach to instructional technology research in higher education. *Journal of computing in higher education*, *16*(2), 96-115.

Reinders, H. (2020). A framework for learning beyond the classroom. In M. J. Raya & F. Vieira (Eds.), *Autonomy in language education: Theory, research and practice* (pp. 63-73). Routledge.

Reinders, H., & Balçikanli, C. (2011). Do classroom textbooks encourage learner autonomy? *Novitas-ROYAL, 5*(2), 265-272.

Reinders, H., & Benson, P. (2017). Research agenda: Language learning beyond the classroom. *Language teaching*, *50*(4), 561-578.

Rose, M. (2016). *Use of music in adult second language instruction: A Canadian perspective* [Doctoral dissertation, Université du Québec à Chicoutimi]. https://constellation.uqac.ca/id/eprint/4081/

Ross, B. (2017). The Causal-manipulative Approach to Musical Meaning. *International Review of the Aesthetics and Sociology of Music*, 3-17.

Ruel, E., Wagner III, W. E., & Gillespie, B. J. (2015). *The practice of survey research: Theory and applications*. Sage Publications.

Rukholm, V. N. (2011). *Facilitating lexical acquisition in beginner learners of Italian through popular song* [Doctoral Disseration, University of Toronto]. https://library-archives.canada.ca/eng/services/services-libraries/theses/Pages/item.aspx?idNumber=1019488034

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*(5294), 1926-1928.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of memory and language*, *35*(4), 606-621.

Salcedo, C. S. (2010). The effects of songs in the foreign language classroom on text recall, delayed text recall and involuntary mental rehearsal. *Journal of College Teaching & Learning (TLC)*, *7*(6). 19-30.

Saldaña, J. (2021). *The coding manual for qualitative researchers.* Sage.

Salimpoor, V. N., Benovoy, M., Larcher, K., Dagher, A., & Zatorre, R. J. (2011). Anatomically distinct dopamine release during anticipation and experience of peak emotion to music. *Nature neuroscience*, *14*(2), 257-262.

Salimpoor, V. N., van den Bosch, I., Kovacevic, N., McIntosh, A. R., Dagher, A., & Zatorre, R. J. (2013). Interactions between the nucleus accumbens and auditory cortices predict music reward value. *Science*, *340*(6129), 216-219.

Salimpoor, V. N., Zald, D. H., Zatorre, R. J., Dagher, A., & McIntosh, A. R. (2015). Predictions and the brain: how musical sounds become rewarding. *Trends in cognitive sciences*, *19*(2), 86-91.

Sarsarabi, S., & Sorouri, Z. (2016). The effect of music on second language pronunciation: Based on multiple intelligences theory. *International Journal of Humanities and Cultural Studies (IJHCS) ISSN 2356-5926*, *3*(1), 1475-1487.

Sauro, S. (2005). English language learning and technology: Lectures on applied linguistics in the age of information and communication technology. *Studies in Second Language Acquisition*, *27*(3), 486-487.

Sauro, S., & Zourou, K. (2019). What are the digital wilds? *Language Learning & Technology, 23*(1), 1-7. https://doi.org/10125/44666

Schön, D., Boyer, M., Moreno, S., Besson, M., Peretz, I., & Kolinsky, R. (2008). Songs as an aid for language acquisition. *Cognition*, *106*(2), 975-983.

Schön, D., & Tillmann, B. (2015). Short-and long-term rhythmic interventions: Perspectives for language rehabilitation. *Annals of the New York Academy of Sciences*, *1337*(1), 32-39.

Shuter-Dyson, R. (1999). Musical ability. In D. Deutsch (Ed.), *The Psychology of Music* (2nd ed., pp. 627-651). Academic Press.

Slater, J., Skoe, E., Strait, D. L., O'Connell, S., Thompson, E., & Kraus, N. (2015). Music training improves speech-in-noise perception: Longitudinal evidence from a community-based music program. *Behavioural Brain Research*, *291*, 244-252.

Slevc, L. R., & Miyake, A. (2006). Individual differences in second-language proficiency: Does musical ability matter? *Psychological Science*, *17*(8), 675-681.

Sposet, B. A. (2008). *The role of music in second language acquisition: A bibliographical review of seventy years of research, 1937-2007*. E. Mellen Press.

Stickler, U., & Emke, M. (2011). Tandem Learning in Virtual Spaces: Supporting Non-formal and Informal Learning in Adults. In P. Benson & H. Reinders (Eds.), *Beyond the Language Classroom* (pp. 146-160). Palgrave Macmillan UK. https://doi.org/10.1057/9780230306790_12

Sundberg, R., & Cardoso, W. (2016). Aligning out-of-class material with curriculum: Tagging grammar in a mobile music application. *CALL communities and culture–short papers from EUROCALL*, 440-444.

Sundberg, R., & Cardoso, W. (2019). Learning French through music: the development of the Bande à Part app. *Computer Assisted Language Learning*, *32*(1-2), 49-70.

Sundberg, R., & Cardoso, W. (2022). The Feasibility of Using Bande à Part to Aid French Language Learners. *CALICO Journal, 39*(2), 196-218.

Sundqvist, P. (2019). Commercial-off-the-shelf games in the digital wild and L2 learner vocabulary. *Language Learning & Technology*, *23*(1), 87-113.

Sundqvist, P., & Sylvén, L. K. (2016). *Extramural English in teaching and learning: From theory and research to practice*. Springer.

Tegge, F. (2017). The lexical coverage of popular songs in English language teaching. *System*, *67*, 87-98.

Thorne, S. L., Hellermann, J., & Jakonen, T. (2021). Rewilding language education: Emergent assemblages and entangled actions. *The Modern Language Journal*, *105*(S1), 106-125.

Tost, G. (2013). Bettering pronunciation through reading aloud and peer appraisal. *Bellaterra Journal of Teaching & Learning Language & Literature*, *6*(1), 35-55.

Trofimovich, P., & Isaacs, T. (2012). Disentangling accent from comprehensibility. *Bilingualism: Language and Cognition*, *15*(4), 905-916.

Trofimovich, P., Kennedy, S., & Blanchet, J. (2017). Development of second language French oral skills in an instructed setting: A focus on speech ratings. *Canadian Journal of Applied Linguistics/Revue canadienne de linguistique appliquée*, *20*(2), 32-50.

Tsai, Y.-R. (2015). Applying the Technology Acceptance Model (TAM) to explore the effects of a Course Management System (CMS)-Assisted EFL writing instruction. *CALICO Journal*, *32*(1), 153-171.

Tseng, W.-T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: The case of self-regulation in vocabulary acquisition. *Applied Linguistics*, *27*(1), 78-102.

Turnbull, D., Gupta, C., Murad, D., Barone, M., & Wang, Y. (2017). Using music technology to motivate foreign language learning. 2017 International Conference on Orange Technologies (ICOT),

Van Belle, G. (2011). *Statistical rules of thumb* (Vol. 699). John Wiley & Sons.

Vandergrift, L., & Goh, C. (2012). *Teaching and Learning second langauge listening: Metacognition in action*. Routledge.

Venkatesh, V., & Davis, F. D. (1996). A model of the antecedents of perceived ease of use: Development and test. *Decision sciences*, *27*(3), 451-481.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management science*, *46*(2), 186-204.

Vygotsky, L. S. (1978). Socio-cultural theory. *Mind in society*, *6*, 52-58.

Wells, G. (2000). Dialogic inquiry: Towards a sociocultural practice and theory of education. *Harvard Educational Review*, *70*(2), 228-230.

Werner, R. (2018). Music, movement and memory: Pedagogical songs as mnemonic aids. *TESOL Journal*, *9*(4), 1-11.

Werner, V. (2020). "Song-advantage" or "cost of singing"? A research synthesis of classroom-based intervention studies applying lyrics-based language teaching (1972–2019). *Journal of Second Language Teaching & Research*, *8*(1), 138-170.

Werner, V., & Tegge, F. (2020). Introduction: Learning languages through pop culture/learning about pop culture through language education. In V. Werner & F. Tegge (Eds.), *Pop Culture in Language Education* (pp. 1-30). Routledge.

Wilcox, W. B. (1995). *Music cues from classroom singing for second language acquisition: Prosodic memory for pronunciation of target vocabulary by adult nonnative English speakers* University of Kansas].

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. *Self-regulated learning and academic achievement: Theoretical perspectives*, *2*, 153-189.

Yanguas, I. (2009). Multimedia glosses and their effect on L2 text comprehension and vocabulary learning. *Language Learning & Technology*, *13*(2), 48-67.

Yih, M. B., & Nah, E. A. (2009). Writing web logs in the ESL classroom: A study of student perceptions and the Technology Acceptance Model. *Asian Journal of University Education Vol. 5 No. 1 June 2009 ISSN 1823*, *7797*.

Yoshii, M. (2006). L1 and L2 glosses: Their effects on incidental vocabulary learning. *Language Learning & Technology*, *10*(3), 85-101.

Zatorre, R. J., & Gandour, J. T. (2008). Neural specializations for speech and pitch: moving beyond the dichotomies. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *363*(1493), 1087-1104.

Zhang, S., Zhao, J., & Tan, W. (2008). Extending TAM for online learning systems: An intrinsic motivation perspective. *Tsinghua science and technology*, *13*(3), 312-317.

Zheng, C., Saito, K., & Tierney, A. (2022). Successful second language pronunciation learning is linked to domain-general auditory processing rather than music aptitude. *Second Language Research, 38*(3), 477-497.

Zimmerman, B. J. (1994). Dimensions of academic self-regulation: A conceptual framework for education. *Self-regulation of learning and performance: Issues and educational applications*, *1*, 33-21.

Zimmerman, B. J. (1998). Developing self-fulfilling cycles of academic regulation: An analysis of exemplary instructional models. In D. H. Schunk & B. J. Zimmerman (Eds.), *Self-regulated learning: From teaching to self-reflective practice* (pp. 1-19). Guilford Press.

Appendix A: Self-regulation Survey

Follow-up

Below is a series of statements about your pronunciation learning experience. We would like to know how far these statements match your own perceptions, that is, your personal view. There are no 'right' or 'wrong' answers. Moreover, the data we collect are for research purposes and your opinions will be respected and kept confidential.

There are twenty items in total in the questionnaire. Please tick the appropriate box concerning your personal pronunciation learning experience. Thank you very much for your cooperation!

Indicate below the degree to which you agree or disagree with each statement. *

Please choose the appropriate response for each item:

	Strongly Disagree	Disagree	Slightly Disagree	Partly Agree	Agree	Strongly Agree
Once the novelty of learning pronunciation is gone, I easily become impatient with it.	◯	◯	◯	◯	◯	◯
When I feel stressed about pronunciation learning, I know how to reduce this stress.	◯	◯	◯	◯	◯	◯
When I am studying pronunciation and the learning environment becomes unsuitable, I try to sort out the problem.	◯	◯	◯	◯	◯	◯
When learning pronunciation, I have special techniques to achieve my learning goals.	◯	◯	◯	◯	◯	◯
When learning pronunciation, I have special techniques to keep my concentration focused.	◯	◯	◯	◯	◯	◯
I feel satisfied with the methods I use to reduce the stress of pronunciation learning.	◯	◯	◯	◯	◯	◯
When learning pronunciation, I believe I can achieve my goals more quickly than expected.	◯	◯	◯	◯	◯	◯

During the process of learning pronunciation, I feel satisfied with the ways I eliminate boredom.	◯	◯	◯	◯	◯	◯
When learning pronunciation, I think my methods of controlling my concentration are effective.	◯	◯	◯	◯	◯	◯
When learning pronunciation, I persist until I reach the goals that I make for myself.	◯	◯	◯	◯	◯	◯
When it comes to learning pronunciation, I have my special techniques to prevent procrastination.	◯	◯	◯	◯	◯	◯
When I feel stressed about pronunciation learning, I simply want to give up.	◯	◯	◯	◯	◯	◯
I believe I can overcome all the difficulties related to achieving my pronunciation learning goals.	◯	◯	◯	◯	◯	◯
When learning pronunciation, I know how to arrange the œ environment to make	◯	◯	◯	◯	◯	◯

learning more efficient.						
When I feel stressed about my pronunciation learning, I cope with this problem immediately.	◯	◯	◯	◯	◯	◯
When it comes to learning pronunciation, I think my methods of controlling procrastination are effective.	◯	◯	◯	◯	◯	◯
When learning pronunciation, I am aware that the learning environment matters.	◯	◯	◯	◯	◯	◯
During the process of learning pronunciation, I am confident that I can overcome any sense of boredom.	◯	◯	◯	◯	◯	◯
When feeling bored with learning pronunciation, I know how to regulate my mood in order to invigorate the learning process.	◯	◯	◯	◯	◯	◯
When I study pronunciation, I look for a good learning environment.	◯	◯	◯	◯	◯	◯

Appendix B: Background Language Questionnaire

Listening to Yourself TS/T

Thank you for choosing to participate in this research. If you have questions at anytime, please email Ross Sundberg at ross.sundberg@mail.concordia.ca or call (438) 822-0300.
There are 158 questions in this survey.

Consent

Please read the following Consent Form (click here (JavaScript:window.open("https://www.rsundberg.ca/Consent_Form.pdf");)). By continuing to the next section you are acknowledging that you have read and understood this form. Only **Firefox** or **Chrome** browsers are compatible.

*

Please choose **all** that apply:

☐ I agree to participate in the research under the conditions described

Background Questionnaire

Name: *

Please write your answer here:

Email: *

Please write your answer here:

Age: *

Please write your answer here:

City: *

Please write your answer here:

Gender *

Please choose **only one** of the following:

- ◯ Male
- ◯ Female
- ◯ Prefer Not To Say
- ◯ Other

Languages Spoken

A few questions on the languages you know.

You will now be asked to self-rate your ability in the languages you know. Fill in the name of your first language. *

Please write your answer here:

Rate your ability in {PQ4A}, from 1-7 *

Please choose the appropriate response for each item:

	Very poor	2	3	4	5	6	Very Good
Listening	◯	◯	◯	◯	◯	◯	◯
Reading	◯	◯	◯	◯	◯	◯	◯
Speaking	◯	◯	◯	◯	◯	◯	◯
Writing	◯	◯	◯	◯	◯	◯	◯

Fill in the name of the next language you know. If you have already rated all the languages you know, leave the box empty.

Please write your answer here:

Rate your ability in {PQ4B}, from 1-7

Please choose the appropriate response for each item:

	Very poor	2	3	4	5	6	Very good
Listening	◯	◯	◯	◯	◯	◯	◯
Reading	◯	◯	◯	◯	◯	◯	◯
Speaking	◯	◯	◯	◯	◯	◯	◯
Writing	◯	◯	◯	◯	◯	◯	◯

Fill in the name of the next language you know. If you have already rated all the languages you know, leave the box empty.

Please write your answer here:

Appendix C: Sample Lyrics

READ ALOUD or IMITATE (Note: Song titles are in blue; they are not part of the lyrics)

Si Tu Es Un Homme

N'y va pas par quatre chemins
Je ne mène pas à Rome
Mais ne craint d'aller trop loin
Si tu es un homme
N'essaie pas de m'impressionner
Je me passerai du décorum
Mais n'aie pas peur de pleurer
Si tu es un homme

Tous les mêmes

Vous les hommes,
Vous êtes tous les mêmes
Macho mais cheap,
bande de mauviettes infidèles
Lorsque je n'serai plus belle
Ou du moins au naturel
Arrête je sais que tu mens
Il n'y a que Kate Moss qui est éternelle
Tous les mêmes et y en a marre

Laisse tomber les filles

La chance abandonne
Celui qui ne sait
Que laisser les coeurs blessés
Tu n'auras personne
Pour te consoler
Tu ne l'auras pas volé

Parlez-moi de lui

Alors dites moi
Si elle est jolie
Plus jolie que moi
Et lui dans ses yeux
Était-il heureux?
Oh, je vous en prie
Même si j'ai mal
Parlez-moi de lui

Je veux

Je veux de l'amour, de la joie, de la bonne humeur

Ce n'est pas votre argent qui fera mon bonheur
Moi je veux crever la main sur le coeur
Allons ensemble, découvrir ma liberté
Oubliez donc tous vos clichés
Bienvenue dans ma réalité
J'en ai marre de vos bonnes manières,
C'est trop pour moi
Moi je mange avec les mains
Et je suis comme ça
Je parle fort et je suis franche, excusez-moi
Finie l'hypocrisie, moi je me casse de là
J'en ai marre des langues de bois
Regardez-moi, de toute manière je vous en veux pas
Et je suis comme ça !

Adieu

Tu me menaces de partir
Parce que je hurle quand
Tu chantes tes souvenirs
Eh bien, chéri, prends donc la porte
Car tu sais que plus rien ne m'importe
Je n'aurai donc plus à t'entendre
Rentrer la nuit quand j'attends l'aube
Qui arrive en poussant les heures
Moi je me lève et toi

La terre est ronde

Tu as besoin d'une voiture
pour aller travailler,
Tu travailles pour rembourser
La voiture que tu viens d'acheter.

Tout le monde en même temps

Si on soulevait la poussière
Debout la face au vent
Est-ce que on retrouverait tous nos repères

hymn a l'amour

Le ciel bleu sur nous peut s'effondrer
Et la terre peut bien s'écrouler
Peu m'importe si tu m'aimes
Je me fous du monde entier
Tant que l'amour inondera mes matins
Tant que mon corps frémira sous tes mains

Peu m'importent les problèmes
Mon amour puisque tu m'aimes

Jersey

J'aimerais. J'aimerais
Prendre la mer vers
Mon Hawaï à moi. Mon Hawaï à moi.
J'aimerais... J'aimerais
Prendre le temps
De t'emmener avec moi.

Papaoutai

Monsieur Je-sais-tout en aurait hérité, c'est ça
Faut l'sucer d'son pouce ou quoi ?
Dites-nous où c'est caché, ça doit
Faire au moins mille fois que on a bouffé nos doigts

Slow

J'aurais pu jouer à la Sophie Marceau
J'aurais pu tomber dans tes bras pour un slow
Être la plus belle les mains sur ton dos
Mais je ne suis pas comme ça
Marilou et les autres elles ont l'habitude
Toutes ces filles-là misent tout sur l'attitude
Et moi j'ai mes passes temps de solitude
Comme regarder des films muets
M'accordiez-vous cette danse as-tu dis ?
Un autre voudrait que j'y aille avec lui
Serais-je ta haine ou la sienne sur la piste
Je préfère jouer à la console
Ces prises de têtes ces prises de bêtes m'indiffère
Je préfère écouter la chanson de Prévert
Pour rembarrer avec mon plus beau Molière
Je n'ai pas envie de danser

Appendix D: Sample Survey Questions

Practice

Please read and record each of the following sentences. To do so:

1. Open this popup window and click "record". Press "stop" when you have finished reading the phrase.

2. Listen to your recording and rate how accurate your pronunciation sounded to you in the box provided. (You may have to resize the pop-up window in order to see your recordings).

3. After rating your pronunciation, record a second attempt followed by a second rating.

DO NOT CLOSE the pop-up window until you have finished the entire survey. If you close it by accident, reopen it using the link above.

RECORD any technical problems you encounter in the text box at the bottom of the page.

*Read and record the following:

"Et puis"

	1	2	3	4	5
Attempt 1. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○
Attempt 2. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○

*Read and record the following:

"Je le savais"

	1	2	3	4	5
Attempt 1. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○
Attempt 2. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○

*Read and record the following:

"Tout au long"

	1	2	3	4	5
Attempt 1. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○
Attempt 2. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○

*Read and record the following:

"Que tu saurais"

	1	2	3	4	5
Attempt 1. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○
Attempt 2. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○

Please relisten to the recordings in your open pop-up window.

Then press "record" and provide your thoughts on what you notice about your pronunciation.

If you have encountered any technical problems you can record them in the text box below.

T1

Please listen to the following audio files and record yourself imitating each of the phrases. To do so:

1. Click "record". Press "stop" when you have finished imitating the phrase.

2. Listen to your recording and rate how accurate your pronunciation sounded to you in the box provided. (You may have to resize the pop-up window in order to see your recordings).

3. After rating your pronunciation, record a second attempt followed by a second rating.

DO NOT CLOSE the pop-up window until you have finished the survey. If you close it by accident: Open this popup window.

RECORD any technical problems you encounter in the text box at the bottom of the page.

*Listen to then record the following:

▶ 0:00 / 0:02

	1	2	3	4	5
Attempt 1. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○
Attempt 2. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○

*Listen to then record the following:

▶ 0:00 / 0:02

	1	2	3	4	5
Attempt 1. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○
Attempt 2. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○

*Listen to then record the following:

"Mais ne craint d'aller trop loin"

	1	2	3	4	5
Attempt 1. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○
Attempt 2. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○

*Listen to then record the following:

"Si tu es un homme"

	1	2	3	4	5
Attempt 1. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○
Attempt 2. How was the accuracy of your pronunciation? 5 = excellent, 1 = poor	○	○	○	○	○

Please relisten to the recordings in your open pop-up window.

Then, open this popup window and record your thoughts on what you notice about your pronunciation. You will have 30 seconds to record.

You can close your windows when you have finished the think aloud and rating your speech samples. I you have encountered any technical problems you can record them in the text box below.