

Language learners' perceptions of automatic speech recognition as a writing tool:

A Technology Acceptance Model analysis

Carol Johnson

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

Montreal, Quebec, Canada

June 2021

© Carol Johnson, 2021

CONCORDIA UNIVERSITY
School of Graduate Studies

This is to certify that the thesis prepared

By: Carol Johnson

Entitled: Language learners' perceptions of automatic speech recognition as a writing tool:
A Technology Acceptance Model analysis

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

MA 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:

D. Liakin Examiner

W. Cardoso Examiner

W. Cardoso Thesis Supervisor

Approved by _____
J. Capobianco Chair of Department or Graduate Program Director

Pascale Sicotte Dean

ABSTRACT

Language learners' perceptions of automatic speech recognition as a writing tool:

A Technology Acceptance Model analysis

Carol Johnson

Automatic speech recognition (ASR) has the potential to mitigate the cognitive burden of L2 writing by facilitating the text input process (using a skill most humans possess: speaking) and offering assistance in terms of linguistic form, thus allowing writers to focus on other aspects of the task (e.g., cohesion, content). ASR is accessible, easy to use, and free; more importantly, it fulfills Chapelle's (2001) criteria of an effective CALL tool (e.g., it promotes authenticity, has potential for language learning). Despite these affordances, there is a dearth of studies examining the affordances of ASR for writing, and none examining the use of ASR with adult ESL writers.

This mixed-methods one-shot study examines L2 writers' perceptions of using ASR to write using the Technology Acceptance Model (TAM; Venkatesh & Davis, 2000), based on three criteria: usefulness, ease of use, and intention to use. The participants were ESL students at an English-medium university (N=17). They were provided with training on Google Voice Typing in Google Docs and, as part of the treatment, carried out two ASR-based writing tasks over a two-hour period. To measure their perceptions of the target criteria, participants filled in a TAM-informed survey after completing the treatment. To further explore the participants' perceptions of using ASR to compose their texts, semi-structured interviews followed the writing tasks.

Findings indicate positive perceptions of ASR in terms of usefulness (language learning potential) and its ease of use (e.g., user-friendly voice commands). As observed in the literature (e.g., Hsu., 2016; Tsai, 2015), these positive perceptions will lead to an intention to continue to use ASR, suggesting that the technology has L2 pedagogical potential.

Acknowledgements

I would like to express my enormous gratitude to my supervisor, Dr. Walcir Cardoso. Through this entire process, you have been simultaneously my biggest critic and my greatest fan. You have constantly pushed me to go beyond, and this research project is better than I had ever imagined it would be because of your support and your belief in me. It is only the first of many more to come - I am looking forward to continuing our collaboration together. I would also like to thank Dr. Denis Liakin for taking the time to read my proposal and thesis for your invaluable feedback. Merci!

I must give special thanks to my daughter, Anne-Marie Sénécal. You made all this possible. I would never have done an M.A. if you had not encouraged me to go back to university. Because of your encouragement and support, I have entered a new world of experiences and opportunities that I never realized existed. Beyond that, it has been such a unique and wonderful experience being in the program with you. Thank you for sharing this experience with me.

I also owe a special thanks to my boys, Andrew and Ian. They have put up with me being away from home many evenings, and even when I am home, I am often lost in my studies or my research. Their understanding has allowed me to finish this degree (relatively) guilt-free!

And finally - Dave, I did it!

Table of Contents

LIST OF FIGURES	viii
LIST OF TABLES	ix
Chapter 1	1
Writing in the 21 st Century	3
ASR and L2 Writing	4
Technology Acceptance Model	4
This study: Google Voice Typing, Goals, and Hypotheses	5
Chapter 2	7
Background	10
Automatic Speech Recognition (ASR)	10
ASR in L2 Learning	11
Google Voice Typing: ASR and CALL Affordances	14
Technology Acceptance Model (TAM)	17
TAM in CALL	19
Current Study	21
Method	22
Design	22
Research Context and Participants	24
Materials	25
Training Session	25
Pronunciation Accuracy Test/Training Activity	25
Writing Tasks	26

TAM Survey	26
Interview	27
Data-analysis Measures	27
Results.....	28
Quantitative Results	28
Qualitative Results	31
Perceived usefulness, academic relevance, output quality, and result demonstrability.....	31
Academic Relevance.....	31
Output Quality.	32
Result Demonstrability.	32
Perceived Ease of Use.....	33
Intention to Use.....	33
Discussion.....	33
Perceived Usefulness, Academic Relevance, Output Quality, and Result Demonstrability	34
Perceived Ease of Use.....	37
Intention to Use.....	38
Conclusion	38
Chapter Three	41
Summary of Goals and Findings.....	41
Implications for Education.....	42
Implications for Research	43
Conclusion	44
References	45

Appendices	52
Appendix A.....	52
Appendix B.....	53
Appendix C.....	58
Appendix D.....	60

LIST OF FIGURES

<i>Figure 1. Automatic Speech Recognition Workflow</i>	3
<i>Figure 2. Automatic Speech Recognition Architecture.....</i>	11
<i>Figure 3. Using Google Voice Typing</i>	16
<i>Figure 4. Proposed Technology Acceptance Model for Current Study</i>	19
<i>Figure 5. Study Design.....</i>	23
<i>Figure 6. Sample Training Activity.....</i>	25

LIST OF TABLES

Table 1. <i>Sample Statements From the TAM Survey</i>	27
Table 2. <i>Descriptive Statistics of Survey Results</i>	30

Chapter 1

When I went to university for the first time in the 1980s, I would hand-write the rough draft of an essay and then type the final copy on an electric typewriter. Then, in the last year of my degree, we got a PC. I did not understand how it worked and was terrified of it; however, I still used it to type the final copy of my essays because, despite my fear and lack of understanding, I realized that it made the task of writing much easier. Mistakes could be easily corrected, and I could change something after it had been written, which meant that I could save precious time. I continued to write draft copies on paper, though, and I was not alone, as many had difficulty composing new ideas and typing on the keyboard at the same time. This was not an issue related to typing skills – I had taken typing classes in high school and was a proficient typist. It was because my brain would freeze when I sat in front of this new and overwhelming technology.

After I graduated from university, I bought a book and taught myself how to properly use WordPerfect; I eventually became less stressed and nervous when sitting at the keyboard. Computers had become more ubiquitous, and I lost my fear of them. Hand-written drafts became an anachronism, a relic of the past, as I learned to compose and type at the same time. And again, I was not alone - it was the early 1990s and word processing software was ushering in a new paradigm in writing as people started to embrace this new technology for use at home, at work, and for school.

Word processing software was not the first technology to cause a paradigm shift in writing. As explained by Dennis Baron (2009) in *A better pencil: Readers, writers and the digital revolution*, from the time of clay tablets to the time of digital tablets, there have been many technologies that have revolutionized how we write. With each technology change, the

affordances offered by the new tool led to a tenacity shown by writers to use the new technology, despite the difficulties and expenses that may come with it. Writing on clay tablets was difficult, not only in terms of physically etching out what the writer wanted to say, but also in terms of sourcing and forming the tablet. Yet people persisted because it helped them to put down information that needed to be remembered. As pens and parchment took over, people had to prepare their own pens and ink, and at the same time source expensive parchment (or dry skins to make their own). They persisted, though, because it allowed them to communicate stories and ideas that needed sharing. The original word processing software was not user-friendly nor inexpensive. Writers still persisted because it allowed them to not only write documents but format them in ways that were previously unimaginable and to share their work with the world.

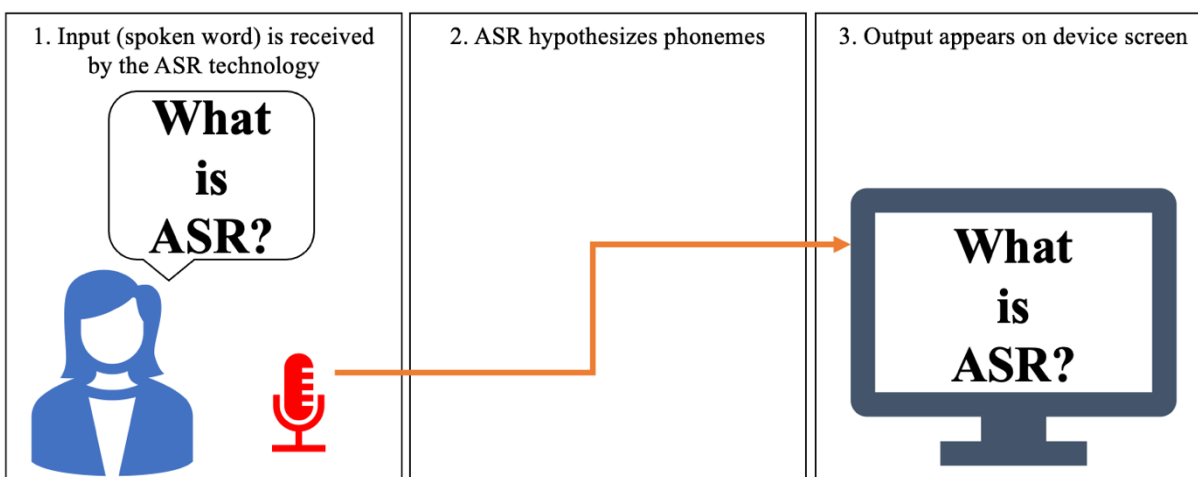
In each paradigm shift, writers overcame the difficulties presented by the new technology, to the point that these difficulties seem unimaginable to us today. Who would think that a pencil was difficult to use when first put to paper? With each paradigm shift, writers are initially slowed down until design changes allow for advancement (Baron, 2009). However, one constant across the paradigm shifts is that no technology has allowed writers to put their ideas into words *instantaneously* (Baron, 2009).

A “new” technology has emerged that just might allow writers to write as fast as they can think. Or at least as fast as they can *speak*. Automatic speech recognition (ASR) allows writers to speak their ideas and have the words almost instantly appear on a computer screen, with no physical action needed on the part of the writer. Writers simply activate the ASR feature in their work processor and speak. The spoken input is modified into textual output on the computer screen (See Figure 1). No slowing down of the stylus in the mud, no sharpening of pencils that are dull, no stopping to do a carriage return on the typewriter, and no searching for the key on a

keyboard. The writer simply speaks, which is a skill most of us learn long before we learn to write or keyboard. ASR may just be a “better pencil” for 21st-century writers.

Figure 1

Automatic Speech Recognition Workflow



Writing in the 21st Century

Technology has changed the way that audiences interact with written texts. Readers now engage with multimodal texts, often reading genres that have only recently come into existence, such as blogs and listicles. Web 2.0 tools have also allowed writers to reach potentially huge audiences instantly in a previously unimaginable manner (National Council of Teachers of English, 2018). To meet the expectations of these audiences, writers more than ever need to focus not only on linguistic form when writing, but also genre, content, context, and the writing process (Hyland, 2011). However, university-level English Second Language (ESL) students focus almost exclusively on form when writing in ESL classes (Hyland, 2011), perhaps because this is what most teachers provide feedback on when correcting (Menke & Anderson, 2019). This does not allow these language learners to focus on meaning (or other aspects of writing) to the extent that is necessary, limiting their potential as writers. This is crucial for the career

aspirations of these students as they enter a work market where employers increasingly value the ability to write when hiring new employees (Allen, 2018). ASR could offer second language (L2) writers affordances that would allow them to focus on meaning to a greater degree, thus improving their writing.

ASR and L2 Writing

As an ESL teacher who spent more than a decade teaching intermediate to advanced high school students the basics of academic writing, the affordances of the tool struck me immediately when I experimented with it for the first time. However, despite the studies that indicate that ASR is a useful tool for improving the pronunciation of L2 learners (e.g., Liakin et al., 2014; McCrocklin, 2018; Mroz, 2018), no studies have been done focusing on using ASR to dictate thoughts into a word processing software (in other words, to write) with L2 learners. Indeed, there exists only extremely limited research into its use with L1 writers with learning difficulties. For instance, Quinlan (2004) found that young native English speakers with writing difficulties were able to leverage the affordances of ASR to produce longer, more accurate, and more narratively complex texts than when writing by hand. This tool has great potential for writers, but its suitability for L2 writing must be determined before attempting to bring about its acceptance by users (Cardoso, in press). Suitability can be determined using the Technology Acceptance Model (TAM), a model for examining user perceptions of a technology.

Technology Acceptance Model

The Technology Acceptance Model (TAM) (Venkatesh and Davis, 2000), a well-established model that analyzes the adoption potential of technology based on user perceptions, was chosen as the framework for this study. In the TAM framework, perceived usefulness and perceived ease of use are seen as predictors of intention to use the technology. By measuring the

perceptions of my participants in a mixed-methods study, I will be able to understand their beliefs about the usability of the technology. This will allow me to determine the suitability of ASR as a writing tool.

TAM is not widely used in CALL research but has been increasingly used to determine user perceptions in studies involving university-level participants (e.g., Dizon & Thanyawatpokin, 2018; Haghghi et al., 2019; Soleimani et al., 2014; Tan, 2019). Overall, these studies show that English L2 learners have positive perceptions of technologies that help them learn English, often because they see the technology as useful in helping them attain higher grades (e.g., Tsai, 2015) and/or improve their language skills (e.g., Soleimani et al., 2014).

This study: Google Voice Typing, Goals, and Hypotheses

This study examines L2 user perceptions of using ASR to write academic texts. Google Voice Typing, the ASR embedded in Google Docs, is a robust system that has been shown to be quite accurate when used by highly proficient English L2 speakers. Indeed, recognition rates of L2 speaker utterances are only 3 – 5% lower than those of English L1 speakers (McCrocklin & Edalatishams, 2020). As it is free of charge to anyone using Google Docs (via the Google Chrome browser), it is easily accessible. Google Voice Typing is also user-friendly as it is simply activated with the click of a mouse. In addition, this technology fits Chapelle's (2001) criteria for evaluating computer-assisted language learning tools; that is, the use of ASR has potential for language learning through the feedback provided, presents opportunities to engage with language at an appropriate level as it is the learner's own text, allows for an attention to meaning as the learner constructs a text, provides for an authentic use of language in a text type that the learner will frequently encounter, has a positive impact on the writer's academic results, and is practical since it is readily available at no cost.

This study aims to use the TAM framework to determine the perceptions of L2 writers using ASR to write their own texts. Based on the affordances offered by ASR, the ease of use of the technology, and the results of other TAM-based studies on CALL tools, it is hypothesized that the participants will have positive perceptions of the technology. As history has shown, writers will embrace new technologies that promise to improve the creation and distribution of their texts, a promise that is inherent in the simplicity of ASR.

This research could be at the ground floor of a paradigm shift in writing technology; by measuring L2 writers' perceptions of ASR, I am taking the first steps of exploring its suitability as a writing tool. Automatic speech recognition could be the "better pencil" that is the catalyst for another paradigm shift in writing.

As per the guidelines for a manuscript-based MA thesis, the next section constitutes "a full submittable draft of a manuscript" that presents the full literature review, methodology, results, and discussion of the abovementioned research.

Chapter 2

Writing in the 21st Century has become a multi-modal skill, largely in thanks to the advent of hand-held devices that are central to our social and professional lives (Allen, 2018). For example, writing is no longer restricted to traditional genres (e.g., academic writing), audiences, and media since it now also encompasses genres that appeared with the advent of the Web 2.0 (e.g., blogs and wikis), is carried out for a more widely varied audience (e.g., co-workers, family and friends, strangers), and is integrated within new media (e.g., social media posts, web sites) (National Council of Teachers of English, 2018). Writing such highly varied text types requires the writer to not only master linguistic form, but to also pay skillful attention to text content, text context, genre, and writing processes (Hyland, 2011). One domain in which the mastery of writing in all of its varied forms is of essence is the workplace. Employers value employees who have developed effective writing skills and expect university graduates to possess this skill upon entering the workforce (Allen, 2018). For this reason, it is vital that during their education, university students learn to become effective 21st century writers, not only in their mother tongue (L1), but also in any second or subsequent language(s) (L2). This is especially true in the context of English as a *Lingua Franca* (or *International Language*; Jenkins, 2000), as university students who speak English as a *second* language (ESL) need to develop the multiple skills needed to write effectively in English in order to succeed professionally after finishing university (American Academy of Arts and Sciences, 2017).

However, despite the need for advanced writing skills among learners of English, writing is often the orphan skill in many ESL university-level classes. Magnan et al. (2014) found that most L2 learners and instructors undervalue the importance of developing writing skills in a classroom in comparison with speaking. One reason this occurs is that many students tend to

focus on linguistic form when writing, rather than on developing content, as they believe that this is the element that teachers will evaluate the most (Barkaoui, 2016; Ding & Zhao, 2019). As a result, students spend inordinate amounts of time verifying grammar and spelling rules rather than focusing on the other elements of effective writing (Hyland, 2011), resulting in texts that may be grammatically correct, but lack content and depth (Ding & Zhao, 2019). This is further complicated by the fact that teachers may be inadvertently reinforcing this perception. In a study of teacher and student perceptions of second language writing, Menke and Anderson (2019) found that while professors of a university-level Spanish L2 writing class claimed that content was more important than linguistic form, most feedback provided for written texts focused on grammatical and spelling errors. As a result, many L2 learners do not fully develop their writing potential, as the other elements of effective writing are sacrificed for linguistic form.

There has been a call to address the limitations that affect the teaching and learning of L2 writing, leading teachers and researchers to increasingly turn to computer-assisted language learning (CALL) in order to bring about these necessary changes. Studies in the last decade have focused on the use of various technological platforms – such as wikis, blogs, and Google Docs – to help students improve L2 writing skills (e.g., Aslan & Ciftci, 2018). These tools allow for peer and teacher feedback (Alharbi, 2020) and authentic sharing of texts (Lee, 2017), which is beneficial after a student writes a text. However, little research has been done into removing the burden of focusing on linguistic form while writing a text in order to allow students to shift their focus to the other elements of effective writing.

Incorporating the use of automatic speech recognition (ASR) into the writing of texts may be one way to reduce this burden. Quinlan (2004) found that young native English speakers with learning difficulties produced texts using ASR that were longer and more narratively complex

than texts written by hand. As well, there were fewer errors in the texts as the ASR program provided error-free spelling of most words. The increased complexity in the texts and the automatic error-free production by the ASR show the potential for this technology to allow students to go beyond form when writing. Furthermore, high quality ASR is easily accessible by language learners for free, on devices that are already part of their day-to-day lives. Indeed, the use of ASR on smartphones is increasing sharply. While only 25% of Americans with smartphones used ASR features at least once a month in September 2018, the number had increased to 33% by early 2019 (Voicebot Research, 2019). Not only is this technology accessible and free, but ASR fulfills all of Chapelle's (2001) criteria of an effective CALL tool. As will be discussed later, the pedagogical use of ASR has potential for language learning through the feedback provided, presents opportunities to engage with language at an appropriate level as it is the learner's own text, allows for an attention to meaning as the learner constructs a text, provides for an authentic use of language in a text type that the learner will frequently encounter, contributes to the development effective writing strategies, and is practical since it is readily available at no cost.

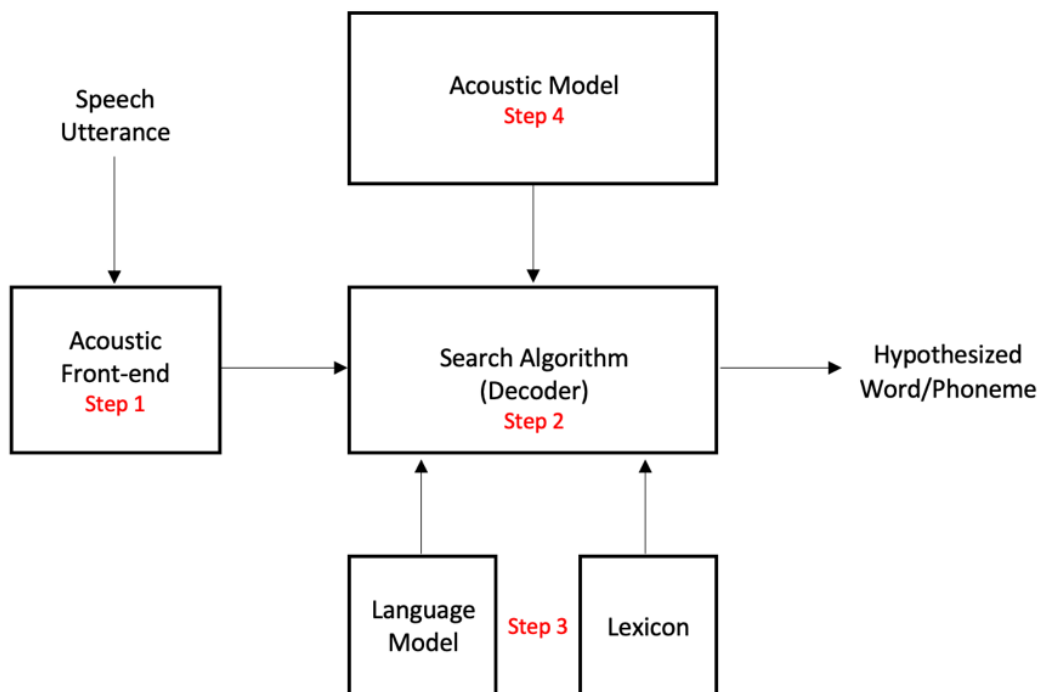
Despite the multiple affordances that ASR offers to language learners who want to develop their writing skills, this technology has not been mined for its potential in ESL writing. However, before any large-scale study on ASR's affordances and pedagogical potential, it is important to determine if language learners, who may already use ASR on their mobile devices (e.g., to send informal texts or perform internet searches), are willing to accept this technology in the context of writing in an academic setting. Thus, the aim of this study is to determine if language learners would accept to use ASR to write academic texts using the framework adapted from Venkatesh and Davis' (2000) revised Technology Acceptance Model (TAM). User

perceptions of using ASR to write academic texts will be measured quantitatively and qualitatively. Quantitative data will be measured via surveys to determine perceived usefulness (PU), perceived ease of use (PEU), and intention to use (IU), the three main constructs of TAM. Qualitative data will be measured via interviews to determine the affordances of ASR when writing L2 texts.

Background

Automatic Speech Recognition (ASR)

Automatic speech recognition (ASR) is a process through which spoken word is encoded by a computer and then subsequently decoded and transcribed as written text (Levis & Suvorov, 2012). Input that is generated through a speech act is analyzed by an acoustic front-end processor which extracts relevant information (Figure 2, Step 1). This information is processed by the decoder (Figure 2, Step 2) which uses algorithms that search the lexicon and language model databases to hypothesize the phonemes and/or words that were said (Figure 2, Step 3). These are analyzed by the acoustic model, which statistically analyzes the probability of different phonemes/words occurring together (Figure 2, Step 4). The most probable phonemes/words are then output in the form of written text (Filippidou & Moussiades, 2020), illustrated in Figure 2.

Figure 2*Automatic Speech Recognition Architecture*

Note. Adapted from Filippidou and Moussiades (2020, p. 75).

ASR in L2 Learning

ASR has been used to help L2 learners improve grammar (e.g., Bodnar et al., 2011, 2017; de Vries et al., 2015), reading abilities (e.g., Poulsen et al., 2007), listening skills (e.g., Mirzaei et al., 2017), and meaningful interactions (e.g., Chen, 2011; Chiu et al., 2007). However, the vast majority of research has occurred in the areas of pronunciation (e.g., Levis & Suvorov, 2012) and speaking (e.g., Cardoso, in press; Walker et al., 2011). The potential affordances of ASR in pronunciation practice were recognized as early as the 1990s. In fact, Eskenazi (1999) explained that ASR not only allows learners to produce a large quantity of output, but also to receive helpful corrective feedback (e.g., via the orthographic output of ASR). However, he noted that ASR technology, as it stood at the time, was problematic due to its inability to always correctly

interpret speech, leading to errors being missed or correct utterances being labelled as erroneous (Eskenazi, 1999). Incorrect interpretation of speech was even more evident if the speaker had an accent, as was shown in Derwing et al. (2000), in a study that compared L1 English speaker recognition rates in ASR to high proficiency L2 speaker recognition rates. The same English sentences were read into Dragon System's Naturally Speaking ASR dictation program by L1 speakers of English, Cantonese, and Spanish. They found that the recognition rate for L1 English speakers was 90.25%, but it was only 72.45% for L1 Cantonese speakers dictating in English, and 70.75 for the L1 Spanish speakers, despite the fact that the Cantonese and Spanish speakers were highly proficient in English. The researchers concluded that until the technology behind ASR dictation programs improved, ASR "cannot be considered to be of benefit to ESL speakers either in the classroom or in business and personal contexts" (Derwing et al., 2000, p. 602).

Due to this issue with recognition rates, researchers and developers focused on creating ASR-based computer-assisted pronunciation tools (CAPT) that were pre-programmed to recognize limited discrete items (e.g., My Pronunciation Coach in Cucchiarini et al., 2012; GeCALL in Young & Wang, 2014) rather than using generic, off-the-shelf ASR dictation tools. Research seems to indicate that these systems are beneficial in improving pronunciation (e.g., Elimat & AbuSeileek, 2014; Wang & Young, 2014).

However, with improvements in the ASR technology, researchers have once more begun investigating the effectiveness of ASR dictation programs in L2 contexts (McCrocklin, 2018). Liakin et al., (2014) found that participants learning French as an L2 who practiced pronouncing the vowel /y/ (as in *tu* "you") using Nuance Dragon Dictation, a free dictation app downloadable on mobile devices, outperformed participants in a control group that practiced in class with a teacher. In a study of the perceptions of sixteen ESL university students using Windows Speech

Recognition (WSR), McCrocklin (2018) found that 81% of the participants found using ASR for pronunciation practice useful because of the feedback provided (implicitly, via orthography). However, students reported being frustrated by instances when the system could not understand them, even when they thought they were pronouncing a word correctly. One participant went as far as having his L1 English roommates try, and they were still unable to make the ASR application understand.

In hopes of avoiding the issues inherent in WSR, Mroz (2018) used the ASR in Gmail on the participants' smartphones in her study of sixteen L2 learners of French in an American university. The participants reported an appreciation of the feedback available on their phone screens and did not report the same levels of frustration with comprehensibility as the participants in McCrocklin's (2018) study. Indeed, Mroz's (2018) participants reported increased confidence in their French speaking skills due to the fact that the system often understood what they were saying. The findings of Mroz's (2018) study seem to indicate that some ASR dictation tools have advanced to the point where L2 speakers can use them without experiencing frustration, unlike when the technology was first released. Because of its ability to accurately understand accented L2 speech (see also Moussalli & Cardoso, 2019 involving Intelligent Personal Assistants), ASR is possibly ready to be used by L2 writers.

Research has shown that advancements in ASR technology have made it possible for L2 speakers to use it without experiencing the frustration of being unjustifiably misunderstood. However, the focus of research into using ASR in language learning remains on pronunciation practice (e.g., Liakin et al., 2014; McCrocklin, 2018; Mroz, 2018). No studies have explored L2 learner use of ASR dictation technology for the purpose for which it was originally designed – to

transform spoken word into written text. An interesting ASR tool designed with dictation in mind that should be further explored is Google Voice Typing.

Google Voice Typing: ASR and CALL Affordances

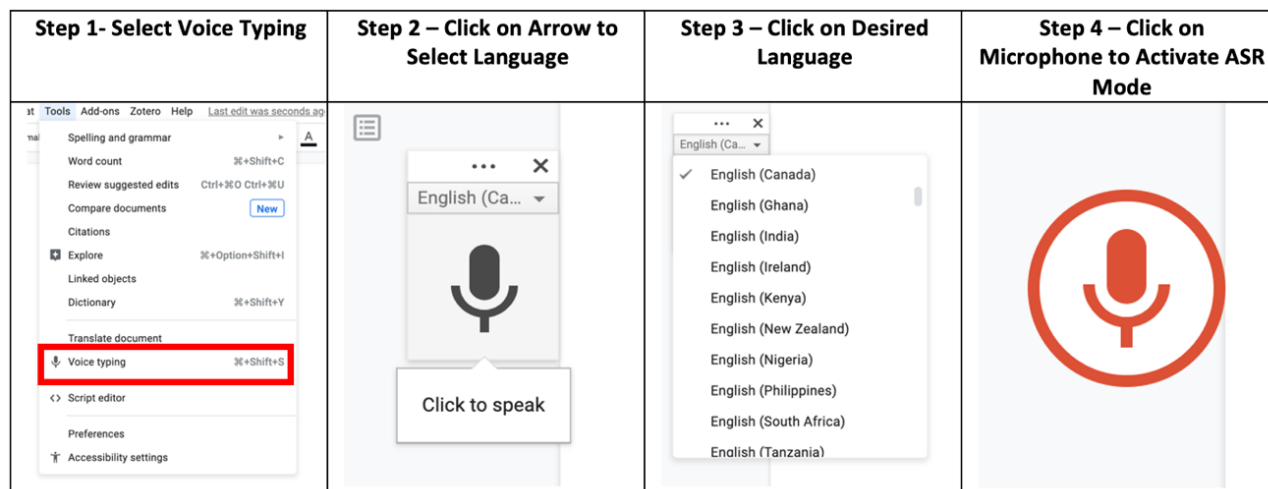
Google has developed a robust ASR system that is based on an extensive database that is processed in a cloud-based setting (McCrocklin & Edalatishams, 2020). Filippidou and Moussiades (2020) compared Google's ASR system with that of IBM Watson and the ASR program Wit in order to determine which system would be the optimal for an L2 pronunciation program currently under development. They found that Google had the lowest system-produced word error rate and the smallest average error rate of the three programs, in addition to adding/removing fewer words. As well, it had the highest number of error-free sentences, which they felt was essential for their program as the researchers did not want do not “an approximation but an accurate recognition” of the users speech for their pronunciation program (Filippidou & Moussiades, 2020, p. 79).

McCrocklin and Edalatishams (2020) compared recognition rates among L1 and L2 users of Google Voice Typing, which is based on Google's ASR product and is available free of charge in all applications in Google Drive on the Google Chrome browser. The authors re-used sentences from Derwing et al. (2000), who had found ASR recognition of highly proficient L2 speakers' utterances to be 15 - 20% less than those of L1 speakers (see discussion above). However, in McCrocklin and Edalatishams (2020), the recognition rate difference was only 3 – 5%. Indeed, the authors (2020) noted that Google Voice Typing was often “forgiving of pronunciation errors” (p. 9), as there were 78 pronunciation errors that did not result in transcription errors. One reason for this is that ASR technology is not only acoustic in nature, but also predictive. According to Ashwell and Elam (2017), pronunciation errors are overridden

“when the input conforms to canonical and typical sentence patterns due to the machine learning algorithms” (p. 73).

Ashwell and Elam (2017) used a customized version of Google’s ASR to design a tool for testing English grammar knowledge among Japanese learners. They found that this predictive nature, which can greatly facilitate an L2 speaker’s use of the system, can also create problems when less common collocates (e.g., “first coat” - less common vs. “first time” - more common) and ambiguous grammatical structures (e.g., tag questions) are used as they were often incorrectly transcribed. Thus, Google’s ASR, which is readily available through Google Voice Typing, while not perfect, has shown great advances in its usability for L2 speakers of English, with recognition rates approaching those of native speakers.

Accessing Google Voice Typing to transcribe a document is simple, as illustrated in Figure 3. When in Google Docs, the user clicks on Tools and Voice Typing (Step 1). A microphone appears to the left of the screen (Step 2) from which more than 120 languages can be selected, including more than a dozen varieties of English (Step 3). Once the language has been chosen, the user clicks on the microphone, which turns red (Step 4), indicating that the ASR mode has been activated and dictation can begin.

Figure 3*Using Google Voice Typing*

Considering all six of Chapelle’s (2001) evaluation criteria for CALL-based tasks, Google Voice Typing is an ideal CALL tool for L2 learners to compose academic writing tasks:

- Language learning potential. Although it is being suggested that ASR be used to alleviate a writer’s focus on linguistic form, at the same time, the output generated by the technology can provide feedback on many linguistic forms that the writer may notice incidentally, which is essential to language learning (Schmidt, 1995). For example, writers are shown the correct spelling of words they can produce orally but may be unsure of in terms of orthography.
- Learner fit. Learners can use ASR to create texts appropriate for their level of language but that require the application of new language skills (e.g., using academic vocabulary and a specific text structure).
- Meaning focus. Using ASR may allow writers to focus on the content and meaning in their texts rather than on producing correct linguistic form.

- **Authenticity.** ASR can be used to write academic texts in English, a task that learners carry out frequently in all of their university classes. As well, with the increasing prevalence of ASR in everyday life, L2 learners will come into increasing contact with the technology.
- **Positive impact.** Learners may have increase accuracy and improved content in their texts, leading to higher grades and confidence in their writing. Moreover, learners may develop new strategies for writing texts that can be used both inside and outside of an academic context, leading to more opportunities to use and improve their English.
- **Practicality.** ASR is available on many word processing applications (e.g., Google Docs) and is readily available on devices the learners already own and use on a daily basis. Minimal training is required and access to help is available on online support pages.

Google Voice Typing offers all the features of ASR noted above, and consequently has great pedagogical potential. However, before adopting a new technology, it must be determined if it is suitable to be used for writing (Stage 2 in Cardoso, in press), particularly among L2 users, a group that has not received the attention it deserves. One way of assessing the suitability of ASR (as found in Google Voice Typing) as an L2 writing tool is via a well-established model, The Technology Acceptance Model, discussed below.

Technology Acceptance Model (TAM)

Davis (1989) developed the Technology Acceptance Model (TAM), the dominant theoretical model used in the field of information systems to measure user acceptance, which has well-established reliability and validity (Dizon, 2016). TAM posits that users' intention to use

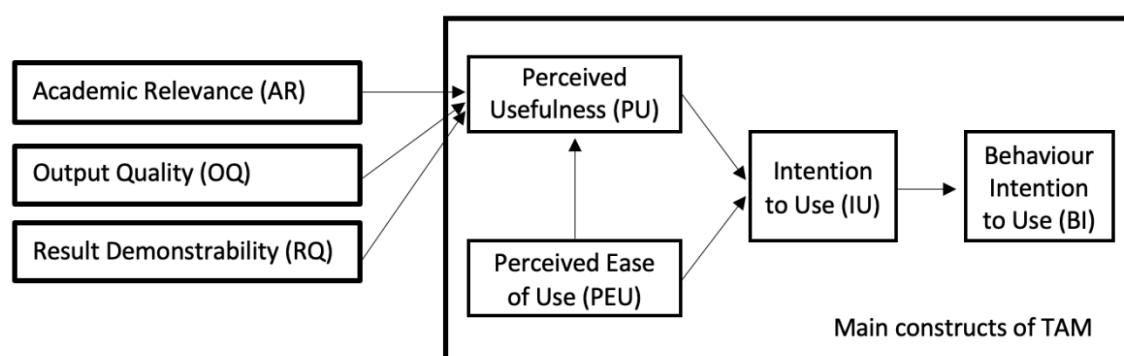
and eventual behavioural intention to use new technologies are based on two determinants: perceived usefulness and perceived ease of use. Perceived usefulness is conceptualized as “the degree to which a person believes that using a particular system would enhance his or her job performance”, while perceived ease of use is the degree which learning a new technology is perceived as “free from effort” (Davis, 1989, p. 320). Davis et al. (1989) used TAM to investigate university students’ acceptance to use an early work processing program, WriteOne. They found that perceived usefulness was the main determinant affecting intention to use when participants were first introduced to the technology, with perceived ease of use being a significant, but secondary in nature. After fourteen weeks of use, the relative importance of perceived ease of use was even less, with perceived usefulness being the only significant determinant of intention to use. Thus, the usefulness of a technology has longer-term implications for its acceptance as “users may be willing to tolerate a difficult interface in order to access functionality that is very important” (Davis et al., 1989, p. 1000).

Venkatesh and Davis (2000) revised the original TAM framework through an analysis of factors that influence perceived usefulness. In a study of the implementation of new technologies in multiple corporate settings, they found two groups of determiners that affect perceived usefulness: social and cognitive. Social determiners involve concepts such as subjective norm, voluntariness and compliance, image, and experience, while cognitive determiners include job relevance, output quality, and result demonstrability. The authors found that social determiners significantly impact perceived usefulness only in contexts in which the usage of the technology is mandatory. Thus, these factors will not be examined in this study as the context is not one of mandatory workplace usage. However, since the three *cognitive* determiners were significant across all contexts, they will be measured as part of perceived usefulness. For the sake of this

study, job relevance will be referred to as “academic relevance” since the technology will be used in an educational context. These three constructs of perceived usefulness can be defined as follows: Academic relevance is the degree to which the technology is applicable to the task, output quality is how well the technology performs the task, and result demonstrability is the tangible results of using the technology to complete a task (see Figure 4 for the proposed model for this study).

Figure 4

Proposed Technology Acceptance Model for Current Study



Note. Adapted from Venkatesh and David (2000, p. 197).

One criticism of the TAM framework is that it is solely based on self-reports (Scherer & Teo, 2019). The current project will adopt a mixed-methods approach to follow their recommendation that TAM studies include other data measurements.

TAM in CALL

The use of TAM in exploring learner perceptions and acceptance of technologies used in language learning is becoming more common and studies show that L2 learners have positive perceptions of the technology they use to improve their language skills. Most studies have taken place in English Foreign Language (EFL) contexts, within university-level courses, and with both a smaller number of participants (e.g., $N = 23$ in Dizon & Thanyawatpokin, 2018) or larger-

scale studies (e.g., $N = 398$ in Tan, 2019). Following TAM protocols for data analysis, positive perceptions have been found when using tools specifically designed for language learning (e.g., MyET for pronunciation practice in Hsu, 2016; UVEL for vocabulary acquisition in Huang et al., 2012), when using well-known social media sites for L2 learning in a classroom context (e.g., Facebook and blogs in Dizon & Thanyawatpokin, 2018; Telegram in Haghighi et al., 2019), and when using technology for autonomous language learning (Soleimani et al., 2014). The existing research seems to suggest that, under the auspices of TAM, learners perceive using technology to learn English in a positive manner and are willing to accept using technology to improve their language skills regardless of the size of the study, the language skill, or the technology,

This positive perception may stem from the fact that learners feel that the technology is useful as it may lead to higher grades or increased language learning. For example, Dizon (2016) and Tan (2019) found a significantly high correlation between perceived usefulness and intention to use in studies that focused on the use of Internet-based texts and e-tutoring websites. In another study involving 341 English students in Taiwan, Hsu (2016) examined not only the relationships between the various TAM constructs, but also the effect of participants' learning styles on their perceptions of an ASR-based CAPT tool. His results seem to indicate that perceived usefulness had a more direct correlation with intention to use than perceived ease of use, regardless of learning style. Therefore, it would seem that participants in these studies found the technology to be useful as they believed it helped them succeed in their classes.

The studies above examined only user perceptions; however, other researchers have measured TAM constructs along with performance gains. These studies seem to indicate that the technology is as useful as students perceive since students using the technology outperform control groups that do not (e.g., Chen Hseih et al., 2017; Haghighi et al., 2019). In a two-year

study of 247 Taiwanese English majors, Tsai (2015) went a step further and calculated the correlation between the TAM constructs (perceived usefulness and perceived ease of use) and writing performance scores. The relationship between both constructs of TAM and writing performance was found to be significant, with perceived usefulness having a higher significance than perceived ease of use. This gives weight to the idea that students perceive technology as being useful when they recognize that it can lead to higher learning outcomes.

Despite being a dominant theoretical framework in the field of information technology, TAM is not yet widely used in CALL research. Of the existing studies, most were conducted in an EFL context (Hsu, 2016; Haghghi et al., 2019) and none have examined user acceptance of ASR dictation technology as a possible language-learning tool. Thus, this current study will use the TAM framework to examine user acceptance of ASR (as found in Google Voice Typing) to determine the feasibility of using ASR to write academic texts in an ESL context.

Current Study

Due to new technologies that have altered the manner in which audiences access and consume written texts, writers must go beyond accurate linguistic form and incorporate other important elements such as content, context, and genre. However, L2 writers in university-level ESL classes still focus on linguistic accuracy when writing (Ding & Zhao, 2019), and instructors tend to reinforce this by providing feedback on form more than on the other aspects of a text (Menke and Anderson, 2019). It is possible that if L2 writers use ASR to compose texts, they can shift their focus from linguistic accuracy to the other constructs of written texts. However, it is important to determine if language learners will accept this technology to write in an academic setting.

Thus, the goal of this study was to use the TAM framework to examine learners' perceptions of using ASR to write texts in order to assess its usability as an L2 writing tool. By determining its usability, pedagogical recommendations can be made and a foundation for follow-up studies on usage of this technology can be laid. The research question (RQ) that guided this project was:

- How do university-level English L2 students perceive ASR (as found in Google Voice Typing) to write academic texts in terms of:
 - a. Usefulness: academic relevance, output quality, result demonstrability;
 - b. Ease of use; and
 - c. Intention to use?

Based on the literature discussed (e.g., Hsu., 2016; Tsai, 2015), and factoring in the quantitative (TAM ratings) and qualitative (semi-structured interviews) nature of the data collected in this study, it was hypothesized that the participants would perceive the three main constructs positively as they would see the usefulness of ASR in terms of their language learning and course grades. For example, features such as ASR's potential to allow writers to focus on meaning rather than linguistic accuracy would be perceived as useful. The participants would also perceive the user-friendly voice commands as easy to use, leading to an overall positive attitude towards ASR, which would contribute to their intention to continue to use the proposed technology to compose their texts.

Method

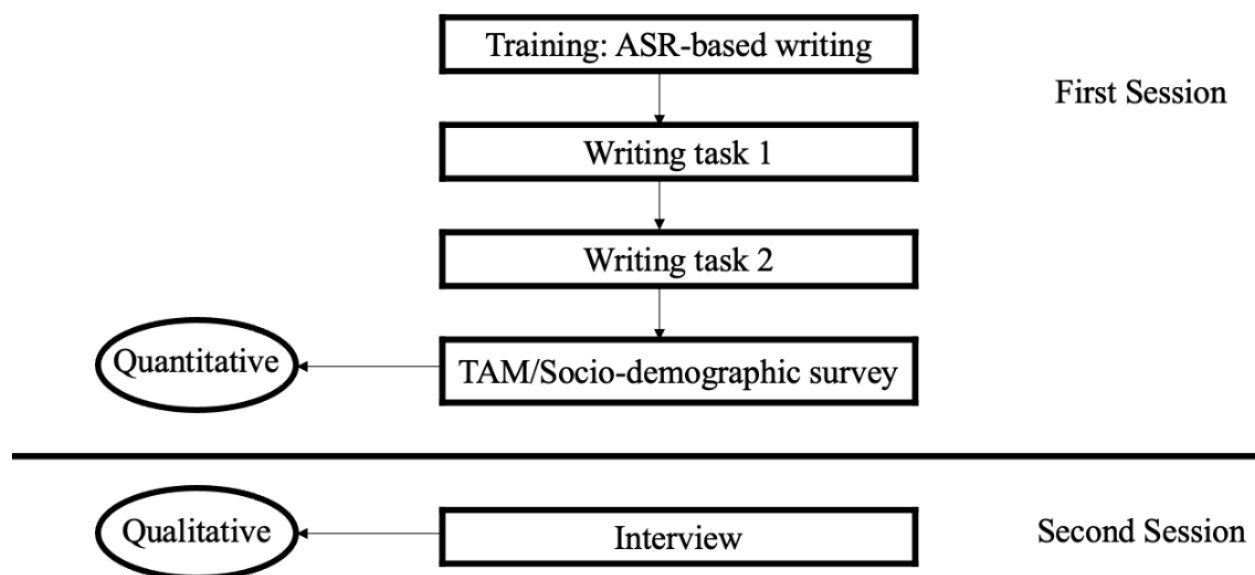
Design

This was a mixed-methods study with quantitative data collected from a survey and qualitative data collected from semi-structured interviews with the participants. Individual

research sessions began with an initial training on using Google Voice Typing, the ASR application within Google Docs. Two short writing tasks were then carried out following the procedure used in the participants' writing class, as will be described below. After the writing tasks, a TAM-informed survey was administered to the participants regarding their perceptions of writing using ASR in terms of its perceived usefulness (including academic relevance, output quality, and result demonstrability), perceived ease of use, and their intention to use the application for composing texts, along with a socio-demographic survey about their language background. At a later date, semi-structured interviews were conducted with the participants to discover their perceptions of writing using ASR and why they perceived the technology the way they did. The design of the study is shown in Figure 5.

Figure 5

Study Design



Research Context and Participants

A convenience sampling method was employed to select the participants. The study was conducted with 17 English L2 students enrolled in a university-level English for Academic Purposes class at a university in Quebec, Canada. All participants were paid \$20 for participating in the study. The students had fulfilled the minimal English Proficiency requirement for acceptance at the university. However, the course was mandatory for all participants based on their TOEFL iBT scores (between 75 and 89 or equivalent). One participant was excluded from the results due to technical difficulties with the device they were using to compose the texts. Thus, the final participants were 16 adults ($N = 16$; 7 male, 9 female; average age: 23.8) from fifteen different L1s.

Participants were given a “pronunciation accuracy test” to verify that they met a minimum level of ASR recognition accuracy. L2 speakers with a high rate of pronunciation errors (and thus a low rate of ASR accuracy) would not be the target population for using ASR to write, as the number of errors would offset any potential affordances of the technology. To determine the level of accuracy, participants’ errors in the ASR output were calculated, with a cut-off point of 90% accuracy to be used. This corresponds to the lowest speech recognition rate of L2 speakers in McCrocklin and Edalatishams’ (2020) study on recognition rates of Google Voice Typing (see discussion above in Background). Since their participants were comparable to the participants in my study (ESL students at an English university), 90% was considered to be an attainable level for most potential participants. However, during the research, it became obvious that a test based upon a *scripted* (not spontaneous) text was not valid as it did not accurately recreate the treatment condition of the proposed ASR-based approach to writing. Thus, it was decided to not use the pronunciation accuracy test to determine eligibility. However,

since some participants had already completed the test, it was completed by all participants and was simply considered to be an additional training activity.

Materials

Due to the COVID-19 pandemic, all data collection and interactions with the participants were virtual, using the Zoom video/audio conferencing tool.

Training Session

Training sessions were developed and carried out by the researcher, after obtaining ethics clearance. The training session was piloted among 2 volunteers representing the same level of English and education as the targeted participants (see Figure 6 for a sample activity).

Figure 6

Sample Training Activity

Becoming Familiar with Google Voice Typing

1. Click in the box below.
2. Make sure Voice Typing is activated, as illustrated in Step 4.
3. Read the sentence above the box.
4. Make sure you say the punctuation as you speak (e.g. question mark, comma, period, colon, semi-colon, exclamation mark, **new line**, **new paragraph**).

Where are you going? You can't **wear** those clothes **to** go out. **They're too** old. Maybe **there** are some nicer clothes in your closet.

Pronunciation Accuracy Test/Training Activity

In Google Voice Typing, participants were asked to carry out a pronunciation accuracy test (later subsumed under “training session” and utilized for training purposes, as discussed above). The test consisted of a reading of Weinberger’s (2021) *Please Call Stella* text, a short

passage in English comprised of common words and most of the sounds in the language (including problematic segments, syllable structures, and sound sequences that can be challenging for L2 speakers; visit <http://accent.gmu.edu> for details). Participants were allowed to review the text to inquire about any unfamiliar words before reading it aloud (see Appendix A for the text).

Writing Tasks

Two writing tasks were designed by the researcher, following the process used in the students' English class to reflect the ecology of the environment in which the participants study: participants were given a reading to provide content and vocabulary; writing prompts were used to elicit structured paragraphs, etc. The topics chosen for the tasks were validated with course instructors to avoid overlap with the course material. This ensured that participants wrote texts that were not based on previous writing assignments, thus maintaining the academic integrity of future writing evaluations. The writing prompts evoked personal responses to the readings rather than summarizing the texts or finding the main ideas to prevent the dictation of information from the articles rather than the composition of an original text (see Appendix B for the writing tasks).

TAM Survey

The perceived usefulness (and the constructs that influence it), perceived ease of use, and intention to use related to using ASR when writing were measured using a survey based upon Venkatesh and Davis' (2000) TAM scales, adapted by the researcher for an L2 educational context and the goals of the study. According to Venkatesh and Davis (2000), each item on the scale has been measured across studies and time periods, and shows high internal validity, with the lowest Cronbach's α for an item being 0.80 for two determiners of perceived usefulness: job relevance (*academic relevance* in this study) and result demonstrability. Participants scored

statements about their experience with ASR on a 7-point Likert scale, with 1 indicating *strongly disagree* and 7 indicating *strongly agree*. Table 1 provides a sample of a statement for each construct (see Appendix C for the survey).

Table 1

Sample Statements from the TAM Survey

Construct	Statement
1. Perceived Usefulness (PU)	Using ASR to write texts enables me to write texts more quickly
Academic Relevance (AR)	Using ASR when writing texts for school is relevant to my studies
Output Quality (OQ)	The quality of the text I write when I use ASR is high
Result Demonstrability (RD)	The results of using ASR to write texts are clear to me
2. Perceived Ease of Use (PEU)	I find it easy to use ASR to write texts
3. Intention to Use (IU)	I intend to continue writing using ASR

Interview

A post-treatment semi-structured interview was used to further explore the participants' perceptions of using ASR to compose their texts. Open ended questions such as *Did ASR help you write the text more quickly? Why or why not?* were used to elicit the participants' perceptions (see Appendix D for the complete list of interview questions). Interviews were conducted in English and audio-recorded for transcription and analysis.

Data-analysis Measures

The participant responses to the survey were entered into the statistical program SPSS and descriptive statistics were run for calculating means, standard deviations, medians, and modes. Positive perceptions of using ASR for writing was indicated by a mean above 4 (out of 7), as that is when participants shift to a positive perception on the Likert scale. Reliability of the

survey items was assessed by for each construct. According to Field (2017), acceptable reliability is reached when Cronbach's alpha (α) = .70. The Cronbach's alpha for the items for each construct was above this level (with scores ranging from .82 to .96), indicating acceptable internal reliability of the survey, except for Result Demonstrability (α = .52) and Intention to Use (α = .60). However, Cronbach's α is sensitive to the number of items and is susceptible to decrease when there are fewer items (Field, 2017). Thus, the results obtained for the reliability of those two constructs may be due more to the small number of items rather than an actual problem with reliability. The Cronbach's α for each construct can be found in Table 2.

The responses to the interview questions were transcribed using Otter Voice Meeting Notes and subsequently coded. Deductive thematic coding (Creswell, 2014) were used, with the codes based upon the three main constructs being measured: perceived usefulness (and its cognitive influences: academic relevance, output quality, and result demonstrability), perceived ease of use, and intension to use. Each statement that had been coded for a construct was then further analyzed to determine if it indicated positive, negative, neutral, or mixed perceptions (Saldaña, 2009).

Results

Quantitative Results

The results indicate that the participants perceived the use of ASR as a writing tool as positive in each of the categories investigated. Table 2 summarizes the results.

In terms of perceived usefulness, participants had an overall positive perception ($M = 4.71$, $SD = 1.69$). Each of the items on the Likert-scale ratings were above the level of neutrality (established at 4 on the scale of 1 to 7). *Using ASR to write texts will improve my performance in my English class* was the item that scored highest ($M = 5.50$, $SD = 1.41$), while *Using ASR*

makes it easier to write good texts scored lowest ($M = 4.06$, $SD = 1.56$). Positive perceptions were also indicated with means above the level of neutrality for the three constructs that influence perceived usefulness: academic relevance ($M = 4.16$, $SD = 1.80$); output quality ($M = 4.13$, $SD = 1.76$); and result demonstrability ($M = 5.84$, $SD = 1.30$). The individual items for each of these constructs were all above the level of neutrality.

Participants also indicated overall positive perceptions of perceived ease of use of ASR as a writing tool ($M = 5.25$, $SD = 1.58$), with each item also scoring above the level of neutrality. *Interacting with ASR to write texts is clear and understandable* ($M = 5.69$, $SD = 1.57$) scored the highest and *I find it easy to make ASR do what I want when I write a text* ($M = 4.81$, $SD = 1.51$) scored the lowest within this construct.

Intention to use ASR-based writing was also positively perceived ($M = 5.78$, $SD = 1.10$). The individual items, *I intend to continue writing texts using ASR* ($M = 5.50$, $SD = 1.17$) and *I intend to show others how to write texts using ASR* ($M = 6.06$, $SD = 0.90$) both scored well above the level of neutrality.

Table 2*Descriptive Statistics of Survey Results*

Survey items	Mean/7	SD	Median	Mode	Min	Max
Perceived Usefulness ($\alpha = .82$)	4.71	1.69				
Using ASR to write texts enables me to write texts more quickly.	5.25	1.68	6	6	1	7
Using ASR to write texts improves the quality of my texts.	4.38	1.69	5	5	1	7
Using ASR to write texts enhances my writing.	4.38	1.58	4.4	4	1	7
Using ASR makes it easier to write good texts.	4.06	1.56	4	4	1	7
Using ASR to write texts will improve my performance in my English class.	5.50	1.41	6	7	4	7
Cognitive Influences ($\alpha = .83$)	4.71	1.77				
Academic Relevance ($\alpha = .92$)	4.16	1.80				
Using ASR when writing texts for school is important to my studies.	4.13	1.76	4.13	5	1	7
Using ASR when writing is important to my academic success.	4.19	1.78	4	4	1	7
Output Quality ($\alpha = .96$)	4.10	1.79				
The quality of the text I write when I use ASR is high.	4.06	1.85	4	6	1	7
I like the quality of the texts I write when using ASR.	4.19	1.66	4.19	5	1	7
Result Demonstrability ($\alpha = .52$)	5.84	1.30				
I will tell others about the results of using ASR to write texts.	6.00	0.94	6	7	4	7
The results of using ASR to write texts are clear to me.	5.69	1.53	6	7	1	7
Perceived Ease of Use ($\alpha = .88$)	5.25	1.58				
Interacting with ASR to write texts is clear and understandable.	5.69	1.57	6	7	1	7
I find it easy to use ASR to write texts.	5.31	1.40	6	6	1	7
Interacting with ASR to write texts does not require a lot of my mental effort.	5.19	1.67	6	6	1	7
I find it is easy to make ASR do what I want when I write a text.	4.81	1.51	5	5	1	7
Intention to Use ($\alpha = .60$)	5.78	1.10				
I intend to continue writing texts using ASR.	5.50	1.17	5.5	5	3	7
I intend to show others how to write texts using ASR.	6.06	0.90	6	7	4	7

Qualitative Results

The qualitative aspect of the RQs, obtained via individual interviews with the participants, offers explanations for the positive perceptions of ASR as a writing tool.

Perceived usefulness, academic relevance, output quality, and result demonstrability

Most participants ($n = 12$) felt their texts had more errors when written with ASR rather than typed, yet at the same time, many ($n = 7$) felt that their overall spelling was better when using their voices to type. Jo explained, “Sometimes when I don’t know how to write the word, I don’t use it. But if I am writing by voice, maybe it’ll give me the correct spelling.” A majority of participants also felt their texts took longer to write ($n = 10$) since corrections had to be made to misunderstood words. As noted by Blair, “I can, like put everything together quickly. But again, I would have to go back and edit and do everything”. When asked how that differed from an assignment typed in a traditional manner, Blair explained that due to misinterpretations caused by accented speech, “some things are completely wrong, completely different” and would require more extensive editing than when typing with a keyboard. Despite these issues, more than two-thirds of the participants ($n = 11$) felt it was a useful tool for writing. Of these participants, most ($n = 6$) felt it would be even more useful when their English pronunciation had improved. Additionally, some ($n = 3$) participants saw an added usefulness in writing with ASR, as it could help improve pronunciation and self-efficacy when speaking. Blair explained, “I will be using it more to try to learn the pronunciation of words. To try to have confidence when I’m speaking.”

Academic Relevance. Many participants saw a relevance to using ASR when writing for their ESL class ($n = 10$). However, five participants explained that they majored in programs that do not require written assignments (e.g., Math), and did not see it as relevant to their other

course. Two students who regularly submit written assignments in their courses felt that ASR's value was limited since the quality of language is not taken into consideration by their professors. As explained by Jordan: "I'm in civil engineering. And I don't think the professors are going really to concentrate on grammar." However, most participants ($n = 14$) indicated they would show friends how to use ASR to write texts for university, with three of those students having already done so in the time between the research session and the interview. As Sasha explained, it would help their friends "eliminate mistakes and organize ideas."

Output Quality. The participants were divided in terms of their beliefs about the quality of texts written using ASR. Only one participant felt the quality of the text written with ASR was better than texts they would type, while five felt the quality of texts written with ASR was worse. The remaining participants ($n = 11$) felt the quality was the same as a first draft typed using the keyboard. Skyler saw a benefit in using it during the brainstorming stage, as she could "write the keywords, like speak it. I can save the time and then have more time to prepare for the formal documents," resulting in a better-quality final text. A third of the participants spontaneously brought up the idea that the quality of the text would be improved when writing in an L1 rather than in an L2 ($n = 6$), even though none of the questions in the interview were related to using the technology in an L1. When asked if using ASR could result in better quality texts, Terry, an L1 French speaker, responded "In French, okay. In English, I don't think so."

Result Demonstrability. There was no consensus among the participants regarding the ability of writing with ASR to demonstrate tangible results for them. Just over a third of the participants ($n = 6$) felt a text written with ASR would get not get a better grade than a text typed using a keyboard. Taylor explained that the grade would not be better because "as for the content, it's my thoughts, right? So, if I type them or record them with my voice, it doesn't really

like make a difference.” However, five participants felt the grade would improve due to improved spelling or content. As noted by Charlie, a photography student, “when I describe my photo, I prefer to talk because I can express better my feelings because I don't think about the grammar.”

Perceived Ease of Use

The ease of interacting with the technology was highlighted unanimously ($n = 17$) by all participants. Taylor declared, “It's very simple. Yeah, I just click the record button and start speaking. That's it.” However, a common complaint about the technology was the fact that the speaker must say punctuation marks as they are dictating the text ($n = 8$). Charlie discussed this difficulty: “I don't know very well the words like the question mark or the comma.”

Pronunciation was also brought up as an issue ($n = 6$) that made the tool harder to use.

Intention to Use

Most participants ($n = 14$) intend to use ASR to write in the future, and some participants had already started to use ASR to write assignments before being interviewed ($n = 3$). Some students ($n = 5$) felt they would use it at a future time when their English will be better, such as Leslie, who said they would use ASR when they become “more proficient in English”. Many participants had recommended it to others or had plans to ($n = 10$), such as Alex, who said, “I already recommended to a few friends. I just talked with them with my experience in your research.”

Discussion

This study set out to answer the following research question: How do university-level English L2 students perceive ASR to write academic texts in terms of usefulness (including academic relevance, output quality, and result demonstrability), ease of use, and intention to use?

The means of the survey items for each category were above the level of neutrality, indicating positive perceptions of ASR as a writing tool. In fact, all measures of central tendency (mean, median, and mode) were above the level of neutrality for each survey item. This confirms the hypothesis that the participants would appreciate the affordances offered by ASR when writing academic texts. The positive perceptions of ASR as a writing tool can be better understood when examining the constructs of perceived usefulness, perceived ease of use, and intention to use through the lens of Chapelle's (2001) criteria for evaluating CALL tools, discussed earlier: language learning potential, learner fit, attention to meaning, authentic use of language, positive impact, and practicality.

Perceived Usefulness, Academic Relevance, Output Quality, and Result Demonstrability

Perceived usefulness and its three moderating cognitive influences of academic relevance, output quality, and result demonstrability were positively perceived by participants, but they were the constructs with the lowest positive perceptions overall. The fact that these constructs behave similarly validates Venkatesh and Davis' (2000) assumption that the three cognitive influences moderate the construct of perceived usefulness.

The results in the current study corroborate Soleimani et al. (2014), who found that all constructs were perceived positively, but that perceived usefulness had lower positive perceptions than perceived ease of use. This pattern, however, contradicts most findings in the TAM literature (e.g., Dizon, 2016; Hsu, 2016; Tsai, 2015), possibly because of the different context of the use of the technology in both this study and Soleimani et al. (2014). In these studies, researchers did not examine users' perception of a technology being mandated for use in a course, but rather a tool that students could use independently, in the future, to improve their English skills. As such, our participants did not see tangible results in the form of grades (which

would impact academic relevance and result demonstrability), which may have slightly moderated the participants' perceptions of the usefulness of the tool for writing academic texts.

It was hypothesized that perceived usefulness would have positive measures, as it would reduce the burden of producing correct form, and consequently address one of Chapelle's (2001) criteria for CALL adoption: opportunity to focus on meaning. This held partially true.

Participants appreciated being able to correctly produce words that they did not know how to spell, which allowed them to improve the lexical diversity and complexity of their texts.

However, the assumption did not hold true in terms of inflectional morphology (grammar), mostly because the participants needed to know it correctly when they spoke for their attempts to appear correctly in their writing. For example, if a participant did not correctly conjugate a verb when speaking (e.g., omitted the *-ing* inflection from a verb in the continuous tense), ASR would not correct the error for them. Therefore, ASR allowed participants to focus on meaning more than form in terms of spelling, but not for grammar. This partial relief from focusing on form was enough, though, to allow the students to positively perceive the technology in terms of its usefulness.

Another factor that may have led to positive measures for perceived usefulness can be found in ASR's language learning potential, another of Chapelle's (2001) criteria. It was noted by some participants that ASR would be an effective tool for improving pronunciation and self-efficacy when speaking, allowing them to simultaneously improve two skills at the same time: speaking and writing. These affordances helped the participants perceive the usefulness of the technology, as they allowed them to improve their English language skills.

However, ASR did not meet some of Chapelle's (2001) criteria for some participants, which may explain why perceived usefulness was perceived relatively less positively than the

other constructs. First, some participants felt that the tool was not fully appropriate for their current level of language because they had not attained a suitable level of intelligibility to be consistently understood by the ASR. Thus, there was a lack of learner fit. However, many felt it would be useful at a future time when their pronunciation had improved – when it would be a more optimal fit with their level of English. Additionally, there was a lack of learner fit and positive impact for participants that did not write academic texts outside of their ESL classes. For these students, improving academic writing was not relevant to their academic success, which may have lowered their perceptions of the tool's academic relevance and result demonstrability. However, participants that did write academic texts outside of English class perceived it as useful. For these participants, there was a sense of learner fit, as ASR was seen as useful at many steps in the writing process, such as note-taking, brainstorming, writing the draft, and dictating citations. They also felt ASR could improve their grades, highlighting its positive impact. This highlights the importance of using CALL tools that truly fit the needs of learners (Bañados-Santana, 2018).

Finally, interview responses hint at another reason for relatively lower positive perceptions of usefulness beyond the scope of Chapelle's criteria: lack of practice. Although the technology is easy to use, practice is needed to be able to optimize the usefulness of the tool. Indeed, two of the participants had taken time to experiment with using ASR to write in the time between the research session and the interview and had found that it became more accurate with use and hence more useful as a writing tool. This may explain why the construct of output quality, which moderates perceived usefulness, while positive, was relatively less so than the other main TAM constructs. Sustained use of ASR as a writing tool over a longer period may be needed for output quality to improve, thus making ASR's usefulness more evident.

Perceived Ease of Use

In terms of perceived ease of use, positive perceptions largely resulted from the user-friendly design of this technology, as reported in the qualitative results. However, the qualitative data indicate that the perception of ease of use goes beyond ASR's user-friendliness. One of Chapelle's (2001) criteria, that of practicality, is fulfilled by using ASR in Google Docs since every participant was already a user of Google Docs. As such, they did not have to download and learn any new application or navigate a new word processing system to write using ASR. Additionally, ASR was familiar for many participants; as reported, most participants had already used the technology for sending short texts on their mobile phones (texting in their L1s), or by interacting with virtual assistants, such as Siri, in English. This authentic use of technology, another Chapelle's (2001) criterion, added to the sense of ease of use.

Interestingly, the ease of use of ASR allowed some participants to focus on what they wanted to express (content) rather than worry about how to express it (form), fulfilling Chapelle's (2001) criterion of focus on meaning. Indeed, the adopted technology allowed them to effortlessly put ideas on their screen. As Sasha noted, "It gives you the space for your ideas to just ... come up. And it gives the space for your mind to throw your ideas I think it's really useful for those who want their ideas to expand, to give a free place to their ideas." For those who are not proficient typists, like Billie, it simplified the physical process of writing as "sometimes I have to find out the keys and write down. But whenever I speak, it automatically write [sic] down everything." Since L2 typists tend to make more typographical errors than L1 typists due to a lack of automaticity of orthography (Stevenson et al., 2006), it is not surprising that the participants felt unburdened when not having to focus on spelling and typing skills, which may be partially behind their positive perceptions of ease of use.

Intention to Use

The construct of intention to use had the highest positive perceptions of the main TAM constructs. This corroborates the findings in the CALL literature using the TAM framework (e.g., Dizon, 2016; Tan, 2019; Tsai, 2015). These results also reflect the premise of TAM, as the participants found ASR to be both useful and easy to use, and thus, showed an intention to continue using the technology (Venkatesh & Davis, 2000). These findings highlight that the acceptance of any technology is determined by two beliefs: perceived ease of use and perceived usefulness which, together, as discussed above, may contribute to increasing learners' usage intention. The findings also highlight the relevance of Chapelle's (2001) criteria for evaluating CALL tools. For instance, in the discussion above of perceived usefulness and perceived ease of use, it was demonstrated that ASR met four of the six criteria and partially met the other two. It could be argued that technology that meets Chapelle's (2001) criteria are, by definition, useful and easy to use, making them tools that language learners will want to engage with and use in their future language learning endeavours.

Conclusion

The purpose of this study was to examine ESL learners' perceptions of using automatic speech recognition (ASR) to write academic texts in English to assess its usability as a writing tool. Both quantitative and qualitative data indicate that users had positive perceptions of the constructs measured using the Technology Acceptance Model framework (Venkatesh & Davies, 2000), suggesting that this technology has the potential to be used by and to benefit L2 English writers.

Certain limitations of this study must be taken into consideration. This research was completed during the COVID pandemic, which has multiple implications. During the research

period, the university from which the participants were recruited was completely shut down and all learning was online. This resulted in a small sample size as recruiting was complicated by pandemic restrictions and the realities and stresses that went along with it (e.g., potential participants declined to participate due to high-stress levels, the snowball effect of recruiting could not be used as students did not know and befriend people in their classes). As well, the fact that the students were studying from home, often through asynchronous classes, meant that they had very limited interactions in English. Indeed, two of the participants had not even left their home countries and were not in Canada during the entire school year. As a result, they had very few opportunities to speak English, which may have affected their fluency and pronunciation, thus diminishing the accuracy of the ASR output and their positive perceptions of the usefulness of the tool. On the other hand, many participants were so happy to have a chance to speak English with the researcher, not only for linguistic but for social reasons, which may have resulted in more positive perceptions of ASR than if the research had been done in a non-pandemic situation. Finally, this was a one-shot intervention of a technology that is simple to use, but that might fundamentally change how a person writes. As noted by the participants who took time to use it between the research session and the interview, effective use of ASR when writing may take time and practice. This lack of practice may have limited the participants' perceptions of the usefulness of the technology.

Further research is required into the pedagogical aspects of using ASR. Studies are needed to determine the true pedagogical potential of the tool, as well as the strategies that support its use and the best practices for implementing it in a classroom. As well, research into using ASR to write in L2s other than English should be carried out (e.g., French). French is an interesting target language because of its opaque orthography (Brissaud & Chevrot, 2011). For

example, there are multiple spellings for the phoneme /ɛ/ (e.g., *-ait*, *-ê*, *-aient*). ASR may offer affordances for students to learn about grapheme-to-phoneme associations in such a language. Additionally, a longitudinal study is needed to determine the true pedagogical benefits of ASR, since it inherently changes how people write, which requires practice. A longitudinal study could also ensure that the novelty effect of a new – and frankly fun to use – technology does not influence the participants’ perceptions and use of the tool (Clark, 1983). Finally, the use of ASR by native English speakers to write should also be examined, as it was spontaneously identified by the participants as being potentially useful in the writer’s L1.

Writing in the 21st century often involves writing new genres for new audiences that access multi-modal content on new technologies. Writers, both L1 and L2, deserve new technology that will allow them to go beyond producing “correct form” but to focus on other important elements of writing such as text content and genre to meet the needs of their readers. ASR has the potential to be such a technology. This study indicates that using automatic speech recognition yields positive perceptions among ESL learners. Future research should determine the best practices for using this technology so that writers can leverage the affordances it has to offer.

Chapter Three

This chapter will present a review of the result from the previous chapter, as well as their implications in language education. Future directions for research into the use of automatic speech recognition (ASR) as a pedagogical tool will then be discussed.

Summary of Goals and Findings

This mixed-methods study set out to examine L2 writers' perceptions of using ASR as a writing tool to determine if it is a suitable technology for writing. Venkatesh and Davis' (2000) Technology Acceptance Model was used to measure participant's perceived usefulness (and its cognitive determiners: academic relevance, output quality, and result demonstrability) and perceived ease of use, both of which determine a person's intention to use a technology. A 7-point Likert scale survey measuring these constructs was given to 17 ESL students at an English-medium university after an intervention in which participants wrote two academic texts using ASR. In addition, semi-structured interviews were carried out to gain insight into the survey responses and their overall experience with the technology. It was hypothesized that the participants would have positive perceptions of writing with ASR: perceived usefulness would be seen in the tool's ability to facilitate the participants' language learning and consequently improve their grades, and perceived ease of use would be seen in the simplicity of the tools due to its "free from effort" nature (Davis, 1989). This would lead to an *intention to use* the tool.

The quantitative survey results showed positive perceptions of all the constructs measured, indicating that the users saw the technology as useful, easy to use, and consequently they intend to use it again in their future language learning endeavours. Interviews revealed that participants appreciated the lessening of the burden to produce correct form in terms of spelling, the possibility to improve their pronunciation and self-efficacy when speaking, and the extreme

easiness and practicality of the tool. However, some participants felt that their current level of English was not high enough to use ASR, since their utterances were not always intelligible to the technology. Others, however, found that practice helped them overcome this issue. The positive perceptions of ASR seen in this study show that it is a suitable tool to use for writing, as the results reflect the premise of TAM: perceived usefulness and perceived ease of use are valid predictors of the adoption of a new technology by its potential users (Venkatesh & Davis, 2000).

Implications for Education

This study, the first to examine users' perceptions of ASR as a writing tool, determined it is a suitable technology to adapt for composing texts. This has implications for ESL writing courses, as allowing L2 writers to use ASR in an ESL classroom could fundamentally change the dynamic of a writing course. As noted in the introduction, good writing requires attention not only to linguistic form, but also to content, context, genre, and the writing process (Hayland, 2011). However, in ESL writing classes, students often overly focus on linguistic form, to the detriment of the other aspects, as they believe this is what will have the greatest impact on their grades (Barkaoui, 2016; Ding & Zhao, 2019). And they may be correct in that assumption. Despite claiming to be more focused on content than form, L2 teachers provide more feedback on form than any other element when correcting their students' work (Menke & Anderson, 2019). Using a technology like ASR, which can remove some of the burden of producing error-free form, could allow students to pay attention to the other important aspects of writing such as content and the storytelling itself. Teachers would be correcting assignments with fewer errors in form, which would allow them to focus feedback on the other crucial aspects of writing, allowing their students to develop into better writers.

Implications for Research

Now that the suitability of using ASR for writing has been established, the next stage is to assess its pedagogical effectiveness (Cardoso, in press). As noted in Chapter 2's conclusion, studies are needed to explore the effects of using ASR to determine its full potential in terms of improvement of writing. A mixed-method longitudinal study that examines changes in the complexity, accuracy, and fluency of writers over an entire course or program should be carried out to better understand the affordances of the tool. In addition to the complexity, accuracy, and fluency measures, qualitative data in the form of semi-structured interviews will allow for the exploration of the affordances offered by ASR to writers with varying levels of skills. Such a study would be inspired by Chapelle's (2003) appeal that the field of CALL research needs more frequent exploration of the relationship between students' behaviours when using technology (e.g., the strategies they use when writing with ASR) and how these behaviours impact their learning (e.g., their writing outcomes). The information learned in such a study would allow for evidence-based recommendations for the implementation of ASR to write in an ESL learning context.

Participants in the current study who felt that their pronunciation of English impeded their use of ASR as a writing tool often spontaneously stated that the tool would be best used by L1 English speakers or by the participants themselves in their L1. Research into usage of ASR as a writing tool in one's native language is sorely lacking: so far, it has been limited to the writing of children with writing disabilities (e.g., Quinlan, 2004). There are no studies into the use of ASR as a writing tool by adults, an important research population because their literacy skills differ from and are further developed than those of the participants in Quinlan's (2004) study. Thus, further research should be undertaken to examine the benefits of using ASR for adult

writers who have already developed their writing abilities in terms of linguistic form and text structure, but who need to focus on the other aspects of writing.

Finally, studies of the use of ASR when writing in L2s other than English are also needed. As mentioned in the conclusion of Chapter 2, the opaque orthography of French makes it an ideal candidate for such a study, as phonemes often are represented by multiple graphemes (e.g., the phoneme /ɛ/ may be represented orthographically by multiple spellings such as *ait*, *ais*, *ê*). As an L2 French speaker and writer, I would certainly appreciate a tool that would lessen the burden of constantly referring to my imperfect morphosyntactic knowledge of the language to write even the simplest of texts.

Conclusion

Technology in the 21st century has changed how and what audiences read. This change was brought about by advent of personal computers and word processing in the late 20th century, which allowed writers to easily create multimodal texts and Web 2.0 tools and be instantly published to the world (Baron, 2009). It is now time for a new paradigm shift, a new technology worthy of 21st century writers. This study found that ASR is perceived by L2 writers as a suitable tool for writing. The stage is now set to determine if ASR is indeed the new “better pencil”.

References

- Alharbi, M. (2020). Exploring the potential of Google Doc in facilitating innovative teaching and learning practices in an EFL writing course. *Innovation in Language Learning and Teaching, 14*(3), 227–242.
- Allen, H. (2018). Redefining writing in the foreign language curriculum: Toward a design approach. *Foreign Language Annals, 51*(3), 513–532.
- American Academy of Arts and Sciences. (2017, February). *America's languages: Investing in language education for the 21st century*. <https://www.amacad.org/publication/americas-languages>
- Ashwell, T., & Elam, J. (2017). How accurately can the Google Web Speech API recognize and transcribe Japanese L2 English learners' oral production? *The JALT CALL Journal, 13*(1), 59–76.
- Aslan, E., & Ciftci, H. (2018). Synthesizing research on learner perceptions of CMC use in EFL/ESL writing. *CALICO Journal, 36*, 100–118.
- Bañados-Santana, E. (2018). Combining theory, practice and technology in a CALL b-learning environment for EFL learners. *European Journal of Applied Linguistics and TEFL, 7*(2), 53–84.
- Barkaoui, K. (2016). What and when second-language learners revise when responding to timed writing tasks on the computer: The roles of task type, second language proficiency, and keyboarding skills. *The Modern Language Journal, 100*(1), 320–340.
- Bodnar, S., Cucchiarini, C., Penning de Vries, B., Strik, H., & van Hout, R. (2017). Learner affect in computerised L2 oral grammar practice with corrective feedback. *Computer Assisted Language Learning, 30*(3-4), 223–246.

- Bodnar, S., Cucchiarini, C., & Strik, H. (2011). Computer-assisted grammar practice for oral communication. In A. Verbraeck, M. Helfert, J. Cordeiro, & B. Shishkov (Eds.), *Proceedings of the 3rd International Conference on Computer Supported Education (CSEDU)*. (pp. 355-361). Noordwijkerhout.
- Brissaud, C., & Chevrot, J.-P. (2011). The late acquisition of a major difficulty of French inflectional orthography: The homophonic /E/ verbal endings. *Writing Systems Research*, 3(2), 129–144.
- Cardoso, W. (in press). Technology for speaking development. In T. Derwing, M. Munro, & R. Thomson (Eds), *Routledge handbook on second language acquisition and speaking*. Routledge, Taylor & Francis Group.
- Celce-Murcia, M., Brinton, D., & Goodwin, J. (2006). *Teaching pronunciation: A reference for teachers of English to speakers of other languages*. Cambridge University Press.
- Chapelle, C. (2001). *Computer applications in second language acquisition: Foundations for teaching, testing, and research*. Cambridge University Press.
- Chapelle, C. (2003). *English language learning and technology*. John Benjamins.
- Chen, H. (2011). Developing and evaluating an oral skills training website supported by automatic speech recognition technology. *ReCALL*, 23(1), 59–78.
- Chen Hsieh, J., Wu, W., & Marek, M. (2017). Using the flipped classroom to enhance EFL learning. *Computer Assisted Language Learning*, 30(1-2), 1–21.
- Chiu, T., Liou, H., & Yeh, Y. (2007). A study of web-based oral activities enhanced by automatic speech recognition for EFL college learning. *Computer Assisted Language Learning*, 20(3), 209–233.

- Clark, R. (1983). Reconsidering research on learning from media. *Review of Educational Research*, 53(4), 445–459.
- Creswell, J. (2014). *Research design: Qualitative, quantitative, and mixed method approaches*. Sage Publications.
- Cucchiarini, C., Nejjari, W., & Strik, H. (2012). My Pronunciation Coach: Improving English pronunciation with an automatic coach that listens. *Language Learning in Higher Education*, 1(2), 365–376.
- Davis, F. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319–340.
- Davis, F., Bagozzi, R., & Warshaw, P. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science*, 35(8), 982–1003.
- de Vries, B., Cucchiarini, C., Bodnar, S., Strik, H., & van Hout, R. (2015). Spoken grammar practice and feedback in an ASR-based CALL system. *Computer Assisted Language Learning*, 28(6), 550–576.
- Derwing, T., Munro, M., & Carbonaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly*, 34(3), 592–603.
- Ding, Y., & Zhao, T. (2019). Chinese university EFL teachers' and students' beliefs about EFL writing: Differences, influences, and pedagogical implications. *Chinese Journal of Applied Linguistics*, 42(2), 163–181.
- Dizon, G. (2016). A comparative study of Facebook vs. paper-and-pencil writing to improve L2 writing skills. *Computer Assisted Language Learning*, 29(8), 1249–1258.

- Dizon, G., & Thanyawatpokin, B. (2018). Web 2.0 tools in the EFL classroom: Comparing the effects of Facebook and blogs on L2 writing and interaction. *The EuroCALL Review*, 26(1), 29-42.
- Elimat, A., & AbuSeileek, A. (2014). Automatic speech recognition technology as an effective means for teaching pronunciation. *The JALT CALL Journal*, 10(1), 21–47.
- Eskenazi, M. (1999). Using a computer in foreign language pronunciation training: What advantages? *CALICO Journal*, 16(3), 447–469.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics*. Sage Publications.
- Filippidou, F., & Moussiades, L. (2020). A benchmarking of IBM, Google and Wit automatic speech recognition systems. *Artificial Intelligence Applications and Innovations*, 583, 73–82.
- Haghighi, H., Jafarigohar, M., Khoshsima, H., & Vahdany, F. (2019). Impact of flipped classroom on EFL learners' appropriate use of refusal: Achievement, participation, perception. *Computer Assisted Language Learning*, 32(3), 261–293.
- Hsu, L. (2016). An empirical examination of EFL learners' perceptual learning styles and acceptance of ASR-based computer-assisted pronunciation training. *Computer Assisted Language Learning*, 29(5), 881–900.
- Huang, Y., Huang, Y., Huang, S., & Lin, Y. (2012). A ubiquitous English vocabulary learning system: Evidence of active/passive attitudes vs. usefulness/ease-of-use. *Computers & Education*, 58(1), 273–282.
- Hyland, K. (2011). Learning to write: Issues in theory, research, and pedagogy. In R. Manchón (Ed.), *Learning to write and writing to learn in an additional language* (pp. 17-35). John Benjamins.

- Jenkins, J. (2000). *The Phonology of English as an International Language*. Oxford University Press.
- Lee, L. (2017). Learners' perceptions of the effectiveness of blogging for L2 writing in fully online language courses. *International Journal of Computer-Assisted Language Learning and Teaching*, 7(1), 19–33.
- Levis, J., & Suvorov, R. (2012). Automatic speech recognition. In C. Chapelle (Ed.), *The encyclopedia of applied linguistics*. John Wiley & Sons.
- Liakin, D., Cardoso, W., & Liakina, N. (2014). Learning L2 pronunciation with a mobile speech recognizer: French /y/. *CALICO Journal*, 32(1), 1–25.
- Magnan, S., Murphy, D., Sahakyan, N., & Lafford, B. (2014). Goals of collegiate learners and the standards for foreign language learning. *The Modern Language Journal*, 98, i–xxiii.
- McCrocklin, S. (2018). Learners' feedback regarding ASR-based dictation practice for pronunciation learning. *CALICO Journal*, 36(2), 119–137.
- McCrocklin, S., & Edalatshams, I. (2020). Revisiting popular speech recognition software for ESL speech. *TESOL Quarterly*, 54(4), 1–13.
- Menke, M., & Anderson, A. (2019). Student and faculty perceptions of writing in a foreign language studies major. *Foreign Language Annals*, 52(2), 388–412.
- Mirzaei, M., Meshgi, K., Akita, Y., & Kawahara, T. (2017). Partial and synchronized captioning: A new tool to assist learners in developing second language listening skill. *ReCALL: The Journal of EUROCALL*, 29(2), 178–199.
- Mroz, A. (2018). Seeing how people hear you: French learners experiencing intelligibility through automatic speech recognition. *Foreign Language Annals*, 51(3), 617–637.

- Moussalli, S., & Cardoso, W. (2020). Intelligent personal assistants: Can they understand and be understood by accented L2 learners? *Computer Assisted Language Learning*, 33(8), 865-890.
- National Council of Teachers of English. (2018, November 14). *Understanding and teaching writing: Guiding principles*. <https://ncte.org/statement/teachingcomposition/>
- Poulsen, R., Hastings, P., & Allbritton, D. (2007). Tutoring bilingual students with an automated reading tutor that listens. *Journal of Educational Computing Research*, 36(2), 191–221.
- Quinlan, T. (2004). Speech recognition technology and students with writing difficulties: Improving fluency. *Journal of Educational Psychology*, 96(2), 337–346.
- Saldaña, J. (2009). *The coding manual for qualitative researchers*. Sage Publications.
- Scherer, R., & Teo, T. (2019). Editorial to the special section—Technology acceptance models: What we know and what we (still) do not know. *British Journal of Educational Technology*, 50(5), 2387–2393.
- Schmidt, R. (1995). Consciousness and foreign language learning: A tutorial on the role of attention and awareness on learning. In R. Schmidt (Ed.). *Attention and awareness in foreign language learning* (pp. 1-63). University of Hawaii, Second Language Teaching & Curriculum Center.
- Soleimani, E., Ismail, K., & Mustaffa, R. (2014). The acceptance of mobile assisted language learning (MALL) among post graduate ESL students in UKM. *Procedia - Social and Behavioral Sciences*, 118, 457–462.
- Stevenson, M., Schoonen, R., & de Glopper, K. (2006). Revising in two languages: A multi-dimensional comparison of online writing revisions in L1 and FL. *Journal of Second Language Writing*, 15, 201–233.

- Tan, P. (2019). An empirical study of how the learning attitudes of college students toward English e-tutoring websites affect site sustainability. *Sustainability, 11*(6), 1–19.
- Tsai, Y. (2015). Applying the Technology Acceptance Model (TAM) to explore the effects of a course management system (CMS)-assisted EFL writing instruction. *CALICO Journal: San Marcos, 32*(1), 153–171.
- Venkatesh, V., & Davis, F. (2000). A theoretical extension of the Technology Acceptance Model: Four longitudinal field studies. *Management Science, 46*(2), 186–204.
- Voicebot Research. (2019, July). *Voice assistant SEO report for brands*. Voicebot.Ai.
https://voicebot.ai/wp-content/uploads/2019/07/voice_assistant_seo_report_for_brands_2019_voicebot.pdf
- Walker, N., Cedergren, H., Trofimovich, P., & Gatbonton, E. (2011). Automatic speech recognition for CALL: A task-specific application for training nurses. *Canadian Modern Language Review, 67*(4), 459–479.
- Weinberger, S. (2021). *Speech accent archive*. George Mason University. Retrieved from <http://accent.gmu.edu>
- Wang, Y., & Young, S. (2014). Effectiveness of feedback for enhancing English pronunciation in an ASR-based CALL system. *Journal of Computer Assisted Learning, 31*, 493–504.
- Young, S., & Wang, Y. (2014). The game embedded CALL system to facilitate English vocabulary acquisition and pronunciation. *International Forum of Educational Technology & Society, 17*(3), 239–251.

Appendices

Appendix A

Pronunciation Accuracy Test Text (later treated as part of the training session)

Please call Stella. Ask her to bring these things with her from the store: six spoons of fresh snow peas, five thick slabs of blue cheese, and maybe a snack for her brother Bob. We also need a small plastic snake and a big toy frog for the kids. She can scoop these things into three red bags, and we will go meet her Wednesday at the train station.

Note: Text from Weinberger, S. (2021). *Speech accent archive*. George Mason University. Retrieved from <http://accent.gmu.edu>

Appendix B

Text 1 Favourite Food



Photo by Spencer Davis on Unsplash

Part A: Read the following text and answer the questions on the next page.

My Favourite Food By Bhushan Satyal

Pizza always looks remarkably delicious. There is no better feeling in the world than a warm pizza box on your lap. I am continuously hungry for pizza at any time of the day. One of my favourite aspects of pizza is that nearly any ingredient can be put on it. Those diced vegetables and variety of meat make me want to eat more. The ingredients on each pizza are like a unique work of art. I also love that every pizza pie is a different shape and size. There are thin-crust pies, deep dish pies, and everything in between. However, the best part is the cheese. It is the secret ingredient in any food - it makes everything taste delicious. I celebrate every special occasion with pizza. There is something between me and pizza. It is a bond that cannot be broken.

Adapted from <https://medium.com/@bhushansatyal90/my-favorite-food-pizza-26b63387c47b>

Answer the following questions in the space provided below. You can answer by typing or by using Google Voice Type. Answers do not have to be complete sentences.

The answers are based on **YOUR** own preferences. You do not have to refer to what was written in the text above.

1. What is your favourite food?

2. Is it your favourite food because of how it tastes, because of what you associate it with (for example a food you eat at a special event), or both?

3. When do you eat your favourite food?

Part B: Using Google Voice Typing, write a response of 100 - 200 words to the following prompt:

Explain what your favourite food is and why it is your favourite.

You can pause and scroll up to look at the answers you wrote at any time.

If you want, you can take a few minutes to brainstorm on a sheet of paper or in the blank space on the previous page.

Begin your text below this line.

Text 2

Studying in a Pandemic



Richlegg/iStock.com

Part A: Answer the following question in the space provided below. You can answer by typing or by using Google Voice Type. Answers do not have to be complete sentences.

1. How do you feel about learning remotely?

Part B: Read the following text and answer the questions on the next page.

Survey Outlines Student Concerns 10 Months into Pandemic

By Lilah Burke

Nearly one year into the pandemic that disrupted higher education, students continue to be concerned about mental health, catching COVID-19, and the quality and challenges of online education. New polling data collected in December for the non-profit organizations New America and Third Way highlight some of these concerns.

For those students learning online through the pandemic, the modality continues to have its drawbacks, polling showed. Over half of students said they have had to make purchases, such as

computers, microphones or desks, to be able to learn online. For 70 percent of those students, the purchases came at “significant” cost.

Internet access has been another barrier to online learning. Nearly 60 percent of respondents indicated that having access to stable, high-speed internet access was a challenge for them.

While online learning comes with challenges – and over half of students said higher ed is not good quality when conducted online – students learning online overwhelmingly say they want classes to remain in a hybrid or online format, with 76 percent agreeing overall.

For students generally, the pandemic has eroded some trust in university leadership. About half of students said the pandemic made them trust university leadership less than they did before. But on the flip side, over 80 percent said their institutions cared about their safety or had ensured a safe on-campus experience.

University students generally continue to be concerned about their well-being and health. Seventy-nine percent of students surveyed said they were concerned about their mental health during the pandemic. The only greater concern for students was the disease itself, being infected or having loved ones affected.

Adapted from <https://www.insidehighered.com/news/2021/01/22/survey-outlines-student-concerns-10-months-pandemic>

Answer the questions in the space provided below. You can scroll back to re-read the text if you need to. You can answer by typing or by using Google Voice Type. Answers do not have to be complete sentences.

1. Were there parts of the text in which you found you would have:
 - a. responded the same as the students who were surveyed? Which parts and why?
 - b. responded differently than the students who were surveyed? Which parts and why?

a.

b.

2. Overall, do you think your experience as a university student in the pandemic is similar to what most university students are currently experiencing? Why or why not?

Part C: Using Google Voice Typing, write a response of **150 - 300 words** to the following prompt:

Describe your experience as a university student during the global pandemic. You can consider the **positive** and/or the **negative** aspects of your experience.

You can pause and scroll up to look at the text or the answers you wrote at any time.

If you want, you can take a few minutes to brainstorm on a sheet of paper or in the blank space on the previous page.

Begin your text below this line.

Appendix C

TAM Measurement Scales: Using Automatic Speech Recognition (ASR) in L2 Writing

Perceived Usefulness (PU)

1. Using ASR to write texts enables me to write texts more quickly.
2. Using ASR to write texts improves the quality of my texts.
3. Using ASR to write texts enhances my writing.
4. Using ASR makes it easier to write good texts.
5. Using ASR to write texts will improve my performance in my English class.

Academic Relevance (AR)

1. Using ASR when writing texts for school is important to my studies.
2. Using ASR when writing texts for school is important to my academic success.

Output Quality (OQ)

1. The quality of the text I write when I use ASR is high.
2. I like the quality of the texts I write when using ASR.

Result Demonstrability (RD)

1. I will tell other about the results of using ASR to write texts.
2. The results of using ASR to write texts are clear to me.

Perceived Ease of Use (PEU)

1. Interacting with ASR to write texts is clear and understandable.

2. I find it easy to use ASR to write texts.
3. Interacting with ASR to write texts does not require a lot of my mental effort.
4. I find it is easy to make ASR do what I want when I write a text.

Intention to Use (IU)

1. I intend to continue writing texts using ASR.
2. I intend to show others how to write texts using ASR.

Note: All items will be measured on a 7-point Likert scale: 1 = strongly disagree, 2 = moderately disagree, 3 = somewhat disagree, 4 = neutral, 5 = somewhat agree, 6 = moderately agree, 7 = strongly agree.

Appendix D

Semi-structured Interview Questions

Aim 1: [Perceived Usefulness – PU] - To identify the features and affordances that make ASR a **useful** technology when writing academic texts.

Q1: Do you feel there were more mistakes or less mistakes in your text than in a first draft that you write by hand or on the computer?

Follow up: What about grammar? Spelling? Punctuation?

Q2: Did ASR help you write the text more quickly? Why or why not?

Q3: Do you think ASR is a useful tool for writing? Explain your answer.

Aim 2: [Academic Relevance – AR] - To identify the features and affordances that make ASR a **relevant** technology for writing academic texts.

Q1: Would you consider ASR to be an important technology for writing texts in your English class? Why or why not?

Follow up: What about other courses?

Q2: Would you recommend to friends to use ASR to write their university assignments? Why or why not?

Aim 3: [Output Quality – OQ] - To identify the features and affordances of ASR that affect the **quality** of the text.

Q1: Were the texts you wrote using ASR of the same quality as what you usually write?

Follow up: Why do you think your texts were of a better quality? Do you think the content was better or worse? Do you think the language (grammar, spelling, punctuation, vocabulary) was better or worse?

Q2: Do you think that using ASR would help you write better quality texts for your university assignments? Why or why not?

Aim 4: [Result Demonstrability – RD] - To identify the features and affordances of ASR that affect the **results** of their writing.

Q1: Do you think you would get a better grade on the text you wrote with ASR than normal? Why or why not?

Q2: Do you think that using ASR to write your assignments for all of your courses would improve your grades? Why or why not?

Aim 5: [Perceived Ease of Use – PEU] - To identify the features and affordances that make ASR **easy** to use when writing academic texts.

Q1: What features were easy to use? What features were less easy for you to use? Did you continue using them even if they were less easy? Why or why not?

Q2: What did you do when issues arose with ASR? Was it easy or difficult to resolve them? Could you give example?

Aim 6: [Intention to Use – IU] - To identify if participants **intend to use** ASR to write texts after participating in the study.

Q1: Do you plan to continue using ASR to write other texts? Why or why not?

Q2: Can you give examples of other times you would use ASR to write?

Note: Whenever appropriate, the planned questions will be followed by any additional questions needed to obtain a better understanding of the participants' perceptions.