Gamifying High Variability Phonetic Training: The development of the *FineTune* app

Sivan Black-Rotchin

A Thesis

in

The Department

of

Education

Presented in Partial Fulfillment of the Requirements

for the Degree of Master of Arts (Applied Linguistics) at

Concordia University

Montréal, Québec, Canada

April, 2022

CONCORDIA UNIVERSITY

School of Graduate Studies

This is to certify that the thesis prepared

By: Sivan Black-Rotchin

Entitled: Gamifying High Variability Phonetic Training: The development of the FineTune app

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

Master of Arts (Applied Linguistics)

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

Signed by the final Examining Committee:

_____ Examiner

T. Hernandez-Gonzalez

_____ Thesis Supervisor

W. Cardoso

Approved by

_____

J. Capobianco Chair of Department or Graduate Program Director

_____

Pascale Sicotte, Dean

ABSTRACT

Gamifying High Variability Phonetic Training: The development of the *FineTune* app

Sivan Black-Rotchin

High Variability Phonetic Training (HVPT), an established method of improving the learner's perception and potentially production of second language (L2) segments, has received considerable attention by the research community (Barriuso & Hayes-Harb, 2018). However, this technique is unknown to most teachers and rarely translated into useable pedagogical tools (Thomson, 2018a). Thomson notes that one reason for this disconnect may be that in its current lab-based form, HVPT is neither accessible nor engaging to L2 learners. In order to mitigate both issues, he suggests implementing this training into a mobile application and adding game elements to the existing HVPT paradigm. He posits that such changes may help in bridging the gap between pronunciation research and pedagogical practices.

This thesis responds to this recommendation by presenting the blueprint for *FineTune*, a mobile-based application that combines HVPT with research-driven gamification techniques, with the aim of helping learners improve their perception of English speech sounds. In the training, learners hone their perception of vowels and consonants by progressing through levels of gamified forced choice identification tasks. The development of *FineTune* is informed by empirically motivated frameworks and theories in the fields of Mobile-Assisted Language Learning, HVPT, and gamification. For instance, the game elements incorporated in *FineTune* are selected on the basis of van Roy and Zaman's (2017) gamification heuristics for educational contexts, and Deci and Ryan's (2000) Self-Determination Theory (thus chosen based on their ability to satisfy the user's feelings of autonomy, competence, and relatedness).

This thesis represents stage one in Cardoso's (2022) chronological framework for the development of CALL material: outlining the development and rationale of the proposed tool (*FineTune*), and concludes with suggestions for how to proceed with stage two of Cardoso's framework: exploring the pedagogical affordances and user attitudes towards the app.

Acknowledgements

I try to approach all tasks in life, no matter how daunting, with a sense of playfulness. I would like to thank my dear supervisor Dr. Walcir Cardoso, for being right there with me in this approach throughout the thesis process. Beyond providing invaluable knowledge and insight into research and academic writing, your positivity, openness and ever-present support truly made this accomplishment possible. I would also like to thank Dr. Teresa Hernandez-Gonzalez for her feedback and guidance on both my proposal and thesis. It is because of your class that I first became interested in gamification, and you continue to inspire me in so many ways.

My family is the foundation for all that I do, and I can't thank them enough for their support and love. I also owe a very special thanks to my friends at Concordia University, in particular Hamidreza Khademi, Alexander Lamont, the girls at TESL RC, and of course, Nina Le. Nina, you were my partner throughout the entire MA. We did it!

I am grateful for my little feline companions, Dandelion and Sumac. They were always there to distract me when I needed a break from it all. I am also forever grateful for my big human companion, Matthew. Somehow, your presence just makes everything easier.

Table of Contents

LIST OF FIGURES viii

LIST OF TABLES ix

Chapter One 1

Introduction 1

L2 Pronunciation and HVPT 3

Gamification and Mobile L2 Learning 4

This Thesis 5

Chapter Two 7

Background 10

High Variability Phonetic Training 10

HVPT and Computer-Assisted Language-Learning 13

Mobile Assisted Language-Learning 13

Gamification 15

Self-Determination Theory 17

Autonomy and game elements 19

Competence and game elements 20

Relatedness and game elements 21

Van Roy and Zaman's Nine Gamification Heuristics 25

FineTune: Development 27

FineTune: Description 28

User experience 29

HVPT in FineTune 35

Gamification in FineTune 37

Stockwell and Hubbard's (2013) MALL Principles in FineTune 43

Conclusion 46

Chapter 3 48

Summary of Goals 48

Implications for Education and Research 49

The Next Stages 50

References 52

LIST OF FIGURES

Figure 1. Stages of motivation 18

Figure 2. FineTune: Main graphical user interface 30

Figure 3. Badges displayed in FineTune 31

Figure 4. Playing FineTune 33

LIST OF TABLES

Table 1. *Stockwell and Hubbard's (2013) ten principles for the development of MALL applications* 15

Table 2. *Summary of game elements and their connection to SDT* 24

Table 3. *HVPT in FineTune* 37

Table 4. *SDT and gamification in FineTune* 41

Table 5. *Stockwell and Hubbard's (2013) ten principles and their application in FineTune* 45

Gamifying High Variability Phonetic Training: The development of the *FineTune* app

Chapter One

Introduction

I have always loved the times that we would learn a little bit of pronunciation in the language classes that I had taken over the years. They were like little treats in the class – an active break from the heady class material. I remember stretching the corners of my mouth to produce the German palatal fricative /ç/ (as "ch" in "ich"), straining to hear the difference between an aspirated and unaspirated "t" sound in Greek, or putting my hands on my throat to simulate the strangled "ha" sound in Arabic. Sometimes it would feel goofy attempting (and often failing) to get these sounds just right, but it also seemed integral to the language-learning process. In these moments, I felt as though I were honouring and becoming closer to the language and the speech community by making the effort to sound correct with the minute details that the production of these sounds requires. However, these active moments of pronunciation learning were few and far between; more often than not, class time was devoted to other aspects of language, such as grammar and vocabulary.

For this reason, I always felt that the way one sounded was relatively unimportant when it came to being a competent speaker of a language, as compared to one's vocabulary and grammar skills. However, that belief was put into question one small moment in a hostel in San Sebastián, Spain, where I was working as a receptionist after graduating university. I was extremely self-conscious about my Spanish skills and frequently worried about how clients would perceive my poor grammar and bare-bones vocabulary. One evening, I expressed these concerns to a colleague. Surprisingly, my colleague held a very different perspective; while he could not really pinpoint *what* it was about my speech that sounded correct to him, he told me that I was easy to

understand, which signalled to him that I had a respectable command of his native language. This was surprising to me, and after that conversation, I paid special attention to people's speech; specifically, what was it about the way they spoke that communicated the message to me, the listener, that the language belonged to them. I was forever in search of these small details of speech that would usually go unnoticed. However, without the knowledge of what exactly I was looking for, it certainly made it difficult to pinpoint these substantive sounds.

This interest in speech sounds carried over into my classes once I became an ESL teacher. I made a point of devoting some class time to pronunciation, and my students seemed to appreciate this. Oftentimes, they would share with me that even though they had been studying English their whole lives, they were never taught anything related to pronunciation. However, as much as I would try to fit some pronunciation into my lessons, there was rarely enough time to do this in a way that demonstrated tangible learning gains. Moreover, I was often at a loss for where to source material that reflected contemporary research in the field, as I was not trained in teaching pronunciation – an issue that is not unique to me as a teacher (for similar claims, see Foote et al., 2012; Breitkreuz et al., 2001). Occasionally, I would have students ask for material to practice their pronunciation beyond the classroom (especially adult learners who would use English daily in their work lives), and I was unsure where to direct them. Many of my students, especially those whose work involved using English on the phone, giving interviews with English media, and frequent presentations in English, were concerned with how to improve their overall intelligibility and the longstanding "speech habits" that affected their career advancement. As their teacher, I was unsure what tools and material would target precisely these concerns; namely, their ability to intelligibly and consistently produce individual speech sounds

(as opposed to simply learning the correct pronunciation of standalone words), so as to increase their overall confidence when speaking.

L2 Pronunciation and HVPT

When most people think of the skills involved in pronunciation learning, the tendency is to focus on oral production. However, an equally necessary aspect of pronunciation is perception, or one's ability to recognize speech sounds. In fact, a basic premise of pronunciation learning is that perception precedes production (Derwing & Munro, 2015); in other words, accurate production cannot occur without the initial ability to perceive the new sound. Perception of new sounds is a common difficulty for many L2 learners (Barriuso & Hayes-Harb, 2018) and, as such, it is a point of interest for many researchers looking to develop techniques to improve this skill. One such techniques is High Variability Phonetic Training (HVPT), a lab-based training that exposes learners to multiple voices producing the same sound in varied phonetic contexts (Thomson, 2018a). This form of perceptual training most often focuses on segments (i.e., consonants, vowels, and related features), but has also targeted tone perception (Perrachione, Lee, Ha & Wong, 2011).

HVPT studies have shown promising results in terms of perceptual gains, but despite being lab-based and lending itself well to Computer Assisted Language-Learning (CALL) material, it is rarely translated into useable tools that reach ESL teachers and learners. Although online HVPT training programs do exist (Thomson, 2018b; Iverson & Evans, 2009), HVPT has yet to reach a wide audience beyond the research community. Thomson (2018a) discusses several reasons why this may be the case: notably, the academic nature of the name HVPT, teachers' inaccessibility to HVPT studies (through jargon and niche publication), and lack of funds to promote existing tools. He also notes that being computer based, if HVPT were to be

implemented in ESL curricula, it would take up limited lab time. He adds that a possible way to circumvent the issue of time is to create mobile based HVPT tools that can be used beyond the classroom, as this medium is accessible and familiar to most students. He further acknowledges that, if HVPT were to be used beyond the classroom, it would need to be more engaging than its current form and suggests gamification as a means of rendering the training more motivating and accessible to learners. However, he does not elaborate on what this gamified, mobile-based HVPT could look like. With these suggestions in mind, this thesis addresses how these ideas can be implemented in a tangible pedagogical tool, using the mobile as the medium of choice for its accessibility and familiarity amongst learners. This medium, along with gamified learning, will be discussed next.

Gamification and Mobile L2 Learning

With the ubiquity and far-reaching capabilities of smartphones, Mobile-Assisted Language Learning (MALL) has emerged as a promising and popular avenue for language learning that allows users unprecedented freedom and personalization (Kukulska-Hulme, 2016). Language-learning mobile applications have exploded in popularity in recent years with Duolingo at the helm, boasting over 300 million users (Shortt et al., 2021). Many of these language-learning apps, with Duolingo as an exemplar, employ gamification, or the use of game elements in non-game contexts (Deterding et al., 2011) as a means of increasing user motivation and engagement with the tool (Munday, 2016). This could include the use of points, badges, streaks, leaderboards, and challenges. Student levels of engagement and motivation have been shown to increase in such gamified environments (Dehganzadeh & Dehganzadeh, 2020); however, it is often noted that game elements have the tendency to be implemented haphazardly and in a decontextualized manner (Domínguez et al., 2013; e.g., using points or rewards just for

the sake of gamifying), and are rarely examined on the basis of empirical frameworks of motivation (Seaborn & Fels, 2015). Thus, as gamification in L2 learning grows as a point of academic interest, there exists a need to not only assess which game elements contribute to learner engagement based on researched theories of motivation, but also how this empirically driven gamification can be applied to new pedagogical tools.

This Thesis

Given the disconnect between research findings concerning perceptual training and pronunciation practices in the classroom and beyond, the goal of this thesis is to address this gap by presenting the blueprint for a mobile-based application that features a gamified version of HVPT. This application has been named *FineTune*, as this name encapsulates the purpose of the training: to finetune one's pronunciation skills, much the way an individual would tune an instrument to perfection. Through this training, learners are able to hone their abilities perceiving L2 English consonants and vowels through a gamified paradigm that encourages them to collect badges by mastering sounds. While the long-term purpose of this project is to develop a useable pedagogical tool, this study also explores the process of developing gamified pronunciation learning material on the basis of empirically motivated frameworks. To this aim, the development of this tool is informed by Stockwell and Hubbard's (2013) ten research-based principles to guide the development of apps, activities, and tasks for MALL applications, as well as van Roy and Zaman's (2017) nine Heuristics, which guide the design of gamified learning tools based on Deci and Ryan's (2000) Self-Determination Theory.

As per the guidelines for a manuscript-based MA thesis, the next section constitutes "a full submittable draft of a manuscript" that presents a full literature review exploring HVPT, MALL, and gamification, and concludes with a description of the app; specifically, how HVPT is used in

the app, which game elements are incorporated into the training, and how the app development conforms with the aforementioned frameworks.

Chapter Two

Many teachers and researchers know that pronunciation is an oft-neglected aspect of language instruction, despite the growing desire amongst learners for more focus on the topic (Foote et al., 2012). Pronunciation is not only a crucial aspect of intelligible communication (Derwing & Munro, 2005); focusing on pronunciation also helps learners develop language awareness and contributes to their sense of confidence and competence as an additional language speaker (Wrembel, 2002). However, despite the need for more focus in the classroom, pronunciation often takes a backseat to other aspects of language learning such as vocabulary or grammar (Huensch, 2019; Harmer, 2001). In addition, teachers are usually unsure about how to incorporate it into the curriculum in an effective way (Foote et al., 2012; Breitkreuz et al., 2001). Given these circumstances, it is worth considering how pronunciation practice can occur beyond the reach of the classroom, be it extramurally (alongside school instruction) or within the digital wilds (digital contexts separate from formal instruction, such as in video games and social networking (Sauro & Zouro, 2019).

One method of pronunciation training that has been researched extensively is High Variability Phonetic Training (HVPT). HVPT focuses on improving learners' perception of individual speech sounds by exposing them to multiple voices, with the understanding that perception precedes production (Derwing & Munro, 2015). This type of training has been shown to improve learners' perception of consonants (e.g., Cebrian & Carlet, 2014), vowels (e.g., Lambacher et al., 2005) and tones (e.g., Perrachione, Lee, Ha & Wong, 2011); however, it remains unfamiliar to most teachers and rarely is translated into useable pedagogical tools (Thomson, 2018a). Notable exceptions are Thomson's English Accent Coach (Thomson, 2018b) and the Vowel Trainer (Iverson & Evans, 2009), which are both available online. HVPT

researcher Ron Thomson notes this gap between research and practice and suggests that a "collaboration with the gaming industry" (Thomson, 2018a, p. 227) could serve as a potential avenue to shift HVPT from the realm of academia into real-world pedagogical practices in a way that is accessible and appealing to both teachers and learners. He also recommends that HVPT could be mobile based, as a means of increasing accessibility and capitalizing on learners' familiarity with smartphones. However, this is merely a suggestion; he does not elaborate on what this mobile-based, gamified HVPT could look like.

Within the field of education, gamification has been gaining traction in recent years as a topic worthy of research and as a promising method of boosting student motivation (Ramirez & Squire, 2015). As this is a burgeoning area of research interest, many gamified pedagogical techniques and tools are not research-based. Rather, they tend to overuse obvious, decontextualized game elements, such as badges, points, and levels (coined "pointsification"), as opposed to carefully selected game elements chosen on the basis of researched theories of motivation (Domínguez et al., 2013; Seaborn & Fels, 2015). This haphazard gamification tends to produce negative results in terms of student motivation and performance (Domínguez et al., 2013). Thus, while it may be gaining popularity, for gamification to be effective in educational contexts it should be intentional and motivated by researched frameworks.

The current project will contribute to this field of interest by exploring how to transform the aforementioned HVPT-based pronunciation training into a gamified system downloadable as a mobile application. The purpose of this application is to help learners improve their perception of L2 English consonants and vowels, which they will do by progressing through levels that challenge users to identify speech sounds in different phonemic contexts. This gamified HVPT application has been named *FineTune* and will be referred to as such throughout the article. This

name was chosen because it succinctly encapsulates the goal of the app, which is for learners to literally finetune their speech, much the way that a musician would tune an instrument to perfection. There is nothing subtractive about the name *FineTune*: it does not insinuate that the learner's other languages are detracting from their capability to learn English, nor that there is a problem with their accent. The name also denotes qualities of precision and perfection, all which are the purpose of this type of pronunciation training that focuses on individual speech sounds, rather than whole words or sentences. The name *FineTune* is also playful and accessible to everyone; it was very important that the meaning of the name is understandable to those who do not have a background in phonetics or even pedagogy – a possible reason as to why HVPT is unknown to many language teachers, who may not have a background in phonetics (Thomson, 2013).

One of the aims in laying the foundations for *FineTune* is to explore the development of a pedagogical tool that is informed by different empirically motivated frameworks (for another example, see Sundberg & Cardoso, 2019, which outlines the development of a web-based application, *Bande à part*). To this end, *FineTune* is informed by Stockwell and Hubbard's (2013) ten research-based principles to guide the development of apps, activities, and tasks for MALL applications, as well as van Roy and Zaman's (2017) nine gamification heuristics for educational contexts, which guide the design of gamified learning based on Deci and Ryan's (2000) Self-Determination Theory. This theory posits that motivation (and consequently, engagement), is a function of fulfilling three basic psychological needs: the need to feel autonomous, competent, and connected to others around you (often referred to in literature as "relatedness"). Note that as this project represents the convergence of different fields (gamification, pronunciation learning, and MALL), the terms "learner" and "user" (used

respectively in Second Language Acquisition and gamification research centering on technology) will be used interchangeably to refer to the individual who engages with the proposed application.

The goal of this paper is to provide the blueprint and rationale for the *FineTune* app. As such, this project represents the first step in Cardoso's (2022) chronological description of how CALL research is conducted, including the following stages: (1) Development of the tool (e.g., based on research and/or users' needs); (2) Exploration of the pedagogical affordances; (3) Assessment of accessibility and user attitudes; and (4) Assessment of pedagogical effectiveness.

Before describing the application and its features, we will first provide a review of HVPT and MALL, as well as an overview of gamification and Self-Determination Theory, the theoretical framework that informs *FineTune*'s game elements.

Background

High Variability Phonetic Training

High Variability Phonetic Training (HVPT) is an empirically driven form of pronunciation instruction that constitutes the basis for the proposed mobile application. HVPT is a method of improving learners' perception (and consequently production) of speech sounds by exposing language learners to multiple voices producing the same target sound/s. The rationale behind HVPT is that hearing multiple voices will prompt the listener to disregard speaker-specific prosodic differences, and focus on the defining, generalizable phonetic and/or phonemic patterns that are characteristic of the target speech sound (Logan et al., 1991). HVPT diverges from typical pronunciation training in the classroom, as in that setting, students are usually only exposed to the teacher's voice, or the audio that accompanies language-learning textbooks.

The first researchers to explore speaker variability were Logan, Lively, and Pisoni (1991), who focused on Japanese learners' ability to distinguish between /r/ and /l/ contrasts. Post-tests using a forced-choice identification task demonstrated learning gains from the groups that were exposed to speaker variability, which served as the first indication that variability leads to an increase in perceptual abilities. Similar patterns were found in later studies, including Lively, Logan, and Pisoni (1993) and Lively et al. (1994). One of the most influential of the early HVPT studies targeting the /r/ and /l/ distinction was Bradlow et al. (1997), which examined the efficacy of HVPT without additional instruction. Their findings revealed improvements in both perception and production of the target phonemes. Considering the lack of explicit instruction, this study highlights the effectiveness of HVPT alone, and its potential in autonomous learning. Other studies have since focused on other consonant contrasts. Cebrian and Carlet (2014), for instance, examined the effects of short-term HVPT on Catalan/Spanish bilinguals learning the distinction between /b/ and /v/, as well as the /d/ and /ð/ phonemes in L2 English. Their findings also demonstrated gains in perception after the training was completed.

Since then, numerous studies have also centered on the perception of English vowels (Nishi & Kewley-Port, 2007, 2008; Iverson & Evans, 2009) and results have demonstrated improvement across participants with different first languages (although there have been some marginal differences between certain L1s). One of the more influential studies of HVPT has been Lambacher et al.'s (2005) research on Japanese learners of English, and their abilities to distinguish between five L2 vowel sounds. The results showed significant improvement in not just perceptual abilities, but also production, as assessed by twenty-six native speaker judges.

HVPT for vowels has also been explored in conjunction with other variables, such as participants' exposure (or lack thereof) to natural variability in their daily lives. Iverson, Pinet

and Evans (2009) examined the discrepancy in effectiveness of HVPT on native French speakers, with half the total number of participants living in France, and the other half in England (and thus exposed to plenty of naturalistic input). After an identical 8-week session of HPVT, all participants were tested for the target vowels in perception and production tests. The results demonstrated that both groups improved, but that there was no significant difference between the group in France and the group in England. This suggests that focused attention to target language is key in perception gains and can compensate for lack of naturalistic exposure.

However, when it comes to translating HVPT into pedagogical tools, one main challenge is the lack of standard agreed-upon practice. Beyond speaker variability, there are few variables that remain consistent throughout HVPT studies. For example, all HVPT studies employ stimuli spoken by multiple speakers, but the optimal number of speakers remains elusive. In Thomson's (2018) review of 32 HVPT studies, he notes that a range of 2-30 speakers (averaging 7.5) were used across the trainings. In addition, the efficacy of using an authentic or artificial voice is still in question. Studies have overwhelmingly used human voices, with the exception of a recent one which used text-to-speech synthesis (Qian, Chukharev-Hudilainen & Levis, 2018). The duration of the trainings, as well as the number of training sessions, is also seemingly arbitrary. For instance, Thomson's (2011) training involved eight sessions for 15 minutes each, spanning three weeks, whereas Cebrian and Carlet's (2014) training was completed in four sessions, lasting a total of three hours. Oftentimes learners are permitted to complete the trainings at their own leisure (e.g., using their own devices, in the digital wilds), an element of HVPT which lends itself well to autonomous learning.

HVPT and Computer-Assisted Language-Learning

HVPT and Computer-Assisted Language-Learning (CALL) go hand-in-hand, as computer-mediated programs allow for input to completely focus on the target form, in contrast to natural speech which permits more general exposure but risks the target form getting lost in a sea of input. Moreover, CALL is inherent in HVPT procedure, as most HVPT studies are carried out in laboratories, with a computational instrument producing the stimuli. CALL platforms are also beneficial in that they provide immediate feedback and allow the student to focus exclusively on the target sounds, as opposed to the contextual meaning of words (for a review of the benefits of form-focused instruction and language acquisition, see Spada, 1997). However, despite these advantages, these computer programs rarely evolve into pedagogical tools used in the curriculum (Thomson, 2011). While certain studies have examined HVPT in the context of foreign language courses (Thomson, 2012; Wang, Jongman, & Sereno, 2003), the link between research findings and the development of computer-mediated pronunciation training platforms in largely missing. Moreover, as research has demonstrated that computer-mediated HVPT can lead to improved phonemic perception (specifically vowels and consonants) even without explicit instruction (Bradlow et al., 1997; Thomson, 2011), there is definitely potential for the development of HVPT activities that foster autonomous learning, particularly in settings that allow learners to learn “anytime anywhere”, as will be discussed next.

Mobile Assisted Language-Learning

Thomson (2018a) recommended HVPT apps for mobile devices as a means of circumventing limited class time. Indeed, mobile-assisted language learning (MALL) is a popular and accessible option for learners, as phones, tablets, e-readers, and media players offer a portable, affordable, and convenient alternative to computers. MALL allows for learners to

personalize the learning experience to an unprecedented degree, as they can engage in "anytime, anywhere" learning for flexible amounts of time and customize application parameters to fit their individual needs (Kukulska-Hulme, 2016). Learners often engage with language learning apps during their workdays and in both public and private milieus (Read, Bárcena & Kukulska-Hulme, 2016), using them to fill little "pockets of time" (Steel, 2012).

These unique aspects of MALL present both the affordances and limitations of the medium, which need to be considered when designing mobile-based applications. For the proposed application, Stockwell and Hubbard's (2013) Ten Principles for Mobile Language Learning offer a research-informed framework to guide the development of the app in a way that considers the inherent characteristics of mobile devices (see Table 1 for a list of the principles). These principles account for the physical properties of mobile tools, the suitability of the pedagogical task in relation to the medium, and users' perception and attitudes surrounding mobile devices as language-learning platforms. For example, the researchers suggest that mobile language learning systems should "push, but respect boundaries" (principle 3). Users' near-constant connectivity to their mobiles allows for frequent reminders and other such calls to action. However, it should be up to the learner to decide when and how often they receive notifications (Kennedy & Levy, 2008).

Table 1

Stockwell and Hubbard's (2013) ten principles for the development of MALL applications

1. Consider the affordances and limitations of the device, and the affordances and limitations of the environment in which the device will be used.
2. Be focused: limit multi-tasking and distraction.
3. Encourage user engagement but respect their boundaries.
4. Consider the issue of equity. App-design should accommodate all types of devices, as not all phones are created equal.
5. Accommodate learners' individual differences.
6. Acknowledge how the users engage with their mobile devices (whether they use mobile tech for social, personal, or educational purposes.
7. Keep tasks succinct and divide longer tasks into chunks.
8. Ensure that the task fits the environment in which the device will be used.
9. Add clear instructions on how to use the app.
10. Acknowledge the other players involved in the language-learning context, be it peers, teachers, parents, colleagues, or supervisors, and how they may influence engagement with the app.

While these principles account for the limitations and affordances of the mobile medium, they do not address how to engage users with the application. Within *FineTune* this is accounted for through gamification, which will be discussed next.

Gamification

In his conclusion of how to encourage independent HVPT technology use outside the classroom, Thomson (2018) suggests "collaboration with the gaming industry." The rationale is logical: games are engaging, and the engaging elements used in games can be applied to other contexts, such as in the workplace, school, or while shopping. This idea is not new: the Boy

Scouts' use of colourful badges to signify achievement, and stores' loyalty reward points systems (S&H Company stamps are a very early example of using rewards points as a shopping incentive, beginning in 1896!) all represent early instances of gamification. However, the term "gamification" itself is fairly new, as is its legitimacy as an academic focus. Gamification as an area of study has gained significant attention in recent years, especially concerning its potential affordances in pedagogy. Many teachers have instinctively used game elements in their classrooms for years (such as providing points and rewards for desired behaviour, or including competitive and collaborative activities to drum up motivation), but now research and L2 pedagogy are beginning to support and build on these practices.

Academia has yet to decide upon a firm definition for gamification. Deterding et al.'s (2011) definition of "the use of game design elements in non-game contexts" (p. 10) is quite pervasive, but critics have noted that it does not encompass integral aspects of the "gameful" experience. For instance, Huotari and Hamari et al. (2017) characterize gamification as a "process of enhancing a service with affordances for gameful experiences in order to support users' overall value creation" (p. 25). Noteworthy here is the focus on gamification being a *process,* and with the definition hinging on the user's perception, as opposed to pre-defined game elements. Others posit that the definition of gamification should encompass the playful, emotional aspect inherent in the gaming process, as this drives the primary aim of gamification: engagement. This project will adopt the definition put forth by Hernandez-Gonzalez (2021), as it represents a combination of the three aforementioned characteristics: "Gamification is the process of adding game and play elements into non-game/play contexts, in order to foster gameful and playful emotions that can help increase engagement." (Hernandez-Gonzalez, 2021). This definition, alongside van Roy and Zaman's (2017) gamification heuristics and Ryan and

Deci's (2000) Self-Determination Theory, will guide the choice of game elements to be included in the application.

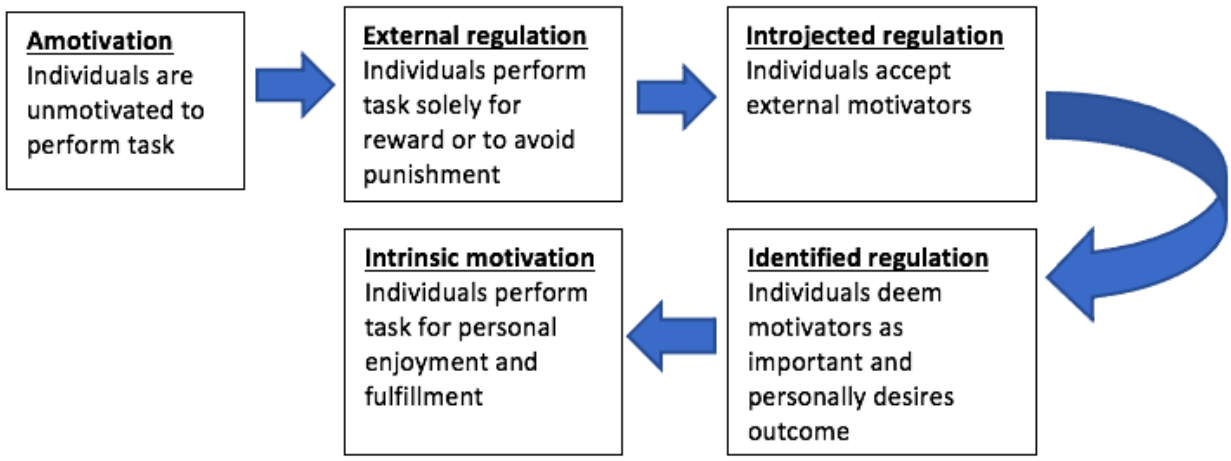
Gamification researchers and architects are thus concerned with how to maintain and increase participant engagement with the task at hand. Addressing this question has led researchers back to foundational theories of motivation, which have come to inform gameful design (i.e., deciding which game elements should be incorporated in the non-game context). One such theory is Deci and Ryan's (2000) Self-Determination Theory, discussed below.

Self-Determination Theory

Self-Determination Theory (Deci & Ryan, 1980; 1985; 1991; 2000) is a broad framework of motivation. One of the key tenets of the theory is that individuals who are extrinsically motivated (i.e., when their behaviour is motivated by factors external to themselves) are inclined to internalize this motivation (Deci & Ryan, 2004), thus bringing the nature of the motivation closer to that of intrinsic. Intrinsic motivation is often looked upon more favourably than its counterpart for several reasons. Intrinsic motivation is associated with higher overall task performance (Ariely et al., 2005), increased creative thinking and problem-solving skills (Glucksberg, 1962), more prosocial behaviour (Ariely et al., 2007), and overall task satisfaction and commitment (Hoffman & Strickland, 2005). The internalization itself contains three substages: external regulation, introjected regulation, and identified regulation (illustrated in Figure 1; adapted from van Roy & Zaman, 2017).

Figure 1

Stages of motivation



The extent to which internalization occurs depends upon the degree to which carrying out the task satisfies three basic psychological needs: the need to feel *autonomous*, *competent*, and *related to others*. Thus, according to SDT, motivation is less concerned with the nature of the goal itself, but rather with the attainment process - and specifically, whether this process addresses these three psychological needs (Deci & Ryan, 2000).

In the context of gamification, the game design implemented in pedagogical tools should support learners' basic psychological needs in order to enhance the motivational quality of the task. This enhancement does not necessarily mean that the end result should be intrinsic motivation – this is likely not realistic in all contexts. Rather, the goal of the gamified system should be to move from external regulation to identified regulation, meaning that the outcome of the task becomes personally important and desirable to the individual. This is achieved by implementing game elements that satisfy the three psychological tenets of SDT. However, most current gamified educational systems do not take into account psychology-based theories of motivation (van Roy & Zaman, 2019); rather, they tend to overuse decontextualized, low-level

game elements, such as badges, leaderboards, and points. This may be because identified regulation is not a priority if students are being directed towards an activity by their teacher (an external regulator). However, in the context of self-directed MALL technology, high levels of motivation are of the utmost importance, as it dictates engagement (Bitrían et al., 2021). Thus, in terms of deciding upon game elements to incorporate in *FineTune*, those chosen are motivated based on their ability to fulfill the needs for autonomy, competence, and relatedness. The following sections will outline how these three needs in SDT can translate to game elements implementable in *FineTune*.

Autonomy and game elements

Game elements that promote feelings of autonomy signal to the user a sense of control and choice (Kim et al., 2015), as well as highlight their inner process. In some games, this could take the form of a quest, or a choose-your-own-adventure scenario, whereby the player can act on their own accord within the game, depending on goals they set for themselves. In the context of *FineTune*, the game elements that promote autonomy need to emphasize the leaner's personal reasons for using the application. One possible way of doing so is to borrow from Locke and Latham's Goal-Setting Theory (1990). Among their principles for successful goal setting, Locke and Latham posit that goals must be sufficiently clear. One way of setting a clear goal is ensuring that it is S.M.A.R.T. (i.e., it should be specific, measurable, attainable, relevant, and time-bound). Setting an initial goal using the S.M.A.R.T. framework not only has the potential to make the goal more achievable; it can also personalize engagement with the application, so that the motivation for using it is no longer an abstract thought of "because the teacher says so," but rather because the goal is personally desirable to the individual (thus shifting the motivator from an external regular, the teacher, to an identified regulator— a personally meaningful objective).

Another game element that promotes autonomy and fits within the HVPT framework is that of choice. Learners, even those with similar linguistic backgrounds, can have different difficulties when learning the sounds of a new language (Derwing & Munro, 2015). For this reason, a gamified HVPT paradigm should provide choices for the leaners to focus on the speech sounds that are difficult to them individually, rather than imposing a training program based on the individual's linguistic repertoire. This element of choice is not only pedagogically sound, but it also personifies the application and provides a sense of control (Kim et al., 2015).

Additional game elements that have been shown to foster feelings of autonomy and control are the use of personal profiles and customizable avatars (Wee & Chong, 2019). These personal representations in the gamified system also may cultivate feelings of relatedness (Mulcahy et al., 2020), the third pillar of SDT.

Competence and game elements

In addition to autonomy, an individual's sense of their own competence plays a role in their engagement with an activity. White (1959) first introduced the notion that an agent's perceived feelings of competence will motivate their behaviour, and later research expounded on this premise by examining what conditions encourage feelings of competence. Research by Boggiano and Ruble (1979) addressed the role of feedback in relation to motivation, concluding that positive feedback increases intrinsic motivation in relation to no feedback, while negative feedback has the adverse effect. More recent research conducted by Burgers et al. (2015) supports these results, suggesting that positive feedback bolsters feelings of competence and reinforces the agent's sense in their ability to move forward and improve at the task at hand. However, it is noteworthy that the benefits of positive feedback on motivation are only felt when

the agent believes they are responsible for the competent act (Fisher, 1978); positive feedback given in a context that appears underserved by the agent has an adverse effect.

Possible game elements signaling competence could be the inclusion of progress bars, points, badges, leaderboards, levels and challenges (Hassan et al., 2020; Peng et al., 2020; van Roy & Zalman, 2019; Sailer et al., 2017) because they act as feedback, signalling to the user that they are capable of achieving a desired outcome. The popular language learning mobile application Duolingo is a strong example of a deeply gamified system that applies a number of these competence-signifying game elements (Shortt et al., 2021).

Relatedness and game elements

Relatedness, or feelings of secure attachment to those around you, serves as an important backdrop for motivation. This connection was observed by Bowlby (1979), who noted that infants who felt a secure attachment to their parents or guardians were more likely to engage in exploratory behaviour. The premises of such findings were transferred to different contexts in later studies: Ryan et al. (1994) observed similar positive outcomes in a classroom setting, whereby the students perceived the teacher to be warm, caring, and supportive. Moreover, a recent study by Akbari et al. (2015) demonstrated that a Facebook English-learning group, which reported high feelings of relatedness, outperformed in learning gains a similar in-person group that did not.

Feelings of relatedness in the context of mobile applications can manifest through the creation of online communities, social networks and systems that promote cooperation or competition (Koivisto & Hamari, 2019). One example of this can be seen in the popular running and cycling application Strava, which allows users to share their progress and pictures of their

routes with the friends in their networks (note that sharing is an option, and in accordance with pillar of Self-Determination Theory #1, the need for autonomy).

However, sometimes connectivity can lead to negative competitive social comparisons that in fact diminish feelings of support (Ryan and Deci, 2000). One recommendation is the possibility of opting out of certain competitive social features (such as leaderboards and visible scoreboards). This personalization of the community features can increase a user's sense of autonomy and accommodate for learner's individual differences (#5 in Stockwell & Hubbard's Ten Principles for Mobile Language Learning). Forums and other types of features that promote the exchange of information and ideas have also been shown to cultivate feelings of relatedness (Hassan et al., 2019). In this setting, users are able to connect by helping one another – a powerful need that strongly supports relatedness. These acts of help also contribute to increased notions of autonomy and competence (Xi & Hamari, 2019), as users feel a sense of ownership and mastery of their skills, as well as the ability to demonstrate their accomplishments while also learning from others. Personalization, and feeling part of a larger narrative or created world can also evoke feelings of relatedness (Mulcahy et al., 2020).

Table 2 provides a summary of the game elements discussed and their connection to SDT, and demonstrates the interplay between game elements. In truth, it is impossible to neatly categorize game elements as exclusively satisfying the psychological need for either autonomy, competence, or relatedness. In reality, not only does individual difference play a role in whether or not a game element satisfies a particular need, but also the game elements themselves do not perfectly correlate to a tenet of SDT. For example, game elements that are typically associated with signalling feelings of competence, such as leaderboards and badges, also have the ability to instill feelings of autonomy (Xi & Hamari, 2019). Games that include social networks, which

allow users to share information, have not only been shown to increase feelings of relatedness, but also those of competence, as users are aware of their increases of knowledge and learning gains (Xi & Hamari, 2019). However, for the purpose of the project at hand, it is important to incorporate game elements that have been shown to satisfy at least one, if not multiple, pillars of SDT.

Table 2

Summary of game elements and their connection to SDT

Game element(s)	SDT tenet(s)	Reference(s)
Badges, medals, challenges, levels, leaderboards, rankings, progress bars	**Competence:** act as feedback to inform users of their learning gains. **Relatedness:** if publicized, connects users to the activities of others.	Hassan et al., (2020) Van Roy & Zaman, (2019) Wee & Chong, (2019) Xi & Hamari (2019)
Avatars, personal profiles	**Autonomy:** offers users a sense of control and choice over representation **Relatedness:** if avatars and personal representations play a role in a gamified world/system, cultivates feelings of connectivity	Wee & Chong (2019) Mulcahy et al. (2020)
Forums	**Relatedness:** cultivates feelings of connectivity amongst users through their ability to help and support one another **Competence:** demonstrates to users their learning gains and promotes information exchange **Autonomy:** the act of helping allows users to feel control and ownership of their knowledge	Hassan et al. (2019) Wee & Chong (2019) Wee & Chong (2019) Xi & Hamari (2019) Xi & Hamari (2019)
Narratives, storylines	**Relatedness:** if users play a role in a gamified world/system, cultivates feelings of connectivity and belonging to a greater whole **Competence:** can divide a task into smaller, themed steps, thereby facilitating goal completion	Hassan et al. (2019) Wee & Chong (2019) Dong et al. (2012)

Customization, choice	**Autonomy:** provides users a sense of control over their actions and freedom from external constraints	van Roy & Zaman (2019)

Van Roy and Zaman's Nine Gamification Heuristics

When it comes to creating a gamified system, it is not just the game elements themselves that should encourage feelings of autonomy, competence, and relatedness, but also the way they are implemented. To provide a framework for selecting game elements and their implementation, van Roy and Zaman (2017) put forth nine gamification heuristics for designing gamified systems for educational contexts, which take into account the tenets of Self-Determination Theory. These heuristics will inform the development of the *FineTune* app. This section will outline each heuristic and elaborate on how it relates to SDT.

1. Avoid forcing the user to use (a part of) the gamified system.

Feelings of autonomy arise when users perceive that they have a sense of control within the gamified system. Moreover, the ability to opt out of gamified systems could negate the effects of game elements that may be discouraging to certain users (such as leaderboards or other inherently competitive elements).

2. Provide a moderate amount of meaningful options.

Although it is encouraged to provide options as a means of promoting feelings of autonomy in the learner, too much choice can be overwhelming (Schwartz, 2009). It is preferable to have some guidance and options in the system that lead to desired outcomes for the learner.

3. Set challenging, but manageable goals

Positive feedback is a crucial aspect of gamified systems for indicating competence and motivating users to engage. This means that tasks themselves need to be designed in a way that

balances challenge with attainability (Malone, 1981). As such, when a task is completed, the learner feels a genuine sense of accomplishment and capability.

4. Provide positive, competence-related feedback

Not all feedback types promote feelings of competence, which is an important consideration when designing gamified systems. Negative feedback should be avoided in gamified systems that abide by SDT, as well as feedback that is performance-oriented (Reeve, 2004). Feedback should be easily attributed to competence and be perceived as such by the learner.

5. Facilitate social interaction

Van Roy and Zaman (2017) note that when learners share a common goal and cooperate, they are more likely to feel bonded and secure with one another. Social interaction in gamified systems can take the form of competition as well as cooperation, although to foster feelings of relatedness it is preferred to design systems that allow users to feel supported by their peers, rather than in competition with them.

6. When supporting a specific psychological need, do not thwart other needs

As was discussed in previous sections, game elements can satisfy multiple tenets of SDT at the same time. However, in certain situations they can also detract from one another. For instance, if a gamified feature involves collaboration within a team (promoting relatedness), the team parameters should be designed to also allow for personal choice within that paradigm (van Roy & Zaman, 2017).

7. Align gamification with the goal of the activity in question

Within the context of education, gamification is used to facilitate particular learning gains and make the process playful. Gamification should not exist at the expense of the educational objective – it should enhance it.

8. The context of the gamified system should be supportive

Gamified systems can exist in a variety of contexts which can influence the way a learner interacts with the system. In the case of the *FineTune*, it is likely that the learner will be using the tool in public spaces, where it may not be comfortable to produce speech sounds out loud. Such environmental factors need to be taken into account when designing gamified systems.

9. Create a system that is flexible and accounts for personal needs and preferences

Individual differences account for how learners interact with gamified systems. Some individuals are naturally more competitive than others and will be highly motivated by game elements that encourage social comparisons, such as leaderboards, whereas others would prefer not to engage. These preferences should be accounted for when designing gamified systems.

The following section will provide a description of the proposed app, *FineTune*, as well as a breakdown of each feature and how they relate to the theoretical frameworks used to inform the app's design.

***FineTune*: Development**

The goal of this project is to motivate and provide a blueprint for the proposed mobile application, *FineTune*. The previous section included a brief overview of HVPT research and introduced Stockwell and Hubbard's (2013) 10 Principles for designing CALL and MALL tools for language learning. It also provided a background of gamification and game elements that satisfy Self-Determination Theory, which is the theoretical framework used to select game elements for the application. The reason for using Self-Determination Theory (SDT) is because game elements that encourage feelings of autonomy, competence, and relatedness lead to a higher degree of engagement (Bitrián et al., 2021; Schreurs et al., 2014). Designing the blueprint for a research-informed application requires combining frameworks from three different fields

(MALL, HVPT, and gamification) in a harmonious way. This is an important consideration, as new variables, such as the incorporation of game elements, have the potential to influence the efficacy of HVPT (Thomson, 2018a). The following section will provide a description of the application, and a summary of each feature and its connection to one of the research-based frameworks used.

FineTune: Description

FineTune is a mobile application under development that uses a gamified HVPT paradigm. The aim of this app is to help learners improve their perception of L2 English vowels and consonants by providing them with input spoken by multiple speakers of L1 Canadian English (8 in total) in differing phonetic contexts throughout the game (non-words, words in isolation, and carrier sentences). For the current stage in the development process, the aim is to develop the game for one variety of Canadian English, although a long-term goal is to apply the same training framework to multiple varieties of English. This app aims to render this form of phonetic development more engaging and accessible to learners, who may not have access to empirically driven pronunciation training, and to teachers, who may not have the time or tools to incorporate pronunciation into their curriculum. The reason for choosing mobile as the platform of choice is because this medium fosters "anytime, anywhere" learning that is easily customizable to the individual.

To use the app, users select a sound (or set of sounds) that they would like to practice, and progress through levels by choosing the token that depicts the target sound as heard in the stimuli (see Figure 4). The reason for using 8 speakers is because the optimal number of speakers is still unknown; 8 is the rounded average of Thomson's (2018a) synthesis of 32 HVPT studies. This number will be equally divided between male and female voices of different ages to ensure

sufficient variability (Lively et al., 1993). In *FineTune*, learners are encouraged to move up levels that provide the target sound(s) in different environments, progressing from non-words, to words in isolation, to words in carrier sentences. The rationale behind this framework for scaffolding is that non-words have been shown to lead to more pronunciation gains among beginner L2 learners rather than real words (Thomson & Derwing, 2016), which is believed to be a function of increased focus on the target sound, as opposed to other lexical information (e.g., meaning, usage). Thus, the levels are scaffolded in such a way to incrementally provide more information beyond the target sound.

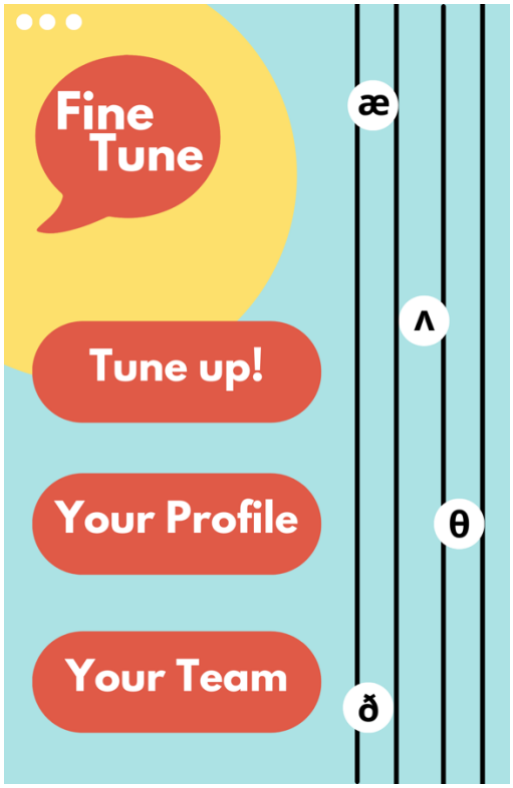
The following section describes the user experience with the application and is followed by an explanation of how gamification is applied in *FineTune*.

User experience

When an individual opens *FineTune*, they are first prompted to take a tour of the application. This is in accordance with Stockwell and Hubbard's (2013) nineth principle for designing MALL tools: include guidance for how to use the training. This guide is available to the learner at all times: they would simply have to tap the three dots at the top of the screen to access it (see Figure 2).

Figure 2

FineTune: Main graphical user interface

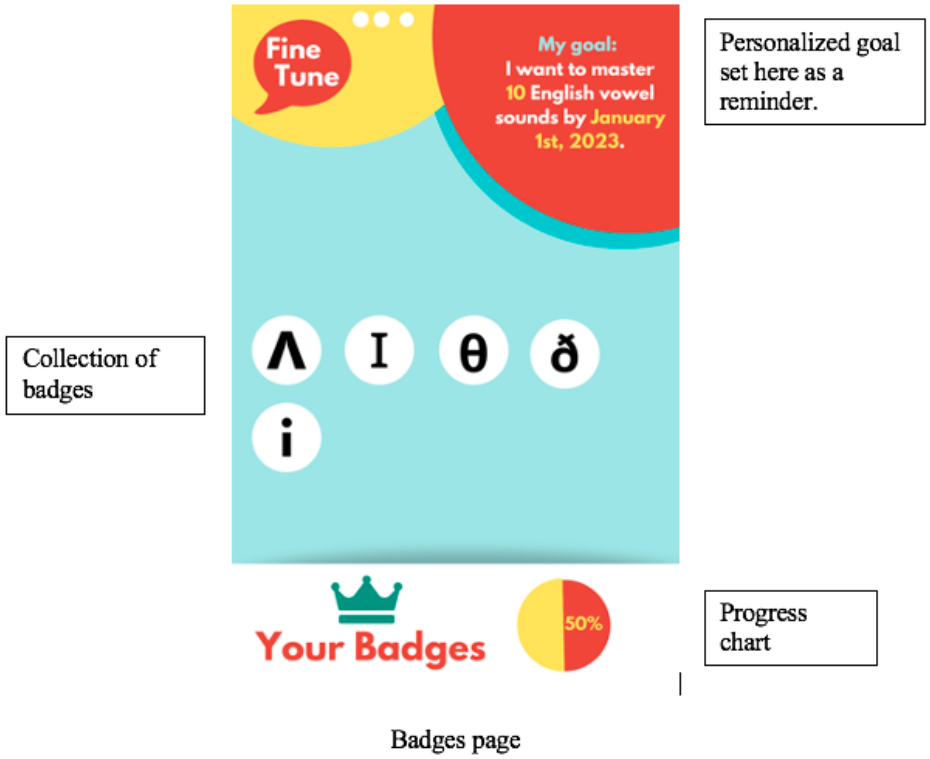


Once the user has taken the tour, they are asked to create a personalized profile and customize their settings for the training. Their settings are saved and applied to their trainings and can be changed at any time. The personalized profile involves creating a username and an avatar based on the flag of the user's home country. The option of uploading an image is also available; however, this is not necessary as a measure of respecting the user's privacy and boundaries. There is also the option to include more information for data collection purposes (such as age, linguistic repertoire, and reason for wanting to partake in the training); however, this is not necessary. The user is also encouraged to add a personal goal for their training based on the S.M.A.R.T. framework (see Figure 3 for an example goal), and options are supplied to help guide the goal-setting process (e.g., the duration of the training, reasons for doing so).

Once the user has taken the tour and completed their profile, they may begin their learning session or training. The aim of the training is to collect as many badges as possible, which are displayed in the user's profile and are used to embellish their avatar (see Figure 3). Each badge represents the mastery of a sound and are awarded once the user has progressed through the three levels with an 100% success rate of that sound.

Figure 3

Badges displayed in FineTune



The gamified HVPT training in *FineTune* is broken down into three levels. The first level involves identifying sounds in monosyllabic non-words (e.g., "stip" for a focus on the lax vowel /ɪ/). The phonetic environment will vary (whenever appropriate, the target sound will be presented at the beginning, middle, and end of the word – see Logan et al., 1991 for the rationale). The user is presented a word with the target sound, and they must identify the sound from a selection of tokens. The sound will be represented as a phonetic symbol, which they must

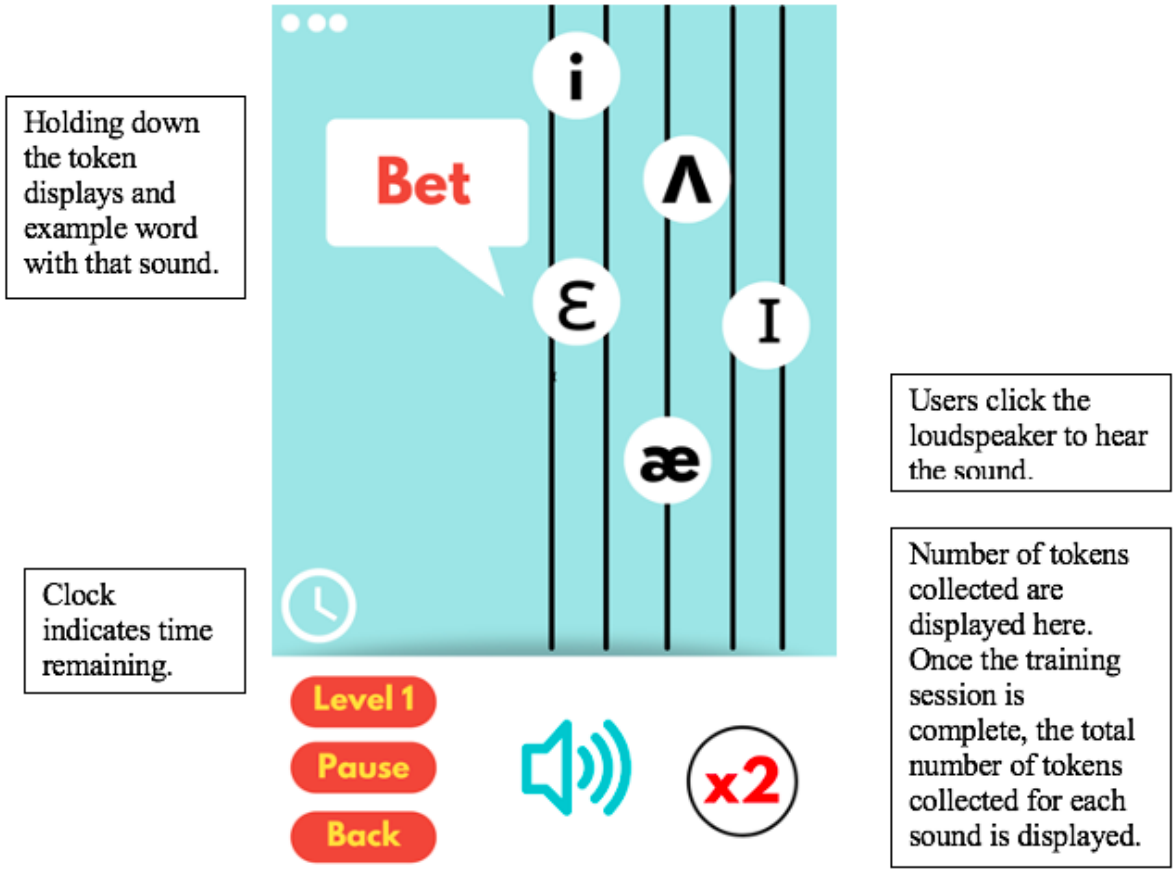
tap. This was a point of consideration— past HVPT studies have differed in terms of how they present certain sounds. Sometimes, sounds are represented using IPA symbols, and other times with icons that do not require previous knowledge of phonetics (such as nautical flags in the case of Thomson, 2012). However, it was decided that IPA symbols would be used, as certain learners may have previous knowledge of them. Moreover, if the user is required to learn to associate a sound with a symbol, it would serve them even beyond the app to become familiar with the symbol that is meaningful and internationally recognized. To mitigate this issue, before training begins, there is the option to consult an IPA chart, whereby the user can tap on a phonetic symbol to hear its sound. Additionally, within the game itself, the user also has the option of holding down on the token to see a written word that contains the sound of that particular IPA symbol (see Figure 4, and Dickerson, 2007, for the rationale). Once the learner has correctly identified the token, it is added to their bank for that particular sound. If they have mastered that sound at all levels, they receive the badge.

Certain tokens will be worth marginally more points at random to introduce an element of surprise and gamble into the game. The user has the option of including a time limit, which they can customize (in accordance with heuristic #1: provide the option of opting out of game elements). For instance, if the user decides to give themselves 5 seconds to select the correct token, once the 5 seconds are up, the tokens will all drop from the screen and no tokens are awarded. This is in line with van Roy and Zaman's (2017) heuristic #4 of avoiding negative feedback: even if the points are lost, there is no subtraction from the user's initial score. In the last two levels of the game, the tokens will move up and down the "strings" of the *FineTune* game interface (see Figure 4 for an illustration of the tokens and strings). This is simply to introduce a tactile element of playfulness into the game by persuading users "to overcome

unnecessary obstacles" (for rationale, see Suits, 1978). The user also has the option to customize the duration of their training session, which is part of the initial goal-setting feature upon downloading the app. The app does provide suggestions (e.g., 5 minutes for 10 days, 15 minutes for 7 days); however, ultimately, it is the choice of the user. It is also the user's decision which and how many sounds they would like to test at once.

Figure 4

Playing FineTune



Level 1

Level one involves identifying the target sound(s) in non-words in varied phonemic environments. Once the user has mastered that level, they move on to level two, which now uses

real words as stimuli (e.g., “tip”, if the intended sound is the lax vowel /ɪ/). The game parameters remain the same. Once the user has completed level two, they move on to the final level, which involves carrier sentences that presents the target sound in words that appear sentence-initially and sentence-finally (e.g., “Tip,” I’d like to say “tip”). While such a sentence may not reflect the contexts in which the learner is likely to hear the word naturalistically, it offers variability in terms of word placement in the sentence.

In addition to the training itself, the option of joining a collaborative training “team” is also available to the user. This adds a social dimension to the learning process (heuristic #5) that is intended to be collaborative and supportive (heuristic #8). This team is selected based on the user’s country or origin/L1, under the belief that this cultural familiarity may translate to smooth communication and an immediate connectivity, and that team members may encounter similar difficulties which they can discuss. Teams can be composed of up to 5 individuals (minimum 2) and have the ability to chat with one another. In this format, users can exchange information and provide support in a setting that is non-competitive and more intimate than a leaderboard. The user does have the option of displaying their mastered tokens if they do wish to share their accomplishments (Sailer et al., 2017). Moreover, users that have mastered certain sounds are invited to join “expert” groups, where they can share their knowledge with others. Users are able to join these groups freely (there is no size limit), and can use them to ask questions to those who achieved that sound badge, and to communicate with other learners targeting the same sound.

While HVPT target perceptual gains, *FineTune* will also provide the option for users to test their pronunciation accuracy with a speech recognition feature, which is accessible by tapping the “Tune up” tab (see Figure 2). This feature is available to users at any time, and can

be used not only as feedback to evaluate their own pronunciation abilities throughout the training, but to assess which sounds require practice from the onset.

HVPT in FineTune

One challenging aspect of concretizing HVPT into an application is that there is no clear consensus amongst researchers as to what parameters lead to the most learning gains (Thomson, 2018). Past HVPT research has little uniformity: studies have differed in terms of numbers of speakers, length of training, whether or not target sounds should be presented in real words, non-words, or carrier sentences, and if learners should be allowed to hear stimuli multiple times before providing answers (to name a few). To address the lack of standard practice in regard to many of these constructs, *FineTune* provides customizable options for the user. Of course, there are some core unchanging aspects of HVPT that are similar across nearly all studies, and as such, they have been applied in *FineTune* in an unchangeable way. For example, the application needs to include multiple voices – this is a core tenet of HVPT, which in the context of the present app appears as 8 separate speakers (see Table 3). Additionally, the majority of previous studies have employed a forced choice identification task (FCID) and provided learners with immediate feedback. This translates to learners hearing the target sounds in either real words, non-words, or in carrier sentences, and then having to choose the sound they heard from a variety of options.

Many of the early HVPT studies have focused on smaller subsets of target sounds; however, research has also supported training many sounds at once (Nishi & Kewley-Port, 2007). Ron Thomson's *English Accent Coach* (Thomson, 2018b), one of the few easily accessible HVPT platforms, also allows users to train multiple sounds simultaneously, while having to choose from a large set of possible answers. The present application also provides this

option. Thus, to summarize, Table 3 lists key HVPT constructs and their application in *FineTune*.

Table 3

HVPT in FineTune

HVPT construct	Application in *FineTune*
Target sounds	Customizable: users may select which sounds they would like to practice or learn (with the default being all at once)
Number of talkers	Non-customizable: 8 different speakers (4 male, 4 female) of varied ages (the average of Thomson's 2018 synthesis of HVPT studies)
Type of stimuli	Non-customizable: speakers are human, as there is limited data on the efficacy of artificial voices
Phonetic contexts	Non-customizable: speech sounds are present in all phonetic contexts (beginning, middle, end of a word, whenever possible) and in non-words, real words, or carrier sentences, depending on the level.
Training length	Customizable, but with options suggested: option to decide for how many days/weeks/months they would like to train, the length of each training, and frequency during the training duration.
Answer identification	Non-customizable: answers are presented as phonetic symbols on tokens.
Feedback style	Non-customizable: feedback is immediate (token is awarded immediately if answer is correct and total points are displayed as game progresses). Once training session is finished, the number of tokens awarded for each sound is displayed.

Gamification in FineTune

The added component of empirically motivated gamification to HVPT is meant not only to engage the learner with the tool, but to add a playful, tactile dimension of the training that makes it accessible to learners who likely would otherwise only view the training in a one-dimensional, pedagogically focused manner. This emphasis on playfulness is part of Hernandez-Gonzalez's (2021) definition of gamification, as "the process of adding game and play elements into non-game/play contexts, in order to foster gameful and playful emotions that can help

increase engagement." Beyond the tactile nature of the game, another sensory element that also adds to the overall playfulness of the tool are the aesthetics. While this is subject to change as development progresses, this is an important consideration as the colours, shapes and fonts presented in the app convey a particular mood and tone. In the present app, the intention is for the aesthetics to express a lighthearted, approachable feel. This is in contrast to the seriousness that accompanies many pedagogical materials, which often display deep colours and stock images. In addition to the focus on playfulness, the gamification of *FineTune* is implemented based on the game's ability to instill feelings of autonomy, competence, and relatedness, as per Deci and Ryan's (2000) Self-Determination Theory, as was expanded upon above. Satisfying these three psychological needs enhances the motivational quality of the task, moving it closer in nature to intrinsic motivation (Deci & Ryan, 2004).

FineTune incorporates gamification in a number of ways. Most evident is the inclusion of obvious game elements, such as the use of badges to signify mastery of a sound, levels to scaffold difficulty and challenge learners to persevere (Peng et al., 2012), and progress pie charts to indicate a user's proximity to achieving their goal (see Figure 3). These are all signifiers of competence (Hassan et al., 2020; Koivisto & Hamari, 2019; Wee & Chong, 2019) that are connected to a defined and personalized learning goal. In this sense, they are not decontextualized, which is a common misstep of other gamification systems in educational settings (van Roy & Zaman, 2017), which can sometimes employ obvious game elements for the sake of gamifying, rather than on the basis of empirical evidence. Moreover, *FineTune* also includes a common game element of *collecting,* as learners are encouraged to accumulate points (indicating time and effort spent on the task), in order to earn and collect badges signifying their mastery and competence. This is an under-researched but common and highly engaging element

of games (anyone who lived through the height of the *PokémonGo* craze is well aware of this, and the insatiable desire to "catch 'em all!"). Previous studies have shown that in the context of digital games, items are collected primarily for their utility and enjoyment purposes (Toups et al., 2016). However, collecting is rarely explored within the context of SDT, although in the context of the present app, it is likely that it would function similarly to accumulating points as a signifier of competence, rather than social relatedness (another common possibility, especially if the items collected serve purposes and are tradeable in a created world).

In addition to these game elements that indicate the user's competence, *FineTune* offers many customizable options that afford the user a sense of control and autonomy. To begin, users are invited to personalize their profiles and set a manageable goal that guides their use of the application (thus concretizing the motivator as an "identified regulator"— something personally meaningful and desirable). They also have the ability to opt out or modify many game elements to address their personal learning styles (heuristic #9 and principle #5), such as setting timers for each level, choosing which sounds they would like to focus on, electing whether or not to join a learning team, and choosing to make use of *FineTune*'s speech recognition feature to assess their production abilities. Users also can decide whether or not they would like to receive notifications, which can be an especially infantilizing (and as such, autonomy-depriving) element of many app-based games.

The need for relatedness is addressed in *FineTune* through the option of a learning team, which is meant to be supportive and collaborative, rather than fostering social comparisons. The presence of a learning team allows for users to connect to one another by sharing their experiences, learning gains, and difficulties in a way that is non-threatening, as opposed to impersonally marking their standing in relation to others. Moreover, the inclusion of the "expert"

groups also allows for those who have mastered sounds the opportunity to help others, an act which supports a sense of relatedness while also encouraging feelings of competence and autonomy. These functions are optional, as a means of promoting autonomy and accounting for learning differences (heuristic #8 and #9 – see Table 4 for a summary of the features and how they relate to SDT and the gamification heuristics).

Table 4

SDT and gamification in FineTune

Gamified *FineTune* feature	Connection to SDT	Connection to Gamification Heuristic
Accumulating points to earn badges	Competence	#4 Provides positive, competence-related feedback, by demonstrating learner's time and effort devoted to the game. #7 Aligns gamification with the pedagogical goal (mastering perception of speech sounds)
Accumulating badges to prove mastery of a sound	Competence	#4 Provides positive, competence-related feedback, by demonstrating successful completion of a task. #7 Align gamification with the pedagogical goal (mastering perception of speech sounds)
Moving up levels of increasing difficulty	Competence	#4 Provides positive, competence-related feedback #7 Align gamification with the pedagogical goal (mastering perception of speech sounds)
Setting personalized username and avatars	Autonomy	Not connected to a heuristic
Setting personalized learning goal(s)	Autonomy	#2 Provides meaningful goal-setting options #3 Set challenging but manageable goals
Customizable game settings (time limits, learning teams)	Autonomy	#1 Avoid forcing user to engage in a game element #9 System is flexible and accounts for personal preference
Customizable learning parameters (choice of speech sounds to target)	Autonomy	#2 Provides meaningful options #9 System is flexible and accounts for personal preference #7 Aligns gamification with the pedagogical goal (mastering perception of speech sounds)
Option of joining a learning team or expert group	Autonomy Relatedness Competence	#5 Facilitate social interaction #8 Create a supportive context by creating community collaboration

Not illustrated in Table 4 are the aesthetic or tactile qualities of the application. While the interfaces shown here (see Figures 4, 5, 6) are prototypes, visuals (and sounds) are gamification features that are of the utmost importance. While they may not directly satisfy any of the 3 basic psychological needs, they do convey feeling and sense of cohesion that bestows the application personality and a sense of being part of its own world, or fantasy. Absent as well from Table 4 are game elements such a time, surprise/gamble, and mystery. *FineTune* does have the option of adding a countdown/timer to the HVPT, and surprise/gamble is included through the random bonus points that are given by some tokens. Mystery is a compelling game element that can be seen in many "soft-built" fantasy worlds. These worlds contain aspects that prompt the user to question why that element is present in the world, or why did it act in a particular way, while not providing any concrete answer. As mystery is most present in games that build fantasy worlds, this element is not especially strong in *FineTune*. However, mystery can be conveyed in subtle, innocuous ways, such as the way tokens are presented and the use of particular sounds and colours. While there may not be evident world-building in *FineTune,* there is an overall musical motif that is present not only in the app name, but also subtly through the use of note-like tokens, which are placed on lines reminiscent of sheet music or guitar strings in the training (see Figures 5 and 6).

A crucial consideration when applying game elements to an established pedagogical training was to do so in a way that did not detract or interfere with the training itself, and to consider the limitations and affordances of the mobile medium. This will be discussed next by outlining application of Stockwell and Hubbard's (2013) principles in *FineTune*.

Stockwell and Hubbard's (2013) MALL Principles in FineTune

The medium of the mobile carries certain affordances and limitations, both of which need to be accounted for in the development of *FineTune*. Stockwell and Hubbard (2013) have divided these features into three categories: physical, pedagogical, and psycho-social (see Table 5 for a summary of the principles). In terms of physical limitations, the most obvious one is the physical size of the device and the screen, and the potential for distraction, given the "anytime, anywhere" nature of MALL (Reinders & Hubbard, 2013). The relatively small size of a mobile screen can be mitigated by having clear, simple visuals, legible font, and sounds that accompany game features. *FineTune* involves a straightforward, easily learnable game that is essentially a forced-choice identification task. Although adding other features may provide more potential for engagement, considering the limited screen space and the likelihood that it will be played in busy environments, the game is kept simple, clear and without small elements that require exceptional scrutiny (see Table 5 – principles #2 and #8). Certain principles here can be seen as connected to the SDT tenet of autonomy, particularly principles #3 and #5, which are accounted for in *FineTune* through the options of customizing certain game elements and learning goals. The customizable learning goals do not only account for individual differences, but the varying language-learning contexts that may involve this app. The application may be used for independent learning, but it can also be incorporated into a class curriculum as an activity assigned as homework. In this context, specific learning goals and parameters could be set with the guidance of a teacher.

In regard to principle #6, Stockwell and Hubbard recommend that the application features should be consistent with "existing cultures of use" (Thorne, 2003). It is notable here that their article was published in 2013, and since then the accessibility and complexity of mobiles has

increased significantly. Nowadays, apps encompass nearly every domain of life, from finance to fitness, social life to education. Language-learning apps are by and large embraced by learners (Cheng & Hyun-Ju, 2019); however, certain features are more in line with common functionalities of mobile apps, and as such are more accessible. For instance, the social dimension of *FineTune* involves a chat function, as opposed to another type of social interaction that may be familiar mainly to users of a specific age or demographic (such as communication solely via images, as in the app *Snapchat*). Games that are straightforward and tactile, such as the forced choice identification used in *FineTune*, are also familiar and intuitive. Certain studies do consider the "novelty effect" of some applications and game elements to be attractive to users (Hamari et al., 2014). However, in terms of user retention, it is preferable to keep features familiar and simple, as this may also increase retention, especially in older users (Bitrián et al., 2021).

Table 5

Stockwell and Hubbard's (2013) ten principles and their application in FineTune

Principle	Application in *FineTune*
1. Consider the affordances and limitations of the device, and the affordances and limitations of the environment in which the device will be used.	Clear, easily discernable visuals
2. Be focused: limit multi-tasking and distraction.	Focused activities
3. Encourage user engagement but respect their boundaries.	Highly customizable, option to not receive notifications or share any personal information
4. Consider the issue of equity. App-design should accommodate all types of devices, as not all phones are created equal.	Not applied in this stage of development
5. Accommodate learners' individual differences.	Options to customize learning parameters (such as target speech sounds) and game elements (such as time limits, team participation)
6. Acknowledge how the users engage with their mobile devices (whether they use mobile tech for social, personal, or educational purposes.	Features that are in line with common mobile uses
7. Keep tasks succinct and divide longer tasks into chunks.	Training session lengths are customizable, and training is scaffolded into levels
8. Ensure that the task fits the environment in which the device will be used.	Training is simple, straightforward, with limited distractions
9. Add clear instructions on how to use the app.	Tour upon downloading and opening the app, available at all times
10. Acknowledge the other players involved in the language-learning context, be it peers, teachers, parents, colleagues, or supervisors, and how they may influence engagement with the app.	Customizable learning goals

Conclusion

The goals of this project are to bridge the gap between pronunciation learning practices and researched methods of instruction and present the development of *FineTune*, an empirically motivated MALL-based tool that uses gamification to address user engagement. As a first step to achieving this aim, we have laid the foundations for the development of the *FineTune* mobile application, a gamified, interactive version of High Variability Phonetic Training, that sharpens users' perception of English L2 speech sounds, with the understanding that aural perception precedes oral production (Derwing & Munro, 2015). The blueprint of this application is informed by empirically driven frameworks, namely Self-Determination Theory (Deci & Ryan, 2000) and van Roy and Zaman's Gamification Heuristics for educational contexts (2017), which guide the process of gamifying the training, and Stockwell and Hubbard's (2013) Ten Principles for MALL applications. As such this project represents the convergence of the disciplines of pronunciation, gamification, and MALL. There are numerous studies that present frameworks for applying gamification to educational contexts and material, but few that provide concrete examples of how these frameworks can be applied to novel, pedagogical tools.

The second stage of developing the present app involves transforming *FineTune* from a blueprint into a concrete application in order to assess its pedagogical affordances, as well as the users' attitudes towards the tool— the second and third steps in Cardoso's (2022) chronological framework for the development of CALL tools. One possible way of carrying out this assessment is by using the Technology Acceptable Model (Davis, 1989), which examines users' perception of the usefulness and useability of a tool. While this method was developed for information systems in general, it has been used to examine CALL-based tools (for a recent

example involving Google-based ASR, see Johnson, 2021). Moreover, we foresee certain features being modified over the course of the development phase, specifically concerning interface and other design principles and possibly the implementation of more game elements. These modifications will likely be a function of app-building capabilities and emerging research findings in the fields of gamification, pronunciation, and mobile application design. These changes will be chronicled and hopefully published in future articles, so as to serve as a detailed record of the development of a MALL-based tool, from inception to completion.

Chapter 3

This chapter will present a review of the goals of this study and their implications for language education and research in the field of Applied Linguistics. It will also discuss the next steps in the development of *FineTune*, in accordance with Cardoso's (2022) chronological framework for the development of CALL tools.

Summary of Goals

The aim of this thesis was to present the rationale and blueprint for the development of the *FineTune* mobile app. This application uses a gamified version of HVPT to help L2 learners hone their aural perception of Canadian English, with the understanding that perception precedes production (Thomson & Derwing, 2015). This thesis is in a part a response to the remarks of HVPT researcher Ron Thomson, who suggests that a gamified, mobile-based HVPT paradigm could shift this form of pronunciation training from the realm of academia into the practices of both ESL teachers and learners (Thomson, 2018a).

The design of *FineTune* is informed by van Roy and Zaman's (2017) gamification heuristics for educational contexts, which in turn is guided by Deci and Ryan's (2000) Self-Determination Theory. As such, the game elements included in the application were selected based on their ability to satisfy the learner's need to feel autonomous, competent, and related to others around them, in order to enhance the motivational quality of the game. The application design is also aligned with Stockwell and Hubbard's (2013) ten principles for the design of MALL-based tools, which considers the affordances and limitations of the mobile medium. Thus, an overarching goal of this thesis was to explore the development of MALL/CALL-based pedagogical tools that combine established trainings with gamification techniques, so as to

render these research-backed training methods accessible to the public, as well as provide a model for how the development of such tools could be carried out in the future.

Implications for Education and Research

One of the impetuses in creating *FineTune* was to find a way to circumvent the issue of limited class time and lab time in ESL curricula, by creating an avenue for pronunciation learning in the "digital wilds" (Sauro & Zauro, 2019). *FineTune* allows for pronunciation training that can be used to complement classroom instruction, or completely independent of it. For instance, teachers can assign the practice of certain sounds as homework and can help students set parameters to guide the training (e.g., setting the number and duration of training sessions, and target sounds). However, this guidance is not necessary; the application features are simple, and the training is procedurally straightforward. Such an application has the potential to allow for further pronunciation practice within and beyond educational contexts to individuals of varying ages, L1's, and educational backgrounds; all that is required is access to a mobile device. Moreover, *FineTune* presents a method of pronunciation practice that focuses on aural perception so that users can achieve overall intelligibility, rather than accent reduction or modification. The latter is a pervasive and harmful framing of pronunciation training that is still used in courses and material, implying that certain speakers must negate their natural speech patterns (and linguistic identity) in order to be viewed as a more competent and acceptable communicator (Grover et al., 2021). *FineTune* presents pronunciation training as an additive skill and contributes to the growing number of pedagogical tools and materials that promote this ethos.

In terms of the research implications, *FineTune* presents an example of how empirically motivated pronunciation techniques, gamification, and MALL frameworks can be applied

harmoniously in the development of tools or material that address a need in real-world pedagogical practices. Many researchers note the gap between SLA findings and pedagogical practices in the classroom and beyond (Rahman & Pandian, 2016) and, as such, it is a much-needed endeavour to explore *how* to create tools and material that reflect the findings of research that otherwise may never reach audiences beyond the academic community. Papers that outline the development of such tools are few and far between; however, a notable exception is Sundberg and Cardoso's (2019) article, which describes the development of the *Band à part* app, a web-based application that teaches French vocabulary through music. Hopefully, by chronicling the development of pedagogical tools and material, this type of endeavour will become more prevalent in our field and, as a result, bridge the gap between theory, research, and the practices of L2 teachers and learners.

The Next Stages

FineTune is currently in the first stage of Cardoso's (2022) chronological framework for the development of CALL-based tools. This stage involves designing the tool on the basis of insights from CALL and SLA, as well as explaining its rationale. As this thesis provides the blueprint for *FineTune*, the next step in development will involve concretizing it into an actual prototype. This step will involve interdisciplinary collaboration, possibly with other researchers in the fields of computer science and educational technology. It is likely that in creating the prototype, certain features outlined in the present article will be omitted or amended, based on either building capabilities or up-to-date research findings in either the fields of pronunciation training or gamification. As changes in technology occur so rapidly, it is necessary that the development process be flexible and open to modifications. For research purposes, these changes should be monitored and noted.

Once a prototype has been created, the next stages in development are to assess the pedagogical affordances and user attitudes towards the app (Cardoso, 2022). One possible model for addressing the latter would be to use the Technology Acceptable Model (Davis, 1989), which examines users' perception of the usefulness and useability of a tool. While this method of assessment was developed for information system in general, it has been used to examine CALL-based tools (for a recent example involving Google-based ASR, see Johnson, 2021).

In sum, this thesis represents the beginning of an exciting journey to create a new pedagogical tool that has the potential to reach users world-wide. The development of *FineTune* invites interdisciplinary collaboration and the perspectives of researchers in different domains, with the overall aim of improving learner practices in the field of pronunciation and gaining insight into the development of pedagogical tools in general.

References

Akbari, E., Pilot, A., & Simons, R-J. (2015). Autonomy, competence, and relatedness in foreign language learning through Facebook and Computers. *Human Behavior*, *48*, 126–134.

Ariely, D., Bracha, A., & Meier, S. (2007). Doing good or doing well? Image motivation and monetary incentives in behaving prosocially. *American Economic Review, 99*(1), 544–555.

Ariely, D., Gneezy, U., Lowenstein, G., & Mazar, N. (2005). Large stakes and big mistakes. *Review of Economic Studies, 76,* 459–469.

Barcomb, M., Grimshaw, J., & Cardoso, W. (2017). I can't program! Customizable mobile language-learning resources for researchers and practitioners. *Languages, 2*(8). 10.3390/languages2030008.

Barriuso, T. & Hayes-Harb, R. (2018). High variability phonetic training as a bridge from research to practice. *CATESOL Journal, 30*(1), 177–194.

Bitrián, P., Buil I., & Catalán, C. (2021) Enhancing user engagement: The role of gamification in mobile apps. *Journal of Business Research, 132,* 170–185.

Boggiano, A. K. & Ruble, D. N. (1979). Competence and the overjustification effect: A developmental study. *Journal of Personality and Social Psychology, 37*, 1462–1468.

Bowlby, J. (1979). *The making and breaking of affectional bonds*. Tavistock.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. (1997). Training Japanese listeners to identify english /r/ and /l/: Some effects of perceptual learning on speech production. *Journal of the Acoustical Society of America, 101*, 2299–2310.

Brandl, K. (2002). Integrating internet-based reading materials into the foreign language curriculum: From teacher to student-centered approaches. *Language, Learning & Technology, 6*(3), 87–107.

Breitkreuz, J. A., Derwing, T.M., & Rossiter, M.J. (2001). Pronunciation teaching practices in Canada. *TESL Canada Journal, Revue TESL Du Canada, 19*(1), 51–61.

Buckley, J., DeWille, T., Exton, C., Exton, G., & Murray, L. (2018). A gamification–motivation design framework for educational software developers. *Journal of Educational Technology Systems, 47*. 004723951878315. 10.1177/0047239518783153.

Burgers, C., Eden, A., van Engelenburg, M., & Buningh, S. (2015). How feedback boosts motivation and play in a brain-training game. *Computers in Human Behavior, 48*, 94–103.

Cardoso, W. (2022). Technology for Speaking Development. In T. Derwing, M. Munro, & R. Thomson (Eds.), *Routledge Handbook on Second Language Acquisition and Speaking* (p. 299-313). Routledge, Taylor & Francis Group.

Cardoso, W., Smith, G., & Garcia Fuentes, C. (2015). Evaluating text-to-speech synthesizers. In F. Helm, L. Bradley, M. Guarda, & S. Thouësny (Eds.), *Critical CALL—Proceedings of the 2015 EUROCALL conference*, Padova, Italy (pp. 108–113). Researchpublishing.net.

Cebrian, J., & Carlet, A. (2014). Second language-learners' identification of target-language phonemes: A short-term phonetic training study. *The Canadian Modern Language Review/La Revue canadienne des langues vivantes, 70*(4), 474–499.

Cerasoli, C. P., Nicklin, J. M., & Ford, M. T. (2014). Intrinsic motivation and extrinsic incentives jointly predict performance: A 40-year meta- analysis. *Psychological bulletin, 140*(4), 980.

Cheng, J. & Kim, H. (2019). Attitudes towards English language learning apps from Korean and Chinese EFL students. *English Teaching, 74*(4), 205–224.

Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340.

Deci, E. L., & Ryan, R. M. (1980). The empirical exploration of intrinsic motivational processes. In L. Berkowitz (Eds.), *Advances in experimental social psychology* (pp. 39–80). Academic Press.

Deci, E. L., & Ryan, R. M. (1985). The general causality orientations scale: Self-determination in personality. *Journal of Research in Personality, 19,* 109–134.

Deci, E. L., & Ryan, R. M. (1991). A motivational approach to self: Integration in personality. In R. Dienstbier (Eds.), *Nebraska symposium on motivation: Perspectives on motivation* (pp. 237–288). University of Nebraska Press.

Deci, E. L., & Ryan, R. M. (2000). The "what "and "why" of goal pursuits: Human needs and the self determination of behavior. *Psychological Inquiry, 11*(4), 37–41.

Deci, E. L., & Ryan, R. M. (2004). *Handbook of Self-determination Research*. University Rochester Press.

Deci, E. & Cascio, W. (1972). *Changes in Intrinsic Motivation as A Function of Negative Feedback and Threats*. University Rochester Press.

Dehganzadeh, H., & Dehganzadeh, H. (2020). Investigating effects of digital gamification-based language learning: A systematic review. *Journal of English Language Teaching and Learning, 12*(25), 53–93.

Derwing, T. M., & Munro, M. J. (2005). Second language accent and pronunciation teaching: A research-based approach. *TESOL Quarterly, 39(3),* 379–397.

Derwing, T. M., & Munro, M. J. (2015). *Pronunciation fundamentals: Evidence-based perspectives for L2 teaching and research*. John Benjamins.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From Game Design Elements to Gamefulness: Defining Gamification. *Proceedings of the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 9–15.

Dickerson, W. (2007). Orthography as a pronunciation resource. *World Englishes, 6,* 11–20.

Domínguez, A., Saenz-de-Navarrete, J., de-Marcos, L., Fernandéz-Sanz, L., Pagés, C., & Martínez-Herráiz, J.-J. (2013). Gamifying learning experiences: Practical implications and outcomes. *Computers & Education, 63*, 380–392. doi:101016/j.compedu.2012.12.020

Dong, T., Dontcheva, M., Joseph, D., Karahalios, K., Newman, M. W., & Ackerman, M. S. (2012). Discovery-based games for learning software. *Proceedings of the Conference on Human Factors in Computing Systems,* 2083–2086.

Egbert, J., & Shahrokni, S. (2018). CALL principles and practices. Open educational resources (OER). https://opentext.wsu.edu/call.

Fisher, C. D. (1978). The effects of personal control, competence, and extrinsic reward systems on intrinsic motivation. *Organizational Behavior and Human Performance, 21*, 273–288.

Foote, J., Holtby, A., & Derwing, T. (2012). Survey of the teaching of pronunciation in adult ESL programs in Canada, 2010. *TESL Canada Journal, 29*(1), 1–22. https://doi.org/10.18806/tesl.v29i1.1086

Glucksberg, S. (1962). The influence of strength of drive on functional fixedness and perceptual recognition. *Journal of Experimental Psychology, 63,* 36–41.

Grover, V., Namasivayam, A., & Nidhi, M. (2021). A viewpoint on accent services: Framing and terminology matter. *The American Journal of Speech-Language Pathology,* 1–10. https://doi.org/10.1044/2021_AJSLP-20-00376

Harmer, J. (2001). *The Practice of English Language Teaching*. Longman.

Hamari, J., Koivisto, J. & Sarsa, H. (2014). Does gamification work? – A literature review of empirical studies on gamification. *Proceedings of the HICSS 2013 Conference.*

Hassan, L., Dias, A., & Hamari, J. (2019). How motivational feedback increases user's benefits and continued use: A study on gamification, quantified-self and social networking. *International Journal of Information Management, 46,* 151–162.

Hassan, L., Xi, N., Gurkan, B., Koivisto, J., & Hamari, J. (2020) Gameful Self-Regulation: A Study on How Gamified Self-Tracking Features Evoke Gameful Experiences. *Proceedings of the 53rd Hawaii International Conference on System Sciences*.

Hernandez-Gonzalez, T. (2021). Lecture 1: What is Gamification [OneNote Slides]. Concordia University, Gamification in Teaching and Research.

Hofmann, D. A., & Strickland, O. J. (1995). Task performance and satisfaction: Evidence for a task- by ego-orientation interaction. *Journal of Applied Social Psychology, 25*(6), 495–511. https://doi.org/10.1111/j.1559-1816.1995.tb01764.x

Huensch, A. (2019). The pronunciation teaching practices of university-level graduate teaching assistants of French and Spanish introductory language courses. *Foreign Language Annals, 52*, 13– 31. https://doi.org/10.1111/flan.12372

Huotari, K., & Hamari, J. (2017). A definition for gamification: Anchoring gamification in the service marketing literature. *Electronic Markets, 27*(1), 21–31.

Iverson, P., & Evans, B. G. (2009). Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. *Journal of the Acoustical Society of America, 126*, 866–877.

Iverson, P., Hazan, V., & Banister, K. (2005). Phonetic training with acoustic cue manipulations: A comparison of methods for teaching English /r/-/l/ to Japanese adults. *The Journal of the Acoustical Society of America, 118*, 3267–3278.

Johnson, C. (2021) *Language learners' perceptions of automatic speech recognition as a writing tool: A Technology Acceptance Model analysis.* Masters thesis, Concordia University. Spectrum Research Repository.

Kennedy, C., & Levy, M. (2008). L'italiano al telefonino: Using SMS to support beginners' language learning. *ReCALL, 20*(3), 315–330.

Kim, K., Schmierbach, M. G., Chung, M. Y., Fraustino, J. D., Dardis, F., & Ahern, L. (2015). Is it a sense of autonomy, control, or attachment? Exploring the effects of in-game customization on game enjoyment. *Computers in Human Behaviour, 48*, 695–705.

Koívísto, J., & Hamari, J. (2019). The rise of motivational information systems: A review of gamification research. *International Journal of Information Management, 45,* 191–210.

Kukulska-Hulme, A. (2016). Mobile assistance in language learning: A critical appraisal. In Palalas, A. and Ally, M. (Eds.), *The international handbook of mobile-assisted language learning* (pp. 138–160).

Lambacher, S., Martens, W., Kakehi, K., Marasinghe, C., & Molholt, G. (2005). The effects of identification training on the identification and production of American English vowels by native speakers of Japanese. *Applied Psycholinguistics, 26*, 227–247.

Levis, J. M. (2018). Research into practice: How research appears in pronunciation teaching materials. *Language Teaching, 49*(3), 423–437.

Lively, S. E., Logan, J. S., & Pisoni, D. B. (1993). Training Japanese listeners to identify English /r/ and /l/. II: The role of phonetic environment and talker variability in learning new

perceptual categories. *The Journal of the Acoustical Society of America, 94*(3), 1242–1255.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y. I., & Yamada, T. (1994). Training Japanese listeners to identify English /r/and/l/. III: Long-term retention of new phonetic categories. *The Journal of the Acoustical Society of America, 96*(4), 2076–2087.

Logan, J. S., Lively, S. E., & Pisoni, D. B. (1991). Training Japanese listeners to identify English /r/ and /l/: A first report. *The Journal of the Acoustical Society of America, 89*(2), 874–886.

Locke, E. A., & Latham, G. P. (1990). *A theory of goal setting & task performance*. Prentice Hall, Inc.

Malone, T. W. (1981). Toward a theory of intrinsically motivating instruction. *Cognitive Science, 5*(4), 333–369.

Mulcahy, R., Russell-Bennett, R., & Iacobucci, D. (2020). Designing gamified apps for sustainable consumption: A field study. *Journal of Business Research, 106,* 377–387.

Munday, P. (2016). The case for using DUOLINGO as part of the language classroom experience. *Revista Iberoamericana de Educación a Distancia, 19*(1), 83–101. doi:10.5944/ried.19.1.14581

Munro, M. J., & Derwing, T. M. (2008). Segmental acquisition in adult ESL learners: A longitudinal study of vowel production. *Language Learning, 58,* 479–502.

Neri, A., Cucchiarini, C., & Strik, H. (2006). Selecting segmental errors in L2 Dutch for optimal pronunciation training. *International Review of Applied Linguistics in Language Teaching, 44,* 357–404.

Nishi, K., & Kewley-Port, D. (2007). Training Japanese listeners to perceive American English vowels: Influence of training sets. *Journal of Speech, Language, and Hearing Research, 50*, 1496–1509.

Nishi, K., & Kewley-Port, D. (2008). Nonnative speech perception training using vowel subsets: Effects of vowels in sets and order of training. *Journal of Speech, Language, and Hearing Research, 51,* 1480–1493.

Nicholson, S. (2012). A user-centered theoretical framework for meaningful gamification. *Presented at Games+Learning+Society 8.0.* Madison, WI.

Peng, W., Lin, J.-H., Pfeiffer, K. A., & Winn, B. (2012). Need satisfaction supportive game features as motivational determinants: An experimental study of a self-determination theory guided exergame. *Media Psychology, 15*(2), 175–196.

Perrachione, T., Lee J., Ha, L.Y., & Wong, P.C. (2018). Learning a novel phonological contrast depends on interactions between individual differences and training paradigm design. *Journal of Acoustical Society of American, 130*(1), 461.

Pink, D. (2011). Chapter 2: Seven Reasons Carrots and Sticks (Often) Don't Work. In *Drive: the surprising truth about what motivates us* (pp. 32–57). Riverhead Books.

Qian, M., Chukharev-Hudilainen, E., & Levis, J. (2018). A system for adaptive high-variability segmental perceptual training: Implementation, effectiveness, transfer. *Language Learning and Technology, 22*(1), 69–96.

Rahman, M., & Pandian, A. (2016). The gap between research and practice in the field of SLA: The rationale behind it and bridging the gap. *Journal of Applied Linguistics and Language Research, 3,* 162–172.

Ramirez, D., & Squire, K. (2015). Gamification and Learning. In S.P. Walz & S. Deterding (Eds.), *The Gameful World. Approaches, Issues, Applications* (pp. 629–652). MIT Press.

Read, T., Bárcena, E. & Kukulska, Hulme, A. (2016). Exploring the application of a conceptual framework in a social MALL app. In Pareja-Lora, A., Calle-Martínez, C. and Rodriguez Arancón, P. (Eds.), *New perspectives on teaching and working with languages in the digital era* (pp. 223–232). https://doi.org/10.14705/rpnet.2016.tislid2014.436

Reeve, J. (2004). Self-Determination Theory Applied to Educational Settings. In E.L. Deci & R. M. Ryan (Eds.), *Handbook for Self-Determination Research* (pp. 2–183). University of Rochester Press.

Reinders, H. & Hubbard, P. (2013). CALL and learner autonomy: Affordances and constraints. In M. Thomas, H. Reinders, and M. Warschauer (Eds.). *Contemporary computer assisted language learning* (pp. 359–375). Continuum Books.

Reitz, J.G., & Sklar, S.M. (1997). Culture, race, and the economic assimilation of immigrants. *Sociological Forum, 12*, 233–277.

Ryan, R. M., & Deci, E. L. (2000). The darker and brighter sides of human existence: Basic psychological needs as a unifying concept. *Psychological Inquiry, 11*(4), 319–338.

Ryan, R. M., Stiller, J., & Lynch, J. H. (1994). Representations of relationships to teachers, parents, and friends as predictors of academic motivation and self-esteem. *Journal of Early Adolescence, 14*, 226–249.

Sailer, M., Hense, J., Mayr, S., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior. 69,* 371–380. 10.1016/j.chb.2016.12.033.

Saito, K. (2013). The acquisitional value of recasts in instructed second language speech learning: Teaching the perception and production of English /ɹ/ to adult Japanese learners. *Language Learning, 63,* 499–529. https://doi.org/10.1111/lang.12015

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

Schreurs, B., Van Emmerik, I. J., Van den Broeck A., & Guenter, H. (2014). Work values and work engagement within teams: The mediating role of need satisfaction. *Group Dynamics, Theory, Research, and Practice, 18,* 267–281.

Schwartz, B. (2009). *The Paradox of Choice*. Harper Collins.

Seaborn, K., & Fels, D. I. (2015). Gamification in theory and action: A survey. *International Journal of Human-Computer Studies, 74,* 14–31. doi: 10.1016/j.ijhcs.2014.09.006.

Shortt, M., Tilak, S., Kuznetcova, I., Martens, B. & Akinkuolie, B. (2021). Gamification in mobile-assisted language learning: A systematic review of Duolingo literature from public release of 2012 to early 2020. *Computer Assisted Language Learning*. DOI: 10.1080/09588221.2021.1933540

Spada, N. (1997). Form-focused instruction and second language acquisition: A review of classroom and laboratory research. *Language Teaching, 29*, 73–87.

Steel, C. (2012). Fitting learning into life: Language students' perspectives on benefits of using mobile apps. In Brown, M., Hartnett, M. and Stewart, T. (Eds.), *Ascilite 2012: Future challenges, sustainable futures: Proceedings* (pp. 875–880).

Stockwell, G., & Hubbard, P. (2013). Some emerging principles for mobile-assisted language learning. Monterey, CA: *The International Research Foundation for English Language*

Education. Retrieved from http://www.tirfonline.org/english-in-the-workforce/mobile-assisted-language-learning

Suits, B. (1978). *The grasshopper: Games, life, and utopia*. University of Toronto Press.

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.

Thomson, R. (2011). Computer assisted pronunciation training: Targeting second language vowel perception improves pronunciation. *CALICO Journal, 28*(3), 744.

Thomson, R. (2012). Improving L2 listeners' perception of English vowels: A computer mediated approach. *Language Learning, 62*(4), 1231–1258.

Thomson, R. (2013). ESL teachers' beliefs and practices in pronunciation teaching: Confidently right or confidently wrong? In J. Levis & K. LeVelle (Eds.) *Proceedings of the 4th Pronunciation in Second Language Learning and Teaching Conference. Aug. 2012*. (pp. 224-233).

Thomson, R., & Derwing, T. (2016). Is phonemic training using nonsense or real words more effective? In J. Levis, H. Le, I. Lucic, E. Simpson, & S. Vo (Eds.), *Proceedings of the 7th Pronunciation in Second Language Learning and Teaching Conference* (pp. 88- 97).

Thomson, R. (2018a). High Variability [Pronunciation] Training (HVPT): A proven technique about which every language teacher and learner ought to know. *Journal of Second Language Pronunciation, 4,* 208–231. 10.1075/jslp.17038.tho.

Thomson, R. (2018b). English Accent Coach [Computer Program]. Version 2.3. Retrieved from < https://www.englishaccentcoach.com/>

Thorne, S. (2003). Artifacts and cultures-of-use in intercultural communication. *Language Learning & Technology, 7*(2), 38–67.

Toups, Z., Crenshaw N., Wehbe, R., Tondello, G., & Lennart. E. (2016) "The Collecting Itself Feels Good": Towards Collection Interfaces for Digital Game Objects. In *Proceedings of the 2016 Annual Symposium on Computer-Human Interaction in Play*. Association for Computing Machinery, 276–290. https://doi.org/10.1145/2967934.2968088

van Roy, R., & Zaman, B. (2019). Unravelling the ambivalent motivational power of gamification: A basic psychological needs perspective. *International Journal of Human Computer Studies, 127,* 38–50.

van Roy, R., & Zaman, B. (2017). Why gamification fails in education and how to make it successful: Introducing nine gamification heuristics based on self-determination theory. In M. Minhua, & A. Oikonomou (Eds.), *Serious Games and Edutainment Applications* (pp. 485–509).

Wang, Y., Jongman, A., & Sereno, J. A. (2003). Acoustic and perceptual evaluation of Mandarin tone productions before and after perceptual training. *The Journal of the Acoustical Society of America, 113*, 1033–1043.

Wee, S. C., & Chong, W. W. (2019). Gamification: Predicting the effectiveness of variety game design elements to intrinsically motivate users' energy conservation behaviour. *Journal of Environmental Management, 233,* 97–106.

White, R. W. (1959). Motivation reconsidered: The concept of competence. *Psychological Review, 66*(5), 297–333. https://doi.org/10.1037/h0040934

Wrembel, M. (2002). New perspectives on pronunciation teaching. In W. Sobkowiak & E. Waniek- Klimczak (Eds). *Dydaktyka Fonetyki Języka Obcego na Poziomie Licencjackim, Neofilologia II* (pp. 173–183).

Xi, N., & Hamari, J. (2019). Does gamification satisfy needs? A study on the relationship between gamification features and intrinsic need satisfaction. *International Journal of Information Management, 46*, 210–221.