October 2015

Training the Dragon: Facilitating English Language Learner (ELL) Students' Persuasive Writing through Dictation

Nina Arcon
The University of Western Ontario

Supervisor
Dr. Perry Klein
The University of Western Ontario

Graduate Program in Education

A thesis submitted in partial fulfillment of the requirements for the degree in Master of Arts

TRAINING THE DRAGON:
FACILITATING ENGLISH LANGAUGE LEARNER (ELL) STUDENTS' PERSUASIVE WRITING THROUGH DICTATION

(Thesis format: Monograph)

by

Nina Arcon

Graduate Program in Education

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Arts

The school of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

Abstract

Writing is working memory intensive for all students, including English language learners (ELLs). Cognitive processes in writing such as transcription compete for limited resources in working memory (Bourdin & Fayol, 1994; Hayes, 2012). Previous research has shown that, when compared to handwriting, students who dictated produced better quality compositions (De La Paz & Graham, 1997; Higgins & Raskind, 1995; MacArthur & Cavalier, 2004). The goal of the present study was to investigate whether dictation would also facilitate better compositions in elementary ELL students. Using a within-subjects design, the effects of handwriting, dictation to a scribe, and dictation to a speech-to-text software were investigated on the persuasive writing of 16 elementary ELL students. Analyses of variance (ANOVAs) revealed that students had higher holistic text quality, better writing mechanics, more persuasive elements and lower cognitive load when in one or both of the dictation conditions when compared to the handwriting condition.

Keywords

English language learner (ELL); dictation; speech-to-text; persuasive writing; cognitive load.

Dedication

This thesis is dedicated to my family for their continued support, encouragement and love.

Acknowledgments

I would like to express sincere gratitude to my thesis supervisor, Dr. Perry Klein, for his insightful advice, useful guidance and immeasurable support in making this thesis possible. I would also like to thank my thesis advisory committee member, Dr. Stephen Bird, for his input and assistance throughout this process. I thank my fellow colleagues, Katrina Haug and Samanta Baker, for their continued encouragement and support. Lastly, a big thank you to all the teachers, administrators and students who participated and made this research possible.

Table of Contents

List of Tables

List of Appendices

Chapter 1: Introduction and Literature Review

Introduction

Writing is challenging for many young students. It involves multiple processes that compete for the attention of the writer including idea generation, sentence formulation, and transcription. A lack of fluency in any part of the writing process limits the availability of cognitive resources during writing (McCutchen, 1996). Young students who struggle with transcription (the process of handwriting and spelling) usually underperform on other aspects of writing because cognitive resources are limited (Bourdin & Fayol, 2000; McCutchen, 1996). Students composing in a second language face additional tasks during writing. Texts composed in a second language are often found to be more laborious and less fluent than texts composed in the primary language (Silva, 1993). Previous research with struggling writers, other than English language learners (ELLs), has shown that when transcription was removed from the writing process via dictation, students were able to compose better quality texts (De La Paz & Graham, 1997; Higgins & Raskind, 1995; MacArthur & Cavalier, 2004). The use of dictation as an alternative composition modality has not been entirely investigated with ELL students despite their difficulties with text generation. The present study investigated the effects of handwriting, dictation to a scribe, and dictation to a speech-to-text software on the persuasive writing of elementary ELL students. It was hypothesized that, when compared to handwriting, students dictating to a scribe and a speech-to-text software would: (a) compose texts with higher holistic text quality, (b) report lower cognitive load, (c) compose texts with better writing mechanics, and (d) compose arguments with more persuasive elements.

Cognitive Process Model of Writing

Writing is a complex process. A review of Flower and Hayes' (1981) Cognitive Process Model demonstrates the various processes involved in writing. Flower and Hayes (1981) developed the Cognitive Process Model using protocol analyses to identify the structures and processes that underpin writing. The model begins with the task environment, which includes elements beyond the writer's control, such as the assigned rhetorical topic. The second element of the model is the writer's long-term memory; where the writer stores their knowledge about writing in general (i.e., who the audience is). The third element is of particular significance to this paper; the set of writing processes. Flower and Hayes (1981) identified three processes that occur and interact during writing: (a) planning, (b) translating, and (c) reviewing. Planning refers to the formation of internal representations of knowledge through idea generation, organization, and goal-setting. Translating requires the writer to generate his or her ideas into language. Here, the writer must translate a meaning or an idea into the visible form, which can be a cognitively demanding task for unskilled writers like children (Flower & Hayes, 1981). The final process is reviewing, which refers to the evaluation of written content and revision. All of these processes can be used at any point, can be embedded within another process, and can interrupt one another during writing. With some modifications, these processes have continued to play a key role in cognitive theories of writing (Hayes, 2012; Kellogg, Whiteford, Turner, Cahill & Mertens, 2013).

In Hayes' (2012) recent adaptation of the Cognitive Process Model, he categorized writing tasks at the control, process, and resource levels. At the control level, processes such as goal setting take place, in which writers must plan, write and revise

their ideas. At the process level, writers propose, translate, transcribe, and evaluate their ideas. Finally, at the resource level, writers utilize individual resources that will help them with their writing task such as working memory and attention. If students have less facility in any of these processes, performance on the remaining tasks during writing is weaker (Hayes, 2012). In this revised model, he also included transcription as a writing process because it competes with the other writing processes. Additionally, transcription plays an important role in children's writing development. A closer look at transcription as a writing process will further demonstrate its role in writing.

Transcription as a Writing Process

There are many students who struggle with transcription. Transcription is the process of translating language representations into text (Berninger, 1999). It entails the physical act of forming letters via handwriting and spelling (McCutchen, 1996). Students who lack automaticity in transcription struggle with the remaining writing processes (Bourdin & Fayol, 1994; Hayes, 2012; Hayes & Chenoweth, 2006).

The cognitive demands of transcription, and their impact on other cognitive processes, were demonstrated in a series of studies by Bourdin and Fayol (1994). The researchers invited adults and children to recall series of words by dictating them aloud as well as writing them across different experimental conditions. The first experiment utilized free-rate recall by asking participants to simply recall word lists in the oral and written modes. The results indicated that children, not adults, recalled significantly fewer words in the written mode compared to the oral mode. The second experiment used fixed-rate recall by inviting participants to recall word lists in oral and written modes as in experiment one, but in a timely manner (every three seconds for adults and every six

seconds for children). The results from the second experiment ruled out the hypothesis that handwriting speed affected performance because children, not adults, again performed significantly better in the oral mode over the written mode.

The third experiment was split into two parts. Researchers first assessed whether graphic transcription contributed to lower performance in the written mode for children. The researchers added an additional recall mode, dictation, using the researcher as a scribe and found that children again recalled fewer words in the written mode compared to the oral and dictation modes. The second part assessed whether graphic execution is partly responsible for the interference of transcription with composition by asking adults to change their handwriting style to all lower-case or cursive capitals. Results indicated that adults underperformed in the cursive capitals condition when compared to oral and lower-case modes. The result of changing the handwriting style demonstrated that handwriting can also be cognitively demanding for adults.

Lastly, in the fourth experiment, the researchers investigated whether orthographic difficulties such as spelling increased working memory load. Adults and children were asked to recall familiar and unfamiliar words in oral and written modes. As expected, children performed better in the oral mode. Adults also performed better in the oral mode when they were given unfamiliar words, suggesting that orthographic difficulties were cognitively demanding.

Overall, the results of these experiments indicated that written composition is more cognitively demanding than oral language production, especially in children. The researchers explained that cognitive load was higher in children because written language production (i.e., text production, graphic execution, and control) placed a greater load on

working memory. In other words, the lack of automaticity in text production left fewer cognitive resources for higher-level writing processes, resulting in poor overall performance (Bourdin & Fayol, 1994).

Hayes and Chenoweth (2006) further investigated whether transcription in fact demands cognitive resources. They had 20 university students type texts from one computer window to another with and without articulatory suppression. Articulatory suppression refers to the repetition of a syllable or word aloud during a task, limiting the availability of working memory resources to attend to the particular task. Hayes and Chenoweth (2006) had students in the articulatory suppression condition say "tap" aloud in time to a metronome during transcription while students in the other condition simply tapped their foot to a metronome. They found that participants in the articulatory suppression condition had significantly lower transcription rates and produced writing with increased uncorrected errors. Thus, when working memory was limited, the rate of transcription was drastically slower. This study demonstrated that transcription competes for cognitive resources with the other writing processes (Hayes & Chenoweth, 2006; Hayes, 2012).

Persuasive Writing

The persuasive writing genre also presents many challenges to young writers. Often times when students are struggling with lower-level processes like transcription, they have fewer cognitive resources to attend to other higher-level processes such as argumentation in the persuasive genre. In their experimental study, Felton and Kuhn (2001) had young adults and teens participate in argumentative discourse. They found that the teens were less likely to achieve the goal of including various elements in the

argumentative discourse, such as counterarguments and rebuttals, when compared to adults.

Goal-setting strategies effectively guide students through persuasive writing. The present study used the self-regulated TREE strategy to assist ELL students during persuasive compositions. The TREE strategy has been recognized as a powerful writing strategy for many young students, guiding them through persuasive elements such as a topic sentence, reasons, explanation of reasons and a conclusion (Harris, Graham, Mason, & Friedlander, 2008).

ELL Writing

ELL writers use similar writing processes as native-speaking English writers; however, their compositions appear more laborious and less effective (Silva, 1993). In his review of research comparing English as a second language (ESL) writers and native-speaking writers, Silva (1993) found that many ESL writers also struggle with transcription demands, often producing less fluent compositions. Silva (1993) also found that ESL writers tend to do less goal-setting during writing and achieve less writing goals. This is particularly true when writing is in the persuasive genre. Often times, ESL compositions lack many argumentative elements (Silva, 1993).

Ferris (1994) also studied features associated with second language writing and found that when students have higher proficiency in the second language, they are able to not only produce longer texts, but use more writing features including synonyms, antonyms, relative clauses and conjuncts.

More recently, Fitzgerald (2008) conducted a research synthesis on multilingual writing and found similarities between native speakers and those composing in a second

language on writing processes. She also found that young bilingual writers were not as good at spelling unfamiliar words and complex phonemes as monolinguals. Young multilingual writers also had difficulties with the writing task when it did not depend on personal knowledge.

The lack of available cognitive resources during writing may be a contributing factor to ELL writing performance. Piolat, Barbier, and Roussey (2008) studied the note-taking strategies of French-speaking undergraduate students by assessing notes written in French and English as well as cognitive effort during note-taking. The researchers held two lectures; one in English (the second language) and one in French (the primary language). For each lecture, students took notes and wrote summaries of their notes, all while responding to sound signals from a computer. Students were also asked to complete a questionnaire on lecture comprehension and cognitive effort following each lecture. The researchers analyzed students' notes, summaries, and responses to questionnaires. Overall, mastery of note-taking in English was not as good as note-taking in French. Additionally, cognitive effort was greater in English. Thus, text generation for these students appeared to be more cognitively demanding in the second language when compared to their primary language.

Second language proficiency is another important factor in writing performance. Chenoweth and Hayes (2001) set out to explore the relationship between writing fluency and language experience in individuals writing in a second language. They had undergraduate, native speakers of English with varying language experience in French and German think aloud while composing essays in English and in the second language. Written compositions, think-aloud transcripts, and videotaped writing sessions were

analyzed. They found that individuals with more experience in the second language had significantly higher writing fluency. Thus, less proficiency in the second language limited the availability of cognitive resources to translate ideas into written text, resulting in lower writing fluency. Alleviating the demand on working memory resources during writing for these students could potentially enable them to produce better quality writing.

Working Memory and Writing

Working memory is responsible for the allocation of cognitive resources during writing. It consists of three parts, each with different roles: (a) the visuospatial sketchpad that stores visual information; (b) the phonological loop that stores verbal information; and (c) the central executive that manages these two parts (Baddeley, 2003). Working memory accounts for many individual differences amongst students due to its limited capacity (Just & Carpenter, 1992; McCutchen, 1996).

Vanderberg and Swanson's (2007) studied the relationship between writing processes and working memory. They invited 160 grade ten students to participate in a variety of writing and working memory tasks such as essay writing, planning, revising, written recall, etc.,. The researchers then conducted a confirmatory factor analysis (CFA) and found that measures of the central executive component of working memory significantly predicted planning, writing, and revision. They concluded that the central executive is responsible for controlling attention, which is necessary for both information storage and processing during writing.

Researchers have also attempted to pinpoint which components of working memory are in demand during text composition. Olive, Kellogg, and Piolat (2008) investigated whether writing would demand primarily verbal, visual or spatial working

memory resources during composition. They asked adults to compose persuasive texts under two conditions. In the first condition, adults wrote persuasive texts while responding to visual stimuli. Their compositions were interrupted with visual tasks (detecting changes between visual shapes), verbal tasks (responding to "ba" and "da" syllables), and spatial tasks (detecting changes between shape positions). In the second condition, adults wrote persuasive texts while responding to the same tasks, but in an aural presentation. They found that demands on verbal and visual working memory were higher than demands on spatial working memory (Olive et al., 2008). This study demonstrated that text composition exerts high demands on certain components of working memory.

McCutchen (1996) reviewed writing research that focused on working memory capacity. Here, capacity is understood as the maximum amount of activation in one's working memory during a given task that can support processing and storage components. She found that all of the writing processes are affected by capacity limitations in both adults and children. For children, the transcription process is not yet fluent; therefore, the processes of handwriting and spelling letters demand considerable resources from limited working memory capacity. Limitations in working memory capacity also contribute to overall poor writing performance. Seeing as it is difficult for children to meet the demands imposed by the writing processes, alleviating working memory resources for these students could potentially enhance their writing performance.

Supporting Composition through Dictation Methods

Dictation. As previously mentioned, the results from Bourdin and Fayol's (1994) study demonstrated that younger students performed better in the oral modes when

compared to the written mode. This suggests that inviting students to dictate their ideas to a scribe or a speech-to-text software could potentially enhance their compositions.

In their study on the effects of dictation to a scribe and persuasive planning instruction, De La Paz and Graham (1997) randomly assigned older elementary students with learning and writing difficulties to four conditions: (a) essay structure and dictation to a scribe, (b) essay structure and handwriting, (c) advanced planning and dictation to a scribe, and (d) advanced planning and handwriting. Students in the essay structure conditions learned about essay structure, reviewed essays, and practiced writing essays. Students in the advanced planning conditions learned specific strategies related to planning a persuasive essay. Researchers used self-regulated STOP and DARE strategies to teach persuasive writing. Similar to the TREE strategy used in the present study, STOP and DARE strategies guide students to form an opinion, organize ideas, and plan as they write. Overall, the researchers found that students in the advanced planning and dictation condition outperformed students in the essay structure conditions on length, completeness, cohesiveness, and quality.

In a more recent study, MacArthur and Cavalier (2004) assessed the impact of dictation as a potential test accommodation for secondary students with learning disabilities (LD). Using a repeated measures design, they assessed the essays of 31 high school students (21 of whom were identified with LD). Students composed essays under three conditions: handwriting, dictation to a scribe, and dictation to speech-to-text software. All compositions were measured for overall quality of writing, length, vocabulary, and word errors. The researchers found that both dictation conditions enabled students to produce better essays than the handwriting condition. They also found that

students composed the best essays under the dictation to a scribe condition. The present study extended MacArthur and Cavalier's (2004) research to ELL students by testing similar conditions to see if dictation would help alleviate the burden of transcription and enhance their persuasive writing skills. A closer look at assistive technologies including speech-to-text software will demonstrate the need to assess its impact on elementary ELL students' compositions.

Computer assistive technology. A variety of computer applications such as word prediction, text-to-speech, and speech-to-text help writers produce more fluent writing by reducing mechanical demands (De La Paz, 1999; MacArthur, 2009).

Silió and Barbetta (2010) studied the effects of word prediction and text-to-speech on elementary students who were culturally and linguistically diverse and had specific learning disabilities (SLD). Word prediction software works by offering users suggestions to words as they type onto a word processor. Text-to-speech is a software that dictates already transcribed text on a word processor to the user. In their study, the researchers conducted a multiple baseline design assessing the narrative compositions of six fifth-grade students with SLD who were previously ELLs. In baseline conditions, students composed narrative texts on a word processor without help from assistive software. In intervention conditions, students were separated into two cohorts. The first cohort group composed narrative texts on a word processor using word prediction alone and with text-to-speech. The second cohort group composed narrative texts on a word processor with text-to-speech alone and with word prediction. Overall, researchers found that the use of word prediction alone and in combination with text-to-speech resulted to compositions with better organization, fewer spelling errors, increased syntactical

maturity, and increased writing fluency. This study demonstrated that assistive technology could be beneficial to culturally and linguistically diverse students like ELLs. The only limitation to word prediction and text-to-speech software is that is still relies on transcription via typing; however, dictation via a speech-to-text software can remove the burden of transcription altogether.

Speech-to-text technology. Speech-to-text or speech recognition technology enables users to dictate their ideas through a microphone to receive word-processed output (Forgrave, 2002).A well-recognized speech-to-text software available in schools is Dragon NaturallySpeaking. With this program, students are not only able to have their ideas transcribed on-screen, but they are also able to control computer functions with their voices (Nuance Communications, 2015). Users must learn special commands when using the software, such as dictating "correct that" or "new line" to guide the software. Also, the program does not automatically insert punctuation, so users must dictate the appropriate punctuation. Previous research has shown that this software enabled students with writing difficulties to produce better texts (Higgins & Raskind, 1995; MacArthur & Cavalier, 2004; Quinlan, 2004).

Higgins and Raskind (1995) investigated the effectiveness of speech-to-text software on post-secondary student compositions. They compared compositions written under three modalities: (a) handwriting without assistance; (b) dictating to a scribe; and (c) dictating to speech-to-text software. Students were trained on the software in advance of participating in the writing conditions. Student compositions were rated using a single holistic measure. Researchers found that compositions written under both dictation conditions received significantly higher holistic scores than compositions written under

the handwriting without assistance condition.

In a more recent study, Quinlan (2004) assessed the impact of speech recognition on the writing performance of less fluent writers using a between-subjects, repeated measures design. A total of 41 children between the ages of 11 and 14 with varying writing proficiencies participated. Prior to writing, students were trained on Dragon NaturallySpeaking and they were taught advanced planning writing strategies for narrative writing. Students composed narrative texts under four conditions: (a) handwriting, (b) handwriting with advanced planning, (c) dictation to speech-to-text, and (d) dictation to speech-to-text with advanced planning. In the advanced planning conditions, participants had five minutes prior to the start of their condition to plan their narrative texts using advanced planning strategies, such as rehearsing "who, what, where, when, and how." Analysis of compositions included surface errors (misspelled and grammatically or semantically inconsistent words), text length, text quality, and t-unit length. Quinlan (2004) found that less fluent writers composed longer narratives and narratives with fewer surface errors in the speech-to-text conditions than in the handwriting conditions. Despite being widely available in schools, the effectiveness of speech-to-text has seldom been investigated with ELL students.

ELL and speech recognition. Coniam (1999) assessed the speech recognition accuracy of a very early version of Dragon NaturallySpeaking. The researcher invited ten Hong Kong Chinese teachers of English to read passages of text into the software. He compared their output to that of native speakers, obtained from an earlier study, with respect to t-units, clausal units, sub-clausal units, and single words. He found that outputs received by individuals with accented speech had significantly lower accuracy ratings. In

a similar study, Derwing Munro, and Carbonaro (2000) assessed the speech recognition accuracy of 30 native and non-native speakers of English. Using a sentence-by-sentence analysis, the researchers found that the speech-to-text software was not as successful at recognizing accented speech.

However, speech-to-text technology has significantly improved since the early 2000s when these studies with ELLs were conducted. With each new version of Dragon NaturallySpeaking, recognition accuracy increased (Zumalt, 2005). Additionally, current versions of the program offer users the opportunity to personalize their user profiles by indicating what variation of accented English they speak for several languages, such as British or Spanish (Nuance Communications, 2015). The present study used the most recent version of Dragon NaturallySpeaking available which was version 11. The software features in this version, relative to previous versions, included a faster and more accurate speech recognition system, an easier user profile creation, a useable toolbar and sidebar for access and commands, and a more efficient training process that was not as time consuming (Nuance Communications, 2015). Overall, the limited number of studies in this area suggest that more empirical research is necessary in order to determine the effectiveness of speech-to-text on ELL students' composition.

Present Study

The present study examined the persuasive writing and cognitive load of elementary ELL students under three modalities: handwriting (HW), dictation to a scribe (DS), and dictation to a speech-to-text software (STT). In the HW condition, students composed persuasive arguments by hand in response to an assigned rhetorical question. In the DS condition, the researcher acted as a scribe and typed out students' dictated

responses. In the STT condition, students dictated their persuasive responses to Dragon NaturallySpeaking version 11. Following each condition, students completed a cognitive load questionnaire where they rated how difficult the task was and how much effort they exhausted on 9-point Likert scales (Sweller, Ayres, & Kalyuga, 2011). This subjective rating has been established as one of the most sensitive measures available for rating cognitive load (Sweller et al., 2011). Aside from cognitive load, the present study also assessed holistic text quality, writing mechanics, and number of persuasive elements.

Hypotheses

When compared to students in the handwriting condition, it was hypothesized that students in the dictation to a scribe and speech-to-text conditions would:

- Compose texts with higher holistic text quality;
- Report lower cognitive load;
- Compose texts with better writing mechanics;
- Compose arguments with more persuasive elements.

Chapter 2: Method

Research Design

The present study used a repeated measures design to assess the effects of handwriting (HW), dictating to a scribe (DS), and dictating to a speech-to-text software (STT) on ELL students' holistic text quality, cognitive load, writing mechanics, and persuasive elements.

Participants

This study was conducted in an elementary school in a mid-sized city. It was located in a neighborhood that served lower socioeconomic status (SES) and middle SES students. The student population was ethnically and linguistically diverse.

Students were invited to participate in the study if they were receiving in-school ESL support services at the time of the study. Participants had to be within nine to 14 years of age and they were required to have had at least one year of education in English to participate in the study to ensure that at the onset, they had learned general English vocabulary relevant to operating Dragon NaturallySpeaking. None of the students had severe speech impediments. Only one student had a mild lisp and remained in the study analysis.

Students received an explanation of the study details from the researcher at the time that they individually arrived to receive support in the ESL/ELD classroom. Following this introduction, letters of information and assent forms were distributed (see Appendices A and B). Information about students' age, sex, backgrounds and computer usage was collected from the students via a take-home demographic questionnaire (see Appendix C). Information about students' most recent report card grades in writing, type

of ESL program, classroom computer use, and whether they were on an Individual Education Plan (IEP) was collected from the ESL/ELD teacher upon consent via a teacher questionnaire (Appendix D). For type of ESL program, the teacher indicated whether students were in a regular ESL program or a modified one. The regular ESL/ELD program provided instruction as identified by the curriculum whereas the modified one was more individualized to suit different learning needs. For parents and students who did not wish to participate, data was not collected and students simply carried on with regular ESL/ELD instruction. No financial compensation was provided for participation, but a small gift valuing approximately five dollars (i.e., school supplies) was handed out to each student who participated along with a participation certificate (see Appendix E).

Sixteen ELL students between the ages of nine and 14 years (M = 11.06, SD = 1.34) participated from one elementary school setting. There were five students in the regular ESL program and 11 in the modified one. All students had at least one year of experience at an English-speaking school (M = 3.17, SD = 1.83). The average writing grade was 2.44 (SD = 0.73). This can be interpreted as a "C" grade in academia. Two students had IEPs because they were receiving special education programs in addition to ESL services.

For 11 out of 16 students, their home countries were in the Middle East. The remaining students were from Somalia (n = 2), Afghanistan (n = 2), and Columbia (n = 1). Half of the students in the study, including their parents, spoke Arabic as a first language, so a translated letter of information was sent out to these students to ensure they understood study details (see Appendix F). The ESL teacher indicated that parents of

other students understood sufficient English to read the consent letter.

According to questionnaires, all of the students had used a computer before; nine students used it every day and seven students used it at least three to four times a week. The majority of students were able to type with various levels of proficiency (N = 14), with only two being unfamiliar with typing. None of the students had used Dragon NaturallySpeaking prior to the commencement of the study.

Confidentiality

Any information obtained from students, teachers, and guardians in connection with this study remained confidential. Upon consent, each student was randomly assigned a three-digit identification number. All data obtained from students throughout the study including questionnaires and compositions was saved under their assigned identification numbers. The master list linking names to their corresponding identification numbers was stored separately from the remainder of the data on a password protected hard drive. All data obtained from this study was stored and locked in a filing cabinet in a locked institution. Additionally, all electronic files (i.e., student compositions) were kept on an encrypted hard drive and stored in the locked filing cabinet when not in use.

Setting

All training sessions and writing activities took place in the ESL/ELD classroom during school hours at the time that individual students received ESL services. When there was more than one student present in the classroom, students sat with one vacant chair between them. When there were too many students in the ESL/ELD classroom to conduct the study effectively, some students were taken to the school computer lab to conduct training sessions or conditions. Completing writing activities in the ESL/ELD

classroom and the computer lab ensured that students were tested under normal conditions.

Procedure

Once consent forms were returned, students were randomly assigned to a sequence of conditions (STT, DS, and HW) that were counterbalanced with respect to order and writing prompt. A schedule was organized over the course of one month with dates for the two training sessions on Dragon NaturallySpeaking and three writing activities for each student.

Training. Prior to conducting the conditions, the primary researcher trained the students individually on a laptop equipped with Dragon Naturally Speaking v. 11 on two separate occasions. In the first training session, students spent 15 to 20 minutes creating their user profiles (see Appendix G for first lesson). Part of this process included training the software to accurately recognize each student's voice. Students had to dictate several passages of text provided by Dragon NaturallySpeaking into the software. Students were able to rehearse the texts prior to dictating. For students who struggled with decoding during this process, the researcher whisper-read the texts to them as they dictated. The training session ended when the following prompt appeared: "Congratulations! You have finished training." There were four students who did not successfully train their voices in the first training session because of technical issues (user profiles were not saving). These students were able to successfully train their voices during the second training session.

The second training session took place approximately one to three days following the first training session. Recall that scheduling was dependent on student availability in the ESL/ELD classroom. In the second session, students practiced dictating to Dragon

NaturallySpeaking alongside the researcher for approximately 20 minutes (see Appendix H for second lesson). They were given a tip sheet to help them remember common dictation commands (see Appendix I). The researcher read over the tip sheet and explained all the commands necessary for Dragon NaturallySpeaking to work. Students then practiced dictating the following three sentences until accurate recognition was successful: "I saw a dragon today. It had big green wings. It looked a little scary, but it was very friendly." When these sentences were dictated successfully, students were then asked to complete a dictation activity that measured the accuracy of speech recognition (MacArthur & Cavalier, 2004). For this activity, students dictated a narrative writing passage into Dragon NaturallySpeaking (see Appendix J). Prior to dictating the passage, the researcher read the text aloud and invited the student to practice. The researcher then turned on the microphone and instructed students to dictate the passage. The students were told not to correct any mistakes and to leave the output from Dragon NaturallySpeaking as it was. The output received from Dragon NaturallySpeaking was saved, and recognition accuracy was later calculated.

Composing. Following the two training sessions, in three subsequent sessions, students composed persuasive arguments under three conditions: handwriting (HW), dictation to a scribe (DS), and dictation to speech-to-text software (STT). Students had 20 minutes to compose each persuasive text in response to an assigned topic (see Appendices K to M for condition instructions). The three topics were: (a) Do you think students should have more time for recess? Why? (b) What is the best subject in school? Why? and (c) Imagine you can choose to be five years older. Would you want to be five years older? Why? These topics were chosen because they were the least culturally

biased of a variety of topics considered; they appeared to be comparable to each other in comprehensibility and difficulty; and they were likely to be of interest to elementary students because they relied on experiences that most children have. The ESL/ELD teacher also verified that the topics were appropriate for all students. Students were not in more than one condition per day so as not to exhaust or bore them.

At the start of each condition, students were given the following general instructions:

> Today you will be writing (or speaking) your opinion on the following topic. You will have 20 minutes to write (or speak). You may take a break at any time. If you make any mistakes, you are able to edit and correct them. I will tell you when you have two minutes left. If you wish to stop and discontinue writing (or speaking) at any time during the 20 minutes, please let me know. When you are done, you will answer two survey questions. The first question asks you how easy or difficult the activity was from 1 (very very easy) to 9 (very very hard). The second question asks you how much effort you had to put into this activity or how hard did you try from 1 (very very little effort) to 9 (very very much effort).

Following the general instructions, students were handed their randomly assigned persuasive topic. The researcher then explained the TREE strategy that was listed at the top of each topic handout:

> Now you will write your opinion on the following topic. You can use the TREE strategy to help you persuade the reader. The first step is, "T," come up with a topic sentence; tell the reader your opinion. Next, "R," give three reasons for your opinion; why is your opinion right? Next, "E,"

explain why those reasons are right; say more about each reason to persuade your reader. Lastly, "E," give a good ending sentence. Remember to try and convince your reader that your opinion is right.

After composing in each condition, students were asked to fill out cognitive load surveys by completing the Likert scales (see Appendix N).

***Handwriting (HW) condition*.** In the HW condition, students wrote their responses to the persuasive writing topic on the handout provided. Students were given a pen, a pencil, an eraser, and extra lined paper. They were asked to make any revisions directly on the page.

Dictation to a scribe (DS) condition. In the DS condition, students dictated their ideas to a scribe (the researcher). The scribe typed student dictations verbatim onto a laptop in front of the student, showing them their transcribed ideas on-screen. When students wished to make corrections, they notified the researcher by pointing on-screen to the location of the error and they dictated the revision. Students were also responsible for dictating punctuation. Once students finished dictating, transcriptions were saved under student identification numbers on the password protected hard drive.

***Dictation to speech-to-text (STT) condition*.** In the STT condition, students dictated their persuasive responses onto Dragon NaturallySpeaking. In this condition, the researcher opened the student's profile on Dragon NaturallySpeaking, opened Microsoft Word, and went over the tip sheet to remind students of speech recognition commands. The researcher then ensured that the headset was set up properly and that the software was ready to use. As students dictated, the researcher was in charge of turning the microphone on and off for the students, which they signaled by raising their hand. This

was done to ensure that the program did not pick up any additional words while students thought aloud. If the software misinterpreted a word that the student dictated after three attempts, then the researcher typed out the word. The researcher kept a tally of the number of typed words for each student. Just as students revised in other conditions, they revised in the STT condition by voice commands or by typing. It is important to note that spell check was turned on, but that grammar check was turned off. The textual output received from this condition was saved under the student's identification number onto the password protected external hard-drive.

Materials

Dragon NaturallySpeaking version 11. The researcher's laptop was equipped with Dragon NaturallySpeaking for the students to use for training and composing.

Headset for Dragon NaturallySpeaking. Headsets distributed by the school were used for training sessions and the STT condition. The headsets were Plantronics .Audio 655 DSP, with features including an adjustable and noise-canceling microphone.

Watch. In order to assess writing fluency across all three conditions, the researcher used a watch to time students in each condition. The researcher wrote down the start and end times for each composition.

Measures

The writing measures in this study assessed aspects of the writing process that are typically taught in schools, such as spelling and persuasive genre elements. All handwritten texts were typed in order to mask compositions with respect to condition. As described in the section below, the primary researcher calculated the scores on the recognition accuracy of Dragon NaturallySpeaking and the writing fluency measures.

Two research assistants independently rated the holistic text quality and counted the number of surface errors and persuasive elements. Identity of the participants and condition of each text was masked and identified by a random alphanumeric code.

Speech recognition accuracy. The accuracy of speech recognition was assessed to determine how well Dragon NaturallySpeaking recognized elementary ELL students' speech. Recall that students dictated a short narrative text into Dragon NaturallySpeaking during the second training session on the software. The textual output produced by Dragon NaturallySpeaking was compared to the original narrative text on word and punctuation accuracy. This was done by dividing the total number of words and punctuation elements recognized accurately, by the total number of words and punctuation elements in the original text. In order to ensure objectivity, a research assistant was asked to calculate word and punctuation accuracy for half of the texts. For both word and punctuation accuracy, inter-rater agreement was 100%.

Holistic text quality. Compositions were scored using a holistic rating of text quality (see Appendix O). A holistic criterion refers to the overall subjective rating of the written product. According to Graham and Perin (2007), holistic measures are the most common and useful method for evaluating writing quality.

To measure holistic quality in this study, a rater was asked to sort all 48 compositions into seven piles ranging from (1) very low quality to (7) very high quality with (4) being average quality. This rater was then asked to go through each pile and select a composition that was most representative of that pile to be the index text. Once this process was complete, two raters were brought together to discuss the seven chosen index texts. They practiced rating an additional three texts and reached agreement on all

three. The raters were then asked to independently use the seven index texts to rate the remaining 38 compositions holistically while ignoring surface errors and keeping in mind criteria such as ideas, content, organization, and overall persuasiveness. Inter-rater reliability was strong across all three conditions (see Table 1).

Table 1.

Inter-rater Reliabilities

Measure	Condition	*n*	*r*	*p*
Holistic Quality	Dragon	12	.96	< .001
	Scribe	12	.95	< .001
	Writing	14	.96	< .001
Surface Errors	Dragon	15	.88	< .001
	Scribe	14	.98	< .001
	Writing	15	.98	< .001
Persuasive Elements	Dragon	15	.83	< .001
	Scribe	14	.79	< .001
	Writing	15	.86	< .001

Cognitive load. To test the hypothesis that students would report lower cognitive load in the STT and DS conditions than in the HW condition, students were asked to complete a cognitive load survey following each condition (see Appendix N). Students rated difficulty and effort on two 9-point Likert scales (Sweller et al., 2011).

Writing mechanics. To test whether students in DS and STT conditions composed texts with better writing mechanics than in the HW condition, the proficiency

of students' writing mechanics was evaluated based on two measures: surface errors and writing fluency.

Surface errors. Initially, the count of surface errors included misspelled words, semantically or grammatically inconsistent words, beginning of sentence capitalization errors, and end of sentence capitalization errors (MacArthur & Cavalier, 2004; Quinlan, 2004). Two raters were trained on coding surface errors. In the initial training, the coding scheme was discussed, examples of surface errors were provided, and samples of texts were coded. It became apparent throughout the training that the raters were not coding surface errors reliably. The problem appeared to have been with the differentiation between spelling errors and semantically or grammatically inconsistent word errors in the coding scheme. MacArthur and Cavalier (2004) had similar issues in their study, finding that separating errors in different categories led to lower reliability. Thus, a new coding scheme was provided to raters and a second training session was held (see Appendix P). In the second training, raters were asked to code surface errors, which included capitalization errors (missing and incorrect capitalization), punctuation errors (missing and incorrect punctuation), and word errors (spelling errors, homophones, semantic errors, missing words, double or unnecessary words, pronoun errors, verb/subject disagreements, and misuse of apostrophes). Following a description of the new coding scheme, raters were asked to independently count the number of surface errors in four sample texts. After raters agreed on coding the four texts, they were given the remaining 44 texts to code. Inter-rater reliability for total count of surface errors per text was very strong across the three conditions (see Table 1). For texts in which the raters disagreed on the number of errors, a resolution rating was reached by averaging the two ratings.

Writing fluency. Writing fluency measures included total composition time (measured in minutes), text length (measured in words), mean length of words (measured in letters), total sentence count, and words per minute. The mean length of words was calculated by dividing the total number of characters in a text by text length. Words per minute were calculated by dividing the total time by text length. The remaining measures were determined using the word count feature of a word processor.

Number of persuasive elements. It was hypothesized that students in DS and STT conditions would compose arguments with more persuasive elements than when in the HW condition. To test this, raters counted the presence of persuasive elements in all compositions (see Appendix Q). The two raters were asked to code the following four persuasive elements, which corresponded to the elements of the TREE strategy: (a) topic sentence, (b) reasons, (c) explanation of reasons, and (d) conclusion. Additionally, raters were asked to code if other persuasive elements appeared in the texts including alternative claims, reasons for alternative claims, and rebuttals against alternative claims. Raters were trained by coding four compositions chosen by the researcher to illustrate the diversity of student responses. Raters coded the texts independently, with the conditions masked. The number of persuasive elements was summed for each text, to yield an approximately normally distributed variable. Inter-rater reliability for the total number of persuasive elements was strong across all conditions (see Table 1). To resolve differences between ratings, the totals were averaged.

Quality

To ensure the present study could contribute to the growing body of writing intervention research, Graham and Harris' (2014) 12 recommendations for conducting

high quality writing research were used as a guide (see Table 2).

Table 2.

Graham and Harris' (2014) Twelve Recommendations and their Approaches in the Present Study

Recommendation	Approach in the Present Study
1. Ask meaningful questions	What are the effects of dictation to a scribe and dictation to a speech-to- text software on elementary ELL students' persuasive writing and cognitive load?
2. Test writing interventions that are well-founded and designed	Study design was carefully thought-out and well-founded based on previous research using dictation as a writing intervention (De La Paz & Graham, 1997; Higgins & Raskind, 1995; MacArthur & Cavalier, 2004; Quinlan, 2004).
3. Compare targeted writing intervention to a credible control/comparison condition	Within-subjects design enabled participants to serve as their own comparisons across conditions.
4. Apply psychometrically sound assessments	The validity of each measure had been tested in previous published studies, and inter-rater reliability was high in the present study.

5. Make the study as representative of the real world context as possible	Students were tested at school under normal conditions in their ESL/ELD classroom and/or computer lab.
6. Apply a rigorous design to answer research questions	A within-subjects design was imperative because ELL students could not be easily grouped due to demographic differences including primary language.
7. Make certain the study is properly powered	Use of within-subjects design provided statistical power because student served as their own comparisons (see a discussion of this issue in limitations).
8. Properly analyze the data	Data analysis included planned, repeated measures of analyses of variance (ANOVAs) testing the effects of holistic text quality, cognitive load, writing mechanics, and persuasive writing elements across conditions. Assumptions for each measure were also tested.
9. Ensure the study is conducted in an ethical manner	Ethical responsibilities were met. Also, the activities of this study and the time devoted to them were valuable to these students because many of them have not used speech-to-text technology before.

10. Take steps to ensure that what is supposed to happen does happen	The researcher ensured that delivery of instructions, lesson plans, instructional handouts were consistent across each participant. Additionally, conditions and topics were counterbalanced across conditions.
11. Provide a clear, cogent, and full description of the study	Full description of the present study meets APA criteria for complete reporting of experimental studies.
12. Design a series of studies to refine and test the writing intervention	Seeing as this was one study, the following criterion was not possible to meet.

Chapter 3: Results

Analysis

One-way, repeated measures analyses of variance (ANOVAs) were conducted to assess the effects of composition modality (HW, DS, and STT) on holistic text quality, cognitive load, writing mechanics, and persuasive elements. Planned ANOVAs were used for all of the analyses to provide more statistical power to hypothesized differences between STT and HW conditions, and DS and HW conditions. Thus, post hoc tests were not reported. It is important to note that differences between group means that were not statistically significant, but that showed a medium effect size statistic, were reported because sample size likely reduced statistical power for these differences. Additionally, for the purposes of this paper, a partial eta squared of .03 was interpreted as a small effect, a partial eta squared of .06 was interpreted as a medium effect, and a partial eta squared of .14 was interpreted as a large effect.

Speech recognition accuracy. If you recall, the results for speech recognition accuracy were analyzed by comparing the percentage of accurate word and punctuation recognition from the original narrative text to the one that students dictated into Dragon NaturallySpeaking. ELL students reached a mean word recognition accuracy of 78% (*SD* = .13) and a mean punctuation recognition accuracy of 98% (*SD* = .05) on Dragon NaturallySpeaking (see Table 3). Eleven students reached over 80% word recognition accuracy, two students had 70% to 79% word recognition accuracy and three students had 52% to 59% word recognition accuracy. For punctuation recognition accuracy, 14 out of 16 reached 100% accuracy. During STT conditions, Dragon NaturallySpeaking

could not accurately recognize on average two to three words (see Table 3). In these cases, the researcher typed out the word(s) for the student.

Table 3.

Results from Speech Recognition Accuracy

Variable	M*(SD)*	Min.	Max.
Word recognition accuracy on Dragon	.78*(.13)*	.52	.94
Punctuation recognition accuracy on Dragon	.98*(.05)*	.86	1
Count of times researcher typed during STT	2.44*(2.42)*	0	9

Holistic text quality. It was hypothesized that compositions in the STT and DS conditions would have higher holistic text quality than compositions in the HW condition. All data for holistic text quality met assumptions of normality; therefore, the tests of within-subjects effects with sphericity assumed were interpreted. The means and standard deviations across each condition are presented in Table 4. Overall, holistic quality ratings differed significantly as a function of the three modality conditions with a large effect size, $F(2, 30) = 6.45$, $p < .05$, $\eta_p^2 = .30$. The texts composed under the DS condition had significantly higher holistic quality ratings than the texts composed under the HW condition, $F(1, 15) = 18.90$, $p = .001$. Effect size was large ($\eta_p^2 = .57$). There were marginal differences in holistic ratings between texts composed under the STT condition and the HW condition with a large effect size, $F(1, 15) = 3.65$, $p = .08$, $\eta_p^2 = .20$.

Table 4.

Descriptive Statistics for Main Dependent Variables

		Dragon	Scribe	Writing
Category	Variables	M*(SD)*	M*(SD)*	M*(SD)*
Holistic	Text Quality	3.66*(1.71)*$_{ab}$	4.31(*1.57*)$_{a}$	2.91(*.90*)$_{b}$
Cognitive Load	Perceived Difficulty	3.25*(2.60)*$_{ab}$	3.06*(1.61)*$_{a}$	4.69*(2.36)*$_{b}$
	Mental Effort	6.13*(2.27)*$_{a}$	6.31*(1.99)*$_{a}$	7.06*(1.44)*$_{b}$
Writing Fluency	Total Time	10.44*(3.97)*$_{ab}$	6.75*(3.15)*$_{a}$	12.81*(5.76)*$_{b}$
	Text Length	63.44*(32.79)*$_{ab}$	91.50*(42.88)*$_{a}$	60.31*(26.46)*$_{b}$
	Word Length	3.92*(.33)*$_{ab}$	3.94*(.35)*$_{ab}$	3.87*(.41)*$_{ab}$
	Sentence Count	4.06*(1.61)*$_{ab}$	4.69*(1.96)*$_{ab}$	3.94*(2.05)*$_{ab}$
	Words Per Minute	7.23*(4.55)*$_{ab}$	15.28*(7.85)*$_{a}$	6.06*(4.35)*$_{b}$
Surface Errors	Total Surface Errors	6.94*(3.82)*$_{a}$	10.69*(6.64)*$_{a}$	18.50*(11.96)*$_{b}$
	Word Errors	5.13*(2.96)*$_{a}$	7.75*(5.39)*$_{ab}$	13.75*(10.73)*$_{b}$
	Cap. Errors	.28*(.45)*$_{a}$	.28*(.36)*$_{a}$	2.56*(2.43)*$_{b}$
	Punc. Errors	1.53*(1.37)*$_{ab}$	2.66*(1.71)*$_{ab}$	2.19*(1.54)*$_{ab}$
Persuasive Elements	Number of Elements	6.19*(1.59)*$_{ab}$	7.53*(2.16)*$_{a}$	6.03*(1.45)*$_{b}$

Note. Means sharing common subscript do not differ significantly, $p > .05$.

These were planned comparisons that compared STT and HW, and DS and HW.

Cap. = Capitalization; Punc. = Punctuation.

Cognitive load. It was hypothesized that students' cognitive load would be higher in the HW condition when compared to DS and STT conditions. The means and standard

deviations for students' cognitive load ratings across the three conditions are presented in Table 4.

Perceived difficulty. Although some variables showed kurtosis, it was overall sufficiently normal to allow the application of ANOVA. The ANOVA results indicated that students' ratings of perceived difficulty did not significantly differ as a function of the three modalities, $F(2, 30) = 2.52$, $p = .10$, $\eta_p^2 = .14$; however, the effect size was large. Students perceived the HW condition as significantly more difficult than the DS condition with a large effect size, $F(1, 15) = 5.29$, $p < .05$, $\eta_p^2 = .26$. The difference between perceived difficulty in the STT condition and the HW condition was not statistically significant, but was large in effect size, $F(1, 15) = 2.34$, $p = .15$, $\eta_p^2 = .14$.

Mental effort. Data met assumptions of normality across all conditions for self-reported mental effort. The results from the ANOVA on mental effort indicated that the mental effort ratings differed significantly as a function of the three modality conditions with a large effect size, $F(2, 30) = 5.34$, $p < .05$, $\eta_p^2 = .26$. The mean mental effort ratings were significantly lower in the DS condition when compared to the HW condition with a large effect size, $F(1, 15) = 7.94$, $p < .05$, $\eta_p^2 = .35$. The mean mental effort ratings were also significantly lower in the STT condition than in the HW condition with a large effect size, $F(1, 15) = 10.08$, $p < .05$, $\eta_p^2 = .40$.

Writing mechanics. Writing mechanics included measures of writing fluency and surface errors. Recall that it was hypothesized that texts composed in the DS condition and the STT condition would have better writing mechanics than texts composed in the HW condition.

***Writing fluency*.** The six different measures of writing fluency were: (a) total composition time, (b) text length, (c) word length, (d) sentence count, and (e) words per minute. The means and standard deviations for all writing fluency measures are presented in Table 4.

***Total time*.** Data met assumptions of normality for total time across all conditions. The ANOVA results indicated that total time differed significantly as a function of the three modalities with a large effect size, $F(2, 30) = 10.35$, $p < .001$, $\eta_p^2 = .41$. Total time in the DS condition was significantly less than the total time in HW condition with a large effect size, $F(1,15) = 19.48$, $p < .05$, $\eta_p^2 = .57$. Total time in the STT condition was not significantly less than total time in HW condition, although the effect size statistic was medium in size, $F(1,15) = 2.28$, $p = .15$, $\eta_p^2 = .13$.

***Text length*.** Although some variables showed skewness and kurtosis, it was overall sufficiently normal to allow the application of ANOVA. The results indicated that text length differed significantly as a function of the three modalities with a large effect size, $F(2, 30) = 8.24$, $p = .001$, $\eta_p^2 = .36$. Text length in the DS condition was significantly greater than text length in HW condition with a large effect size, $F(1, 15)= .11.33$, $p < .05$, $\eta_p^2 = .43$. There was no significant difference between text length in the STT condition and the HW condition and effect size was small, $F(1, 15)= .20$, $p = .66$, $\eta_p^2 = .01$.

***Word length*.** Despite there being slight skewness and kurtosis across the three conditions for word length, data was assumed to be normal for the application of ANOVA. The ANOVA results indicated that word length did not significantly differ as a function of the three conditions, and the effect size statistic was small, $F(2, 30) = .14$, $p =$

.87, η_p^2 = .01. Words composed in the DS condition were not significantly longer than words composed in the HW condition, and there was a small effect size, $F(1,15) = .23$, $p = .64$, η_p^2 = .01. Words composed under the STT condition were also not significantly longer than words composed in the HW condition, and the effect size was small, $F(1,15) = .18$, $p = .68$, η_p^2 = .01.

Sentence count. All data met assumptions of normality across all variables. The results from the ANOVA indicated that sentence count did not significantly differ as a function of the three modalities, $F(2, 30) = 1.36$, $p = .27$, $\eta_p^2 = 08$; however, the effect size statistic was medium. Sentence count was not significantly different in the DS condition when compared to the HW condition, but there was a medium effect size, $F(1,15) = 2.21$, $p = .16$, $\eta_p^2 = .13$. Sentence count in the STT condition was also not significantly different than sentence count in the HW condition and there was no effect, $F(1, 15)= .06$, $p = .82$, $\eta_p^2 = .00$.

***Words per minute*.** There was slight skewness and kurtosis across the three conditions for words per minute, but overall it was assumed to be normal for the application of ANOVA. Overall, words per minute significantly differed as a function of the three conditions with a large effect size, $F(1.50, 22.54) = 25.03$, $p < .001$, $\eta_p^2 = .63$. Planned comparisons showed that there were significantly more words per minute generated in the DS condition when compared to the HW condition with a large effect size, $F(1, 15) = 26.62$, $p < .001$, $\eta_p^2 = .64$. There was no significant difference between words per minute generated in the STT condition when compared to the HW condition, but the effect size statistic was medium, $F(1, 15) = 1.50$, $p = .24$, $\eta_p^2 = .09$. .

Surface errors. The measure of surface errors included counts of word errors, capitalization errors, and punctuation errors. The means and standard deviations for these variables are presented in Table 4. For several surface error measures, the homogeneity of variances assumptions were violated because variances differed substantially between groups; therefore, the Huynh-Feldt test of within-subject effects was interpreted because these tests corrected the degrees of freedom in order to estimate sphericity.

Total surface errors. Overall, data for total surface errors was slightly skewed and kurtotic. Results from the ANOVA indicated that Mauchly's test of sphericity was significant ($p < .05$); thus, the assumptions for homogeneity of variances were violated. Therefore, the degrees of freedom were corrected using the Huynh-Feldt estimate of sphericity. Overall, surface errors differed significantly as a function of the three conditions with a large effect size, $F(1.54, 23.04) = 9.78$, $p < .01$, $\eta_p^2 = .40$. Planned comparisons showed that there were significantly fewer surface errors in the DS condition when compared to the HW condition with a large effect size, $F(1,15) = 5.86$, $p < .05$, $\eta_p^2 = .28$. There were also significantly fewer surface errors in the STT condition when compared to the HW condition with a large effect size, $F(1,15) = 16.47$, $p < .01$, $\eta_p^2 = .52$.

***Word errors*.** There was skewness and kurtosis for word errors across conditions. The results from the ANOVA indicated that Mauchly's test of sphericity was significant ($p < .05$), indicating that variances differed significantly between groups; therefore, the degrees of freedom were corrected using the Huynh-Feldt estimate of sphericity. Results indicated that the number of word errors differed significantly as a function of the three conditions with a large effect size, $F(1.31, 19.70) = 6.48$, $p < .05$, $\eta_p^2 = .30$. There were

marginally fewer word errors in the DS condition when compared to the HW condition with a large effect size, $F(1,15) = 3.61$, $p = .07$, $\eta_p^2 = .19$. There were significantly fewer word errors in the STT condition when compared to the HW condition with a large effect size, $F(1,15) = 11.92$, $p < .01$, $\eta_p^2 = .44$.

***Capitalization errors*.** Datum was skewed and kurtotic in STT and HW conditions, but assumed to be normal for the DS condition. It is important to note that most students in this measure scored zero; thus, variability of datum points was low. Results from the ANOVA indicated that Mauchly's sphericity test was significant ($p < .001$); therefore, the degrees of freedom were corrected using the Huynh-Feldt estimate of sphericity. Results indicated that the number of capitalization errors significantly differed as a function of the three conditions with a large effect size, $F(1.07, 16.06) = 17.00$, $p < .01$, $\eta_p^2 = .53$. Planned comparisons showed that when compared to the HW condition, there were significantly fewer capitalization errors in the DS condition with a large effect size, $F(1, 15) = 16.43$, $p < .01$, $\eta_p^2 = .52$. There were also significantly fewer capitalization errors in the STT condition when compared to the HW condition with a large effect size, $F(1, 15) = 18.37$, $p < .01$, $\eta_p^2 = .55$.

***Punctuation errors*.** There was skewness and kurtosis across the three conditions, but Shapiro-Wilk tests of normality were not significant $p > .05$; therefore, data was assumed to be normal for the application of ANOVA. The results from the ANOVA indicated that punctuation errors marginally differed as a function of the three modalities with a large effect size statistic, $F(2, 30) = 2.59$, $p = .09$, $\eta_p^2 = .15$. Punctuation errors in the DS condition were not significantly different than punctuation errors in the HW condition, $F(1,15) = .91$, $p = .35$. This is despite there being a medium effect size, $\eta_p^2 =$

.06. Punctuation errors in the STT condition were also not significantly different than those in the HW condition, but there was a medium effect size, $F(1, 15)= 2.07$, $p = .17$, $\eta_p^2 = .12$.

Persuasive elements. The number of persuasive elements included counts of topic sentences, reasons, explanation of reasons, conclusions and other persuasive elements such as alternative claims or rebuttals. It was hypothesized that compositions would include more persuasive elements when students were in DS and STT conditions than when in the HW condition. An ANOVA was conducted for total persuasive elements across all three conditions. Means and standard deviations are present in Table 4. Overall, the number of persuasive elements differed significantly as a function of the three modality conditions with a large effect size, $F(2, 30) = 5.75$, $p < .05$, $\eta_p^2 = .28$. The mean number of persuasive elements was significantly higher in the DS condition when compared to the HW condition with a large effect size, $F(1, 15) = 10.29$, $p < .05$, $\eta_p^2 = .41$. The mean number of persuasive elements was very similar between the STT condition and the HW condition, so differences were not significant and small in effect size, $F(1, 15) = .14$, $p > .05$, $\eta_p^2 = .01$.

Correlations. To further understand the relationship between student characteristics and main dependent variables, correlations were calculated. Three different correlation analyses were carried out based on modality (see Tables 5 to 7). Overall, there were some significant, medium to large relationships worth mentioning.

Student predictors of word recognition accuracy. Interestingly, word recognition accuracy on Dragon NaturallySpeaking positively correlated with years in an English-speaking school, $r(16) = .57$, $p < .05$. There was also a significant positive correlation

between word recognition accuracy and writing grade, $r(16) = .61$, $p < .01$. There was a significant positive relationship between holistic text quality in STT and word recognition accuracy, $r(16) = .74$, $p < .01$. There were also correlations between some writing fluency measures and word recognition accuracy. There were significant positive relationships between word recognition accuracy and text length in STT, $r(16) = .57$, $p < .05$, sentence count in STT, $r(16) = .72$, $p < .01$ and words per minute in STT, $r(16) = .61$, $p < .01$. Lastly, there was a significant positive relationship between word recognition accuracy and number of persuasive elements in STT, $r(16) = .52$, $p < .05$.

Student predictors of holistic text. There were several student predictors of holistic text quality across conditions. Holistic text quality positively correlated with words per minute in STT, $r(16) = .70$, $p < .01$ and DS, $r(16) = .68$, $p < .01$, but not in the HW condition, $r(16) = .14$, $p > .05$.

There was a significant, positive relationship between holistic text quality and text length in the STT condition, $r(16) = .89$, $p < .001$, in the DS condition, $r(16) = .74$, $p < .01$, and in the HW condition, $r(16) = .47$, $p < .05$.

There were also significant positive correlations between holistic text quality and number of persuasive writing elements in the STT condition, $r(16) = .64$, $p < .01$, in the DS condition, $r(16) = .63$, $p < .01$, and in the HW condition, $r(16) = .44$, $p < .05$.

Interesting correlations that add to these relationships are that of text length and number of persuasive elements in the STT condition, $r(16) = .70$, $p < .01$, in the DS condition, $r(16) = .62$, $p < .01$, and in the HW condition, $r(16) = .72$, $p < .01$.

Table 5.

Correlation Matrix for Variables and Speech-to-Text Condition

	AG	YR	GR	WR	PR	CT	TD	HD	C1	C2	TLD	WLD	SCD	WMD	WED	CED	PED	SED	RED
AG	1																		
YR	-.03	1																	
GR	-.44*	.22	1																
WR	.05	.57*	.61**	1															
PR	.31	.20	-.30	-.06	1														
CT	.61**	.04	-.49*	-.18	.15	1													
TD	-.11	-.22	-.33	-.35	-.40	.26	1												
HD	.36	.61**	.40	.74**	.15	.33	-.27	1											
C1	.09	-.06	-.24	.17	.04	.31	.35	.20	1										
C2	-.09	-.27	.01	-.11	-.32	.22	.17	.11	.41	1									
TLD	.43*	.56*	.10	.57*	.24	.40	-.11	.89***	.31	.27	1								
WLD	-.03	0	.62**	.42	-.18	-.32	-.26	.24	-.02	-.18	-.09	1							
SCD	-.25	.52*	.66**	.72**	-.23	-.18	-.14	.69**	-.02	0	.52*	.30	1						
WMD	.16	.43*	.45*	.61**	.32	-.04	-.71**	.70**	-.09	.11	.67**	.18	.52*	1					
WED	.41	-.06	-.06	.14	.31	.26	.23	.40	.49*	.35	.55*	.06	-.07	.14	1				
CED	-.03	-.49*	.11	-.27	.03	-.06	-.06	-.26	-.47*	0	-.23	-.10	-.21	0	-.06	1			
PED	.11	.02	-.22	-.36	.37	.31	.04	.07	.11	.38	.26	-.44*	-.36	.11	.47*	.04	1		
SED	.35	-.10	-.11	-.06	.38	.31	.19	.31	.37	.41	.49*	-.12	-.21	.15	.94***	.08	.73**	1	
RED	.35	.43*	.01	.52*	.11	.10	-.11	.64**	.20	-.08	.70**	-.01	.52*	.48*	.31	-.34	.02	.21	1

Note. AG = Age; YR = Years in English-speaking school; GR = Writing grade; WR = Word recognition accuracy STT; PR = Punctuation accuracy STT; CT= Count of typing during STT; TD = Time in STT; HD = Holistic score in STT; C1 = Cognitive difficulty in STT; C2 = Cognitive mental effort in STT; TLD = Text length in STT; WLD = Word length in STT; SCD = Sentence count in STT; WMD = Words per minute in STT; WED = Word errors in STT; CED = Capitalization errors in STT; PED = Punctuation errors in STT; SED = Surface errors in STT; RED = Rhetorical elements in STT.

* $p < .05$, one-tailed; ** $p < .01$, one-tailed; *** $p < .001$, one-tailed.

Table 6.

Correlation Matrix for Variables and Dictation to a Scribe Condition

	AG	YR	GR	TS	HS	C3	C4	TLS	WLS	SCS	WMS	WES	CES	PES	SES	RES
AG	1															
YR	-.03	1														
GR	-.44*	.22	1													
TS	.32	-.20	-.65**	1												
HS	.23	.21	.16	.01	1											
C3	-.13	.06	-.31	.40	-.22	1										
C4	-.28	-.26	.08	-.31	-.03	-.11	1									
TLS	.43	.11	-.01	.31	.74**	-.14	.15	1								
WLS	-.21	.29	-.26	.18	.22	.16	.35	.20	1							
SCS	.06	.47*	.29	.28	.58	.09	-.01	.75***	.31	1						
WMS	.14	.12	.50*	-.50*	.68**	-.37	.36	.59**	.10	.38	1					
WES	.36	-.26	-.29	.32	.33	-.11	.09	.66**	.08	.27	.25	1				
CES	.30	-.25	-.12	.15	.13	-.26	-.18	.04	-.37	-.15	-.12	.06	1			
PES	.64**	-.39	-.33	.26	.31	-.06	.13	.52*	-.07	.01	.30	.62**	.33	1		
SES	.47*	-.32	-.33	.33	.36	-.12	.10	.67**	.03	.21	.27	.97***	.19	.78***	1	
RES	.14	-.10	.25	-.08	.63**	-.31	.13	.62**	.02	.29	.66**	.35	.22	.59**	.45*	1

Note. AG = Age; YR = Years in English-speaking school; GR = Writing grade; TS = Time in DS; HS = Holistic score in DS; C3 = Cognitive difficulty in DS; C4 = Cognitive mental effort in DS; TLS = Text length in DS; WLS = Word length in DS; SCS = Sentence count in DS; WMS = Words per minute in DS; WES = Word errors in DS; CES = Capitalization errors in DS; PES = Punctuation errors in DS; SES = Surface errors in DS; RES = Rhetorical elements in DS.

* $p < .05$, one-tailed; ** $p < .01$, one-tailed; *** $p < .001$, one-tailed.

Table 7.

Correlation Matrix for Variables and Handwriting Condition

	AG	YR	GR	TH	HH	C5	C6	TLH	WLH	SCH	WMH	WEH	CEH	PEH	SEH	REH
AG	1															
YR	-.03	1														
GR	-.44*	.22	1													
TH	.13	-.02	-.41	1												
HH	.03	.09	.27	.06	1											
C5	-.12	-.12	.16	.21	-.25	1										
C6	-.14	.02	.10	.05	.42	.18	1									
TLH	.42	.22	.10	.26	.47*	-.25	.29	1								
WLH	.55*	.01	-.30	.15	.19	-.11	-.06	.15	1							
SCH	.27	-.07	.20	-.04	.56*	-.10	.07	.49*	.35	1						
WMH	.05	.30	.63**	-.74**	.14	-.25	-.04	.23	-.11	.28	1					
WEH	.25	.16	-.16	.43*	-.31	-.07	-.14	.56*	0	-.24	-.10	1				
CEH	.38	-.37	0	.19	.30	.02	-.04	.45*	.05	.56*	-.07	.03	1			
PEH	.11	.19	-.02	.43*	.11	.21	.21	.52*	-.24	-.06	-.24	.48*	.29	1		
SEH	.32	.09	-.15	.48*	-.21	-.03	-.11	.66**	-.02	-.11	-.14	.97**	.27	.62**	1	
REH	.34	-.03	-.01	.15	.44*	-.22	-.08	.72**	.10	.66**	.16	.31	.41	.37	.41	1

Note. AG = Age; YR = Years in English-speaking school; GR = Writing grade; TH = Time in HW; HH = Holistic score in HW; C5 = Cognitive difficulty in HW; C6 = Cognitive mental effort in HW; TLH = Text length in HW; WLH = Word length in HW; SCH = Sentence count in HW; WMH = Words per minute in HW; WEH = Word errors in HW; CEH = Capitalization errors in HW; PEH = Punctuation errors in HW; SEH = Surface errors in HW; REH = Rhetorical elements in HW.

* $p < .05$, one-tailed; ** $p < .01$, one-tailed; *** $p < .001$, one-tailed.

Chapter 4: Discussion

General

Removing transcription from the writing process and offering students the opportunity to dictate their ideas via speech-to-text technology or via a scribe has been shown to lead to better quality writing in struggling writers (De La Paz & Graham, 1997; Higgins & Raskind, 1995; MacArthur & Cavalier, 2004). This finding may be due to cognitive load being intensified during written language production when compared to oral language production, especially in children (Bourdin & Fayol, 1994). The goal of the present study was to extend research on this topic to elementary ELL students by assessing the effects of dictation to a scribe, dictation to a speech-to-text software, and handwriting on their persuasive composition and cognitive load.

Overall, results from speech-to-text recognition accuracy indicated that students had an average word recognition accuracy of 78% and an average punctuation accuracy of 98%. Three students had recognition accuracies between 50% and 60%. The researcher had to type on average two to three words during speech-to-text conditions when the software could not accurately recognize the student's dictation.

Results from analyses of variances revealed that when compared to handwriting, students in one or both dictation conditions composed texts with higher holistic text quality, reported lower cognitive load, composed texts with stronger writing mechanics, and composed arguments with more persuasive elements.

Results from correlation analyses revealed some significant relationships amongst variables. The larger correlations included word recognition positively correlating with number of years in an English-speaking school, writing grades, holistic text quality in the

speech-to-text condition, text length in the speech-to-text condition, and persuasive writing elements in the speech-to-text condition. Additionally, holistic text quality positively correlated with text length, number of persuasive elements and words per minute. All of these findings will now be interpreted and their implications relative to current research literature will be discussed.

Speech recognition accuracy and ELLs. Overall, students achieved moderate to strong recognition accuracies on Dragon NaturallySpeaking. Students reached an average of 78% word recognition accuracy and 98% punctuation recognition accuracy. Only three students had less than average recognition accuracies (between 52% and 59% accuracy). These ratings are comparable to the accuracy ratings attained with English language speakers with earlier versions of speech-to-text software (MacArthur and Cavalier, 2004). For example, MacArthur and Cavalier (2004) attained 77% to 80% recognition for sentence probes and 79% recognition for word lists during initial recognition sessions. The one confound in the present study occurred when the program would not recognize a specific word or a string of words during speech-to-text conditions. When this occurred, the researcher typed out the word or string of words. Thus, although recognition accuracy was moderate to strong, the software still needed to be monitored for inaccurate recognition.

Previous studies assessing speech recognition software with second language learners were based on earlier, less developed versions of software. These had relatively lower accuracy, particularly for recognizing accented speech when compared to native English speech (Coniam, 1999; Derwing et al., 2000). Dragon NaturallySpeaking now offers many options to personalize the user profile to increase recognition accuracy

including details on type of accented English in selected languages; however, version 11 of Dragon NaturallySpeaking did not offer Arabic as an option, so some participants in the present study did not personalize their profiles for more accurate recognition. Despite this, the software did a good job at recognizing ELL students' speech and the students were able to use the software during their speech-to-text conditions.

Holistic text quality. The first hypothesis was that students would compose texts with higher holistic quality when dictating to a scribe and when dictating to a speech-to-text software than when writing by hand. This was true of the dictation to a scribe condition, with a large effect size. Students received significantly higher holistic quality ratings when dictating to a scribe than when writing by hand. Previous research can be extended to the present findings with ELL students. MacArthur and Cavalier (2004) also found that for students with learning disabilities, the compositions with the highest quality were produced when dictating to a scribe. Overall, these findings suggest that cognitive resources were limited during writing and may have contributed to lower text quality. Thus, the removal of transcription from the writing process via dictation to a scribe invited students to allocate working memory resources to higher-level writing processes that contributed to text quality including idea generation, organization, and argumentation in the persuasive genre.

Students had marginally higher text quality when composing in the dictation to speech-to-text software condition when compared to the writing by hand condition with a large effect size, but results did not reach significance. Significance was likely not reached for this comparison because limitations in sample size reduced statistical power. Previous research has found that dictation to a speech-to-text software helps students with

learning disabilities produce better quality essays (MacArthur & Cavalier, 2004) as well as students who were previously identified as having a learning disability (Higgins & Raskind, 1995). These findings were comparable to the results found in the present study with ELL students. Overall, the trend towards significance for this effect suggests that speech-to-text could potentially enable ELL students to produce better quality texts.

Cognitive load. The second hypothesis was that students would report lower cognitive load in the dictation to a scribe and dictation to a speech-to-text conditions than in the handwriting condition. Recall that there were two measures of cognitive load: perceived difficulty and mental effort. For perceived difficulty, students reported that the handwriting condition was significantly more difficult than the dictation to a scribe condition with a large effect size. For mental effort, students reported that they put significantly more effort in the handwriting condition than in the dictation to a scribe condition with a large effect size. The present study is the first to investigate cognitive load in dictation to a scribe or dictation to a speech-to-text conditions. Previous researchers have only conducted interviews with participants to gather general opinions on dictation modalities. The results from the present study are somewhat consistent with the post-test interviews of De La Paz and Graham (1997). They found that 90% of their sample enjoyed dictating to a scribe, saying that they preferred dictating because it removed handwriting, spelling and punctuation difficulties. Thus, transcription difficulties associated with writing by hand increased cognitive load whereas this wasn't an issue when dictating to a scribe.

When dictation to speech-to-text was compared to handwriting on ratings of perceived difficulty, students rated handwriting as more difficult with a large effect size;

however significance was not reached likely due to sample size limitations. Significance was reached, however, when dictation to speech-to-text was compared to handwriting on ratings of mental effort. Students reported that they put significantly more mental effort into the handwriting condition than into the speech-to-text condition with a large effect size. In their post-test interviews, MacArthur and Cavalier (2004) also found that the majority of students thought that speech-to-text helped them compose better quality texts. Students explained that it helped them with spelling and fluency. Thus, speech-to-text likely eased load on cognitive resources during composition, contributing to lower cognitive load.

Overall, dictation via a scribe and a speech-to-text software reduced difficulty and mental effort during composition for elementary ELL students. These results are consistent with previous research that found cognitive load to be consistently higher for children when they were transcribing than when they were dictating orally (Bourdin & Fayol, 1994). The limited capacity theory of working memory could be used to interpret the present findings; the removal of transcription via dictation to a scribe and dictation to speech-to-text provided sufficient working memory and attention resources to be allocated to other writing processes like idea generation and persuasive argumentation. Thus, the findings from the present study suggest, but do not conclusively prove, that dictation improved the quality of compositions because it reduced cognitive load. That is, reduction in cognitive load may be the mechanism that allowed dictation to improve the quality of writing.

Writing mechanics. The third hypothesis was that students would compose texts with stronger writing mechanics when dictating to a scribe and speech-to-text software

than when writing by hand. As previously mentioned, writing mechanics included measures of writing fluency (total time, text length, word length, sentence count, and words per minute) and count of surface errors (word errors, capitalization errors, and punctuation errors).

Writing fluency. Total time during composition was significantly less in the dictation to a scribe condition than in the handwriting condition. Text length was also significantly longer in the dictation to a scribe condition than in the handwriting condition; thus, students composed significantly more words per minute in the dictation to a scribe condition than in the handwriting condition. There were no significant differences between the dictation to a scribe condition and the handwriting condition for word length and sentence count, but there were small to medium effect size. The present findings can be supported with previous research on writing fluency measures across modalities. MacArthur and Cavalier (2004) also found that total time was less for students composing in the dictation to a scribe condition than in the handwriting condition. With respect to text length, De La Paz & Graham (1997) also found that elementary students in the advanced planning condition who dictated produced longer essays than those who were in the comparison writing condition. Overall, dictation to a scribe enabled elementary ELL students to write longer texts in a shorter amount of time when compared to writing by hand.

There were no significant differences between the dictation to speech-to-text condition and the handwriting condition for total time, text length, and words per minute; however effect sizes were medium to large for these findings. Differences between dictation to the speech-to-text condition and the handwriting condition for word length

and sentence count were also not significant and the effects were small to null. MacArthur and Cavalier's (2004) assessment of writing fluency across speech-to-text and handwritten compositions supports the present findings. They were also not able to find significant differences with respect to text length between speech-to-text and handwriting conditions. In contrast, Quinlan (2004) was able to find that narratives composed under speech recognition were longer than handwritten ones, but only in a sub-sample of less fluent writers. One possible interpretation is that speech-to-text software was new to the students, so working memory was not reduced enough to influence writing fluency. Thus, it appears that there might be a trend towards speech-to-text software improving some measures of writing fluency for elementary ELL students, but not all.

Surface errors. There were significantly fewer surface errors in the dictation to a scribe condition than in the handwriting condition with a large effect size. This included marginally fewer word errors and significantly fewer capitalization errors with large effect sizes. There were no significant differences for punctuation errors despite there being a medium effect size. Statistically significant differences were likely not reached between dictation to a scribe and handwriting with respect to word errors and punctuation errors because of limitations in sample size. Comparisons between present findings and previous research could not be made because researchers did not investigate the presence of surface errors in texts that were dictated to a scribe.

Elementary ELL students also had significantly fewer surface errors in the dictation to speech-to-text condition when compared to the handwriting condition, including fewer word and capitalization errors, with large effect sizes. Despite there being a medium effect size, differences between speech-to-text and handwriting for

punctuation errors was not significant. Limitations in sample size likely contributed to this result as well. MacArthur and Cavalier (2004) found similar results in their sample of students with learning disabilities, who composed texts with significantly more word errors when handwriting than when dictating to speech-to-text. Quinlan (2004) also found that handwritten narratives contained significantly more surface errors than speech recognition narratives for less-fluent writers. Thus, previous research on surface errors and speech-to-text technology can be extended to the present findings with elementary ELL students.

These findings allow at least two interpretations. First, McCutchen's (1996) review of working memory capacity research could be extended to these findings, by proposing that limitations in working memory capacity during writing may have contributed to students' surface errors. Once transcription was removed, there were more working memory resources available to attend to spelling or grammar, thus reducing errors. A second possible interpretation is that the two dictation conditions simply provided students with spelling and grammatical knowledge that they did not have in long term memory.

Persuasive elements. The final hypothesis was that texts composed in dictation to a scribe and dictation to speech-to-text conditions would have more persuasive elements than texts composed in the handwriting condition. This was true of the dictation to a scribe condition, with a large effect size. De La Paz & Graham (1997) similarly found that essays composed by students in the advanced planning and dictation condition had significantly more elements than those in the comparison writing condition at post-test. The present results extend these findings to ELL students. A possible interpretation of

this finding is based on competition between writing processes for working memory resources (Chenoweth & Hayes, 2006; Hayes, 2012). It is possible that handwriting consumed working memory resources, reducing those that were available to attend to higher-level processes like idea generation and argumentation. Dictation to a scribe reduced working memory load, leaving increased resources available for generating rhetorical elements.

There were no significant differences in number of persuasive elements between texts composed under dictation to a speech-to-text software and texts that were written by hand, and computed effect sizes were small. It is possible that because speech-to-text software was new to the students, it did not reduce working memory load to a great enough extent to affect the number of rhetorical elements in text.

Correlations. There were many interesting relationships between student characteristics and measures in the present study. Groups of correlations that were medium to large, and that involved the dependent variables will be discussed in this section.

One set of correlations pointed to the nature of text quality. Across all three conditions, holistic text quality correlated most strongly with the following text features: text length in words and number of rhetorical elements. Thus, as expected, texts were perceived to be higher in quality to the extent that they were more fully developed as arguments.

A second set of correlations pointed to the processes that gave rise to quality texts. In both the speech-to-text condition and the dictation condition, holistic quality correlated strongly with words produced per minute. Interestingly, there was only a slight

non-significant correlation between words per minute and the handwriting condition. One possible interpretation is that in both dictation conditions, students could rely on a strategy of rapid production, perhaps allowing them to "dump" a clause or sentence from working memory in one "spurt." Conversely, handwriting may elicit a strategy in which students generate a phrase, draft it, and reread it, to remember "where they are" in the sentence. This may make rapid production less important.

In the speech to text condition, a set of correlations suggests a relationship between student characteristics and the effectiveness of this modality. Accuracy correlated strongly with holistic quality. In turn, accuracy was strongly predicted by the students' writing grade and years in an English speaking school. Thus, these correlations appear to tell a story in which writing grade and years of English contribute to speech-to-text accuracy, which in turn supports text quality. This suggests that future research could examine the interaction between student variables, and modality of production, in affecting the quality of student texts.

Educational Implications

Overall, elementary ELL students consistently composed better texts and reported lower cognitive load when dictating via a scribe and/or speech-to-text software than when writing by hand.

Dictating to a scribe can help ELL writers with their persuasive compositions. In the ESL classroom, ESL/ELD teachers can act as scribes for students. Acting as a scribe could ease transcription difficulties for these students and enable them to compose stronger persuasive arguments. In terms of practicality, there are not enough teachers to

act as scribes for all students during composition; thus, training students to use speech-to-text software can encourage independence from personal assistance during writing.

Current speech-to-text software can also help elementary ELL writers compose better quality texts with fewer surface errors during persuasive composition. Students can be paired up with a computer in the school computer lab and use Dragon NaturallySpeaking to assist them with their compositions. Accessibility to the software is attainable; the ESL/ELD teacher would ensure that the computers in the school are equipped with the software.

A few practical considerations for the implementation of speech-to-text technology in elementary classrooms should be addressed. Current speech-to-text software works best in quieter environments, where students can dictate in their normal voice to attain accurate recognition. As we have seen in the results, word recognition accuracy predicted many variables including holistic text quality. Ensuring participants compose in a quiet environment could potentially contribute to better word recognition accuracy. Teachers need to consider whether they have the space to offer this to their students. Another practical issue is time. An elementary ELL student would need several sessions to learn how to use the software and to learn how to train it to recognize his or her voice, as well as time to compose. In the present study, students had minimal training time on the software. This affected their independent performance during speech-to-text conditions because for every unrecognized word or string of words, the researcher had to intervene and transcribe for the student. Practically, teachers do not always have time to conduct individual training sessions on the software for each student as well as monitor their performance. One practical solution is for teachers to encourage peer assisted

training sessions, where students take turns dictating and monitoring each other's performances. Another practical solution is for teachers to train lessons on the software while teaching students the persuasive writing unit. Thus, students would be practicing their persuasive writing strategies while dictating on speech-to-text technology. Overall, introducing dictation modalities in the ESL/ELD classroom during persuasive composition would help elementary ELL students compose better quality texts with more persuasive elements, fewer surface errors, and lower cognitive load.

Limitations

There were several study limitations. First, the initial plan was to include 24 students from two schools. However, after the first school agreed to participate, the Elementary Teachers' Federation of Ontario (ETFO) held a work-to-rule campaign in which elementary teachers withdraw certain services, and the Board of Education chose not to request that schools participate in research projects. Thus the study was conducted on a small sample of elementary ELL students (n = 16). This affected the results of the analyses of variance, including several comparisons between speech-to-text and handwriting. Several of these comparisons produced medium or large effect sizes, but no statistical significance. Additionally, the small sample size potentially contributed to the violation of homogeneity of variances assumptions for several writing mechanics measures. With more participants, these variances may be more equal.

Another limitation was the lack of variety in first language and ethnic backgrounds amongst participants. Half of the participants in the present study (n = 8) spoke Arabic as a first language and the majority of students were from the Middle East

(n = 11). Thus, study findings could not be entirely generalized to all elementary ELL students.

Third, students had a minimal number of training sessions on Dragon NaturallySpeaking. The present study only included two 20-minute training sessions on Dragon NaturallySpeaking due to time and scheduling constraints. The lack of additional sessions on the software may have contributed to the poor recognition accuracy for the three students who attained 52% to 59% word recognition accuracy. It may have also contributed to unrecognized words during the speech-to-text condition, which the researcher had to type out.

Future Research

For future research, the present study should be replicated with more participants to include elementary ELL students with varying demographic characteristics. The inclusion of more participants would also increase statistical power and may balance out variances between groups on several measures to successfully meet homogeneity of variance assumptions.

Researchers using Dragon NaturallySpeaking in their study should aim to include at least three to four training sessions. Recognition accuracy increases with each of the first several uses because it gives the software an opportunity to further develop and recognize vocabulary and speech patterns.

The present study yielded strong correlations between student characteristics and dependent measures. Specifically, the results suggest that there is a floor at a certain level of English knowledge, below which speech-to-text software may not be substantially effective. The floor may be at approximately two years of experience in an English

language school. Future investigation into relationships between speech-to-text and student characteristics would further contribute to composition strategies for elementary ELL students.

Conclusion

This study was motivated by previous research showing that transcription imposes a significant cognitive load on young writers. It investigated the effects of two modes of dictation on elementary ELL students' persuasive writing and cognitive load. Students composed texts with higher holistic text quality, more persuasive elements, fewer surface errors, and higher writing fluency in one or both dictation conditions when compared to handwriting. Additionally, students reported lower cognitive loading in both dictation conditions compared to the handwriting condition. These results suggest that cognitive resources were limited for these students during handwriting and once transcription was removed from the writing process via dictation, students composed better persuasive texts. Overall, dictation to a scribe and speech-to-text software are promising tools in reducing cognitive load during text composition for some elementary ELL students.

References

Baddeley, A. (2003). Working memory and language: An overview. *Journal of Communication Disorders, 36,* 189-208.

Berninger, V. W. (1999). Coordinating transcription and text generation in working memory during composing: Automatic and constructive processes. *Learning Disability Quarterly, 22,* 99-112.

Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International Journal of Psychology, 29,* 591-620.

Bourdin, B., & Fayol, M. (2000). Is graphic activity cognitively costly? A developmental approach. *Reading and Writing, 13,* 183-196.

Chenoweth, N. A., & Hayes, J. R. (2001). Fluency in writing: Generating text in L1 and L2. *Written Communication, 18,* 80-98.

Coniam, D. (1999). Voice recognition software accuracy with second language speakers of English. *System, 27,* 49-64.

De La Paz, S. (1999). Composing via dictation and speech recognition systems: Contemporary technology for students with learning disabilities. *Learning Disability Quarterly, 22,* 173-182.

De La Paz, S., & Graham, S. (1997). Effects of dictation and advanced planning instruction on the composing of students with writing and learning problems. *Journal of Educational Psychology, 89,* 203-222.

Derwing, T. M., Munro, M. J., & Carbonaro, M. (2000). Does popular speech recognition software work with ESL speech? *TESOL Quarterly, 34,* 592-603.

Felton, M., & Kuhn, D. (2001). The development of argumentative discourse skill. *Discourse Processes, 32,* 135-153.

Ferris, D. R. (1994). Lexical and syntactic features of ESL writing by students at different levels of L2 proficiency. *TESOL Quarterly, 28,* 414-420.

Fitzgerald, J. (2008). Multilingual writing in preschool through 12th grade. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Ed.), *Handbook of Writing Research* (pp. 337-354). New York, NY: The Guildford Press.

Flower, L., & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication, 32,* 365-387.

Forgrave, K. E. (2002). Assistive technology: Empowering students with learning disabilities. *The Clearing House, 75,* 122-126.

Graham, S., & Harris, K. R. (2014). Conducting high quality writing intervention research: Twelve recommendations. *Journal of Writing Research, 6,* 89-123.

Graham, S., & Perin, D. (2007). A meta-analysis of writing instruction for adolescent students. *Journal of Educational Psychology, 99,* 445-476.

Harris, K. R., Graham, S., Mason, L. H., Friedlander, B. (2008). *Powerful writing strategies for all students*. Baltimore, MD: Paul H. Brookes Publishing Co.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication, 29,* 369-388.

Hayes, J. R., & Chenoweth, N. A. (2006). Is working memory involved in the transcribing and editing of texts? *Written Communication, 23,* 135-149.

Higgins, E. L., & Raskind, M. H. (1995). Compensatory effectiveness of speech recognition on the written composition performance of postsecondary students with learning disabilities. *Learning Disability Quarterly, 18,* 159-174.

Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review, 99,* 122-149.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Merlens, A. (2013). Working memory in written composition: An evaluation of the 1996 model. *Journal of Writing Research, 5,* 159-190.

MacArthur, C. A. (2009). Reflections on research on writing and technology for struggling writers. *Learning Disabilities Research & Practice, 24,* 93-103.

MacArthur, C. A., & Cavalier, A. R. (2004). Dictation and speech recognition technology as test accommodations. *Exceptional Children, 71,* 43-58.

McCutchen, D. (1996). A capacity theory of writing: Working memory in composition. *Educational Psychology Review, 8,* 299-325.

Nuance Communications. (2015) *Dragon NaturallySpeaking*. Retrieved from http://www.nuance.com/dragon/index.htm

Olive, T., Kellogg, R. T., & Piolat, A. (2008). Verbal, visual, and spatial working memory demands during text composition. *Applied Psycholinguistics, 29,* 669-687.

Piolat, A., Barbier, M. L., & Roussey, J. Y. (2008). Fluency and cognitive effort during first-and second-language notetaking and writing by undergraduate students. *European Psychologist, 13*, 114-125.

Quinlan, T. (2004). Speech recognition technology and students with writing difficulties: Improving fluency. *Journal of Educational Psychology*, *96*, 337-346.

Silió, M. C., & Barbetta, P. M. (2010). The effects of word prediction and text-to-speech technologies on the narrative writing skills of Hispanic students with specific learning disabilities. *Journal of Special Education Technology, 25*, 17-32.

Silva, T. (1993). Toward an understanding of the distinct nature of L2 writing: The ESL research and its implications. *TESOL Quarterly*, *27*, 657-677.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory (Explorations in the Learning Sciences, Instructional Systems and Performance Technologies)*. New York, NY: Springer.

Vanderberg, R., & Swanson, H. L. (2007). Which components of working memory are important in the writing process? *Reading and Writing, 20*, 721-752.

Zumalt, J. R. (2005). Voice recognition technology: Has it come of age? *Information Technology and Libraries, 24*, 180-185.

Appendix A. Letter of Information

Project Title: Training the Dragon: Facilitating English Language Learner (ELL) Students' Persuasive Writing through Dictation

Principal Investigators:

Nina Arcon, M.A., Faculty of Education, Western University

Perry Klein, Ph.D., Faculty of Education, Western University

Letter of Information

1. **Invitation to Participate**

   Your child is being invited to participate in a research study about persuasive writing and speech-to-text technology.

2. **Purpose of the Letter**

   The purpose of this letter is to provide you with information that you need to make a decision about whether your child may participate in this study.

3. **Purpose of this Study**

   The purpose of this study is to learn how different writing methods, such as handwriting, speaking to a person, and using speech-to-text technology, affect students' persuasive writing.

4. **Who can be in this study?**

   English Language Learners (ELLs) between the ages of 9 to 14, who have attended an English-speaking school for at least one year will be invited to participate.

5. **Who cannot be in this study?**

   Students who have a severe speech impediment will not be recruited because the computer will not be able to recognize their speech. If your son or daughter has difficulty speaking in English **AND** in his or her primary language, then please indicate this on question #7 of the 'Take-Home Demographic Questionnaire' that came with the consent forms.

6. Study Procedures

If you agree that your child may participate, he or she will be asked to complete:

(a) a take-home questionnaire to answer general information about age, gender, and primary language;

(b) two training sessions lead by the researcher on how to use the speech-to-text computer program, Dragon NaturallySpeaking. These training sessions will take place in school at the time that your son or daughter is receiving ESL services;

(c) three persuasive writing activities during school at the time that your son or daughter is receiving ESL services.

The writing activities and training sessions will take place in the ESL/ELD classroom when it is convenient for the ESL/ELD teacher and your son or daughter. Additionally, after each writing activity, your child will be asked to answer two questions about the difficulty of the activity.

If you agree that your child may participate, their ESL/ELD teacher will be asked to provide some information about your child's computer usage at school, the amount of time they have been receiving ESL support at school, whether they have an Independent Education Plan (IEP), and their most recent writing grade. If you do not agree that your child may participate, your child will not be in the study and he or she will carry on with the regular classroom activities.

7. Time Commitment

The researcher will work with the ESL/ELD teacher to create a schedule over the course of one month with dates for the two training sessions on the computer program (approximately 30 minutes each) and the three writing activities (approximately 20-30 minutes each).

8. Possible Risks and Harms

There are no known or anticipated risks or discomforts associated with participating in this study. Your child's results on the writing activities will not affect his or her report card grades.

9. Possible Benefits

This study will benefit your son or daughter by teaching them to use speech-to-text software and by having them practice persuasive writing.

10. Compensation

No financial compensation will be provided for this voluntary participation; however, your child will be offered a small gift for their participation.

11. Voluntary Participation

Participation in this study is voluntary. Your child may refuse to participate, refuse to answer any questions or withdraw from the study at any time with no effect on his or her current or future education. If your son or daughter starts the study, but is unable or unwilling to complete study procedures then his or her data will be removed from the study.

12. Confidentiality

All writing activities and questionnaires that we collect will remain confidential and accessible only to the investigators of this study. We will do our best to protect your child's information by providing your child with a 3-digit identification number upon participation. The data from our study will be stored in an electronic file that we will provide to other researchers on request, but no personal information, such as your child's name, initials or age, will be included. If you choose to withdraw your child from this study, or he or she chooses to withdraw, his or her data will be removed from our database and destroyed. Representatives of Western University's Ontario Non-Medical Research Ethics Board may contact you or require access to your child's study-related records to monitor the conduct of the research.

13. Contacts for Further Information

If you require any further information or clarification regarding this research project or your child's participation in the study you may contact Ms. Nina Arcon or Dr. Perry Klein

If you have any questions about your child's rights as a research participant or the conduct of this study, you may contact The Office of Research Ethics.

14. Publication

If the results of the study are published, your child's name will not be used. If you would like to receive a copy of the study results, please contact Ms. Nina Arcon or Dr. Perry Klein

15. Consent

Your child may participate in the study if he or she completes the attached letter of assent, and you sign the attached parental consent form.

Consent Form

Project Title: Training the Dragon: Facilitating ELL Students' Persuasive Writing through Dictation

Study Investigator's Name: Ms. Nina Arcon

Study Supervisor's Name: Dr. Perry Klein

I have read the Letter of Information, have had the nature of the study explained to me and I agree to participate. All questions have been answered to my satisfaction.

Child's Name: (if applicable)

_____

Date:

_____

Parent / Legal Guardian / Legally Authorized Representative (if applicable)

Print: _____

Parent / Legal Guardian / Legally Authorized Representative (if applicable)

Sign: _____

Date: _____

Person Obtaining Informed Consent (please print): _____

Signature: _____

Date: _____

Appendix B. Student Assent Letter

Project Title: Training the Dragon: Facilitating English Language Learner (ELL) Students' Persuasive Writing through Dictation

Principal Investigators:

Nina Arcon, M.A., Faculty of Education, Western University

Perry Klein, Ph.D., Faculty of Education, Western University

Assent Letter

1. **Why we are here.**
   Ms. Arcon wants to tell you about a study that will look at students' persuasive writing. She wants to see if you would like to be in the study.

2. **Why are they doing this study?**
   Ms. Arcon is doing this study because she wants to see if some kinds of writing activities help you with your persuasive writing more than others.

3. **What will happen to you?**
   If you want to be in the study four things will happen:
   1. You will fill out a take-home questionnaire telling me about yourself
   2. Ms. Arcon will train you on a computer program called Dragon NaturallySpeaking
   3. Ms. Arcon will ask you to complete three persuasive writing activities
   4. Ms. Arcon will ask you to complete a small questionnaire after each writing activity

   Your work on these activities will be collected and kept as a copy for my study. Your teacher will also be asked to tell us a little about yourself including your writing grade.

4. **Will there be any tests?**
   There will be no tests in this study and there will be no marks on your report card from this study.

5. **Will the study help you?**
   This study will help you practice your persuasive writing skills in English and teach you to use a computer program called Dragon NaturallySpeaking.

6. **What if you have any questions?**
   You can ask questions at any time, now or later. You can talk to the teachers, your family or the researcher.

7. **Do you have to be in the study?**
   You do not have to be in the study. No one will be mad at you if you do not want to do this. If you do not want to be in the study, just say so. Even if you say yes, you can change your mind later. If you do not want to finish all of the activities in this study, then your information will not be used. This will not affect your schooling and you will also receive a participation present. If you choose not to be in the study, the ESL/ELD teacher will give you different reading and writing activities.

I want to participate in this study.

Name of Child _____
Date_____

Signature of Child _____

Age _____

Signature of Person Obtaining Consent

_____

Appendix C. Student Demographic Questionnaire

Getting to Know You

Here are a few questions that you can answer to help us get to know you better. Please check [√] off the boxes that apply to you.

1. Are you:

☐ Male
☐ Female

2. How old are you? _____

3. Were you born in Canada?

☐ Yes
☐ No

3b. **If NO**, how many months or years have you been in an English-speaking school?

_____

4. What is your family's home country (or countries)?

_____

5. What is your first language? _____

6. What language do you speak at home? _____

7. Do you have trouble speaking in English AND in your own language?

☐ Yes
☐ No
☐ Other:_____

8a. Have you used a computer or laptop or tablet before?

☐ Yes

☐ No

8b. If **YES**, how often do you use a computer/laptop/tablet?

☐ Not that often (several times a month or less)

☐ Often (several times a week)

☐ Very often (once or more a day)

9. Do you know how to type on the computer?

☐ Yes

☐ No

10a. Have you ever used a computer program called "Dragon NaturallySpeaking" before?

☐ Yes

☐ No

10b. If **YES**, how often have you used Dragon NaturallySpeaking?

☐ Not that often (several times a month or less)

☐ Often (several times a week)

☐ Very often (once or more a day)

Appendix D. Teacher Questionnaire

Student Study ID Number: : _____

1. **Please indicate the student's most recent report card grade in Writing:**_____

2. **How long has the student been receiving in-school ESL support?**

- ☐ Less than 1 year
- ☐ 1-2 years
- ☐ 2 - 3 years
- ☐ 3 years +

3. **Does the student have an Independent Education Plan (IEP)?**

- ☐ Yes
- ☐ No
- ☐ Other:_____

4. **How often does the student use a computer/laptop/tablet at school?**

- ☐ Always (every day)
- ☐ Very often (3-4 times a week)
- ☐ Sometimes (1-2 times a week)
- ☐ Rarely (a few times a month)
- ☐ Never (student does not use a computer at school)
- ☐ Other:_____

Appendix E. Certificate of Participation

Great Participation!

Thank you

for your participation in this

University of Western Ontario

research study.

Appendix F. Arabic Letter of Information

طلب الموافقة الاهلية

عنوان المشروع: تدريب التنين : تسهيل الكتابة المقنعة لمتعلمي اللغة الانجليزية
(ELL) من خلال الإملاء

المحققون الرئيسيون :

نينا أركون ، ماجستير ، كلية التربية ، جامعة وسترن

بيري كلاين ، دكتوراه، كلية التربية ، جامعة وسترن

رسالة الإعلام

1. دعوة للمشاركة

هذه الدعوة موجهة لطفلك للمشاركة في دراسة بحثية حول الكتابة المقنعة persuasive writing وتكنولوجيا تحويل الخطاب إلى نص.

2. الغرض من الرسالة

الغرض من هذه الرسالة هو تقديم المعلومات التي تحتاجها لاتخاذ قرار بشأن ما إذا كان طفلك يستطيع المشاركة في هذه الدراسة.

3. الغرض من هذه الدراسة

الغرض من هذه الدراسة هو معرفة الى اي درجة يمكن لاساليب الكتابة المختلفة، مثل
الكتابة اليدوية ، التحدث إلى شخص ، او استخدام تكنولوجيا تحويل الحديث إلى نص ، ان تؤثر على اسلوب الكتابة المقنعة لدى الطالب.

4. من الذي يمكن أن يكون في هذه الدراسة؟

ثلاثون من متعلمي اللغة الإنجليزية ELL من إعداد مدرسة ابتدائية واحدة،الذين تتراوح أعمارهم بين 9 إلى 14 ، والذين داوموا في مدرسة ناطقة باللغة الانكليزية لمدة عام واحد على الأقل.

5. الذين لا يمكن ان يشاركوا في هذه الدراسة؟

لن يتم تجنيد الطلاب الذين لديهم عائق خطابي شديد لأن الكمبيوتر لن يستطيع التعرف على خطابهم. لو ابنك أو ابنتك لديها صعوبة في التحدث باللغة الانكليزية او لغته الأساسية ، الرجاء الاشارة لذلك في السؤال رقم 7 من الاستبيان الديموغرافي الذي جاء مع استمارات الموافقة، واذي يؤخذ مع الطالب الى المنزل.

6. إجراءات الدراسة

إذا كنت توافق على مشاركة طفلك ، سوف يطلب هنه او منها إكمال:
(أ) استبيان للرد على معلومات عامة حول طفلك مثل العمر، والجنس ، و اللغة الأساسية .

(ب) دورتين تدريبيتين بقيادة الباحث عن كيفية استخدام برنامج تحويل الخطاب إلى النص - التنين NaturallySpeaking . وهذه الدورات التدريبية تجري في المدرسة في الوقت الذي يتلقى فيه ابنك أو ابنتك خدمات ال ESL .

(ت) ثلاثة أنشطة تتعلق بالكتابة المقنعة اثناء المدرسة في الوقت الذي يتلقى ابنك او ابنتك خدمات اللغة الإنجليزية كلغة ثانية ESL .

سوف تتم أنشطة الكتابة والدورات التدريبية في فصل ال ESL/ELD عندما يكون الأمر مريحا للمعلم و ابنك أو ابنتك . بالإضافة إلى ذلك ، وبعد كل نشاط للالكتابة ، سيطلب من طفلك الاجابة على سؤالين حول صعوبة هذا النشاط.

إذا كنت توافق على مشاركة طفلك، سوف يطلب من معلم ال ESL/ELD توفير بعض المعلومات حول استخدام طفلك للكمبيوتر في المدرسة ، والمدة التي تلقى ابنك دعم ال ESL في المدرسة ، وما اذا كان لديك خطة التعليم المستقل (IEP) ، و أحدث درجة لهم في الكتابة. إذا كنت لا توافق على مشاركة طفلك، فإن طفلك لن يكون في الدراسة و سوف يتم استمراره او هي في الأنشطة الصفية العادية.

7. الالتزام بالوقت

سيعمل الباحث مع معلم ال ESL/ELD لانشاء جدول على مدة شهر واحد يحتوي على مواعيد الدورتين التدريبيتين على الكمبيوتر(30 دقيقة تقريبا لكل منهما) ، وأنشطة الكتابة الثلاث (حوالي 20-30 دقيقة لكل منهما) .

8. الاضرار والمخاطر المحتملة

لا توجد مخاطر أو مضايقات متعلقة بالمشاركة في هذه الدراسة. نتائج طفلك في انشطة الكتابة لن تؤثر على درجاته في بطاقة التقرير.

9. الفوائد المحتملة

هذه الدراسة ستفيد ابنك أو ابنتك عن طريق تعليمهم استخدام برمجيات تحويل الكلام إلى نص ،و مساعدتهم على ممارسة الكتابة المقنعة.

10. تعويضات

لن يتم تقديم أي تعويض مالي عن هذه المشاركة الطوعية؛ ولكن سيتم تقديم هدية صغيرة لطفلك على مشاركتهم ، مثل مقلمة.

11. المشاركة الطوعية

المشاركة في هذه الدراسة طوعية. باستطاعة طفلك رفض المشاركة ، رفض الإجابة عن أي أسئلة أو الانسحاب من الدراسة في أي وقت بدون التأثير على تعليمه الحالي أو في المستقبل . إذا بدأ طفلك في الدراسة ، واصبح غير قادر أو غير راغب في استكمال إجراءات الدراسة ، سيتم إزالته من الدراسة.

12. خصوصية

جميع الأنشطة الكتابية والاستبيانات التي نجمعها ستبقى سرية، و لا يمكن الوصول إليها إلا بواسطة المحققين في هذه الدراسة. وسوف نبذل قصارى جهدنا

لحماية معلومات طفلك من خلال توفير طفلك ب 3 ارقام مميزة عند المشاركة, PIN .
وسيتم تخزين بيانات الدراسة في ملف إلكتروني سيتوفر لباحثين آخرين حسب الطلب، ولكن لن تحتوي على معلومات شخصية لطفلك، مثل الاسم او الحروف الأولية من اسمه أو العمر.
إذا اخترت سحب طفلك من هذه الدراسة، أو اختار هو ان ينسحب، سيتم إزالة البيانات الخاصة به وتدميرها.
ممثلي أونتاريو للجنة البحووث الاخلاقية الغير طبية لجامعة وستيرن قد يتصلوا بك أو يطلبوا الوصول إلى السجلات المتعلقة بالدراسة لمراقبة سير الأبحاث.

13. اتصالات لمزيد من المعلومات

إذا كنت بحاجة إلى مزيد من المعلومات أو التوضيحات بخصوص هذا البحث أو مشاركة طفلك في الدراسة، بامكانك الاتصال بالسيدة نينا أركون عل، أو ، او الدكتور بيري كلاين،

إذا كان لديك أي أسئلة حول حقوق الطفل الخاصة كمشارك فالبحث أو في إجراء هذه الدراسة ، بامكانك الاتصال بمكتب أخلاقيات البحث (Office of Research Ethics)

او بالبريد الإلكتروني:

14. النشر

إذا تم نشر نتائج هذه الدراسة ، فلن تحتوي على اسم طفلك . وإذا كنت ترغب في الحصول على نسخة من نتائج الدراسة المحتملة ، يرجى الاتصال بالسيدة نينا أركون على ، أو ، او الدكتور بيري كلاين،

15. الموافقة

يمكن لطفلك ان يشارك في الدراسة باستكمال رسالة الموافقة المرفقة وتوقيع الاهل على استمارة موافقة الوالدين المرفقة.

هذه الرسالة تحفظ للرجوع إليها في المستقبل .

نموذج الموافقة

عنوان المشروع : تدريب التنين : تسهيل الكتابة المقنعة لمتعلمي اللغة الإنكليزية (ELL students) من خلال الإملاء.

اسم الباحث في الدراسة : السيدة نينا أركون

اسم المشرف على الدراسة : الدكتور كلاين بيري

لقد قرأت رسالة المعلومات ، و قد شرحت لي طبيعة الدراسة ، وأنا أوافق على المشاركة . ولقد تم الرد على جميع اسئلتي بوضوح .

اسم الطفل : _____

التاريخ : _____

الوالد / الوصي القانوني / الممثل المفوض قانونا (إن وجد):

طباعة: _____

الوالد / الوصي القانوني / الممثل المفوض قانونا (إن وجد):

توقيع: _____

التاريخ : _____

Appendix G. Training Lesson One

Lesson Topic: Introduction to Dragon NaturallySpeaking

SPECIFIC EXPECTATIONS:
By the end of the lesson the student will be able to:
Successfully train their voice to Dragon NaturallySpeaking

MATERIALS:
Laptop equipped with Dragon NaturallySpeaking
Microphone headset

PROCEDURE:
Introduction: ~5 mins
Introduce myself and explain to students that they will be learning how to use Dragon NaturallySpeaking. Next, provide students with a description of the software: *Dragon NaturallySpeaking is a computer program that types out your ideas, but it can't read your mind. You have to talk into the microphone very clearly so that Dragon NaturallySpeaking can type out what you say. Today, we will train Dragon NaturallySpeaking to know your voice by doing 3 steps: (1) answer a few questions about yourself for the computer program so that it knows a little bit about you, (2) read aloud a few times to make sure the microphone can hear you, (3) read a short story to fully train the computer to know your voice.*

Lesson: ~20 mins
The investigator will open up Dragon NaturallySpeaking. Next, the investigator will complete the following steps for the student:

(a) Launch Dragon > **Profile Creation** (Or, choose **New Profile** in the Dragon Bar Profile menu)

(b) Answer the questions on the subsequent screens, including the profile/user name, age, language, region of origin and accent.

(c) Indicate the microphone type and verify the sound system used. Choose **Mic-In-Jack speech device**.

(d) Review the subsequent screen and ensure that all of the selections for the user profile are accurate. Click **Create Profile.**

(e) **Check Microphone**: In this step, Dragon will adjust the volume to better understand the student's voice. The microphone's listening side must face the corner of the student's mouth (not the front) about an inch away. It must not touch the student's hair or catch breathing sounds. Explain this step to students: *Dragon NaturallySpeaking needs to listen to you read aloud with a clear voice. Please speak into the microphone as if you are talking to a friend. I will press **start volume check** once your headphones are on and you are ready to read. When you hear a beep, it means Dragon does not need you to read*

anymore. I can whisper read the text with you, if you prefer.

(f) Click **Start Volume Check for the Quality Check.** This stage is similar to the one previously completed. Explain to the student: *Just as before, please read aloud until you hear a beep. If you do not hear the beep, start at the beginning and read again.*

(g) Investigator will click **Next** when audio quality check indicates: PASSED or ACCEPTABLE. Explain to student: *Now, you will need to read to Dragon for a little longer. This activity may take you up to 10 minutes, so take your time.*

(h) Choose **Show Text with Prompting** to highlight the words as the student is reading.

(i) Press **Go** to have the student read the two sentence prompts that appear on screen: "*Welcome to general training. Training is about to begin."*

(j) Once complete, click **Next** to get to the Read Training Text screen. This screen indicates that students are about to read an extended text in order to finish off the training.

(k) Click **Select Text** and choose a text from Reading for Children or Easier Reading: Instructional and click **OK**.

(l) Click **Next** to start reading. Explain to the student: *When I click Next you will start reading. The words will turn gray once the computer has heard them. Make sure you speak normally. If the computer needs to hear you read something again, a yellow arrow will show you what to read. Remember to speak clearly.*

(m) Click **OK** on the popup screen: "Congratulations! You have finished training…" Explain to students: *Great job! Dragon NaturallySpeaking now knows your voice*!
Note: Saving the user profile may take several minutes.

(n) On the "Let Dragon Search for Words" screen, the investigator will uncheck the options **Search through Emails** and **My Documents** then click **Next.**

(n) Ensure that "Automatically Improve Accuracy" is checked off then click **Next.**

(o) On the "Help Us Improve" screen, investigator will check off the box that says: "**Don't run data collection"**

Considerations:
If the student is having difficulty reading the text provided by Dragon NaturallySpeaking, then the researcher will turn off the student's microphone and rehearse the text with them or whisper read the text to them.

Appendix H. Training Lesson Two

Lesson Topic: Dragon NaturallySpeaking Practice and Accuracy

SPECIFIC EXPECTATIONS:
By the end of the lesson the student will be able to:
1. Successfully train their voice to Dragon NaturallySpeaking if they have not done so already in the first training session.
2. Practice dictating a few sentences into Dragon NaturallySpeaking using the Training the Dragon tip sheet for guidance.
3. Complete the Accuracy of Dragon NaturallySpeaking activity.

MATERIALS:
Training the Dragon tip sheet
Laptop equipped with Dragon NaturallySpeaking
Microphone headset

PROCEDURE:
Introduction to the activity: ~5 mins
Explain to students that they will be learning how to practice dictating into Dragon NaturallySpeaking: *Today we will practice speaking to Dragon NaturallySpeaking. First, we will practice saying a few sentences into it - focusing on reading clearly and adding in punctuation. Then, you will read a short little story into Dragon NaturallySpeaking. I will also give you a tip sheet that you can use to help you when you are working on Dragon NaturallySpeaking.*

LESSON: ~20 mins
If the student has successfully trained their voice to the program from the initial training session, then open up their profile on Dragon NaturallySpeaking. Click **Launch Dragon** > **Open profile**. Additionally, open up Microsoft Word. If the student has not successfully trained their voice to the program, then open up the student's profile using the same settings and complete it starting from the step they left off at.

(a) Once the student has successfully trained their voice to the program, hand out the Training the Dragon tip sheet. The researcher will then explain to the student: *This tip sheet will help you use Dragon NaturallySpeaking. Let's read it together.*

i) Read through the tip sheet with the student and explain all the commands as noted on the tip sheet.
ii) Once the student has familiarized themselves with the tip sheet, the researcher will proceed with the remainder of the training.

(b) Proceed to practice activity. Explain to students: *Next, we are going to practice speaking to Dragon. You will say the following three sentences into your microphone word-for-word. Don't forget to say the punctuation marks. We can practice these*

sentences together. I can read the sentences before we begin to show you how to say them in a clear voice. Then, you can try saying them before we start.

i) The investigator will then read the following three sentences aloud, including the punctuation marks: *"I saw a dragon today. It had big green wings. It looked a little scary, but it was very friendly."*

ii) Next, the investigator will have the student practice the three sentences. Once the student is ready, the investigator will turn on the student's microphone and have the Microsoft Word page opened and ready for dictation.

c) After the student has successfully dictated the three sentences, they will be asked to complete an activity that measures the accuracy of Dragon NaturallySpeaking (see Appendix J). Explain to the student: *Great job reading those sentences. Now you will read a short story into Dragon NaturallySpeaking. Just as before, you will read the sentences into your microphone. We can practice these new sentences together. I will read the sentences before we begin to show you how to say them in a clear voice. Then, you can try practicing before we start. Don't forget to say the punctuation marks.*

d) The investigator will then read the activity aloud and have the student practice the text aloud afterwards (without turning on the microphone): *One day, a fish was swimming around the pond when it saw bread in the water. It swam to the bread and bit it. The fish did not know this was a trap. Just before it was pulled onto a fisherman's boat, it let go of the bread. The fish swam happily ever after.*

e) After the student has rehearsed the story aloud and is familiar with it, he or she will dictate it to Dragon NaturallySpeaking. Explain to the student: *Great job reading! Now you will read the story to Dragon. Make sure you read it word-for-word, just as before. If Dragon does not type out the story correctly that is okay. Just skip the mistakes and keep reading.*

f) The investigator will then turn **on** the microphone and tell the student to begin. When the student has finished dictating, the word document will be saved under the student's three digit identification number on the encrypted hard drive. This document will be later assessed for the percentage of accurate recognition.

f) Let the student know that they have successfully completed the second training session: *Great job! You have finished practicing with Dragon NaturallySpeaking. The next time we meet you will be working on persuasive writing activities. See you then!*

Appendix I. Training the Dragon Tip Sheet

Dragon NaturallySpeaking Tip Sheet

Here are some tips to help you with Dragon NaturallySpeaking:

- **Before you start speaking, make sure you click on your word document. This will show Dragon where to type your ideas.**

Say these commands to Dragon:

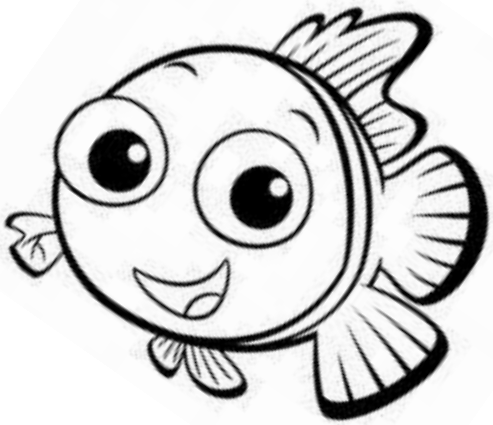
Key Word(s)	What it means to Dragon
"Microphone Off" **Press the + key to turn your microphone on.**	This will turn off your microphone. *Make sure you always say this when you are finished talking to Dragon.*
"Erase That" **OR** "Scratch That"	This will erase the last thing that you said.
"New Paragraph"	This will start a new paragraph.

- "period" **OR** "full stop" ! "exclamation mark"

Appendix J. Text for Testing Speech Recognition Accuracy

Instructions: Please read this story into your microphone. Make sure you speak clearly as you read. If the computer makes a mistake on a word, skip it. Now, let's hear your read!

One day, a fish was swimming around the pond when it saw bread in the water. It swam to the bread and bit it. The fish did not know this was a trap. Just before it was pulled onto a fisherman's boat, it let go of the bread. The fish swam happily ever after.



Appendix K. Topic One Instructions

<u>Question:</u> Do you think students should have more time for recess? Why?

In this activity, you will argue your opinion to this question. Use the TREE strategy to help you.

T - Topic Sentence

- Tell what you believe

R - Reasons

- Why do you believe this? Give 3 reasons

E - Explain

- Say more about each reason

E - Ending

- Finish up and write an ending sentence

Appendix L. Topic Two Instructions

<u>Question:</u> What is the best subject in school? Why?

In this activity, you will argue your opinion to this question. Use the TREE strategy to help you.

T - Topic Sentence

- Tell what you believe

R - Reasons

- Why do you believe this? Give 3 reasons

E - Explain

- Say more about each reason

E - Ending

- Finish up and write an ending sentence

Appendix M. Topic Three Instructions

<u>Question:</u> Imagine you can choose to be five years older. Would you want to be five years older? Why?

In this activity, you will argue your opinion to this question. Use the TREE strategy to help you.

T - Topic Sentence

- Tell what you believe

R - Reasons

- Why do you believe this? Give 3 reasons

E - Explain

- Say more about each reason

E - Ending

- Finish up and write an ending sentence

Appendix N. Cognitive Load Survey

How easy or difficult was this writing activity? Please circle a number.

1	2	3	4	5	6	7	8	9
Very very easy		Easy		Medium		Difficult		Very very difficult

How much effort did you put into this writing activity? Please circle a number.

1	2	3	4	5	6	7	8	9
Very very little effort		Little effort		Medium effort		Much effort		Very very much effort

Appendix O. Holistic Text Quality Criteria for Raters

Please rate the overall quality of texts as a pieces of persuasive writing by following these steps:

1. **Rater 1:** Browse through the texts.

2. **Rater 1:** Read the texts again, and choose seven papers to represent each point of a 7-point rating scale ranging from (1) very low quality to (7) very high quality , with (4) being average quality. Base your selection on ideas, content, organization and overall persuasiveness. Please ignore surface errors.

3. Once **Rater 1** has successfully chosen the seven anchor papers, both raters will then sort all of the compositions into seven piles using the anchor papers as indexes of the reflective pile.

4. Read through each pile again to verify that all the texts are similar to their chosen anchor/index text.

4. On a separate piece of paper, please make a list of ratings by indicating the following:

a) the ID number found in the top left-hand corner (i.e., "410R");

b) text quality rating

Appendix P. Surface Errors Criteria for Raters

Surface Errors: Please mark up the text by identifying the following surface errors and then record their presence in the text. Please treat the error types as a hierarchy, that is, code ambiguous errors as being word errors first, or capitalization errors if that is not applicable, or punctuation errors if either of those are not applicable. Any one word can be categorized as only one type of error. Lastly, please use the error (at each point in the text) as the unit of count.

Type of Error	Definitions	Example and Count of Errors	Number of Errors
Word Error	1. Spelling Error (a string of letters that is not a word) 2. Homophones (words that sound the same, but are spelled differently, i.e., accept/except, no/know, through/threw) 3. Semantic errors (meaning of the word is related to intended word, but not appropriate i.e., bigger vs. older) 4. Missing words (key words are missing from the phrase) 5. Double words or unnecessary words (same word repeated twice or unnecessary word added) 6. Pronoun error (unclear pronoun reference) 7. Verb/subject disagreement (verbs and subjects do not agree) 8. Apostrophe Use (misuse of apostrophes in contraction words and possessive nouns)	1. He dose his homework at skool untill he finnishes it. **= 4 errors** 2. The principle was two funny. **= 2 errors** 3. When I am bigger, I will be better at math. **= 1 error** 4. Math is the subject. **= 1 error;** code at how many points there are missing words 5. Students have a lot of **of** homework to do. **= 1 error** 6. The pencil broke, so I fixed them. **= 1 error** 7. I like school because I can take books out of the library by himself. **= 1 error** 8. I missed school today because my moms car wasnt starting. **= 2 errors**	
Capitalization Error	1. Missing capitalization (the first letter of the word following an appropriate end of sentence punctuation was not capitalized **OR** the first letter of a word of a new sentence **OR** a proper noun is missing capitalization **OR** first-person, "I," and its contractions [I'm, I've, I'll] were not capitalized) 2. Incorrect capitalization (student capitalized the first letter of a word that was not a proper noun, was not first-person, "I," or its contractions, and was not following an end of sentence punctuation)	1. i like school because the teachers are nice. they help me with my homework. **= 2 missing capitalization errors** 2. Today at School I learned to Read. **= 2 incorrect capitalization errors**	
Punctuation Error	1. Missing punctuation (appropriate punctuation mark was not placed) 2. Incorrect punctuation (the appropriate punctuation mark was not used **OR** the student inserted a punctuation mark that does not suit the sentence)	1. I enjoy school because it will help me with my future another reason I like school is because it is fun = **2 missing punctuation errors** 2. During gym. we played indoor soccer and, hockey? **= 3 incorrect punctuation errors**	
TOTAL			

Appendix Q. Persuasive Elements Criteria for Raters

Please indicate whether the following persuasive elements were present within the text. Record the number of elements.

Persuasive Elements	Present (√)	Number	Definitions
Topic Sentence	☐	_____	The writer's proposition of their argument: "Children need to go to school." In other words, the student's claim of their opinion.
Reasons	☐	_____	Evidence that the writer presents to support their claim(s): "School is a lot of fun"
Explanation of Reasons	☐	_____	Explanation of reasons ie. using examples: "School is especially fun when we go on field trips." These may appear later on in the text.
Ending/ Conclusion	☐	_____	Another statement of the writer's opinion, this time at the end of the text: "These are the reasons why I believe children should have to go to school"
Other	☐	_____ _____ _____	Student included other persuasive elements such as: a) **Alternative Claim** An opposing argument to the writer's claim: "I know that some kids might think that school is boring." **b) Reasons for Alternative Claim** Reasons for the other claim: "Students think school is boring because there is a lot of homework" **c) Rebuttal to Counter Argument** Writer's refutation of the counter argument: "School is not boring because you can always make friends"
		Total number of elements: _____	

Curriculum Vitae

Name:	Nina Arcon
Post-secondary Education and Degrees:	Ryerson University Toronto, Ontario, Canada 2009 - 2013 Arts and Contemporary Studies Major in English; Minor in Psychology
Honours and Awards:	Dean's List of Honours 2013
Related Word Experience:	Graduate Research Assistant The University of Western Ontario 2013 - 2015 Learning Success Strategist Ryerson University 2012 - 2013 Research Assistant at the Brain and Early Experiences Lab Ryerson University 2012 - 2013

Publications:

Klein, P., Arcon, N., & Baker, S. (Accepted). Writing to Learn. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research*. The Guilford Press.

Klein, P., Arcon, N., & Haug, K. (Submitted). The Effects of Rhetorical and Content Subgoals on Text, Cognitive Load and Learning about Classification. *Learning and Instruction*.

Sugden, N., Arcon, N., Wei, M., & Moulson, M. (Submitted). These are a few of my favourite things: Infants' preference for familiar and unfamiliar faces, hands, and bottles. *Infancy*.

Presentations:

Arcon, N. (2014). *Training the Dragon: Facilitating English Language Learner (ELL) Students' Persuasive Writing through Dictation*. Graduate Research in Education Symposium (GRiES), University of Western Ontario, London, Ontario.

Sugden, N., Arcon, N., Wei, M., & Moulson, M. (2013). *These are a few of my favourite things: Infants' preference for familiar and unfamiliar faces, hands, and bottles*. Canadian Psychological Association (CPA) Annual Convention, Quebec City, Quebec.