Impact of Multimodal Feedback and Formulaic Sequences on Improving Fluency of English Learners on Computer-Based Speaking Assessments

by

By Jennifer François

A Dissertation Submitted in Partial Fulfillment of the Requirements for the Doctor of Philosophy Degree in Literacy Studies

Middle Tennessee State University

June 2021

Dissertation Committee:

Dr. Mohammed Albakry, Chair

Dr. Jwa Kim

Dr. Stacy Fields

ABSTRACT

While recent research has shown that multimodal feedback and use of formulaic sequences (FS) are effective in improving student performance on writing tasks, there is a dearth of studies on the impact of these aspects on computer-based academic speaking assessments. This dissertation seeks to fill this gap by examining the impact of both multimodal feedback and formulaic sequences (FS) on improving performance and fluency of adolescent English Learners (ELs) on computer-based speaking assessments. Students in this interactive sequential mixed methods study were randomly assigned to a group receiving asynchronous feedback or a group receiving synchronous video feedback. Both groups were evaluated using the WIDA Speaking Interpretive Rubric that analyzes speech holistically at the word/phrase level, sentence level, and discourse level. Students then engaged in a second speaking task and were evaluated using the same rubric. Results indicated that both groups showed significant overall improvements: asynchronous feedback (n = 12, Task One Mdn = 3.00, Task Two Mdn = 4.50, Z = 3.07, $p < .001$, r = .89) and synchronous feedback (n = 12, Task One Mdn = 3.00, Task Two Mdn = 3.83, Z = 2.59, p = .008, r = .75), with the asynchronous feedback group out-performing the synchronous feedback group. Furthermore, regression analysis indicated that formulaic sequences significantly predicted speech fluency ($\beta = 1.01$, $t(46) = 9.65$, $p < .001$). Formulaic sequences also explained a significant proportion of variance in fluency scores ($R^2 = .67$, $F(1, 46) = 93.20$, $p < .001$). Results from this study can inform and optimize remote and face-to-face (F2F) instruction in academic speaking and the implications include not only potentially improving EL students' skills on standardized measures of academic speaking performance, but also enhancing their linguistic skills in general education classes and improving their college and career readiness.

Keywords: remote learning, asynchronous learning, synchronous learning, academic feedback, SLA, CALL, speaking assessment, formulaic sequences, academic speech fluency

TABLE OF CONTENTS

LIST OF FIGURES……vi

LIST OF TABLES……vii

LIST OF ABBREVIATIONS……vii

CHAPTER I: INTRODUCTION……..1

Statement of the Problem……1

Instructional Feedback and Types of Multimodal Computer-Based Feedback…….6

Formulaic Sequences (FS) and Fluency……..7

Fluency……9

Connection to the Statement of the Problem……10

Purpose of the Study……11

Research Questions……11

Significance of the Study……12

CHAPTER II: LITERATURE REVIEW……..13

Feedback Stage: Systematic Search of Literature……13

Production Stage: Systematic Search of Literature ……20

Study Purpose Based on Literature……28

CHAPTER III: METHODOLOGY……..29

Research Design……29

Participants……30

Experiment Delivery Technology……31

Study Procedure: Stage One……32

Study Procedure: Stage Two……38

Study Procedure: Stage Three……39

CHAPTER IV: RESULTS AND DISCUSSION……...……44

Results: Research Question One……..……44

Results: Research Question Two…….…….…….47

Discussion……..51

CHAPTER V: CONCLUSION…….63

Instructional Implications……...……..…..64

Limitations and Directions for Future Study……..……....68

Final Thoughts……..…….71

REFERENCES…….…….72

APPENDICES……...……87

APPENDIX A: CODING TABLE FEEDBACK STUDIES….…….……..88

APPENDIX B: CODING TABLE FS AND FLUENCY STUDIES…….90

APPENDIX C: MTSU IRB APPROVAL……..……...92

APPENDIX D: WIDA SPEAKING INTERPRETIVE RUBRIC……...…..….99

APPENDIX E: STUDENT-FRIENDLY RUBRIC…….……100

APPENDIX F: SPEAKING TASK TRANSCRIPTS……101

LIST OF FIGURES

Figure 1: Definition of Formulaic Sequences

Figure 2: Discourse Function of Standard FS

Figure 3: Location of FS Within Runs of Fluent Speech

Figure 4: Location of Single FS Within Sentences

Figure 5: Sentence Tree

LIST OF TABLES

Table 1: Sample Script for Teacher Prompting and Virtual Administration on a Speaking Assessment Task

Table 2: Change in Tennessee WIDA ACCESS 2.0 Domain Scores

Table 3: Student Speech Samples with Scores and Feedback Excerpts

Table 4: Feedback Growth Between Task One and Task Two on Composite Scores

Table 5: Growth by WIDA Rubric Level

Table 6: MLFR and FS Growth

LIST OF ABBREVIATIONS

BICS	Basic Interpersonal Communication Skills
CALP	Cognitive Academic Language Proficiency
COBUILD	Collins Birmingham University International Language Database
COCA	Corpus of Contemporary American English
COVID-19	Corona Virus Disease
dB	Decibels
DOE	Department of Education
EL	English Learner
EFL	English as a Foreign Language
ELL	English Language Learner
ERP	Event-Related Potential
ESL	English as a Second Language
FERPA	Family Educational Rights and Privacy Act
FS	Formulaic Sequences
F2F	Face-to-Face
IELTS	International English Language Testing System
IRB	Internal Review Board
L1	First Language
L2	Second Language
MICASE	Michigan Corpus of Academic Spoken English
MLFR	Mean Length of Fluent Run

MS	Model Sequence
SLA	Second Language Acquisition
TBOCLLA	Technology-based out-of-class language learning activities
TN	Tennessee

CHAPTER I: INTRODUCTION

Statement of the Problem

Of the four domains of acquiring English as an additional language (listening, reading, writing, and speaking), writing and speaking, being the domains that involve active production of language, are often considered more difficult to master (Cummins, 1979; Ellis et al., 2008; Krashen, 1982; Wood, 2002; Woolbert, 1922). While writing encompasses three processes, including thought, language, and typography; speaking, on the other hand, involves four processes including thought, language, voice, and action (Woolbert, 1922). Ellis et al. (2008) explain the difference in complexity between speaking and writing by pointing out that "Speech is constructed in real time and this imposes greater working memory demands compared with writing" (376). Furthermore, because speech is spontaneous communication, it requires the speaker to have a flexible selection of language for output instantaneously as opposed to writing where the writer has time to reflect and restructure output (Wood, 2002). In light of these challenges pertaining to the speaking domain, English Learners (ELs) commonly go through a silent period (Krashen, 1982), where they are receiving language input but are not yet ready to express themselves orally. Nonetheless, assessing ELs' speaking skills is a required part of the overall language proficiency assessment process, and federal regulations annually require state education agencies to test ELs' performance in all the skills, including in the speaking domain, regardless of speaking proficiency level (TN Board of Education, 2003/2020; TN DOE, 2018).

Speaking Domain Proficiency, Instruction, and Assessment

Over the years, the speaking domain is one that has been often neglected in EL instruction (Cummins, 1981). Second language speaking skills in social situations or, Basic

Interpersonal Communication Skills (BICS) (Cummins, 1979), are acquired through communication with peers in social situations and usually develop within one to two years; academic speaking, or Cognitive Academic Language Proficiency (CALP), on the other hand, is much more difficult and time-consuming to master, taking up five to seven years before academic proficiency is acquired (Cummins, 1979). During the 1980s and 1990s, this distinction was not yet recognized on a wide scale; and students achieving proficiency in BICS were regularly exited from ESL programs into mainstream classrooms, where they then had challenges functioning academically alongside their native English-speaking peers (Cummins, 2008). Academic language as Cummins (2008) notes is largely found in academic texts, and the focus of EL instruction had traditionally been on improving reading skills to access the language necessary for processing these texts. However, to improve EL student academic speaking skills, instructional focus needs to not only dwell on reading, but also producing academic speech using the academic language acquired through the reading of academic texts (Biber, 1986; Corson, 1997; Cummins, 2001; Gee, 1990; Vincent, 1996). While scholars have investigated academic speech in second language acquisition (e.g., Biber, 2006; Ellis & Bogart, 2007; Kormos, 2014; Krashen, 1982; Simpson-Vlatch & Ellis, 2010), educational methodologies, including instruction, assessment, and feedback have not been consistently investigated, especially considering the evolution of testing and instructional technologies.

Early iterations of speaking domain assessments for ELs (ACCESS, ELDA, IPT, CELLA, LAS Links) all contained a face-to- face (F2F) individual, test administrator-scored assessment of the speaking domain. Teachers received extensive training to score the speaking assessment in real time using holistic rubrics, and testing scripts provided guidelines for acceptable responses and approved prompting (verbal cues given to the students to help them

produce the targeted speech output). Teachers sat one-on-one with test subjects and administered the assessment as a type of academic dialogue about content area topics and scored the speech performance according to the holistic rubrics.

Advances in technology, however, changed not only the way the speaking assessments are administered and scored, but also the way in which the students participate in the assessment. Moving to a computer-based, virtual test administrator that delivers the stimulus materials with no direct student interaction and requiring students to record their responses electronically fundamentally changed the testing environment for students and teachers alike by removing the F2F aspect. Furthermore, the responsibility of scoring of student responses was removed from the live test administrator and transferred to a remote scorer. These changes often left teachers and students in a vacuum struggling to adjust to the new testing format. While computer-based speaking assessments attempt to simulate academic discourse, the simulated discourse environment can be challenging for the ELs to replicate in the short time given to respond to the speaking tasks, often under one minute. Moreover, the change in format from an academic dialog to an academic monologue has posed challenges for students, as these two forms of discourse require different skills and techniques. In addition, the switch to remote delivery and scoring has removed the teacher from the testing situation, leaving the students on their own with their computers instead of engaging in an academic conversation as in previous iterations of the speaking assessment. Consequently, performing successfully in the new environment involves acquiring a new skill set for students and teachers, not only in adapting to the new testing environment, but also in how academic speech instruction is implemented in the EL classroom. See Table 1 below for examples of teacher prompting for a F2F speaking assessment versus a virtual test administrator-delivered question in a computer-based speaking assessment task.

Table 1

Sample Script for Teacher Prompting and Virtual Administration on a Speaking Assessment Task[1]

Example of Teacher-Administered Speaking Task Prompting	**Example of Virtual Test Administrator Administered Speaking Task**
Teacher: Look at these two pictures of where a family goes on vacation (*points to pictures in the testing book of a family at the beach and in the mountains*). Tell me, do you like the mountains or the beach, and tell me two reasons why you think that way. **Student:** the beach (*Since the student did not produce two reasons, the teacher can prompt the student for more information.*) **Teacher:** (*prompts*). Can you tell me why you like the beach? **Student:** The beach is nice. (*Since the student still has not provided two reasons, the teacher can continue prompting.*) **Teacher:** (*prompts*). Tell me, what is nice about the beach? **Student:** It has sand, and the water is fun. (*Since the student has produced the expected output, the teacher moves on to the next question.)*	**Virtual Test Administrator**: Look at the two pictures on your screen. They show two places a family might go on vacation. Where would you like to go on vacation: to the beach or to the mountains? Be sure to tell me two reasons why you think this way. **Student**: (*presses the record button and speaks into the microphone*) the beach (*student presses the stop button and the screen advances to the next question*).

[1] These questions have been created for exemplification and do not reflect content of actual testing.

This change in format to computer-based standardized testing of speaking performance, coupled with 2017 changes in scoring dynamics implemented by the World-Class Instructional Design and Assessment or WIDA Consortium[2], has caused a decrease in score performance on the Speaking Domain of the WIDA ACCESS 2.0[3] assessment compared to previous years. Data from the Tennessee Department Education illustrate this dramatic drop in scores on the speaking domain portion of this test that is required by the State of Tennessee and the federal government to be administered to all ELs in public schools each year, as shown in Table 2.

Table 2

Change in Tennessee WIDA ACCESS 2.0 Domain Scores

Assessment	Listening	Speaking	Reading	Writing
ACCESS 2016	*M* = 4.69	*M* = 4.13	*M* = 4.12	*M* = 3.38
ACCESS 2017	*M* = 4.49	*M* = 2.93	*M* = 3.34	*M* = 3.14

Note. Data from the Tennessee Department of Education (2019). SD has been requested but not received.

As a public school teacher working with students involved in computer-based academic speaking assessments, the writer of this dissertation has witnessed these testing innovations and their accompanying challenges firsthand. I have experienced the development of speaking

[2] WIDA Consortium is the testing company utilized by 40 states in the US as their federally mandated English Language Proficiency Assessment.
[3] WIDA ACCESS 2.0 is the English Language Proficiency Assessment provided by the WIDA Consortium.

assessments through the evolution of the testing products implemented in public schools in the state of Tennessee for the past 14 years and witnessed how students and teachers have responded and reacted to these changes. The drop in 2017 scores referenced here has been the impetus for me to conduct classroom research and study academic speech instruction and computer-based testing to help find methods of assisting teachers and students to adapt to the developments in speaking assessment conditions and expectations.

Instructional Feedback and Types of Multimodal Computer-Based Feedback

For the first purpose the study, feedback is defined as: "information provided by an agent . . . regarding aspects of one's performance or understanding" (Hattie & Timperley, 2007, p. 81). The term *multimodal feedback* is used to mean multiple modes of feedback being presented to the student at the same time (Ali, 2016; Alvira, 2016; M. Cunningham, 2015; K. Cunningham, 2018, 2019; Elola & Oskoz, 2016; Ghosn-Chelala & Al-Chibani, 2018; Maas, 2017; Özkul & Ortaçtepe, 2017; Silva, 2012).

Multimodal feedback can include both asynchronous and synchronous types of feedback. *Asynchronous feedback* refers to feedback materials that the student can view at a time of his or her choosing (Ali, 2016; Aljaser, 2019). Some examples of asynchronous feedback include static web pages, email communication, storage media, or pre-recorded videos, etc. *Synchronous feedback*, on the other hand, means materials that are provided to the student at a specified date and time and that are viewed by the student at the same time the materials are being shared (Ali, 2016; Aljaser, 2019). Live video meetings using any one of several technologies available such as Zoom, Google Meet, Skype, Facetime, etc. are examples of synchronous feedback, also including discussion forums or chat rooms.

Typically, feedback on speech in Second Language Acquisition (SLA) is provided immediately at the time of speech and is interactional, where speaker and teacher engage with each other during the speech process (Fan, 2019; Long, 1983; Lyster & Ranta, 1997; Mackey, 2006; Swain & Lapkin 1998). However, for the purposes of this study and the nature of computer-based speaking assessments, immediate feedback is not only not allowed (test administrators are not allowed to provide any assistance or feedback while students are participating in the speaking assessment), it is also often impossible because the virtual test administrator within the testing platform is not interactive. That means that different forms of feedback need to be utilized during instruction to help prepare students authentically for participation in these types of computer-based speaking tasks. Additionally, the virtual nature of the test begs for a similar environment for instruction and feedback to provide the most authentic instructional setting as possible.

One alternative is to provide instructional feedback on speaking tasks after the speech has been delivered, either live or pre-recorded. Of course, in the classroom, this feedback can still be provided using older methods of written comments or oral feedback. However, the same advent of technology that has moved speaking assessments to computer-based platforms has also provided computer-based alternatives for providing feedback utilizing the advantages that access to multiple modes provides teachers and students.

Formulaic Sequences (FS) and Fluency

The second purpose of this study is to investigate the connection between FS and fluency in the speaking performance of students participating in the study. As previously discussed, the characteristics of academic speech that set it apart from social and conversational speech are important to investigate to determine their impact on improvement in fluency. Improvements in

the utilization of academic discourse patterns typical of academic speech could support students in lightening the cognitive processing load so they could have more capacity to devote towards improving content and stylistic elements of their responses (Biber & Conrad, 1999; Boers et al., 2006; Ellis, 1996; Ellis, et al., 2008; Eyckmans et al., 2015; Gray & Biber, 2013; McGuire & Larson-Hall, 2017; Mohammadi & Enayati, 2018; Nattinger & DeCarrico, 1992; Pawley & Snyder, 1983; Rafieyan, 2018; Wood, 2002, 2006; Yilmaz & Korban Koc, 2020). Additionally, the onset of computer-assisted corpus examination and study has led to increasing recognition of the effect FS have on oral language production (Biber et al., 2004; Boers, et al., 2006).

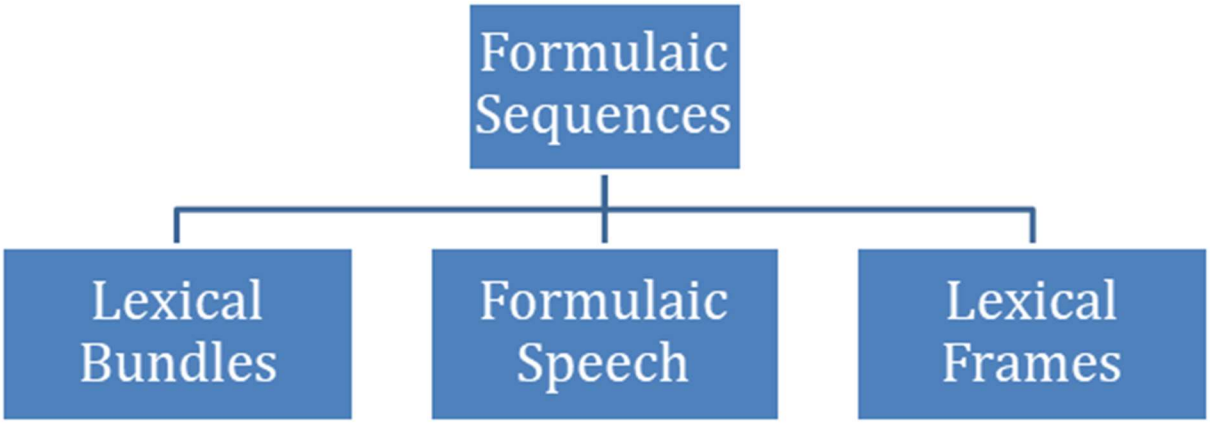
One aspect of fluency that this study seeks to investigate is the relationship between use of FS and improvements in fluency. A clear definition of what constitutes FS is a topic that many authors have been refining and debating since the early 1980s. For example, one of the first attempts to clearly demonstrate the nature of formulaic language is the COBUILD project (Collins Birmingham University International Language Database) at the University of Birmingham in England which developed a corpus of contemporary text from which many dictionaries have been published (Wood, 2002). The COBUILD project rendered word co-occurrence in the English language visible to a wide audience and paved the way for researchers such as Biber and Conrad, 1999; Boers, et al., 2006; Ellis, 1996; Ellis et al., 2008; Nattinger and DeCarrico, 1992; Pawley and Syder, 1983; Wood, 2002, 2006; and Wray, 2002 to expand the study of the role of FS in sequencing and structuring language output.

The study adopts a wide classification of FS (François & Albakry, 2021) as supported by Wray's (2002) definition as "a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language

grammar" (p. 9). An additional factor to consider about defining FS is that if they are defined as being stored holistically and retrieved as prefabricated chunks, it goes without saying that not all speakers will store or retrieve them identically (Biber, et al., 2004; Boers, et al., 2006; Wray, 2002). For the purposes of this study then, FS include lexical bundles such as *the author of* (Biber & Conrad, 1999), formulaic speech such as *you know* (Nattinger & DeCarrico, 1992), and lexical frames similar to *because X and Y* (Wood, 2006), including complex sentence stems like *there are X differences between Y and S* (Pawley & Syder, 1983). See Figure 1.

Figure 1

Definition of Formulaic Sequences



Fluency

Fluency has been studied at length by several researchers and has been shown to have multiple facets, all of which combine to give a picture of what fluent speech looks like. There are subjective (e.g., holistic rubrics or rating scales) and objective (e.g., quantitative measures of speech produced) measures of fluency, and which measure is used depends on the purpose and

goals of the study (N. De Jong & Perfetti, 2011; N. H. De Jong & Mora, 2019; N. H. De Jong & Schoonen, 2013; Ginther, et al., 2010; Levelt et al., 1999; Prefontaine & Kormos, 2016; Rossiter, 2009; Towel, et al., 1996). Temporal measures of fluency, however, are frequently used to evaluate speech performance objectively due to the ability to precisely measure and quantify the measures in speech samples (De Jong & Schoonen, 2013; Ginther et al., 2010; Prefontaine & Kormos, 2016). The quantitative or objective fluency measures utilized most frequently in evaluating speech production include:

1. Quantity of production: the time spent speaking or the number of units produced, measured by counting words, syllables, or morphemes
2. Rate of production: units produced per second or minute
3. Disfluencies in production: length and type of pauses, repetitions, or use of non-fluent units (Ginther et al., 2010)

Mean length of fluent run (MLFR) is one measure that takes each of these factors into consideration. MLFR gauges language density by only measuring units of *fluent* speech, e.g., excluding silent pauses; repetitions; lexical, but non-meaningful fillers such as *eh* or *uh*; and self-corrections. Using MLFR as the measure of fluency for this study is supported by previous studies (De Jong & Perfetti, 2011; Pawley & Syder, 1983; Towel et al., 1996; Ushigusa, 2009). According to this previous research, MLFR is identified as the best measure for FS because MLFR gives a precise measure of fluent language produced in one unit of uninterrupted speech, highlighting the focus on FS and their relationship to fluency. Therefore, for the purposes of this study, MLFR has been chosen as the most accurate, objective measure of fluency.

Connection to Statement of Problem

As mentioned previously, the challenges posed by the shift in computer-based speaking assessments from conversational dialogue about academic topics (similar to BICS) to students

independently producing short, impromptu academic monologues about previously unprepared content area topics (similar to CALP) can be better understood when considering the differences between acquiring BICS and CALP for ELs. The change in format from dialogue to planning and delivering an academic monologue caused deficits across multiple aspects of speech production, including choosing the right register for a speech in contrast to a dialogue, to mastering the fluency required for maintaining an academic monologue independently without a conversation partner.

Purpose of the Study

The purpose of this dissertation is twofold: To examine the impact of both multimodal feedback and formulaic sequences (FS) on improving performance and fluency of adolescent ELs on computer-based speaking assessments. The participants were recruited from the English as a Second Language (ESL) program in high schools (Grades 9-12) of a suburban school system in middle Tennessee.

The study comprises two main sections: feedback and speech production. In the feedback section, the effectiveness of multimodal feedback methods was tested in a randomized, between groups experiment. In the production section, the speech produced in the experiment was analyzed to determine if the linguistic features of FS contributed to noticeable improvements in performance demonstrated by the subjects.

Research Questions

The research questions which this study has addressed are:

1. Which mode of feedback, asynchronous or synchronous, elicits improved performance on computer-based speaking tasks?

Emerging from results from research question one, the following question was subsequently

examined:

2. Does the use of FS predict fluency in academic speaking tasks?

Significance of the Study

The data from this study can contribute to the body of research that investigates the effectiveness of different types of feedback on computer-based speaking tasks, in particular differences between synchronous and asynchronous feedback. The results can provide instructional guidance not only for remote learning situations, but also for brick-and-mortar classrooms. With increasing mandates and reliance on technology tools to structure and implement instruction, educational leaders and classroom practitioners can benefit from the study results by learning about effective and reliable instructional methods. Furthermore, investigating the connection between FS use and fluency can guide future instruction in academic speaking for ELs. Data regarding the discourse functions of FS used in academic speaking can inform speaking domain instruction, as well as provide guidance for the types and structures of FS that demonstrate the greatest impact on fluency. Finally, providing educators not only the data demonstrating the impact multimodal feedback and use of FS have on academic speaking fluency, but also providing them instructional practices for implementation in F2F, hybrid, or remote learning situations will offer practitioners the tools they need to prepare their students for facing computer-based academic speaking tasks, as well as aid in preparing them for academic presentations in content area classes.

CHAPTER II: LITERATURE REVIEW

While there is no dearth of research on feedback in reading and writing, one area of research that has not been adequately explored is the impact that feedback and FS have on SLA speech production in academic speaking tasks. Research that not only covers modes of feedback, but also the effectiveness of different modes of feedback delivered remotely is relevant and, more than ever, necessary to design effective instructional methods in the ever-changing pandemic and post-pandemic educational environment, especially considering the technological advances that make computer-based instruction easy, accessible, and efficient. Additionally, examining the relationship between FS and fluency and the impact this has on improving speaking task performance is necessary to further guide speaking domain instruction for practitioners.

Feedback Stage: Systematic Search of Literature

To ensure that an exhaustive search of the available research on feedback for ELs on computer-based speaking tasks was conducted, a thorough search of 539 databases including Eric, Scopus, and PsychINFO just to name a few, was completed focusing on peer-reviewed articles using the search terms *speaking assessment* and *ESL* or *ELL* or *English as a second language* or *English language learners* and *feedback*. Articles concerned with speech acquisition or speech disabilities as physiological or cognitive phenomena were excluded for being outside the scope of the proposed study. Additional exclusionary criteria included studies concerned with teacher education, peer feedback, and gaming as a form of instruction or feedback to name a few.

After every attempt was made to locate previous research on the effectiveness of multimodal feedback on academic speaking tasks, as mentioned above, only eight closely related studies to the topic of the dissertation study at hand were found: Ali, 2016; Aljaser, 2019; Alvira,

2016; Elola and Oskoz, 2016; Faramarzi et al., 2019; Ghosn-Chelala and Al-Chibani, 2018; Honarzad and Rassaei, 2019; and Özkul and Ortaçtepe, 2017 (see summary in Appendix A).

Several of the studies found investigated the impact of multimodal instruction on performance tasks of foreign language learners at the university and elementary levels, (Ali, 2016; Alvira, 2016; Cunningham, 2017, 2018, 2019; Faramarzi et al., 2019; Ghosn-Chelala & Al-Chibani, 2018; Honarzad & Rassaei, 2019; Hung, 2016; Maas, 2017; Özkul & Ortaçtepe, 2017; Silva, 2012). Two studies (e.g., Hung, 2016; Maas, 2017) were eliminated because they either focused solely on peer feedback using Facebook or gave students a choice about the feedback they received. Studies utilizing forms of social media platforms as instructional tools are not necessarily valid comparisons for public school settings due to FERPA (Family Educational Rights and Privacy Act) laws and the fact that most social media platforms are unable to guarantee the safety of personally identifiable data required for using these platforms, making them unusable in public school settings. Furthermore, student choice is usually more restricted in public school settings, limiting the usefulness and validity of studies which allow students to control the form or content of feedback used by instructors.

A further study by Silva (2012) was excluded due to the focus of the topic being multimodal feedback on the drafting/revising process in writing being too narrow to offer relevance to the topic of this study. Finally, three studies by Cunningham (2017, 2018, 2019) were excluded as a result of their topics focusing on the difference in language structures used when giving multimodal or written feedback, which is, additionally, too narrow to be of benefit for this study. The majority of the other studies found were either too broad or too narrow and were not included for this reason.

Feedback Stage: Multimodal Instruction

Some of the eight studies, including those by Aljaser (2019), Faramarzi et al. (2019), and Honarzad and Rassaei (2019) investigated a variety of multimodal instructional applications. In a quasi-experimental (without random assignment to groups) study, Aljaser (2019) investigated EFL primary students in Saudi Arabia to determine the effectiveness of e-learning as compared to traditional classroom instruction and found significant differences. This study was conducted with two classes of fifth grade students divided into a control group, who received traditional instruction in the classroom, and an experimental group, who received instruction in an e-learning environment via the internet. An achievement test was used to demonstrate growth, and the results indicated an effect size of $\eta^2 = 0.39$ for the e-learning group over the traditional group. This study is particularly relevant because it is one of the few studies to examine differences in achievement of school-aged children based on mode of instruction. Furthermore, it provides evidence that multimodal instruction is effective on younger learners and provides a basis supporting the investigation of secondary students conducted in this dissertation study.

Additionally, in his study, Aljaser (2019) included a learning attitude scale to gauge learner perceptions using a 33-item questionnaire that was administered to the control group (n = 15) and the experimental group (n = 15). A Wilcoxon Rank Sum Test revealed that students in the experimental group had significantly more positive attitudes towards learning English than the students in the control group: control group M = 11.27; experimental group M = 19.73; z = -2.24; p = .008. Similarly, Honarzad and Rassaei (2019) investigated the attitudes of learners in using technology-based out-of-class language learning activities (TBOCLLA) with EFL graduate students at a university in Iran with L1 Persian speakers (N = 100). This study involved exploratory research and collected data via four researcher-developed questionnaires to gauge

learner attitudes towards multimodal learning tasks involving learner motivation, autonomy, and self-efficacy. In all three areas, the TBOCLLAs were significantly positively correlated with attitudes about learning English as a foreign language through technology: motivation $r(98) = .67$, $p < .001$; autonomy $r(98) = .64$, $p < .001$; self-efficacy $r(98) = .59$, $p < .001$.

In Faramarzi et al.'s (2019) pre-experimental study, student attitudes toward technology-based listening instruction were also investigated. Data from EFL students ($N = 120$) at an Iranian university were collected through observational and attitudinal tools in the form of a researcher-designed Learner Engagement Questionnaire (LEQ). Students participating in the English course were presented authentic listening tasks over 12 weeks of instruction using Vodcasting (video podcasting) and subsequently completed the LEQ, containing 32 items and measured using a 6-point scale (with scores of 4 or above indicating positive attitudes as reported by the author). On the questions measuring student attitudes towards the effectiveness of multimodal tasks improving their listening skills in English ($M = 5.04$), results indicated that the overall mean of questions addressing this attitude was higher than the central point and signified that the overall attitude was positive.

In all three of these studies (Aljaser, 2019; Honarzad & Rassaei, 2019; Faramarzi et al., 2019), data demonstrated that students at both elementary and university levels had positive attitudes towards multimodal instructional methods and that elementary students demonstrated more growth in achievement when using multimodal methods over traditional classroom methods (Aljaser, 2019). These studies support the continued investigation of the effectiveness of multimodal instruction and demonstrate the need for additional studies to validate the results with objective measures.

Feedback Stage: Multimodal Feedback

In addition to studies that investigated student attitudes towards multimodal instruction (Aljaser, 2019; Faramarzi et al., 2019; Honarzad & Rassaei, 2019), five earlier studies examined the effectiveness of multimodal feedback on EFL writing tasks at the university level (Ali, 2016; Alvira, 2016; Elola & Oskoz, 2016; Ghosn-Chelala & Al-Chibani, 2018; Özkul & Ortaçtepe, 2017).

Ali (2016) conducted a mixed methods study examining the effectiveness of multimodal screencast feedback in improving writing. In this study, university students were randomly assigned to a control group receiving only written feedback and an experimental group receiving screencast feedback for higher order skills, like content, organization, and structure, and written feedback for lower order skills such as accuracy. Students in the experimental group receiving written and screencast feedback out-performed the control group receiving only written feedback in improving overall writing skill, content, structure, and organization $t(61) = 8.46$, $p < .010$. In addition to objective measures of growth between feedback modes, Ali (2016) also gauged student attitudes towards multimodal feedback by administering an 18-item questionnaire with results indicating that 94% of students participating ($N = 63$) had a positive attitude towards the multimodal feedback.

Similar to Ali (2016), Alvira (2016) completed a qualitative action research with Colombian university EFL students, examining differences between written and multimodal feedback on writing tasks, and results here also indicated that students receiving the multimodal feedback had a higher final average on their writing tasks than students receiving written feedback (Pre-test $M = 2.62$, Post-test $M = 3.70$). Additionally, Alvira (2016) conducted a

student survey to gauge student attitudes towards multimodal feedback and data showed that 80% of respondents felt that multimodal feedback helped them improve their writing.

Similar but varying results were found in a case study conducted by Elola and Oskoz (2016) at a US university with students of Spanish writing ($N = 4$). In this small case study, the authors examined not only differences between student performance based on different modes of feedback, but also differences in the content and quantity of feedback provided by instructor based on feedback mode. Results indicated that the instructor provided lengthier and more detailed feedback using the multimodal method over the written method, but that students incorporated feedback in subsequent writing tasks equally for both types of feedback. An additional finding indicated that while students preferred multimodal feedback for corrections regarding content, structure, and organization; they preferred written feedback for corrections regarding form (e.g., grammatical structures).

A further experimental study by Özkul and Ortaçtepe (2017) examined differences in university student writing between how written and video feedback groups incorporate feedback into subsequent writing attempts to improve performance. Descriptive statistics indicated that the video feedback group incorporated more feedback in their second drafts than the control group, and significant differences were found across the feedback modes in three out of five tasks, with the control group receiving written feedback and the experimental group receiving video feedback. Mann-Whitney U tests showed that the experimental group outperformed the control group: Assignment 2: Treatment ($n = 14$, $M = 18.04$) and Control ($n = 15$, $M = 12.17$), $U = 62.5$, $p = .030$; Assignment 3: Treatment ($n = 11$, $M = 15.15$) and Control ($n = 14$, $M = 7.9$), $U = 24$, $p = .004$; Assignment 4: Treatment ($n = 13$, $M = 17$) and Control ($n = 10$, $M = 9.86$), $U = 33$, $p = .007$ (no r values were reported). These results differ from the results achieved by Elola and

Oskoz (2016) where there was no difference in student performance; however, the sample size in Elola and Oskoz was much smaller and could have accounted for the differing results. Consistent with the other studies, Özkul and Ortaçtepe (2017) also included a student perception survey as part of their study. Results from this indicated that 100% of respondents ($N = 23$) preferred receiving multimodal feedback on future assignments because they felt it provided more information than traditional written feedback.

Finally, a case study by Ghosn-Chelala and Al-Chibani (2018) investigated the differences in student attitudes between written and multimodal feedback on Lebanese Arabic L1 university EFL student writing ($N = 8$). In line with the other studies, the results indicated an overall positive attitude of students toward multimodal feedback over written feedback. Data was collected via a researcher-designed 9-item survey and an informal group discussion was also conducted by the researcher. Thematic analysis indicated that students perceived multimodal feedback as more helpful and preferred to receive it in the future. These findings are consistent with the other studies referenced here, supporting student preference for multimodal feedback.

In summary, the results of these studies involving multimodal instruction and feedback demonstrate that multimodal feedback is significantly effective in improving student performance on writing tasks (Ali, 2016; Alvira, 2016; Elola & Oskoz, 2016; Özkul & Ortaçtepe, 2017). While two studies involved adult learners (Ali, 2016; Özkul & Ortaçtepe, 2017), one study did examine school-aged EFL children and the effect multimodal feedback had on their writing performance (Aljaser, 2019). The questions that remain to be examined, however, are whether multimodal feedback is equally as effective on improving academic speech of learners, especially at the secondary school level. Furthermore, while these studies investigated whether

students preferred written to multimodal feedback, studies involving differences among different modes of multimodal feedback are lacking.

The second stage of this dissertation examined the speech production of the students in the study to investigate if the use of FS could predict fluency as measured by MLFR. Studies available on the subject rarely investigated the use of FS in academic speech; however, studies on the use of FS in academic writing were more prolific. Additionally, there was paucity in studies examining the use of FS by secondary school students. Finally, while most, if not all, studies examined the use of FS by EFL students, very few, if any, were found that investigated the use of FS by ESL students.

Production Stage: Systematic Search of Literature

After the feedback stage of the experiment was conducted and the patterns and themes of FS and fluency emerged from the qualitative analysis of the speech samples, a second exhaustive search of the literature was completed. To conduct a search of the available research on FS and fluency in academic speech for adolescent ELs, a thorough search of the same databases including Eric, Scopus, and PsychINFO, and others., was completed focusing on peer-reviewed articles using the search terms *fluency* and *formulaic sequences* and *speaking skills*. The search criteria yielded only two results (Khodadady & Shamsaee, 2012; Tavakoli, 2011), both of which were applicable to this study. However, to locate additional studies, the search criteria were modified to include the following search terms in varying combinations: *lexical phrases, lexical frames, lexical bundles, formulaic speech, ESL* or *ELL* or *English as a second language* or *English language learners*. The number of results reached 253 articles matching the search criteria. Additional qualifying factors were used to target applicable studies: studies involving only listening, reading, or writing were excluded to ensure that the focus remained on studies

investigating speaking due to the previously discussed differences between the domains of reading, writing, listening, and speech production. Further criteria used to delimit articles included: articles with only adults (as non-academic scholars) since this study concerns adolescents; and articles about idioms since this study investigates academic speech, and according to Simpson and Mendis (2003), "idioms are neither rare nor particularly frequent in academic speech" (p. 427). Finally, articles exclusively about instructional methods or teacher training were also excluded since this study seeks to establish connections between FS and fluency and not necessarily effective instruction.

Additional studies met the initial search criteria but were later eliminated due to either being too broad or too narrow in scope. For example, Puimège and Peters' (2020) study was eliminated because it covered speech from television and was too broad to apply to academic speech and, therefore, outside the scope of this study. Further studies that were eliminated for being too narrow in scope include Alraddadi's (2016) study of discourse markers, Kashiwagi and Ito's (2017) study about grammatical awareness and morphological structures, and Römer's (2019) study about verb constructions found in FS.

While the majority of the studies found involved university EFL students, these were included in the research analysis due to being closely related to secondary school learners and comprised seven studies (Boers et al., 2006; Khodadady & Shamsaee, 2012; McGuire & Larson-Hall, 2017; Rafieyan 2018; Tavakoli, 2011; Yan, 2020; Yilmaz & Korban Koc, 2020). Two additional studies focus on the use of FS in the speech of adolescent English learners, similar to those in the study at hand (Eyckmans, et al., 2015; Mohammadi & Enyati, 2018) (see summary Appendix B).

Four overarching themes emerged from the qualitative review of these applicable studies: Type of instruction/intervention utilized in the study, type of speech investigated, type of source used to identify and classify FS, and type of speech feature analyzed.

Theme One: Type of Instruction/Intervention

The first theme emerging from the nine studies is the type of intervention or instruction focus utilized in the studies. In six of the studies, the authors chose to focus on the difference between recognizing FS as one lexical unit as opposed to examining individual word units (Boers et al., 2006; Eyckmans, et al., 2015; McGuire & Larson-Hall, 2017; Mohammadi & Enyati, 2018; Rafieyan, 2018; Yilmaz & Koban Koc, 2020). In Boers et al. (2006), the authors completed an experimental, exploratory study in which university EFL L1 Dutch speakers were divided into a control group (n = 15), who received stimulus materials in which the grammar-lexis dichotomy was implemented and the experimental group (n = 17), who received materials in which standardized word combinations in the stimulus materials were explicitly taught. Results demonstrated that the experimental group (M = 14.44) outperformed the control group (M = 13.31), U = 70, $p < .050$, on the basis of an evaluation carried out by native English speakers conducting oral proficiency interviews (no r values were reported). These results support the hypothesis that encoding FS as single units help facilitate perceived fluency when utilized in speech.

In Eyckmans et al. (2015), the authors conducted a quasi-experimental study (no randomized groups) that presented L1 Dutch EFL learners (N = 65) in a Belgian secondary school a list of 32 target FS phrases to memorize. They then divided the participants into three groups: the first group received no instructions, the second group received instructions to notice alliterative patterns in the target phrases, and the third group was instructed to notice any

incongruencies between the L1 and English translations of the target phrases. Results demonstrated that alliterative over incongruencies and alliterative over no treatment were both significant (Alliterative over incongruent: X^2 (14, n = 22) = 21.82, p < .001, d = 0.32; Alliterative over no intervention X^2 (14, n = 22) = 7.02, p = .008, d = 0.17), whereas incongruencies between the two languages had little effect on the outcomes. This demonstrated that directing student attention towards salient features may contribute to the ability to learn and process FS more efficiently.

An additional study that focused on the type of instruction or intervention is McGuire and Larson-Hall's (2017) case study with experimental and control groups about the effects of formulaic sequence use on spoken fluency. In their study, EFL students studying in a U.S. university (N = 19), with L1s including Chinese, Thai, and Japanese, were exposed to listening and speaking exercises. The control group (n = 8) received instruction that focused on single words and grammar, while the treatment group was explicitly taught to notice and use FS during instruction. Measures that were taken included speech rate, mean length of run, a subjective rating, and a ratio of FS to speech produced. Results indicated that the treatment group significantly outperformed the control group in all measures except the subjective rating. The explanation the authors give for this is that they suspect that giving the native speaker judges seven criteria to evaluate may have provided too much variance to get valid results. Speech Rate: Treatment: t(10) = -5.3, p = .003, d = 1.3; Control: t(7) = 0.20, p = .090, d = 0.06. Mean Length of Run: Treatment: t(10) = -3.5, p = .006, d = 1.1; Control: t(7) = 1.04, p = .340, d = 0.17. Subjective: Treatment: t(10) = -1.41, p = .190, d = 0.26; Control: t(7) = 1.62, p = .150, d = 0.71. FS Ratio: Treatment: t(10) = -1.41, p = .009, d = 1.2; Control: t(7) = 1.62, p = .800, d = 0.20.

A further experimental study by Mohammadi and Enayati (2018) conducted with L1 Persian EFL secondary school students in Iran ($N = 60$), examined the effect of lexical chunk instruction on students' speaking fluency. The control group ($n = 30$) received instruction using text passages and speaking tasks from a standardized textbook which utilized business as usual instruction focusing on grammar and translation, whereas the treatment group received explicit instruction on lexical chunks, including noticing and utilization during speaking and writing. Results demonstrated that the treatment group outperformed the control group (fluency treatment: $t(29) = 1.97$, $p < .001$; fluency control: $t(29) = 0.58$, $p = .566$). These results confirm the results from previous studies discussed here.

Additionally, a 2018 study by Rafieyan investigated the relationship between FS and language proficiency with EFL students at a Japanese university with L1 Japanese ($N = 42$). This study also focused on an experiment that divided the intervention between groups focusing on FS within context and focusing on FS in isolation. While this study showed a significant positive correlation between proficiency level and performance ($p < .001$, $\eta^2 = 0.82$) indicating that knowledge of FS increased with proficiency, it did not show a significant relationship between the treatment and control groups for the intervention. These results indicated that proficiency level of the participants was positively correlated with increased use of FS in the study; however, it did not show that the method of instruction showed significant differences in the results, e.g., learners in both groups made gains as a result of the intervention. These results differ from previously discussed results which all indicate a significant difference in results based on type of intervention (focus on FS or business as usual focus on single word or grammar instruction). This can be explained by the fact that students in both groups received an intervention that focused on FS, the only difference being how the focus was targeted. Therefore, the results are

still consistent with previous studies indicating that explicit instruction in FS leads to significant improvement in use of FS.

The final study in the first theme involving type of instruction/intervention is Yilmaz and Koban Koc's (2020) quasi-experimental study (with no randomized groups) involving EFL university L1 Turkish learners ($N = 35$) which investigated pragmatic comprehension and production of FS. The experimental group ($n = 19$) received a corpus-based instruction focusing on 19 FS, while the control group received traditional instruction focused on grammar. A Wilcoxon Rank Sum Test indicated that students receiving the experimental treatment showed significant gains over the control group (experimental group pre-test and immediate post-test: $z = 136.00$, $p < .001$; control group $z = 14.00$, $p = 1.00$). The delayed post-test, however, showed that the results, while still significant in the experimental group, showed less of an impact over time (post-test and delayed post-test experimental group: $z = 18.00$, $p = .012$, drop in mean from 2.74 to 2.05). These results demonstrated that gains achieved from an intervention faded over time when the subjects did not continue to practice learned skills.

Theme Two: Type of Speech

Seven of the relevant studies distinguish between the type of speech (academic monologue, academic dialog, or conversational dialog) under investigation (Boers et al., 2006; Khodadady & Shamsaee, 2012; McGuire & Larson-Hall, 2017; Mohammadi & Enayati, 2018; Rafieyan 2018; Tavakoli, 2011; Yilmaz & Korban Koc, 2020). Some studies focused on the type of speech that is represented by conversation, either about academic or conversational topics, (Boers et al., 2006; McGuire & Larson-Hall, 2017; Mohammadi & Enayati, 2018; Rafieyan 2018; Yilmaz & Korban Koc, 2020) and other studies focused on academic speech (Khodadady & Shamsaee, 2012) and monologic speech (Tavakoli, 2011), both similar in some ways to the

speech necessary for production on academic speaking assessments in this study. In her study, Tavakoli (2011) investigated speech between L1 English speakers (n = 40) and L2 learners of English from different language backgrounds (n = 40) by having participants perform monologic narrative tasks based on picture-story narration. This is the only study that investigated monologic speaking tasks, although they required participants to perform narrative discourse as opposed to academic discourse. Furthermore, Khodadady and Shamsaee (2012) in their study of EFL Persian L1 students in Iran (N = 41) focused on academic speech production by incorporating the IELTS Speaking Exam (International English Language Testing System) as a measure of fluency during the study. All of the other studies referenced in Theme Two (Boers et al., 2006; McGuire & Larson-Hall, 2017; Mohammadi & Enayati, 2018; Rafieyan 2018; Yilmaz & Korban Koc, 2020) focused on dialog, either conversational or academic in nature. While these are not the same type of speech as the computer-based speaking assessments under investigation, the speech produced in these studies is still important to consider and relevant to the discussion about academic speech production in educational settings as it encompasses all of these facets.

Theme Three: Type of Source for FS

Another theme that carried across the studies is the source authors used to provide the FS used in their studies. Three authors chose a corpus-based approach (McGuire & Larson-Hall, 2017; Yan, 2020; Yilmaz & Korban Koc, 2020), and all chose the COCA (Corpus of Contemporary American English) as their source. While the COCA is known for its vast collection, it does not limit its sources to academic publications; in fact, it not only includes audio and video examples, but also contributions from multiple print sources. While this does provide a broad base of authentic English language materials to select FS from, it neither focuses

on spoken English, nor does it focus on *academic* spoken English. Two further authors (Mohammadi & Enayati, 2018; Rafieyan, 2018) utilized academic English language textbooks as sources for authentic materials. While these sources provided academic language materials, neither specialized in spoken academic English. These studies are relevant, however, because they demonstrate the wide variety of sources where FS are utilized in both spoken and written English, and in conversational, social, and academic discourses across content areas and entertainment modes.

Theme Four: Type of Unit Investigated

The fourth theme concerns the different units of speech that were under investigation in the studies. Two studies (Tavakoli, 2011; Yan, 2020) focused their investigations on pauses in speech. In Tavakoli (2011), the study compared at what point in speaking an utterance did L1 speakers and L2 speakers of English pause while speaking. The results indicated that more than L1 speakers of English, L2 speakers paused more mid-clause $t(39) = 2.25$, $p = .030$. Additionally, the results indicated that pauses that occurred mid-pause hardly ever interrupted formulaic sequences, demonstrating that participants were able to produce the FS without pausing and lending credibility inferentially to the theory that FS are stored as single units and retrieved and used as a single string from memory. In a quasi-experimental study (without randomized groups), Yan (2020) investigated EFL students studying at a US university with mostly Chinese L1 participants ($N = 252$). Students were required to listen to 24 sentences and then repeat them. Speech samples were calculated by formulaic nature, sentence length, and proficiency across speech rate and silent pauses, and results showed that speech rate was not significant, but a reduction in silent pauses was: silent pauses: $F(1, 1684) = 11.847$, $p < .001$;

speech rate: $F(2, 6157) = 2.635$, $p = .072$. This study further demonstrated that FS help increase fluency by reducing the number of pauses made during speaking.

Study Purpose Based on Literature

With the advancement of integrating technology-based tasks in instruction that is occurring, more research needs to be conducted to observe the increasing impact of technology adoption on the use of multimodal teaching and learning, particularly concerning the effectiveness of different remote learning modes, including asynchronous and synchronous modalities. Furthermore, while there are plenty of studies on the significant impact multimodal feedback has on writing improvement (Ali, 2016; Alvira, 2016; Cunningham, 2017, 2018, 2019; Faramarzi et al., 2019; Ghosn-Chelala & Al-Chibani, 2018; Honarzad & Rassaei, 2019; Hung, 2016; Maas, 2017; Özkul & Ortaçtepe, 2017; Silva, 2012), studies regarding the impact of multimodal feedback on academic speaking improvement are rare to non-existent. The purpose of this proposed study, therefore, is to examine these gaps and attempt to add to the body of research on multimodal feedback on computer-based academic speaking tasks for adolescent ESL learners. Moreover, the dearth in studies involving adolescent EL learners and their use of FS in spoken academic discourse has been demonstrated. Therefore, this dissertation seeks to add to the empirical research on FS and fluency by investigating the use of FS on computer-based academic speaking assessments by secondary school ELs in US public schools.

CHAPTER III: METHODOLOGY

Research Design

An interactive sequential mixed methods design was chosen for this study, consisting of a randomized between groups experiment to examine the effectiveness of multimodal feedback and a qualitative analysis of linguistic features of the speech production collected during the study. These results led to a third level of analysis examining the linguistic and lexical features in the speech samples of the original experiment. After the quantitative, experimental method was adopted, it became clear that the results from the quantitative study, while demonstrating which type of feedback elicited improved performance, neglected the qualitative aspect of which linguistic features might have contributed to the demonstrated growth. Implementing a mixed methods design allowed the researcher additional "breadth and depth of understanding and corroboration" (Johnson et al., 2007, p. 123) of quantitative results from the experiment with qualitative and linguistic analysis.

The purpose of this study can be characterized as a complementary study, whereby the researcher was seeking elaboration of the results from the quantitative feedback portion of the study with the results of the qualitative production portion of the study (Greene et al., 1989). Furthermore, the rationale according to Bryman (2006) can be defined as illustrative, e.g., using the results of the qualitative linguistic analysis of speech production to illustrate the quantitative feedback findings. And by extension, since the qualitative linguistic analysis could only arise after the experimental study was conducted, the design was one of emergence (Creswell & Plano Clark, 2011). Furthermore, the timing of the study is dependent since the speech samples could only be analyzed once they had been collected for the feedback experiment. The point of interface (Morse & Niehaus, 2009) is one of connection between the quantitative feedback data

demonstrating more effective mode of feedback and the qualitative linguistic analysis data demonstrating the salient linguistic features that contributed to the improved performance.

Schoonenboom and Johnson (2017) propose that a mixed methods study need not have a primary and secondary component, choosing to refer to such a study where both quantitative and qualitative components are balanced as interactive. While this study is sequential and dependent, the results from one nevertheless inform the results from the other; therefore, the researcher considers this an interactive sequential mixed methods design, which is consistent with previous research using mixed methods designs where the researcher is encouraged to choose design methodologies that are in line with answering the stated research questions, even as far as to customize design methodologies utilizing features from varying sources (Creswell & Plano Clark, 2011; Greene et al., 1989; Morse & Niehaus, 2009; Schoonenboom & Johnson, 2017).

Before the experiment for this study was conducted, approval from the Middle Tennessee State University Internal Review Board (IRB) was granted, and parent consent and student assent were secured. The IRB approval form, including the parent consent and student assent form, is included in Appendix C. The feedback segment of the study was conducted over a two-week period from the assignment of Task One to the subject, to the completion of the informal student discussion.

Participants

The participants were recruited from the ESL program in high schools (Grades 9-12) of a suburban school system in middle Tennessee where the researcher is an ESL teacher. The students who were chosen to participate have an English proficiency level of intermediate or higher based on most recent scores attained on the WIDA ACCESS 2.0 English language proficiency assessment and are L1 speakers of Japanese. Students at similar levels of English

proficiency were considered to eliminate the confounding factors that might result from differences in proficiency. Likewise, students in one language group were selected for this study to eliminate possible threats to validity that might be caused by differences in first language background.

Some aspects that could impact student speaking performance by students from different cultural backgrounds include the cultural importance of saving face; class distance between the students, peers, and teachers; level of formality used when communicating with peers or teachers; student status in the peer group; and power position of students in relation to each other and to the teacher (Koizumi & Matsuo, 1993; Purdie & Hattie, 1996; Takanashi, 2004; Dörnyei, & Ushioda, 2009). There were 24 students who fit this profile and agreed to participate in the study. The researcher administered an experiment consisting of a set of two speaking tasks and providing asynchronous screencast feedback in Group One and synchronous video feedback in Group Two using the school district Google Classroom and student email accounts for communication purposes.

Experiment Delivery Technology

The researcher created a Google Classroom within the school system's intranet for conducting the between groups experiment portion of the study. This platform protects all student data and is freely accessible to all students and teachers in the district. Students participated in this proposed study remotely for all parts of the study. Due to COVID-19's impact on students and teachers attending in-person school, conducting the study remotely allowed for equal participation by all students. Students participated in the study asynchronously, except when participating in synchronous video feedback meetings.

Study Procedure Stage One

In the first stage of this study, speaking task one was administered to both groups. After the tasks were completed, they were evaluated by the researcher and another ESL teacher, both with over ten years' teaching experience, and feedback was provided to the participants by the researcher according to the random assignment to Group One (asynchronous feedback) or Group Two (synchronous feedback).

Speaking Tasks

After students were randomly assigned to Group One who received asynchronous screencast feedback or Group Two who received synchronous video feedback, participants responded to speaking task one. Both groups participated in identical speaking tasks that required students to respond to a video prompt using compare and contrast skills. These skills require complex reasoning according to Webb's Depth of Knowledge, Level 3 (Webb, 2005) to synthesize information from the prompt. This type of task was chosen since the amount of language required to produce a response to this type of question provided a larger speech sample for analysis than using less-complex questions at lower Depths of Knowledge, such as retelling facts or describing images. Both speaking task one and speaking task two are a compare and contrast task in order to reduce a threat to validity based on differences in the type and difficulty of task and the amount of speech produced for analysis.

This type of task is further in line with the WIDA Can Do Descriptors Key Uses Edition for Grades 9-12 (WIDA, 2016). WIDA identifies four key uses of language: *recount*, *explain*, *argue*, and *discuss*. At the *explain* level in the speaking domain, compare and contrast tasks fall into the language level 4 category and are, therefore, appropriate for the proficiency level for all students participating in the study (WIDA, 2020b).

Speaking tasks in the content area of science were chosen for both tasks. According to the English Language Development Standards established by WIDA, five types of language are included: language for Social and Instructional purposes, language of Language Arts, language of Mathematics, language of Science, and language of Social Studies. Therefore, choosing tasks utilizing the language of Science is in line with best practice for EL instruction (WIDA, 2020a).

Evaluation Tools

After both groups completed speaking task one, the responses were evaluated using the WIDA Interpretive Rubric for Speaking and Writing (Appendix D), hereafter WIDA Rubric. Due to the nature of evaluating second language speech acquisition, this rubric is a holistic rubric and is not based on analysis of discrete language features but, rather, is based on making a judgement of overall language production (O'Malley & Valdez-Pierce, 1996; Tedick, 2002). The responses were evaluated by the researcher and another ESL teacher, both of whom have been teaching ESL in the public school setting for more than 10 years and both of whom have been trained on using the WIDA Rubric to evaluate speech samples for both assessment purposes and for formative classroom instructional purposes. There were 144 cases of scoring per evaluator, and a total of 23 cases were different. In the case of differences in scoring, the evaluators met virtually to come to a consensus; this was the score that was used in analyzing the data and providing student feedback.

For the purpose of sharing the feedback with the students, the WIDA Rubric was modified into simple, student-friendly language (Appendix E). This eliminates a threat to validity caused by student misunderstanding of the feedback provided which could, consequently, influence performance outcomes and negatively impact the results of the study. A small scale informal discussion with a similar group of ELs in secondary school in the same district ($N = 7$)

was conducted to determine the comprehensibility of the student-friendly rubric, and necessary changes were made in the rubric to make it more comprehensible.

WIDA Interpretive Speaking Rubric. The WIDA Rubric is a publicly available rubric for analyzing student performance on speaking tasks and is a tool available for teachers to use for analysis of academic speech performance in the classroom and on assessments. The rubric contains three levels of evaluation: the word/phrase level concerned with vocabulary usage, the sentence level concerned with language forms, and the discourse level concerned with linguistic complexity.

Word/Phrase Level. The word/phrase level involves how students use vocabulary and expressions to respond to a prompt. This level involves analyzing the precision of use of technical or content-specific vocabulary, as well as academic vocabulary and expressions, to clearly fulfill the purpose of speech. For example, if the prompt addresses a topic about a historical molecular biologist, the expectation is that the student uses as precise vocabulary as possible when speaking about this figure in the response. Usage of the term *molecular biologist* is indicative that the student has a high level of precise, content-specific vocabulary use; on the other hand, use of the word *scientist* is less specific, yet still in the correct context. Contrarily, use of the word *woman* or *person* indicates lack of precision and knowledge of content-related vocabulary needed to adequately respond to the prompt. This level also includes any use of transition words, phrases, or expressions used considering the context and content area. For example, the use of the expression *similar to* in order to discuss details in a response that were different would not indicate the correct context; however, if the same phrase were used to describe details that were similar or related, it would indicate a high level of precise vocabulary or expression use.

Sentence Level. The sentence level examines language forms used by the student, including combinations of sentence patterns specific to purpose and content area, as well as use of correct grammatical structures to convey meaning effectively. At the sentence level, the expectation is that students speak in complete and correctly structured sentences that are varied in length and complexity as appropriate for addressing the prompt. A response, for example, that contains only short phrases, chunks of language, or is a generalization of the topic indicates that a student is at a lower level of proficiency at the sentence level.

Discourse Level. The discourse level measures the linguistic complexity of a response and is the measure that takes into account how students connect vocabulary with sentence structure to produce a coherent and fluent speech sample. The discourse level encompasses the connectedness of the response based on a sustained, precise, and fluent expression of ideas that are appropriate to purpose and content of the prompt. As previously discussed, academic speech requires a different discourse than academic classroom conversations or social conversations with peers (Cummins, 1979; Hyland & Tse, 2007; Krashen 1982; Nagy & Townsend, 2012). Additionally, fluency, taken by itself, is a difficult construct to measure due to the profusion of studies that have been done on different variables to consider when measuring fluency including: quantity of speech produced, rate of speech, and whether to factor out disfluencies, pauses, and repetitions (De Jong & Perfetti, 2011; Ginther et al., 2010; Pawley & Syder, 1983; Towell et al., 1996; Ushigusa, 2009). In response to inquiries regarding the method utilized by WIDA in assessing fluency in the speaking assessment responses, their expert replied that no quantitative measure of fluency is used to evaluate speaking assessment responses (A. Traverse, personal communication, July 24, 2019). Finally, the discourse level also measures the register of

speech—to which degree students are able to speak in a very formal manner to present their response to the prompt.

The WIDA Rubric measures the word/phrase, sentence, and discourse level using 6 different levels of language proficiency: Level 1 Entering, Level 2 Emerging, Level 3 Developing, Level 4 Expanding, Level 5 Bridging, and Level 6 Reaching. For each speaking task, students received a score in each of the three levels (word/phrase, sentence, and discourse) and an overall score that reflects an average of all three.

Student-friendly Feedback Rubric and Form. Traditionally, in feedback in second language acquisition (SLA), errors have been viewed as negative, and performance was judged based on the number of errors committed (Edge, 1989; Hyland, 2001; Khoram et al., 2020; Lynch, 2009; Lyster & Ranta, 1997; Price et al., 2010). However, in the last 30 years, approaches to feedback have been evolving to view it as formative and more learner-oriented as opposed to teacher-oriented. These changes in attitudes about language learning from focus on form (grammatical-lexical, error-correction, teacher-centered) to focus on meaning (content, skills-based, student-centered; Mackey, 2006) have created a learning environment with a greater tolerance for error and opportunities for students to use language in authentic ways (Lightbown & Spada, 1990; Nicol & Mcfarlane-Dick, 2006). In fact, highlighting this shift, Edge (1989) refers to feedback in L2 learning as a path towards acquiring proficiency and not to achieving perfection.

Feedback tries to identify the gap between what the student produces and the expected standard of performance (Hattie & Timperley, 2007; Price et al., 2010; Royce, 1989). For providing feedback on monologic speaking tasks, a simple correction of knowledge is rarely the case and is secondary to measuring student performance based on skills. Rather, the gaps in these

types of speaking tasks typically identify the need for skills development, for example, in utilizing vocabulary and expressions that are specific and appropriate to task (Price et al., 2010). More importantly, Price et al. (2010) points out that the pervading view of feedback is that it should "explicitly address future activity, that is, feed-forward rather than feedback" (Price et al., 2010, p. 279). This means, not focusing on errors committed, but rather channeling student focus towards skills that can improve the next instance of speech output.

The simplified student version of the WIDA Rubric was used to provide feedback to both Group One and Group Two. The individual student scores were recorded using a Google Form and it was shared with the students using a screenshare during the feedback delivery in both groups. The Student-friendly WIDA Rubric was explained to the students, and they received feedback comments for each level of evaluation on the rubric (word/phrase, sentence, and discourse). This feedback provided to the students in both groups was specific to the task (Hattie, 1999), citing examples from students' responses that justify the rating (Lysakowski & Walberg, 1982). Furthermore, this feedback also included suggestions for improvement (Hattie, 1999) based on the students' performance.

Feedback Delivery: Asynchronous and Synchronous

Feedback was provided asynchronously using pre-recorded screencasts or synchronously using live video meetings. For students in Group One, the researcher recorded a live screenshare of the student-friendly electronic feedback form and included a reduced-sized webcam audio-video recording of the researcher delivering the feedback embedded in the corner of the screen, lasting about 10-15 minutes. Students were offered the opportunity to email the researcher with questions after viewing the screencast feedback. Students watched this video asynchronously and had the ability to replay the video. For students in Group Two, the researcher scheduled a

synchronous Zoom video meeting with the student. First, the researcher briefly greeted the student and then shared a computer screen with the same student-friendly feedback form as described above. The student had the opportunity to ask questions during the Zoom meeting, lasting approximately 10-15 minutes, but did not have the ability to watch a recording of the meeting after the meeting had ended.

Study Procedure Stage Two

After the first set of feedback was provided to the subjects, a second speaking task using compare and contrast skills was administered to both groups. The responses were evaluated by the same raters as speaking task one, using the same WIDA Rubric. After the second round of feedback was delivered electronically to the participants, an informal follow up discussion was conducted electronically by email with the participants. Not all students participated, leaving a response of $N = 17$. The conversation consisted of two open-ended questions to gauge student perceptions regarding each mode of feedback, including sharing something positive and something negative about the synchronous and asynchronous feedback modes. The term *coaching* was used instead of *feedback* to facilitate understanding of the purpose of the study as far as the students were concerned.

Statistical Analysis

Data from the groups collected after evaluations from speaking tasks one and two had been completed were analyzed using Wilcoxon Signed Rank Tests. This type of statistical analysis is most appropriate due to the small sample size of this study and that the groups compared were dependent (Field, 2018). The overall composite scores were compared, as were each level of production word/phrase level, sentence level, and discourse level to determine if there were any differences overall and at the specific levels.

Study Procedure Stage Three

Role of FS in Speech Production

FS are important to speech production for several reasons. First, their use helps language learners sound like native speakers (Boers et al., 2006; Khodadady & Shamsaee, 2012; McGuire & Larson-Hall, 2017; Mohammadi & Enayati, 2018; Pawley & Syder, 1983; Rafieyan, 2018; Tavakoli, 2011; Yilmaz & Korban Koc, 2020). It has been recognized that one-third to one-half of English L1 (first language) speech production consists of FS (Biber, 1999; Conklin & Schmitt, 2008; Ellis & Simpson-Vlatch, 2009; Hatami, 2015; Nattinger & DeCarrico, 1992; Pawley & Syder, 1983). Therefore, it follows logically that ELs who are able to incorporate more FS into their speech will sound more like native speakers of English.

Second, because FS language chunks are stored as single lexical units, the speaker can retrieve them in their entirety from memory to maximize processing in real time when producing responses for timed computer-based speaking assessment tasks (Boers, et al., 2006; Conklin & Schmitt, 2008; Eyckmans et al., 2015; McGuire & Larson-Hall, 2017; Mohammadi & Enayati, 2018; Khodadady & Shamsaee, 2012; Rafieyan, 2018; Yan, 2020; Yilmaz & Korban Koc, 2020). While it is not possible to look inside a speaker's brain to see how language output is organized, some studies are doing just that by examining event-related potential of brain area activations related to use of FS (Siyanova-Chanturia et al., 2017), word monitoring experiments (Jeong & Jiang, 2019), and online processing of FS (Jiang & Nekrasova, 2007).

Finally, use of FS can help learners structure output with word strings they know to be correct, leaving gaps for the novel or original content to be inserted (Boers, et al., 2006; Eyckmans et al., 2015; Khodadady & Shamsaee, 2012; Pawley & Syder, 1983; Wray, 2002; Yilmaz & Korban Koc, 2012). These word strings or sentence stems serve as matrices that

support structured output without the student necessarily understanding the grammar underlying the sentence stem structures (Pawley & Syder, 1983). The use of FS not only helps speakers sound more native-like, but it can provide this scaffold to help structure more complex language output.

In stage three of the study, a qualitative linguistic analysis was performed on the student speech samples. After the results from the feedback experiment were evaluated, the emergent linguistic features were analyzed to determine if there was a relationship between use of formulaic sequences (FS) and fluency which might have contributed to improved scores within each group and between both groups. The speech samples were assigned anonymized numbers to remove any personally identifiable data and transcribed by hand using Google Voicetyping. Then to ensure the accuracy of the transcription, the speech samples were analyzed through Praat, a speech analysis software that allows users to analyze very small segments of recorded speech (Boersma & Weenink, 2021) to verify the accuracy of the transcription.

First, the FS were identified by using the MICASE (Michigan Corpus of Academic Spoken English) (Simpson et al., 2002) to identify standard FS in the speech samples, counting all examples that appear in the MICASE with a mean token frequency per 10,000 words at or above $M = 0.10$. This frequency threshold is in line with previous studies where mean frequency thresholds of .40 per 10,000 words (Biber et al., 2004) or even .10 per 10,000 words (Bardovi-Harlig et al., 2015; Biber & Conrad, 1999; Ellis & Simpson-Vlatch, 2009) were used. Additionally, model sequences (MS) (Pawley & Syder, 1983) were identified utilizing the language contained in the speaking tasks found in Appendix F.

The MICASE was chosen to identify the standard FS because, while other sources of spoken language, such as the Corpus of Contemporary American English (COCA), also contain

FS, the MICASE has an exclusive focus on speech production in academic settings. Since the speech samples collected during this study were samples of academic speech, the FS identified in the MICASE were relevant to this study. The COCA, on the other hand, while it contains speech samples, they are from news videos. Additionally, the COCA also includes writing samples from a variety of print media, some including academic content, but others not. As a result of this, due to the fact that the MICASE contains only speech samples collected in academic settings, it proved to be a more applicable choice for identifying FS used in academic speech such as the samples examined in this study.

All speech samples were coded by hand for standard FS evidenced in the MICASE. Additionally, MS representing language taken directly from the speaking prompts were identified and coded. The researcher manually coded the standard and model FS based on the above referenced procedures, and this was manually reviewed by the research assistant. Any discrepancies were discussed, and a mutual agreement was reached.

Discourse Function of FS

After the FS were identified and coded as either standard or model sequences, they were categorized by discourse function: focus, sequence, clarify, compare/contrast, and summarize (Biber & Conrad, 1999; Biber et al., 2004; Nattinger, 1980; Nattinger & DeCarrico, 1992; Pawley & Syder, 1983; Wood, 2002; Xu, 2018). Standard FS were coded S, Model FS were coded M, and original student response material sequences were coded O. Coded pauses were noted with an ellipsis (. . .) and interrupted thoughts or utterances coded with a dash (--). Student responses were written in italics in one continuous line, and formulaic language sequences were coded using an interspersed format due to the fact that the codes are short and distinguishable from words (Edwards, 2003). The coding was completed by the researcher and research assistant

who was trained by the researcher. Any discrepancies in the coding were discussed until an agreement was reached.

Fluency

The results from the qualitative analysis of the FS led the researcher to further investigate the relationship between use of the FS and the growth in the word/phrase level and the discourse level found in the original experimental study. Because the WIDA Rubric does not contain an objective, quantitative measure of fluency, it was concluded that the best way to examine the relationship between use of FS and growth in the discourse level was to apply a quantitative measure of fluency to the speech samples.

For the purposes of this study, only fluent lexical syllables spoken between silent pauses were counted towards the fluency score; and dysfluent lexical utterances such as repetitions, self-corrections, or fillers such as *um*, *eh*, and *ah*, were not included in the measures. By using this MLFR as the fluency measure, it allows for the evaluation of FS by eliminating the possibility that pauses and disfluencies are counted towards fluency, when no meaningful speech was produced in those cases (De Jong & Perfetti, 2011; Pawley & Syder, 1983; Towel et al., 1996; and Ushigusa, 2009).

Next, the MLFR for each speech sample was calculated using Praat, (Boersma & Weenink, 2021). Speech files were converted to .WAV files using a free conversion software and then uploaded into the Praat software. After loading the speech file into the software, the speech files were run using the Praat Script: Syllable Nuclei (De Jong & Wempe, 2009/2010). This Script analyzed each speech sample based on a silence threshold of -25 dB, minimum dip between peaks of 2 dB, and minimum pause duration 0.3 seconds (McGuire & Larson-Hall, 2017; Wood, 2006). In order to determine pauses between runs, these default settings in the Praat

Script were used to distinguish between silences and sounds in each sample (De Jong & Wempe, 2009/2010). Additionally, since this Script is unable to determine whether spoken syllables are fluent or not, after runs were identified and manually coded on the speech sample transcriptions using the Praat results, non-fluent runs, those containing non-meaningful utterances such as *eh* or *uh*, those containing repetitions, and those containing self-corrections, were manually removed. In the last step, fluent syllables were then counted manually. Then, FS, both standard and model (Pawley & Syder, 1983), per number of runs in a speech sample were manually coded and calculated. Finally, the research assistant completed a review of the coding to verify procedures and results. Any discrepancies were analyzed and discussed until a consensus could be reached. Then a linear regression was run with fluency (MLFR) as the independent variable and FS/Run as the dependent variable.

CHAPTER IV: RESULTS AND DISCUSSION

Results: Research Question One

The first research question of the between groups randomized experiment examining the effectiveness of multimodal feedback methods was: Which mode of feedback, asynchronous or synchronous, elicits improved performance on computer-based speaking tasks? Both groups in this study received feedback structured identically, with the only difference being feedback modality. Following previous research about best practice for providing feedback (Fan, 2019; Hattie & Timperley, 2007; Long, 1983; Lyster & Ranta, 1997; Mackey, 2006; Swain & Lapkin 1998), student feedback included:

1. concrete examples from the students' speech samples,
2. suggestions for improvement to feed-forward (Price et al., 2010), and
3. timely provision of the feedback (was provided within 24 hours)

(See Table 3 for excerpts from a student speech sample and feedback comments provided.)

Table 3

Student Speech Samples with Scores and Feedback Excerpts

Task One Speech Excerpt	Task One Feedback Excerpt	Task Two Speech Excerpt	Task Two Feedback Excerpt
The main similarity between complete & incomplete metamorphosis is that both of them have same three stages, their egg stage, la- larva stage, and adult stage. And another similarity is that both of them are gonna start from egg stage and end as adult. **Student Scores** Word/Phrase: 3 Sentence: 3 Discourse: 3	You did a good job using the connecting words "both of them" when comparing the two kinds of metamorphosis. You can improve your answer by expanding your vocabulary . . . . For example, you used the words "start" and "end," but you could have expanded those concepts by saying "at the beginning of the life cycle" or "at the end of the life cycle. . . ."	There are two different things the stages of rock cycle are different. One way the stages of the rock cycle are different is all rocks are diff- were created by different ways. For example, . . . Another way the stages of the rock cycle are different is their structure . . . . In conclusion, the stages of their, of the rock cycle are different in several ways.	You did an excellent job expanding your answer by including . a variety of transition words and phrases to structure and connect the different parts of your answer. Some of these I noticed were "one way," "another way," "for instance," and "in conclusion. . . . " **Student Scores** Word/Phrase: 5 Sentence: 5 Discourse: 4

Data were collected from this feedback in both groups on both speaking tasks and the results revealed that both groups achieved significant improvement from Task One to Task Two with large effect sizes, answering research question 1 (See Table 4).

Table 4

Feedback Growth Between Task One and Task Two on Composite Scores

Group	*n*	*Mdn*	*Z*	*p*	*r*
Group One asynchronous	12	3.00 Task One 4.50 Task Two	-3.07	.000	.89
Group Two synchronous	12	3.00 Task One 3.83 Task Two	-2.59	.008	.75

Note: *r* is an estimated effect size.

These results indicated that Group One achieved significant differences between Task One and Task Two, and Group Two also achieved significant differences between Task One and Task Two.

Next, emerging from the results from research question one, a closer analysis of the differences at each linguistic level was conducted to determine exactly where the growth occurred in each group (See Table 5).

Table 5

Growth by WIDA Rubric Level

	Group One Asynchronous					Group Two Synchronous				
	n	*Mdn*	*Z*	*p*	*r*	*n*	*Mdn*	*Z*	*p*	*r*
Word/ Phrase Level	12	3.00 T1 5.00 T2	-2.74	.004	.79	12	3.00 T1 4.00 T2	-2.49	.013	.72
Sentence Level	12	3.00 T1 4.00 T2	-2.60	.008	.75	12	3.00 T1 4.00 T2	-2.00	.078	.58
Discourse Level	12	3.00 T1 4.00 T2	-2.74	.004	.79	12	3.00 T1 4.00 T2	-2.13	.055	.61

Note: T1 = Task One, T2 = Task Two, *r* is an estimated effect size

Groups One and Two showed significant growth at the word/phrase level (Group One p = .004 and Group Two p = .013). Since the word/phrase level includes not only vocabulary, but also expressions, this indicated that FS could have been one cause of the growth and led to a subsequent quantitative and qualitative examination of the speech samples to determine what role FS might have contributed to the observed growth.

Results: Research Question Two

To answer research question two: Does the use of FS in computer-based academic speaking tasks predict fluency as measured by MLFR, a linear regression was conducted with MLFR as the dependent variable and FS/Run as the independent variable. The results indicated that use of FS significantly predicted fluency ($\beta = 1.01$, $t(46) = 9.65$, $p < .001$) and FS also explained a significant proportion of variance in fluency scores ($R^2 = .67$, $F(1, 46) = 93.20$, $p < .001$). However, a Wilcoxon Signed Rank test of the improvement in fluency as measured by MLFR between Task One and Task Two for both groups revealed that the growth in fluency was

only significant for Group One, and the use of FS per run was not significantly different in either group. (See Table 6).

Table 6

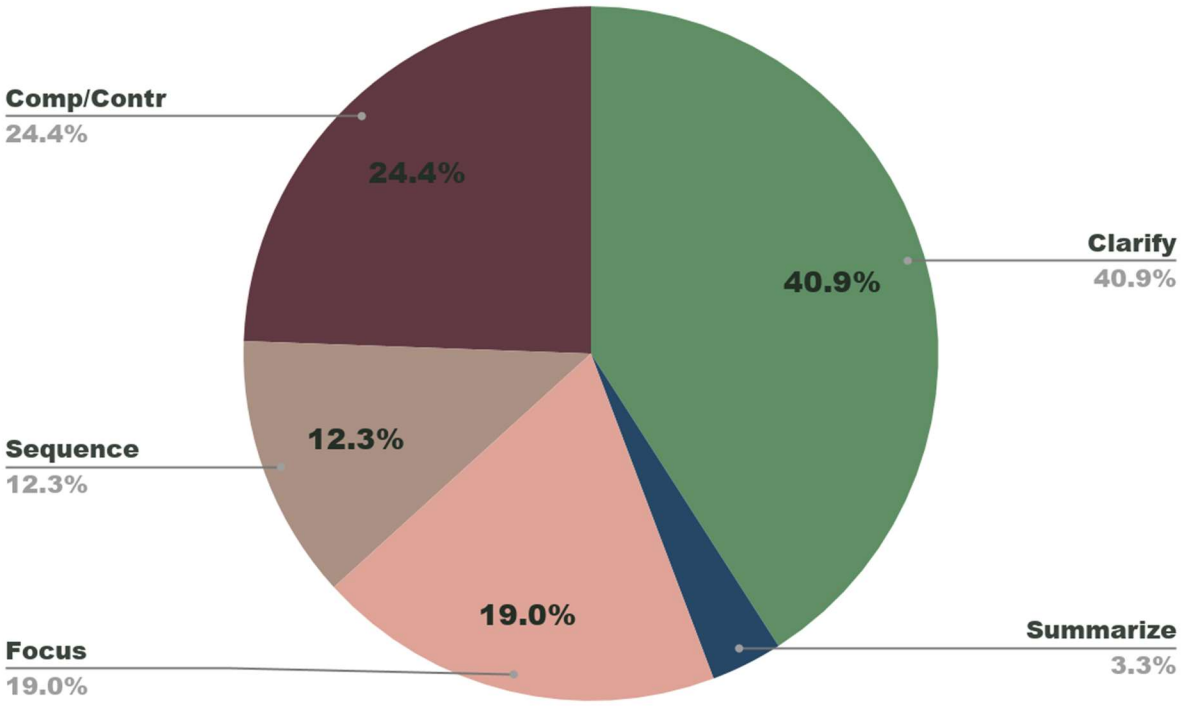
MLFR and FS Growth

	MLFR					FS/Run				
	n	*Mdn*	*Z*	Sig.	*r*	*n*	*Mdn*	*Z*	Sig.	*r*
Group One Asyn-chronous	24	11.37 Task One 10.66 Task Two	-2.43	.015	.70	24	6.90 Task One 7.19 Task Two	0.63	.650	.18
Group Two Syn-chronous	24	8.13 Task One 8.20 Task Two	0.63	.530	.18	24	6.14Task One 5.65Task Two	1.06	.289	.31

To further inform the results of the quantitative analysis, all speech samples were analyzed using Praat and coded for FS according to the MICASE analysis as outlined in CHAPTER III of this document and model sequences (MS) using the language used in the speaking tasks and prompts, the standard FS were analyzed and coded by discourse function in the following categories: **CLARIFY, COMPARE/CONTRAST, FOCUS, SEQUENCE,** and **SUMMARIZE** (see Figure 2).

Figure 2

Discourse Function of Standard FS



Next the location of the Standard FS within the units or runs of fluent speech was examined. Each FS was coded for its location as beginning, middle, or end of the run (see Figure 3). Some FS were self-contained, meaning the only speech in the run was the FS. These were coded as Single FS and were not counted towards beginning, middle, or end locations. They were, however, analyzed to see at which point in the sentence they occurred (see Figure 4). Additionally, some students paused in the middle of the FS, so that the FS started at the end of one utterance and continued through the beginning of the following utterance. These have been labeled as split locations and were not counted towards the beginning or the end locations.

Figure 3

Location of FS Within Runs of Fluent Speech

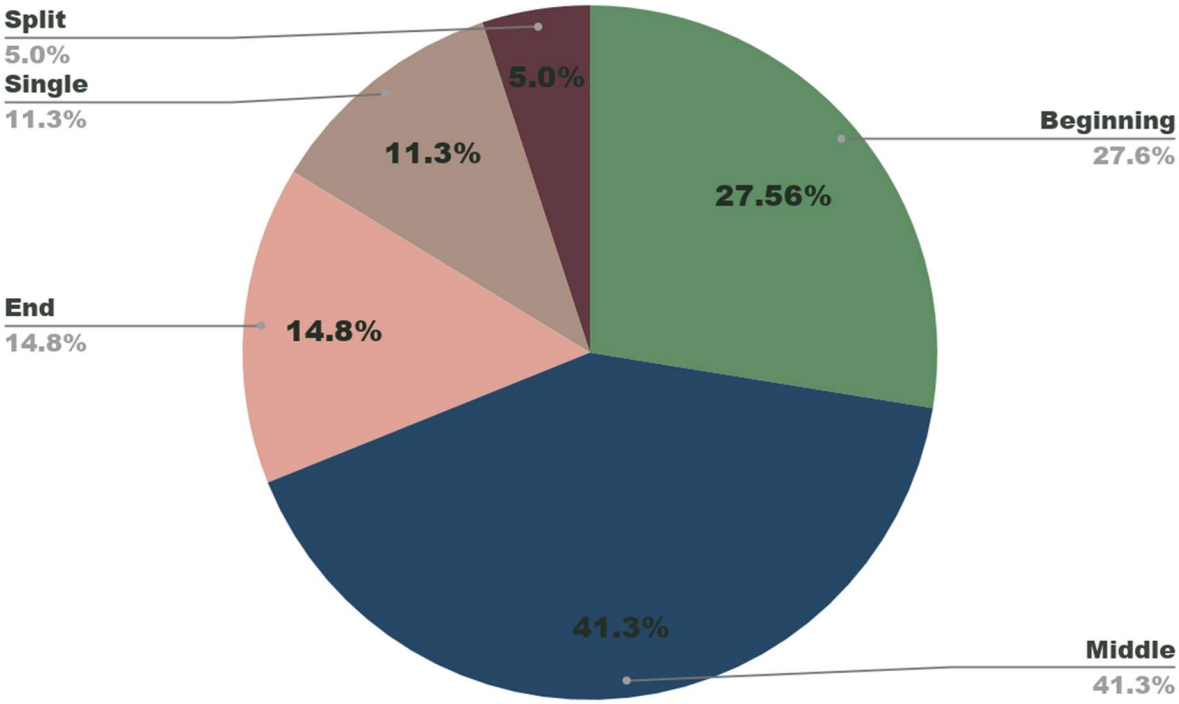
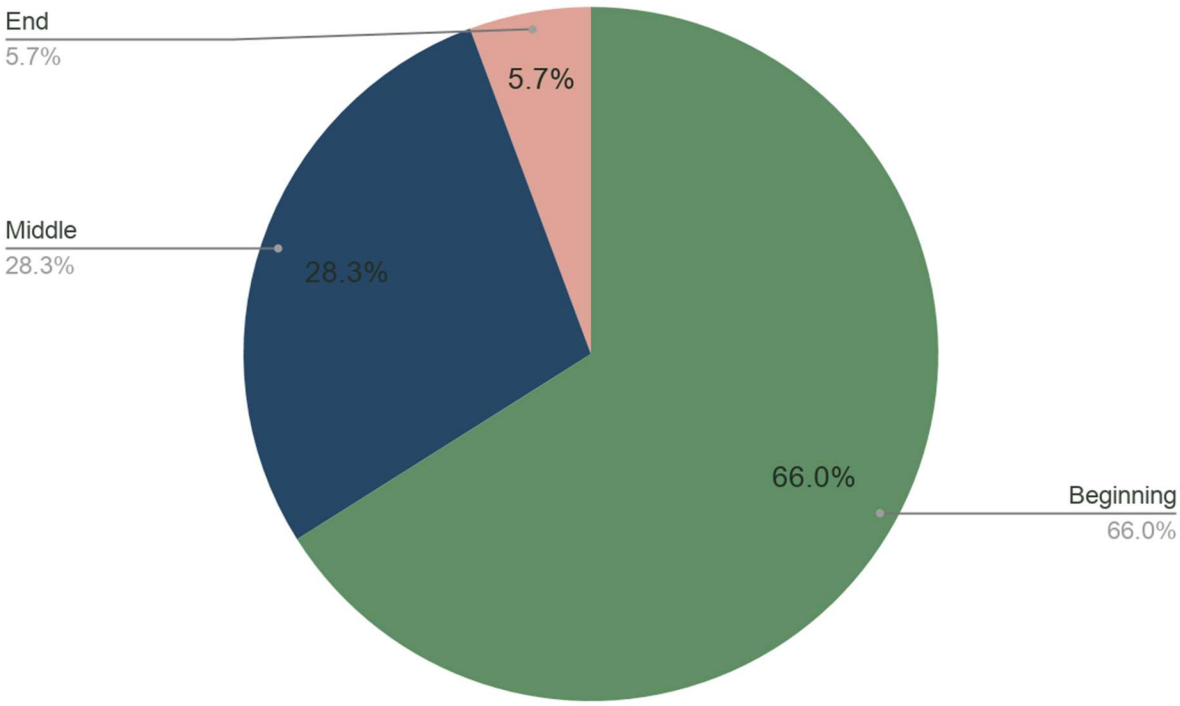


Figure 4

Location of Single FS Within Sentences



A further distinction that needs to be made is that students not only utilized standard formulaic sequences, but they also utilized FS taken from the speaking tasks, called MS in this study. Of the total 815 FS counted in this study, 59% of them were standard FS, while the rest were MS.

Discussion

Feedback Level

Overall growth. The results indicated that students in both groups achieved significant improvement from Task One to Task Two. This demonstrates that multimodal feedback is effective at improving performance when provided on computer-based speaking tasks and confirms previous studies which produced similar results on writing tasks (Ali, 2016; Aljaser,

2019; Alvira, 2016; Elola & Oskoz, 2016; Ghosn-Chelala & Al-Chibani, 2018; and Özkul & Ortaçtepe, 2017). The difference in this study is that comparisons were not made between multimodal feedback and traditional feedback modes, such as written feedback, as in previous studies. On the contrary, this study demonstrated that two different types of multimodal feedback are both effective in improving student performance. With the increased implementation of remote, virtual, or hybrid learning, studies such as these are necessary to continue to ensure that educational best practices are still being met in these formerly non-traditional settings.

Furthermore, these results confirm Aljaser's (2019) study with primary school students (grade 5 in Saudi Arabia) in which he found that younger learners in a technology-based instructional group performed better than learners in the control group which implemented traditional instruction (e.g., without the use of technology tools). In light of prior results confirming that university students perform better in technology-based instructional settings (Ali, 2016; Alvira, 2016; Cunningham, 2017, 2018, 2019; Faramarzi et al., 2019; Ghosn-Chelala & Al-Chibani, 2018; Honarzad & Rassaei, 2019; Hung, 2016; Maas, 2017; Özkul & Ortaçtepe, 2017; Silva, 2012), the finding from this study combined with Aljaser's (2019) findings also confirm that younger learners can benefit from technology-based instruction and feedback modalities.

Additionally, in contrast to previous studies, this study examined differences between different modes of technology-based feedback and demonstrated that asynchronous feedback produced more significant results than synchronous feedback. This could be attributed to a number of different factors. This study was completed near the beginning of the students' first year utilizing Zoom or other synchronous instructional methods from a remote environment. While use of asynchronous screencast technology was already gaining ground in public school

classrooms before the pandemic, the use of Zoom-like synchronous video instruction was previously unheard of or only utilized in rare situations. The students' transition to remote learning coupled with their novel use of this type of synchronous educational technology could have been a factor contributing to the less successful response to the synchronous feedback. Due to their relative unfamiliarity with a remote learning environment and the use of novel technology, their affective filters (Krashen, 1982) could have been elevated, thus preventing them from as efficiently accessing the content of the synchronous feedback.

Another factor that may have played a role in the difference in effectiveness between asynchronous and synchronous feedback can be understood by examining the results of the informal follow-up conversations with the participants which occurred after the completion of the study, which highlighted student agreement on two issues:

1. Students liked synchronous feedback because of the interactive element and having the opportunity to ask questions (74%), but
2. Students also liked asynchronous feedback because they were able to view the feedback multiple times (59%).

Fortunately, both situations can be addressed with one solution: teachers can provide synchronous feedback for the interactive element and the opportunity to ask questions and record the session to share with the students for later viewing.

Another issue that students had with the synchronous feedback was that they had to appear at a specified time and were no longer free to choose when they viewed/participated in the feedback session. Further concerns with the synchronous feedback included technical difficulties participating in the live video meeting, including sound quality, difficulty using the Zoom app, and internet bandwidth issues which interfered with streaming live videos. The

technological issues with using Zoom could stem from this study being conducted early on in the remote learning experience, when most students were not as adept at navigating Zoom meetings.

Finally, when asked which type of feedback they felt benefitted them most, students felt that both types of feedback helped them improve their performance. This is aligned with the other relevant studies that confirming that students' perception of multimodal feedback as helpful in improving their performance (Alvira, 2016; Ghosn-Chelala & Al-Chibani, 2018; Özkul & Ortaçtepe, 2017). In fact, in Hattie's (1999) study of 74 meta-analyses, involving over 7,000 studies and 13,370 effect sizes, the results indicated that Video and Audio Feedback with an overall effect size of 0.64 and Computer-Assisted Instructional Feedback with an overall effect size of 0.52 were among the four most effective feedback types.

In other studies by Ali (2016) and Ghosn-Chelala and Al-Chibani (2018), students had an overall positive attitude towards multimodal feedback. Additionally, students in studies by Ghosn-Chelala and Al-Chibani (2018) and Özkul and Ortaçtepe (2017), overall felt that multimodal feedback was better than written feedback. In this study, however, the results demonstrated that while both types of feedback did help students improve their speaking performance, the asynchronous Screencastify feedback was more effective than the synchronous Zoom feedback. This shows that students can perceive an advantage that is not confirmed by the data.

Elola and Oskoz (2016) results indicated that the students revised their writing according to feedback to the same extent with both types of feedback, written and screencast. The instructor, however, provided more extensive feedback in the screencast feedback treatment. In this study, differences in the amount and content of feedback were limited by the use of the same

feedback form for both types of feedback, as well as limited in time by restricting the length of feedback provided in both groups to 15 minutes.

Results by WIDA Rubric level. After the preliminary results indicated that significant improvements overall were made in both groups, a separate analysis was done for each level analyzed by the WIDA Rubric to determine where the growth occurred. According to the data analysis, the most growth occurred at the word/phrase level and the discourse level of Group One (asynchronous group) and the word/phrase level of Group Two (synchronous group). The word/phrase level involves use of vocabulary and expressions, which includes FS; and the discourse level includes use of effective transitional expressions and fluent speech production. Both of these levels are impacted by use of FS, so these results led to the subsequent linguistic analysis of FS used in the entire collection of speech samples.

Production Level

FS Effect on Fluency and Growth. Data analysis revealed that while FS use in both Task One and Task Two for both groups was a significant predictor of fluency ($\beta = 1.01$, $t\,(46) = 9.65$, $p < .001$), the difference in growth in fluency between Task One and Task Two was only significant for Group One.

An explanation of this effect could be that the students participating in the study were already producing FS in their speech, so that their starting point for FS use was already high. One way to optimize this would be to conduct a pre-test to determine level of FS use before conducting future studies. Another cause could be the proficiency level of the participants. Since all students were at an intermediate or higher level of English proficiency when starting the study, the amount of growth possible was already limited. This is confirmed by Rafieyan (2018), whose results indicated that students with a higher English proficiency level showed more use of

FS in their speech. Additionally, Rafieyan's results demonstrated that explicit instruction in FS, regardless of whether they were taught in or out of context, produced no differences in performance results.

This finding, furthermore, brings several relevant issues to light. To discuss these issues, it is important to keep in mind that FS were a significant predictor of fluency in all speech samples collected in the study. The first issue involves the perceived fluency of the samples. Because student use of FS is significant in all samples, it is clear that their use contributed to the perceived fluency of the raters scoring the speech samples using the WIDA Rubric. This finding is relevant because it highlights the importance of FS use in sounding more fluent as confirmed in Boers et al. (2006). Following from this is the fact that the WIDA ACCESS computer based speaking assessment is evaluated using a holistic rubric; in other words, sounding more fluent will benefit students participating in similar holistically scored computer based speaking assessments. Furthermore, if use of FS increases the perceived fluency of ELs, then oral communication on all levels can be perceived as more fluent, including academic speech in content area classes, as well as other academic and social conversation instances. While the objective measures of fluency taken still indicated growth and large effects in some areas, the growth was not statistically significant; however, the benefits in perceived fluency far outweigh this insignificant finding. On the other hand, in McGuire and Larson-Hall (2017), the only measure that was not significant was the subjective measure of fluency taken in the study. The authors indicated that their rubric may have been too complicated for raters to implement with fidelity; however, in future studies when English L1 speakers are used to rate perceived fluency, simplified rubrics may increase the validity of the findings.

Discourse function of FS. Several different discourse features are important to consider when analyzing second language speech production (De Jong & Perfetti, 2011; De Jong & Mora, 2019; De Jong & Schoonen, 2013; Ginther et al., 2010; Rossiter, 2009; Towel et al., 1996). After the initial data from modes of feedback were collected, the speech samples were transcribed and individually analyzed and coded for the standard formulaic sequences (FS) and the model sequences (MS). Subsequent to that, the FS were then coded by discourse function: CLARIFY, COMPARE/CONTRAST, FOCUS, SEQUENCE, and SUMMARIZE (Biber & Conrad, 1999; Biber et al., 2004; Nattinger, 1980; Nattinger & DeCarrico, 1992; Pawley & Syder, 1983; Wood, 2002; Xu, 2018).

The most frequent discourse function of CLARIFY appeared in 41% of the FS used in the study. This finding is similar to a finding by François and Albakry (2021) that also found that the most frequent discourse function of FS was CLARIFY. In the CLARIFY function, students used the FS to highlight or explain information or add details or description to information in the response. Some examples of CLARIFY from the speech samples include:

> *Sedimentary rocks are rocks that have become thinner due to erosion . . .*
>
> *and metamorphic are changed by heat radiation . . . .*
>
> *that are both used by insects to turn into adult forms from their egg forms .*
> *. . .*

In these utterances, the students used the FSs to clarify information by explaining the reasons or the situation that they are addressing in the prompt. As can be seen in these examples, one feature of the CLARIFY discourse function is that the FS can be *verb + preposition* collocations or lexical bundles, words that frequently appear together. Another example of FS used in the CLARIFY function include:

The number of the life stage is different . . .

what type of mammal has . . .

From here on complete cycle goes through a larva state . . .

In these examples of CLARIFY, the FS constructions are combinations of *prepositions + nouns/pronouns*, which are also, for the most part, collocations or lexical bundles. These FS are being used to elucidate examples, describe details, and provide the content direction of the upcoming speech.

The second most frequent discourse function represented in the speech samples is COMPARE/CONTRAST (24%). These sequences are used when comparing or contrasting two or more elements from the speaking tasks. Since both speaking tasks involve comparing and contrasting, it is not surprising that these are frequent. What is surprising is that they are not the most frequent. This could be an indicator that in spite of all speech samples addressing COMPARE and CONTRAST tasks, the discourse function of CLARIFY is needed more often during the speech events than specific discourse functions like COMPARE and CONTRAST. Some examples of COMPARE can be seen here:

The similarities between . . .

How the stages of the rock cycle are similar . . . .

There are two things that were the same for complete and incomplete metamorphosis . . . .

Another way to show comparisons is the use of *than* + and *both* + *verb*, for example:

also looks rough . . . but it looks . . . smoother than sedimentary . . .

made hotter than metamorphic . . .

First both are born from eggs . . .

They both go through the stages of . . .

Some examples of CONTRAST can be seen below:

I found different points according to the rock cycle video and graph . . . .

The difference between incomplete and complete metamorphosis is . . .

There are several differences . . .

A further way to express differences is by using adverbial clauses, for example:

Although there are some differences associated with those two processes . . . incomplete and complete metamorphosis . . . both are used by insects

However . . . the beginning and end of their life are the same. . . .

These adverbial clauses are a good example of FS that can be used as sentence stems to help students structure more complex speech. As can be seen in the examples, the FS are separated by other sequences from both Model Sequences and other standard FS. These are also examples of Pawley and Syder's (1983) sentence stems that help students structure complex content output by utilizing FS and leaving slots for students to input original material.

The next most frequent discourse function represented in the samples is FOCUS (19%). Students use FOCUS FS to draw attention to facts or content in their speech. Some examples of FOCUS include:

I'm going to talk about . . . how stage of rock cycle are similar. . . .

So . . . this example . . . you can see the similar thing . . .

These are the same points that I found . . .

I noticed the differences of them . . .

The next discourse function is that of SEQUENCE, and it comprised 12.3% of the FS used in the study. Students use sequencing to order details, such as:

The next way the stages of the rock . . . cycle are similar is . . .

The first different point is color . . .

The next similar point is . . .

The second similarity between . . .

Another way of sequencing speech that is found in the samples includes use of adverbial clauses to indicate reference to previous information given or to sequence information in time order, such as:

Therefore . . . they have differences during their growth process . . .

When the larva has come then it changes in the pupa. . . .

if the rocks within the earth are superheated, (then) they turn into a liquid called magma. . . .

As with the adverbial clauses used to indicate compare and contrast, these FS used to sequence information can also be used as sentence stems (Pawley & Syder, 1983) to help students structure more complex speech.

SUMMARIZE is the final discourse function represented in the speech samples (3%). SUMMARIZE FS are used to sum up ideas in the response and typically occur at the end of a speech sample:

In conclusion these are two main differences between each stage. . . .

All in all the different types of rocks formed differently and displayed distinct information . . .

As a result . . . life cycle of rocks are different in various ways. . .

The very small number of SUMMARIZE FS found in these samples indicates that students participating in the study were not very successful at including a closing remark to wrap up their response. This can be remedied by instruction on the structure of a monologic response that needs to include not only an introduction and body, but a conclusion as well. One challenge is that the speech output is timed on the computer-based assessments and students might be nervous about recording their responses. Seeing the time clock tick down might be causing some students stress and causing them to forget their conclusion before clicking the button to stop recording before time runs out.

Location of FS in a Run. A further analysis was conducted to determine at what location in a fluent run did most FS occur. Based on the results, the most frequent occurrences were, rather surprisingly, located in the middle of the run (41%). When one thinks about formulaic sequences, the ones that most frequently come to mind are those that start utterances; however, these results revealed that in this study, more FS were used in the middle of utterances. This demonstrates that FS play a larger role in connecting different parts of responses than simple transition words and phrases. It highlights the importance of explicit instruction of collocations or lexical bundles, especially *verb + preposition* and *noun + preposition* that were found in this study. However, 28% FS were found at the beginning of the run and 14.8% were found at the end of the run.

Beginning of Run: *The first different point is color . . .*

Middle of Run: *sedimentary rocks were created by erosion . . .*

End of Run: *And they turn into . . .*

Furthermore, some FS were self-contained in one run (11%). These Single FS were overwhelmingly found at the beginning of the sentence where the run was located (66%). This

indicates that the FS that are short transition phrases used to connect different parts of response are present but do not represent a very large proportion of FS used in the samples.

> Single FS: *Another way is how the rocks transform into another rock cycle stage. . . . <u>For instance</u> . . . if the rocks transform into metamorphic rocks . . .*

Finally, some students paused in the middle of the FS:

> Split FS: *which is the <u>hatching</u> . . . <u>from</u> their eggs*

This mid-utterance pause is an indication that some students are still developing the capacity to store and retrieve the FS as single units (Biber & Conrad, 1999; Boers et al., 2006; Ellis, 1996; Ellis, et al., 2008; Eyckmans et al., 2015; Gray & Biber, 2013; McGuire & Larson-Hall, 2017; Mohammadi & Enayati, 2018; Nattinger & DeCarrico, 1992; Pawley & Snyder, 1983; Rafieyan, 2018; Wood, 2002, 2006; Yilmaz & Korban Koc, 2020). This highlights a finding in Yilmaz and Korban Koc's (2020) study that demonstrated in a delayed post-test, the treatment started to lose effect. Students not only need explicit instruction in FS, they also need on-going re-enforcement of the FS and on-going practice of FS implementation in academic speech.

In this study, the type of speech investigated was monologic academic speech in response to previously unprepared content area speaking tasks. Previous studies indicate gaps in the research in this area. With a focus on monologic narrative speaking tasks (Tavakoli, 2011) or academic or conversational dialogic tasks (Boers et al., 2006; McGuire & Larson-Hall, 2017; Mohammadi & Enayati, 2018; Rafieyan 2018; Yilmaz & Korban Koc, 2020), previous studies have not examined the effect of FS on fluency in monologic academic speaking tasks. This study thus contributes to the body of research on the influence of FS on monologic academic speaking tasks and demonstrates the need for further research to help practitioners optimize instruction to meet the challenges of producing this type of speech.

CHAPTER V: CONCLUSION

The research questions in this study investigated:

1. Which mode of feedback, asynchronous or synchronous, elicits improved performance on computer-based speaking tasks?

Emerging from results from research question one, the following question was subsequently examined:

2. Does the use of FS predict fluency in academic speaking tasks?

The results indicated that both synchronous and asynchronous feedback modes elicited improved performance on the speaking tasks used in the study, answering research question one. These results showed significant growth by both groups, with the asynchronous group outperforming the synchronous group. This finding can help bridge the research to practice gap by highlighting the effectiveness of multimodal feedback on computer-based speaking tasks for ELs for practicing educators working in F2F, hybrid, or remote instructional environments. Furthermore, results from the student survey provided insight into students" perceptions about the usefulness of asynchronous and synchronous feedback modalities and highlighted student-perceived limitations and benefits of multimodal feedback.

With the dramatic increase in remote learning that the pandemic has instigated, this study also demonstrates the effectiveness of multimodal feedback, particularly considering synchronous and asynchronous methods that are being implemented in classrooms around the globe. It clearly demonstrates that multimodal feedback is effective in improving student performance on computer-based speaking tasks and provides rationale for implementing this type of feedback more regularly moving forward.

Additionally, further investigation into the areas of growth indicated greatest improvement at the word/phrase level and discourse level of the subjective holistic rubric used in the evaluation of speech samples of both groups. This led to a subsequent investigation of FS to determine the answer to research question two: Does FS use predict fluency as measured by MLFR in the samples? Results indicated that use of FS significantly predicted fluency, thereby answering research question two. Moreover, an analysis of discourse functions represented by the FS used by students in the study indicated which discourse functions are most represented in the study samples and the forms of FS used in those functions were also discussed to illustrate the ubiquity and utility of FS in academic speech.

These findings warrant the explicit instruction in knowledge and use of FS for ELs towards improving academic speech not only on speaking assignments in the ESL classroom, but also in general education content area courses, to college or career readiness, and beyond.

Instructional Implications

The results of the analyses conducted during this study have demonstrated the significance that FS play in speech production. Providing explicit instruction in how to learn and use FS is clearly something that could benefit student performance on computer-based speaking assessments in particular and in academic speaking in general. Instruction in FS should cover both types of FS discussed in this study: standard language sequences that take on a variety of forms, and model language sequences, which are chunks of language taken from the speaking tasks themselves. To teach the former, direct instruction in the implementation of standard formulaic language sequences should be taught, while to teach the latter involves training students how to manipulate the language in the speaking prompts to make use of the structures and sequences for use in generating their response.

Instructional Model

One way to provide this instruction is to base it on a noticing, retrieving, and generating model (Hatami, 2015; Rafieyan, 2018; Schmitt, 2004; Wray & Perkins, 2000). Noticing involves calling attention to FS as they appear in authentic texts and speech samples and making students aware of what FS are and how they are used. FS are important because they are stored as single units, so the noticing process should involve calling attention to the salient features of the FS the students will be expected to use (Hatami, 2015). There are a number of sources for teachers to consult when deciding which FS their students need (see Martinez & Schmitt's "A Phrasal Expressions List;" 2012 or Simpson-Vlatch & Ellis' "An Academic Formulas List: New Methods in Phraseology Research;" 2010). These lists sort the FS by discourse function, so knowing what types of speaking tasks students will be expected to master and what discourse functions would be necessary in these types of tasks could help teachers to further limit their FS selection for instruction.

The next step is teaching retrieval, the process that helps solidify the new knowledge gained from noticing FS (Hatami, 2015; Wray & Perkins, 2000). When students are provided a list of previously noticed FS and asked to recognize them and then reproduce them in either written or spoken practice activities, it facilitates the brain's retrieval from short term memory from the previous work done during noticing tasks. Finally, in the generating phase, students not only retrieve knowledge of how to use the FS from short term memory, but they produce FS independently, thereby strengthening the brain's capacity to store FS knowledge in long term memory and to make the FS available for use when needed (Hatami, 2015; Wray & Perkins, 2000).

Instructional Activities

There are several instructional activities that can be implemented following this noticing, retrieving, and generating model.

Dictogloss Activities. These activities, for example, cover the noticing and retrieving steps by providing students with targeted language structure practice by integrating listening, writing, and speaking skills (Lindstromberg et al., 2016). The teacher reads a text and draws student attention to the FS used, for example, by providing a transcript with the FS highlighted for the students to follow along while the text is being read. After the text is read, students' attention is called to the FS and how they are used in the sample. Students subsequently listen to the text multiple times, once without taking notes and additional times with the purpose of noting keywords and language structures. Finally, students work collaboratively in pairs or small groups to reconstruct the contents of the text to provide a summary. This highlights not just the noticing stage, but the retrieval of FS to complete the missing text.

Speech Shadowing Activities. An additional activity to help students retrieve FS are speech shadowing exercises. These teach students to retrieve FS by having students listen to a short text containing target FS and practice repeating the text as close to the sample as possible, matching pronunciation, speed, and prosody. The aids with retrieval by having students focus on the use of FS in the target texts or audio passages and then having them reproduce the language as accurately as possible (Thomson, 2017; Wood, 2009).

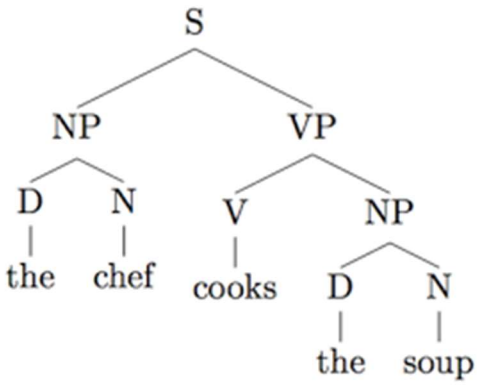
Speech Task Performance Activities. Finally, implementing speaking task performance exercises such as the ones for which the speaking samples in this study were collected is an example of providing students the opportunity to put into practice the skills and knowledge gained from the noticing and retrieving activities. Students can be presented with sample

speaking tasks and required to respond to the task in a similar environment to that of the testing situation, independently generating the FS in authentic speaking tasks.

Sentence Structure Activities. In order to help students learn how to manipulate model sequences found in audio passages and prompt questions for use in the speaking task responses, sentence structure activities can be implemented. One way to accomplish this is with activities directed towards recognizing and manipulating parts of speech such as sentence trees (see Figure 5).

Figure 5

Sentence Tree



Sentence trees help students identify relationships among the words in a sentence, which then allows them to manipulate the language to change questions into statements or replace nouns with participle phrases, for example.

The chef cooks the soup.
Who cooks the soup?
What does the chef cook?
Cooking the soup is done by the chef.
The soup is cooked by the chef.

Understanding what role different words play in a sentence allows students to take those words and manipulate them by changing their tense or part of speech in order to re-use these MS language from the speaking tasks in their spoken responses.

As this chapter indicates, the results of this between groups study have added to the scholarly literature by filling gaps in the research covering the effectiveness of multimodal feedback on and the role FS play in speech fluency on computer-based speaking tasks of secondary school students. The results from the between groups experiment were presented and discussed, demonstrating that multimodal feedback and FS use are significantly effective in improving student performance. Further results demonstrated that the asynchronous feedback group outperformed the synchronous feedback group. Additionally, emerging from the speech samples collected, it was shown that FS use significantly predicted fluency as measured by MLFR in the speech samples.

Limitations and Directions for Future Study

It is important to acknowledge some of the limitations of the study and how they were addressed. While the sample size is relatively small, adapting the statistical analysis method best suited to small sample sizes was implemented. Additionally, to limit the number of threats to validity, the study included only subjects from one L1 background. A broader analysis of students with multiple L1s with an increased sample size, however, would enhance the results to be more applicable to public school EL classrooms. Studies that investigate how students from a variety of L1 backgrounds respond to multimodal feedback and, additionally, how their use of FS compares to the results of this study would be needed to inform instruction for classroom practitioners.

Due to the limited scope of this study, several directions for further study can be suggested.

Pauses and Event-Related Potential Brain Responses

With the increase in event-related potential (ERP) studies making the inner-workings of the brain more visible, a study that investigates brain responses to events that occur when spontaneous speech is produced would shed some light on the role the FS use plays in structuring speech output. Some studies have already been done, for example, about brain area activations related to use of FS (Siyanova-Chanturia et al., 2017), word monitoring experiments (Jeong & Jiang, 2019), and online processing of FS (Jiang & Nekrasova, 2007). Additionally, applying this technology to examine the effects pauses have on speech output or to investigate what happens in the brain when pauses occur and what happens in the brain when it compensates for these pauses would help elucidate the relationship between FS use and fluency. Moreover, investigating the location of FS within units of speech during ERP studies could shed light on the role FS play within speech utterances. While most English speakers use FS, not all speakers will store or retrieve them identically (Biber, et al., 2004; Boers, et al., 2006; Wray, 2002), and shedding light on how FS are stored and retrieved in the brain could inform instruction and practice in the classroom.

FS Types and Discourse Function

Further studies that investigate which types of FS are found most frequently in different discourse functions could help teachers target types of FS by discourse function for instructional purposes. Additionally, administering a pre-test to determine which FS are already known and used could help identify areas of growth more accurately. Furthermore, while studies have already been done to investigate the differences between implicit and explicit instruction of FS

(Boers & Lindstromberg, 2012; Rafieyan, 2018; Wood, 2009), additional studies of FS by discourse function or type (lexical bundles vs. formulaic speech vs. lexical frames) could help shed light on instructional methods that are effective for teaching students how to acquire and implement different types of FS in their speech.

Proficiency Level and Grade Level

In order to reduce the threats to validity, this study was limited to students with similar English language proficiency levels. However, future studies could expand on work already done in this area (Rafieyan, 2018) with students with a variety of proficiency levels to determine which proficiency levels use the most FS and to what degree of accuracy. Furthermore, attempting to identify the turning point between speaking in words and short phrases to transitioning to sentences would help pinpoint the moment in language acquisition when teaching FS could have the greatest impact on academic speech output. Furthermore, examining differences in FS use between elementary and secondary students could shed light on the best age to start instruction in FS. Changes in the brain that happen around puberty could affect the way students use FS in speech, including brain functions such as abstract thinking and cognitive analysis.

L1 Differences

As previously mentioned, this study was limited to L1 speakers of the same language. To provide more depth to the study of FS use among ELs, a larger variety of L1 backgrounds could be investigated. Of great interest would be a comparative examination to see if the use of FS in the L1 match the use of FS in English, either from the perspective of frequency of occurrence, discourse function of FS in L1 and English, or if the use of FS in the L1 translate equivalently to

English. While some studies have been done along these lines (Taguchi, 2008; Eyckmans et al., 2015), additional studies would be needed to form a broader picture for practicing teachers.

Type of Feedback

As was mentioned previously, the timing of this study could have had an impact on the results based on the subjects' familiarity with the technology tools implemented. Additional studies could re-examine the results to determine if increased familiarity with technology tools also improves the effectiveness of technology tools used in delivering multimodal feedback.

Final Thoughts

As can be seen by the research completed and research yet to be done, studies examining multimodal feedback and FS use in computer-based speaking assessments are topics that are forward feeding areas of research. Technological advances are not only facilitating different ways to teach and learn, they are also facilitating different ways to research the teaching and learning processes. This movement towards computer-based testing and instruction has been propelled forward by the global pandemic, moving unprecedented numbers of teachers and students into remote or hybrid educational environments, without the adequate tools, knowledge, or skills necessary for navigating the involuntary immersion in computer-based learning. In the years to come, more studies such as this one will continue to add to the scientific knowledge about best practice for ELs in F2F, hybrid, and remote learning environments.

REFERENCES

ACrameThirtyTwo. (2011, November 14). *Complete vs. Incomplete Metamorphosis* [YouTube Video]. https://youtu.be/nuPCu8IHC8I

Ali, A. D. (2016). Effectiveness of using screencast feedback on EFL students' writing and perception. *English Language Teaching, 9*(8), 106-121. http://doi.org/10.5539/elt.v9n8p106

Aljaser, A. (2019). The effectiveness of e-learning environment in developing academic achievement and the attitude to learn English among primary students. *Turkish Online Journal of Distance Education, 20*(2), 176-194.

Alraddadi, B. M. (2016). The effect of teaching structural discourse markers in an EFL classroom setting. *English Language Teaching*, *9*(7), 16–31.

Alvira, R. (2016). The impact of oral and written feedback on EFL writers with the use of screencasts. *PROFILE Issues in Teachers' Professional Development, 18*(2), 79-92. http://doi.org/10.15446/profile.v18n2.53397

Bardovi-Harlig, K., Mossman, S., & Vellenga, H. E. (2015). Developing corpus-based materials to teach pragmatic routines. *TESOL Journal*, *6*(3), 499. https://doi-org.ezproxy.mtsu.edu/10.1002/tesj.177

Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language, 62*(2), 384-414. doi:10.2307/414678

Biber, D. (2006). *University Language: A corpus-based study of spoken and written registers*. https://doi.org/10.1075/scl.23

Biber, D., & Conrad, S. (1999). Lexical bundles in conversation and academic prose. In H. Hasselgard and S. Oksefjell (Eds.). *Out of corpora: Studies in honor of Stig Johansson.* (pp.181-190). Amsterdam.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at …: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, *25*(3), 371–405. https://doi-org.ezproxy.mtsu.edu/10.1093/applin/25.3.371

Boers, F., Eyckmans, J., Kappel, J., Stengers H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral fluency: Putting a lexical approach to the test. *Language Teaching Research, 10*(3), 245-261.

Boers, F., & Lindstromberg, S. (2012). Experimental and intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics, 32*, 83-110. doi:10.1017/S0267190512000050

Boersma, P., & Weenink, D. (2021). *Praat*: *Doing phonetics by computer* (Version 6.1.37) [Computer program]. http://www.praat.org/

Bryman A. (2006). Integrating quantitative and qualitative research: how is it done? *Qualitative Research*, *6*(1), 97–113. https://doi-org.ezproxy.mtsu.edu/10.1177/1468794106058877

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*(1), 72–89. https://doi-org.ezproxy.mtsu.edu/10.1093/applin/amm022

Corson, D. (1997). The learning and use of academic English words. *Language Learning*, 47, 671-718.

Creswell, J., & Plano Clark, V. (2011). *Designing and conducting mixed methods research* (2nd ed.). Sage.

Cummins, J. (1979). Cognitive/academic language proficiency, linguistic interdependence, the optimum age question and some other matters. *Working Papers on Bilingualism*, 19, 121-129.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and Language Minority Students: A Theoretical Framework*. Los Angeles: Evaluation, Dissemination and Assessment Center California State University.

Cummins, J. (2001). *Negotiating identities: Education for empowerment in a diverse society* (2nd ed.). Los Angeles: California Association for Bilingual Education.

Cummins, J. (2008). BICS and CALP: Empirical and theoretical status of the distinction. *Encyclopedia of language and education*, *2*(2), 71-83.

Cunningham, M. A. (2015). *Using audio screencast for feedback on short written essays* [Doctoral dissertation, Southern Illinois University Carbondale]. ProQuest Dissertations and Theses Global.

Cunningham, K. J. (2017). APPRAISAL as a framework for understanding multimodal electronic feedback: Positioning and purpose in screencast video and text feedback in ESL writing. *Writing & Pedagogy, 9*(3), 457-485. https://doi.org/10.1558/wap.31736

Cunningham, K. J. (2018). *Modes of feedback in ESL writing: Implications of shifting from text to screencast* [Doctoral dissertation, Iowa State University]. ProQuest Dissertations and Theses Global.

Cunningham, K. J. (2019). How language choices in feedback change with technology: Engagement in text and screencast feedback on ESL writing. *Computers & Education, 135*, 91-99. http://doi.org/10.1016/j.compedu.2019.03.002

De Jong, N., & Perfetti, C. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and pronunciation. *Language Learning, 61*(2), 533-567. http://doi.org/10.1111/j.1467-9922.2010.00620x

De Jong, N. H., & Mora, J. C. (2019). Does having good articulatory skills lead to more fluent speech in first and second language? *Studies in Second Language Acquisition, 41*, 227–239. http://doi:10.1017/S0272263117000389

De Jong, N. H., & Schoonen, R. (2013). Second language fluency: Speaking style or proficiency? Correcting measures of second language fluency for first language behavior. *Applied Psycholinguistics*, 1-21. http://doi.org/10.1017/S014271643000210

De Jong, N. H., & Wempe, T. (2009/2010). *Praat* script to detect syllable nuclei and measure speech rate automatically. *Behavior Research Methods*, *41*(2), 385–390. (Script modified by H. Quene, I. Person, & N. De Jong, 09.17.2010)

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

Edge, J. (1989). *Mistakes and Correction*. Longman.

Edwards, J. A. (2003). The Transcription of discourse. In D. Schiffrin, D. Tannen, and H. Hamilton (Eds.). *The Handbook of discourse analysis.* Blackwell Publishing. http://www.blackwellreference.com/subscriber/tocnode?id=g9780631205968_chunk_g978063120596818

Ellis, N. (1996). Sequencing in SLA: Phonological memory, chunking, and points of order. *Studies in Second Language Acquisition, 18*(1), 91-126. Retrieved from: https://www.jstor.org/stable/44487860

Ellis, N., & Bogart, P. (2007). Speech and language technology in education: the perspective from SLA research and practice, In *SLaTE-2007*, 1-8.

Ellis, N., & Simpson-Vlatch, R. (2009). Formulaic language in native speakers: Triangulating psycholinguistics, corpus linguistics, and education. *Corpus Linguists and Linguistic Theory 51*(1), 61-78.

Ellis, N., Simpson-Vlatch, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, Corpus Linguistics, and TESOL. T*ESOL Quarterly, 42*(3), 375-396. Retrieved from: https://www.jstor.org/stable/40264474

Elola, I., & Oskoz, A. (2016). Supporting second language writing using multimodal feedback. *Foreign Language Annals, 49*(1), 58-74. http://doi.org/10.1111/flan.12183

Eyckmans, J., Boers, F., & Lindstromberg, S. (2015). The impact of imposing processing strategies on L2 learners' deliberate study of formulaic language sequences. *System 56*, 127-139. doi:10.1016/j.system.2015.12.001

Fan, N. (2019). An investigation of oral corrective feedback in an ESL listening and speaking class. *Journal of Language Teaching & Research*, *10*(1), 197–203. https://doi-org.ezproxy.mtsu.edu/10.17507/jltr.1001.22

Faramarzi, S., Tabrizi, H., & Chalak, A. (2019). Learners' perceptions and attitudes towards L2 vodcasting tasks in an e-learning project. *Teaching English with Technology, 19*(3), 3-21, http://www.twetjournal.org

Field, A. (2018). *Discovering statistics using IBM SPSS* (5th ed.). Sage Publications

François, J. & Albakry, M. (2021). Effect of formulaic sequences on fluency of English learners in standardized speaking tests. *Language Learning & Technology, 25*(2), 26–41. https://doi.org/10125/73429

Gee, J. P. (1990). *Social linguistics and literacies: Ideologies in discourses*. New York: Falmer Press.

Ghosn-Chelala, M., & Al-Chibani, W. (2018). Screencasting: Supportive feedback for EFL remedial writing students. *The International Journal of Information and Learning Technology, 35*(3), 146-159. http://doi.org/10.1108/IJILT-08-2017-0075

Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing, 27*(3), 379-399. http://doi.org/10.1177/0265532210364407

Gray, B., & Biber, D. (2013). Lexical frames in academic prose and conversation. *International Journal of Corpus Linguistics, 18*(1), 109-136. https://doi.org/10.1075/ijcl.18.1.08gra

Greene, J. C., Caracelli, V. J., & Graham, W. F. (1989). Toward a conceptual framework for mixed-method evaluation designs. *Educational Evaluation & Policy Analysis*, *11*, 255–274. https://doi-org.ezproxy.mtsu.edu/10.3102/01623737011003255

Hatami, S. (2015). Teaching formulaic sequences in the ESL classroom. *TESOL Journal*, *6*(1), 112–129. https://doi-org.ezproxy.mtsu.edu/10.1002/tesj.143

Hattie, J. (1999, August 2). *Influences on student learning*. [Inaugural Lecture]. University of Auckland, New Zealand.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research, 77*(1), 81-112. https://www.jstor.org/stable/4624888

Honarzad, R., & Rassaei, E. (2019). The role of EFL learners' autonomy, motivation and self-efficacy in using technology-based out-of-class language learning activities. *JALT CALL Journal, 15*(3), 23-42.

Hung, S. A. (2016). Enhancing feedback provision through multimodal video technology. *Computers & Education, 98*, 90-101. http://doi.org/10.1016/j.compedu.2016.03.009

Hyland, F. (2001). Providing effective support: Investigating feedback to distance language learners. *Open Learning*, *16*(3), 233–247. https://doi-org.ezproxy.mtsu.edu/10.1080/02680510120084959

Hyland, K., & Tse, P. (2007). Is there an "Academic Vocabulary"? *TESOL Quarterly, 41*(2), 235-253. http://www.jstor.org/stable/40264352

Jeong, H., & Jiang, N. (2019). Representation and processing of lexical bundles: Evidence from word monitoring. *System*, *80*, 188–198. https://doi-org.ezproxy.mtsu.edu/10.1016/j.system.2018.11.009

Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, *91*(3), 433–445. https://doi-org.ezproxy.mtsu.edu/10.1111/j.1540-4781.2007.00589.x

Johnson, R. B., Onwuegbuzie, A. J., & Turner, L. A. (2007). Toward a definition of mixed methods research. *Journal Of Mixed Methods Research*, *1*(2), 112–133. https://doi-org.ezproxy.mtsu.edu/10.1177/1558689806298224

Kashiwagi, K., & Ito, Y. (2017). EFL learners' grammatical awareness through accumulating formulaic sequences of morphological structure (-ing). *English Language Teaching*, *10*(11), 60–75.

Khodadady, E., & Shamsaee, S. (2012). Formulaic sequences and their relationship with speaking and listening abilities. *English Language Teaching*, *5*(2), 39–49.

Khoram, A., Darabi Bazvand, A., & Sarkawt Sarhad, J. (2020). Error feedback in second language speaking: Investigating the impact of modalities of error feedback on

intermediate EFL students' speaking ability. *Eurasian Journal of Applied Linguistics*, *6*(1), 63–80.

Koizumi, R., & Matsuo, K. (1993). A longitudinal study of attitudes and motivation in learning English among Japanese seventh-grade students. *Japanese Psychological Research, 35*(1), 1–11.

Kormos, J. (2014). *Speech production and second language acquisition*. Routledge.

Krashen, S. D. (1982). *Principles and practice in second language acquisition*. http://www.sdkrashen.com/content/books/principles_and_practice.pdf

Levelt, W. J. M., Roelofs, A., & Meyer, A. S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences, 22*, 1–37.

Lightbown, P. M., & Spada, N. (1990). Focus-on-Form and corrective feedback in communicative language teaching: Effects on second language learning. *Studies in Second Language Acquisition*, *12*(4), 429–448.

Lindstromberg, S., Eyckmans, J., & Connabeera, R. (2016). A modified dictogloss for helping learners remember L2 academic English formulaic sequences for use in later writing. *English for Specific Purposes*, 41, 12-21. https://doi-org/10.1016/j.esp.2015.08.002

Long, M. H. (1983). 'Native speaker/non-native speaker conversation and the negotiation of comprehensible input,' *Applied Linguistics 4*:126–41.

Lynch, T. (2009). Responding to learners' perceptions of feedback: The use of comparators in second language speaking courses. *Innovation in Language Learning and Teaching*, *3*(2), 191–203. https://doi-org.ezproxy.mtsu.edu/10.1080/17501220802379109

Lysakowski, R. S., & Walberg, H. J. (1982). Instructional effects of cues, participation, and corrective feedback: A quantitative synthesis. *American Educational Research Journal, 19*(4), 559–578. https://doi.org/10.2307/1162544

Lyster, R., & Ranta, L. (1997). Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, *19*(1), 37–66.

Maas, C. (2017). Perceptions of multimodal learner-driven feedback in EAP. *Writing & Pedagogy, 9*(3), 487-516. https://doi.org/10.1558/wap.33292

McGuire, M., & Larson-Hall, J. (2017). Teaching formulaic sequences in the classroom: Effects on spoken fluency. *TESL Canada Journal*, *34 SI*(3), 1. https://doi-org.ezproxy.mtsu.edu/10.18806/tesl.v34i3.1271

Mackey, A. (2006) Feedback, noticing and instructed second language learning. *Applied Linguistics, 27* (3). pp. 405-430.

Martinez, R., & Schmitt, N. (2012). A phrasal expressions list. *Applied Linguistics*, *33*(3), 299–320. https://doi-org.ezproxy.mtsu.edu/10.1093/applin/ams010

MITK12. (2012, November 1). *Rock cycle* [YouTube Video]. https://www.khanacademy.org/partner-content/mit-k12/mit-k12-science/mit-k12-ea/v/rock-cycle

Mohammadi, M., & Enayati, B. (2018). The effects of lexical chunks teaching on EFL intermediate learners' speaking fluency. *International Journal of Instruction, 11*(3), 179-192. doi:10.12973/iji.2018.11313a

Morse, J., & Niehaus, L. (2009). *Mixed method design: Principles and procedures.* Left Coast Press Inc.

Nagy, W., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly, 47*(1), 91-108. http://doi.org/10.1002/RRQ.011

Nattinger, J. (1980). A lexical phrase grammar for ESL. *TESOL Quarterly*, *14*(3), 337–344. https://doi-org.ezproxy.mtsu.edu/10.2307/3586598

Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford: Oxford University Press.

Nicol, D., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. https://doi-org.ezproxy.mtsu.edu/10.1080/03075070600572090

O'Malley, J., & Valdez-Pierce, L. (1996). *Authentic assessment for English language learners: Practical approaches for teachers*. Longman.

Özkul, S., & Ortaçtepe, D. (2017). The use of video feedback in teaching process-approach EFL writing. *TESOL Journal,8*, 862-877. http://doi.org/10.1002/tesj.362

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and Communication* (pp. 191-227). Longman Group Limited.

Prefontaine, Y., & Kormos, J. (2016). A qualitative analysis of perceptions of fluency in second language French. *Iral-international Review of Applied Linguistics In Language Teaching*, *54*(2), 151–169. https://doi-org.ezproxy.mtsu.edu/10.1515/iral-2016-9995

Price, M., Handley, K., Millar, J., & O'Donovan, B. (2010). Feedback: All that effort, but what is the effect? *Assessment and Evaluation in Higher Education*, *35*(3), 277–289. https://doi-org.ezproxy.mtsu.edu/10.1080/02602930903541007

Puimège, E., & Peters, E. (2020). Learning formulaic sequences through viewing L2 television and factors that affect learning. *Studies in Second Language Acquisition*, *42*(3), 525–549. https://doi-org.ezproxy.mtsu.edu/10.1017/S027226311900055X

Purdie, N., & Hattie, J. (1996). Cultural differences in the use of strategies for self-regulated learning. *American Educational Research Journal, 33*(4), 845–871. https://doi.org/10.3102/00028312033004845

Rafieyan, V. (2018). Role of knowledge of formulaic sequences in language proficiency: Significance and ideal method of instruction. *Asian-Pacific Journal of Second and Foreign Language Education*

Römer, U. (2019). A corpus perspective on the development of verb constructions in second language learners. *International Journal of Corpus Linguistics*, *24*(3), 268–290.

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review, 65*(3), 395–412. http://doi.org/10.3138/cmlr.65.3.395

Royce, S. (1989). Formative assessment and the design of instructional systems. *Instructional Science*, *18*(2), 119–144.

Schmitt, N. (ed.) (2004). *Formulaic sequences: Acquisition, processing and use (Language Learning & Language Teaching, 9)*. John Benjamins Publishing Co.

Schoonenboom, J., & Johnson, R. B. (2017). How to construct a mixed methods research design. *Koelner Zeitschrift Fur Soziologie Und Sozialpsychologie*, *69*(2), 107. https://doi-org.ezproxy.mtsu.edu/10.1007/s11577-017-0454-1

Silva, M. L. (2012). Camtasia in the classroom: Student attitudes and preferences for video commentary or Microsoft Word comments during the revision process. *Computers and Composition, 29*, 1-22. http://doi.org/10.1016/j.compcom.2011.12.001

Simpson, R. C., S. L. Briggs, J. Ovens, and J. M. Swales. (2002) The Michigan Corpus of Academic Spoken English. Ann Arbor, MI: The Regents of the University of Michigan.

Simpson, R. & Mendis, D. (2003). A Corpus-Based Study of Idioms in Academic Speech. *TESOL Quarterly, 37*(3), 419–441. https://doi.org/10.2307/3588398

Simpson-Vlach, R. & Ellis, N. (2010). An Academic Formulas List: New Methods in Phraseology Research, *Applied Linguistics*, *31*(4), 487–512, https://doi-org.ezproxy.mtsu.edu/10.1093/applin/amp058

Siyanova-Chanturia, A., Conklin, K., Caffarra, S., Kaan, E., & van Heuven, W. J. B. (2017). Representation and processing of multi-word expressions in the brain. *Brain and Language*, *175*, 111–122. https://doi-org.ezproxy.mtsu.edu/10.1016/j.bandl.2017.10.004

Swain, M., & Lapkin, S. (1998). Interaction and second language learning: two adolescent French immersion students working together. *Modern Language Journal, 82*, 320–337. doi:10.2307/329959

Taguchi, N. (2008). Building language blocks in L2 Japanese: Chunk learning and the development of complexity and fluency in spoken production. *Foreign Language Annals, 40*(1), 132-156.

Takanashi, Y. (2004). TEFL and communication styles in Japanese culture. *Language Culture and Curriculum*, 17, 1-14. http://doi.org/10.1080/07908310408666678

Tavakoli, P. (2011). Pausing patterns: Differences between L2 learners and native speakers. ELT Journal, 65(1), 71–79. https://doi-org.ezproxy.mtsu.edu/10.1093/elt/ccq020

Tedick, D. J. (2002). Proficiency-oriented language instruction and assessment: Standards, philosophies, and considerations for assessment. In Minnesota Articulation Project, D. J. Tedick (Ed.), *Proficiency-oriented language instruction and assessment: A curriculum handbook for teachers (Rev Ed.)*. CARLA Working Paper Series. Minneapolis, MN: University of Minnesota, The Center for Advanced Research on Language Acquisition.

Tennessee Department of Education. (2018). *Every student succeeds act: Building on success in Tennessee. ESSA state plan*. https://www.tn.gov/content/dam/tn/education/documents/TN_ESSA_State_Plan_Approved.pdf

Tennessee Department of Education. (2019). [Unpublished raw data on the Change in TN WIDA ACCESS 2.0 Scores 2016-2017].

Tennessee State Board of Education. (2003/2020). English as a second language program policy. (Policy No. 3.207).

Thomson, H. (2017). Building Speaking Fluency with Multiword Expressions. TESL Canada Journal, 34(3), 26–53.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*(1), 84-119. http://10.1093/applin/17.1.84

Ushigusa, S. (2009). *The relationships between oral fluency, multiword units, and proficiency scores.* [Doctoral dissertation, Purdue University]. ProQuest Dissertations and Theses Global.

Vincent, C. (1996). *Singing to a star: The school meanings of second generation Salvadorean students.* [Doctoral dissertation, George Mason University]. ProQuest Dissertations and Theses Global.

Webb, N. (2005). *Web alignment tool*. Wisconsin Center of Educational Research. University of Wisconsin-Madison. https://nanopdf.com/download/downloading-wat-web-alignment-tool_pdf

WIDA. (2016). *WIDA Can Do Descriptors: Key Issues Edition, Grades 9-12*. Board of Regents of the University of Wisconsin System, on behalf of the WIDA Consortium.

WIDA. (2020a). *WIDA English language development standards framework, 2020 edition: Kindergarten–grade 12*. Board of Regents of the University of Wisconsin System.

WIDA. (2020b). *WIDA Speaking Rubric Grades 1-12*. Board of Regents of the University of Wisconsin System.

Wood, D. (2002). Formulaic language in acquisition and production: Implications for teaching. *TESL Canada Journal, 20*(1), 1-15.

Wood, D. (2006). Uses and functions of formulaic sequences in second-language speech: An exploration of the foundations of fluency. Canadian Modern Language Review, 63(1), 13–33.

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics*, *12*(1), 39–57.

Woolbert, C. H. (1922). Speaking and writing: A study of differences. *Quarterly Journal of Speech Education, 8*(3), 271.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge University Press.

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language & Communication*, *20*(1), 1–28. https://doi-org.ezproxy.mtsu.edu/10.1016/S0271-5309(99)00015-4

Xu, J. (2018). Measuring “spoken collocational competence” in communicative speaking assessment. *Language Assessment Quarterly, 15*(3), 255-272. doi:10.1080/15434303.2018.1482900

Yan, X. (2020). Unpacking the relationship between formulaic sequences and speech rluency on elicited imitation tasks: Proficiency level, sentence length, and fluency dimensions. *TESOL Quarterly, 54*(2), 460–487. https://doi-org.ezproxy.mtsu.edu/10.1002/tesq.556

Yilmaz, N., & Koban Koc, D. (2020). Developing pragmatic comprehension and production: Corpus-based teaching of formulaic sequences in an EFL setting. *Journal of Language and Linguistic Studies, 16*(1), 474–488.

APPENDICES

APPENDIX A: CODING TABLE FEEDBACK STUDIES

Study	Journal	Type of Study	Intervention	Grade/Age, Students, Sample Size	Treatment Description	Measure	Effect	Notes
Ali. (2016). Effectiveness of using screencast feedback on EFL students' writing and perception	English Language Teaching	Mixed Methods	Control group: written Experimental Group: screencast to higher order skills and written to lower order skills; online questionnaire about perception towards screencast feedback Random assignment to groups	university Freshmen, academic writing, 63	pre and posttest and questionnaire	pre posttest, questionnaire	Experimental group outperformed control group on higher and lower skills. Majority of students perceived screencast as positive. $df = 61$, $t = 8.46$, $p < 0.01$ 94% of students had positive attitudes towards screencast feedback	
Aljaser (2019). The effectiveness of e-learning environment in developing academic achievement and the attitude to learn English among primary students	Turkish Online Journal of Distance Education	Quasi-Experimental	E-learning environment	5th grade, EFL Saudi Arabia, 30	Control group received classroom instruction; treatment group received e-learning over the internet.	Pre and post-Achievement test and attitude Test	$\eta 2 = .39$ (achievement)	Students in e-learning did better and had a better attitude about learning English.
Alvira. (2016). The impact of oral and written feedback on EFL writers with the use of screencasts	Profile	Action Research	150 word paragraph	university, EFL learners in Colombia, 18	150 word paragraphs; pre-and post-study questionnaire and students' writing samples. All students were given screencast and written feedback at the same time.	pre- and post-study questionnaires	Writing improved in measured areas; student perceptions confirmed preference of screencast feedback	Students widely accepted screencasting, positive results in improvement of writing at the paragraph level, motivational strategy. Student autonomy increased.
Faramarzi, S., Tabrizi, H., & Chalak, A. (2019). Learners' perceptions and attitudes towards L2 vodcasting tasks in an e-learning project.	Teaching English with Technology	Case Study	5 vodcast tasks	university undergraduate, EFL Iranian Students learning English, 120	Students were pushed vodcasting tasks and could complete them on their own time. After they completed a learner engagement questionnaire.	Learner Engagement Questionnaire	$M = 4.47$ (above four is positive)	Students were happy with the e-learning aspect, motivated to learn English. The types of tasks allowed them to work on multiple skills simultaneously.
Ghosn-Chelala & Al-Chibani. (2018). Screencasting: supportive feedback for EFL remedial writing students.	The International Journal of Information and Learning Technology	Case Study	Screencast feedback on one essay sample for clarity, learning preferences, and engagement	college, Arabic-speaking English learners, 8	Screencast feedback on one writing sample	Perspectives survey and informal group discussion	Students responded positively to screencast feedback	Students preferred screencast for better engagement, support of learning preferences, clarity
Study	Journal	Type of Study	Intervention	Grade/Age, Students, Sample Size	Treatment Description	Measure	Effect	Notes

Honarzad & Rassaei. (2019). The role of EFL learners' autonomy, motivation and self-efficacy in using technology-based out-of-class language learning activities.	JALT CALL Journal	Case Study	Questionnaires	university graduate students, EFL Iranian-speaking English learners, 100	Four questionnaires administered by the researchers	Technology-based out-of-class language learning activities questionnaire; English learning motivation questionnaire; Learner autonomy questionnaire; The general perceived self-efficacy scale	Motivation p = .000; autonomy p = .001; self-efficacy p < .010	Results consistent with other studies indicating technology-based activities increase learner motivation, autonomy, and self-efficacy.
Özkul & Ortaçtepe. (2017). The use of video feedback in teacher process-approach EFL writing	TESOL Journal	Experimental	Random assignment to group by entire class Control: written feedback Experiment: video feedback	university, EFL writing students, 47	ll	5 assignments, control and experiment feedback groups	p = .030, statistically significant in 3 out of 5 assignments	Students reported positive responses to video feedback

APPENDIX B: CODING TABLE FS AND FLUENCY

Appendix B							
Coding Table							
Study	Journal	Type of Study	Grade/Age, Students, Sample	Treatment	Measure	Effect	Notes
Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M.(2006). "Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test"	Language Teaching Research	experimental, exploratory study	college students *N* = 32, aged 19-22, modern language speakers, upper-intermediate proficiency	listening and reading materials, control group BAU, experiment group extra attention to "phrase-noticing"	oral proficiency interviews; dialogs	Mann Whitney U, experimental group over control group: $p < .050$, $U = 70$, $M = 14.44$	
Eyckmans, J., Boers, F., & Lindstrombeg, S. (2015). "The impact of imposing processing strategies on L2 learnrs' deliberate study of lexical phases."	System 56	Quasi-experimental	*N* = 65, EFL learners in secondary school (age 13-14), L1 Dutch	noticing verb + noun phrases: study list of FS, identify incongruencies, identify alliterations; 3 groups	Memorize 32 target phrases; no conversations/just written test	**Alliterative over incongruent**: $p < .0001$, $d = 0.32$; **Alliterative over no intervention** $p = .008$, $d = 0.17$	
Khodadady, E. & Shamsaee, S. (2012). "Formulaic sequences and their relationship with speaking and listening abilitites."	English Language Teaching		*N* = 41 EFL university females,in Iran, age 18-35	IELTS speaking and listening specimens	number of 8 types of FS, frequency analysis, discriminant function analysis		
McGuire, M., & Larson-Hall, J. (2017). "Teaching formulaic sequences in the classroom: Effects on spoken fluency."	TESL Canada Journal	case study with control and treatment groups	*N* = 19, university in the US, English learners (Thai, Chinese, and Japanese speakers)	control group: focus on single words and grammar, treatment group: focus on FS	speech rate, mean length of fluent run, number of syllables of FS; conversation/dialogs	**Speech Rate**: Treatment: $p = .0003$, $d = 1.3$ control: $p = .090$. $d = 0.06$ **Mean Length of Run**: Treatment: p = .006, d = 1.1; Control: p = .340, d = 0.17 **Subjective**: treatment p = .190, d = 0.26, control p = .150, d = 0.71 **FS Ratio**: treatment p = .0009, d = 1.2, control p = .800, d = 0.20	

Mohammadi, M. & Enayati, B. (2018). "The effects of lexical chunks teaching on EFL intermediate learners' speaking fluency."	International Journal of Instruction	experimental	*N* = 60, EFL L1 Persian, age 12-17	**Treatment**: focus on Lexical Chunks, **control:** focus on translation and grammar; Lexical chunk learning and speaking fluency	Test of Lexical Chunks, 10 minute interview; number of words per T-units; conversation/dialogs	**Fluency treatment**: $p = .000$; **fluency control** $p = .566$	
Rafieyan, V. (2018). "Role of knowledge of formulaic sequences in language proficiency: Significance and ideal method of instruction."	Asian-Pacific Journal of Second and Foreign Language Education	experimental	*N* = 42, university, Japanese-speakers, age 18-20	oral production discourse completion task: 30 scenarios; low, medium, and high proficiency speakers	focus on forms: forms in isolation; focus on form: forms in context; dialogs	**Correlations**: the higher the proficiency, the better the use of FS $p = .000$, *partial eta squared* = 0.92; no difference between groups for treatment $p = .700$	
Tavakoli, P. (2011). "Pausing patterns: Differences between L2 learners and native speakers."	ELT Journal	comparative study	40 L1 English, ages 19-60; 40 L2 English, ages 19-35; university in London	4 oral narrative talks	number of pauses mid-clause; number of pauses end-clause; total silence mid-clause; total silence end-clause	L2 learners paused more frequently mid-clause than L1 speakers $t = 2.25$, $p < .030$	
Yan. (2020). "Unpacking the relationship between formulaic sequences and speech fluency on elicited imitation tasks: Proficiency level, sentence length, and fluency dimensions."	Tesol Quarterly	Quasi-experimental	university freshman, US university, 18-25, Chinese speaking N = 252, and 17 L1 English speakers	elicited imitation task; students had 20 seconds to repeat each of the 24 sentences	Formulicity, proficiency, and difficulty; no dialog/just reading sentences	did not increase speech rate but decreased silent pauses	pauses: $F(1, 1684) = 11.847$, $p < .001$ speech rate: $F(2,6157) = 2.635$, $p = .072$
Yilmaz, N., & Korban Koc, D. (2020). "Developing pragmatic comprehension and production: Corpus-based teaching of formulaic sequences in an EFL setting."	Journal of Language and Linguistic Studies	Quasi-experimental	N = 35, university, Turkish speaking English learners	19 FS for agree, disagree, self-clarify, and other-clarify *experimental group: *corpus-based teaching; control group: traditional instruction both with listening and reading	dialogs	pragmatic comprehension and production; dialogs	experimental group pre-test and immediate post-test: $p = .001$; $r > 0.5$, $r = 1.99$ control group pre-test and immediate post test, $p = 1.00$ post-test and delayed post-test experimental group $p = .012$, drop in mean from 2.74 to 2.05

APPENDIX C: IRB APPROVAL

IRB
INSTITUTIONAL REVIEW BOARD
Office of Research Compliance,
010A Sam Ingram Building,
2269 Middle Tennessee Blvd
Murfreesboro, TN 37129

MIDDLE TENNESSEE STATE UNIVERSITY

IRBN001 - EXPEDITED PROTOCOL APPROVAL NOTICE

Tuesday, October 13, 2020

Principal Investigator	**Jennifer Francois** (Student)
Faculty Advisor	Mohammed Albakry
Co-Investigators	NONE
Investigator Email(s)	*jlm2ad@mtmail.mtsu.edu; mohammed.albakry@mtsu.edu*
Department	Literacy Studies
Protocol Title	***Effectiveness of multimodal feedback in improving speaking assessment performance of middle and high school English learners***
Protocol ID	**19-2276**

Dear Investigator(s),

The above identified research proposal has been reviewed by the MTSU Institutional Review Board (IRB) through the **EXPEDITED** mechanism under 45 CFR 46.110 and 21 CFR 56.110 within the PRIMARY category *(7) Research on individual or group characteristics or behavior* and SECONDARY category *(6) Collection of data from media.* A summary of the IRB action and other particulars in regard to this protocol application is tabulated below:

IRB Action	**APPROVED for ONE YEAR**		
Date of Expiration	**8/31/2021**	Date of Approval	8/23/19
Sample Size	200 (TWO HUNDRED)		
Participant Pool	Primary Classification: **Minors (younger than 18 years of age)** Specific Classification: **English learner students in Williamson County (grades 6-12: Age 12-17)**		
Exceptions	1. Verbal participant recruitment permitted. 2. Voice recording is permitted for data collection (with restriction)		
Restrictions	**1. Mandatory signed parental consent followed by independent child assent; The participants must be given Parts A and B signed by PI/FA.** **2. All identifiable data/artifacts that include audio/video data, photographs, handwriting samples, and etc., must be used only for research purpose and they must be destroyed after data processing.** **3. Mandatory Final report (refer last page).** **4. Spanish translations of the IRB templates have not been approved. .**		
Approved Templates	MTSU templates: Combined parental consent/child assent template Non-MTSU Templates: Verbal recruitment script		
Comments	NONE		

Post-approval Actions

The investigator(s) indicated in this notification should read and abide by all of the post-approval conditions (https://www.mtsu.edu/irb/FAQ/PostApprovalResponsibilities.php) imposed with this approval. Any unanticipated harms to participants, adverse events or compliance breach must be reported to the Office of Compliance by calling 615-494-8918 within 48 hours of the incident. All amendments to this protocol, including adding/removing researchers, must be approved by the IRB before they can be implemented.

Continuing Review (Follow the Schedule Below)

This protocol can be continued for up to THREE years by requesting a continuing review before 8/31/2021. Refer to the following schedule to plan your annual progress report; **REMINDERS WILL NOT BE SENT**. Failure to obtain an approval for continuation will result in cancellation of this protocol.

Reporting Period	Continuing Review	
First year report	10/13/2020 IRBCR2021-071 A CR conducted using the progress report submitted by the PI determined that the protocol can be continued for an additional year. The CR request was delayed due to COVID-19.	
Reporting Period	**Requisition Deadline**	**IRB Comments**
Second year report	8/31/2021	NOT COMPLETED
Final report	8/31/2022	NOT COMPLETED

Post-approval Protocol Amendments:

Only two procedural amendment requests will be entertained per year. *In addition, the researchers can request amendments during continuing review. This amendment restriction does not apply to minor changes such as language usage and addition/removal of research personnel. .*

Date	Amendment(s)	IRB Comments
NONE	NONE.	NONE

Other Post-approval Actions:

Date	IRB Action(s)	IRB Comments
08/23/2019	1. The PI must submit the permission letter from Wilson County 2. Spanish translation of all of the approved templates must be submitted for IRB records; the Spanish version must not be mere one-to-one stranslation but should reflect cultural and ethnic equivalent of the approved English template.	Refer to action on 08/27/2019
08/27/2019	Permission letter from WCS has been submitted. The PI decided to discontinue her efforts to use Spanish translation.	One year approval is granted
10/13/2020	The student PIs name changed from Jennifer Myers to Jennifer Francois.	IRBCR2021-071

Mandatory Data Storage Requirement: All research-related records (signed consent forms, investigator training and etc.) must be retained by the PI or the faculty advisor (if the PI is a student) at the secure location mentioned in the protocol application. The data must be stored for at least three (3) years after the study is closed. Subsequently, the data may be destroyed in a manner that maintains confidentiality and anonymity of the research subjects.

The MTSU IRB reserves the right to modify/update the approval criteria or change/cancel the terms listed in this letter without prior notice. Be advised that IRB also reserves the right to inspect or audit your records if needed.

Sincerely,

Institutional Review Board
Middle Tennessee State University

Quick Links:

- **Post-approval Responsibilities**: http://www.mtsu.edu/irb/FAQ/PostApprovalResponsibilities.php
- **Expedited Procedures**: http://www.mtsu.edu/irb/FAQ/PostApprovalResponsibilities.php

Primary Investigator(s)	Jennifer François		
Contact information	Jlm2ad@mtmail.mtsu.edu, jennifer.francois@wcs.edu		
Department Institution	Middle Tennessee State University		
Faculty Advisor	Dr. Mohammed Albakry	Department	Literacy Studies

Child's Name: _____

The following information is provided to you because your child qualifies to participate in a research project. Please read this disclosure document carefully and feel free to ask any questions before you agree to enroll your child. The university student must adequately answer all of your questions before your child can be enrolled. The researcher MUST NOT enroll your child without an active consent from you. Also, a copy of this consent document, duly signed by the university student, must be provided to you for future reference.

Your child's participation in this project is absolutely voluntary. You or your child can withdraw from this project at any time. In the event new information becomes available that may affect the risks or benefits associated with this project or your willingness to participate in it, you will be notified so that you can make an informed decision whether or not to continue your participation in this project.

For additional information about giving consent or your rights as a participant in this project, please feel free to contact me at jlm2ad@mtmail.mtsu.edu or Jennifer.francois@wcs.edu or Tel. 629-214-3382 or Dr. Mohammed Albakry at mohammed.albakry@mtsu.edu or Tel. 615- 494-8658.

Parental Consent

Please read this section and sign if you wish to enroll your child. The researcher will not enroll your child without your physical signature.

1. **Purpose of the project:**
   Your child is being asked to participate in a research project as a part of the researcher's dissertation. Your child's participation will help provide information about how students respond to academic feedback.

2. **General description of procedures to be followed and approximate duration of the project:**
   Your child will participate in two speaking assessment tasks and a student attitude survey. The project will take 4-5 weeks to complete.

3. **What are we planning to do to your child in this project?**
   Your child will participate in two speaking assessment tasks using school technology to record a response. The researcher will provide feedback to your child about the responses. Then your child will be asked answer questions about the feedback.

4. **What will your child be asked to do in this project?**
   Your child will be asked two speaking assessment questions and will record the answers using school technology. Your child will also be asked to answer questions about the experience.

5. **What are we planning to do with the data collected using your child?**
   I will use the information as part of my dissertation project. I will use the data to analyze which mode of feedback is more effective and which mode of feedback students prefer.

6. **What are your expected costs, effort and time commitment:**
   There will be no cost to you or your child. Your child will be asked to complete two speaking assessment tasks and a survey during regular ESL class.

7. **What are the potential discomforts, inconveniences, and/or possible risks that can be reasonably expected as a result of participation in this project?**
   For the Child: There are no anticipated discomforts or risk for your child
   For you the Parent: There are no anticipated discomforts or risk for you as a parent.

8. **How will you or your child be compensated for enrolling in this project?**
   There will be no compensation offered to your child for participation.

9. **What are the anticipated benefits from this project?**
   This project will help your student prepare for the speaking assessment test given each spring. It will also provide valuable information about how students respond to feedback and help teachers improve how they teach students in the future.

10. **Are there any alternatives to this project such that you or/and your child could receive the same benefits?**
    No

11. **What happens if you choose to withdraw from project participation?**
    You may withdraw your child at any time for any reason with no negative impact on you or your child.

12. **Can you or/and your child stop the participation any time after initially agreeing to give consent/assent?**
    Yes, your child can stop participation at any time.

13. **Contact Information.** If you should have any questions about this research project or possibly injury, please feel free to contact Dr. Mohammed Albakry by telephone 615-494-8658 or by email mohammed.albakry@mtsu.edu.

14. **Confidentiality.** All efforts, within reason, will be made to keep the personal information in your child's research record private but total privacy cannot be promised. Your information may be shared with MTSU or the government, such as the Middle Tennessee State University Institutional Review Board, Federal Government Office for Human Research Protections, if you or someone else is in danger or if we are required to do so by law.

Consent obtained by:

June 26, 2020
Date

Jennifer Francois
University Student's Signature

Jennifer François, ESL Teacher
Print Name and Title of the University Student

CHILD ASSENT

(To be retained by the participating child who is over 12 years of age)

Primary Investigator(s)	Jennifer François		
Contact information	Jlm2ad@mtmail.mtsu.edu, jennifer.francois@wcs.edu		
Department Institution	Literacy Studies		
Faculty Advisor	Dr. Mohammed Albakry	Department	Literacy Studies

Child's Name: _____

The following information is provided to you because your parents/guardians have agreed to enroll in the above identified research project. Please read this sheet carefully and feel free to ask any questions before you agree to enroll. The researcher must answer all of your questions before he/she asks you to do anything. Before you start:

- Make sure this sheet is signed by the researcher.
- Your participation is absolutely voluntary; you can decline any time and your parents/guardians will not be notified.
- You are entitled to decline or withdraw at any time.
- Any new information on this research will be notified to you and you can decide whether to continue your participation based on the new information.

1. **Why is the researcher doing this project?**
   The researcher is conducting a research project as part of her dissertation for her PhD in Literacy Studies at MTSU. The data will be used to analyze which mode of feedback is more effective and which mode of feedback students prefer.

2. **What will the researcher do and how long will it take?**
   The researcher will facilitate the administration of two speaking assessment tasks and a student attitude survey. The project will last 4-5 weeks.

3. **Do I have to be in this research project, and can I stop if I want to?**
   You can stop the project at any time.

4. **Will anyone know that I am in this project?**
   Only myself, your parents, and my professor will know that you are participating in this project.

5. **How will this research help me or/and other people?**
   This project will help teachers learn which types of feedback students prefer and which ones work best. This project will help you practice speaking assessment tasks like you will have to do for spring testing.

6. **Can I do something else instead of this project?**
   You do not have to work with me if you do not want to. It is entirely up to you.

7. **Who do I talk to if I have questions?**
   If you have any questions at any time, please let me know and I will answer them.

June 26, 2020 — Jennifer François

Date — University Student's Signature

Jennifer François, ESL Teacher

Print Name and Title of the University Student

Signature Section

University Student	Jennifer François
Contact information	jfm2ad@mtmail.mtsu.edu, jennifer.francois@wcs.edu
Department Institution	Middle Tennessee State University
Faculty Advisor	Dr. Mohammed Albakry Department: Literacy Studies

Child's Name: _____

PARENT SECTION

☐No ☐Yes I have read the parental consent document pertaining to the above identified research.
☐No ☐Yes The research procedures to be conducted have been explained to me.
☐No ☐Yes I understand each part of the project, and all my questions have been answered.
☐No ☐Yes I received a signed copy, and I am aware of the potential risks of the project.

By signing below, I give permission for my child, whose name is identified above, to participate in this project. I understand I can withdraw my child from this project at any time without facing any consequences.

Date _____ Signature of the Parent _____

CHILD SECTION

☐No ☐Yes I have read this child assent document, and I received a signed copy.
☐No ☐Yes The researcher explained what they planned to do, and all my questions were answered.
☐No ☐Yes I understand what I was told.
☐No ☐Yes I know the risks, and I also know I can withdraw at any time.

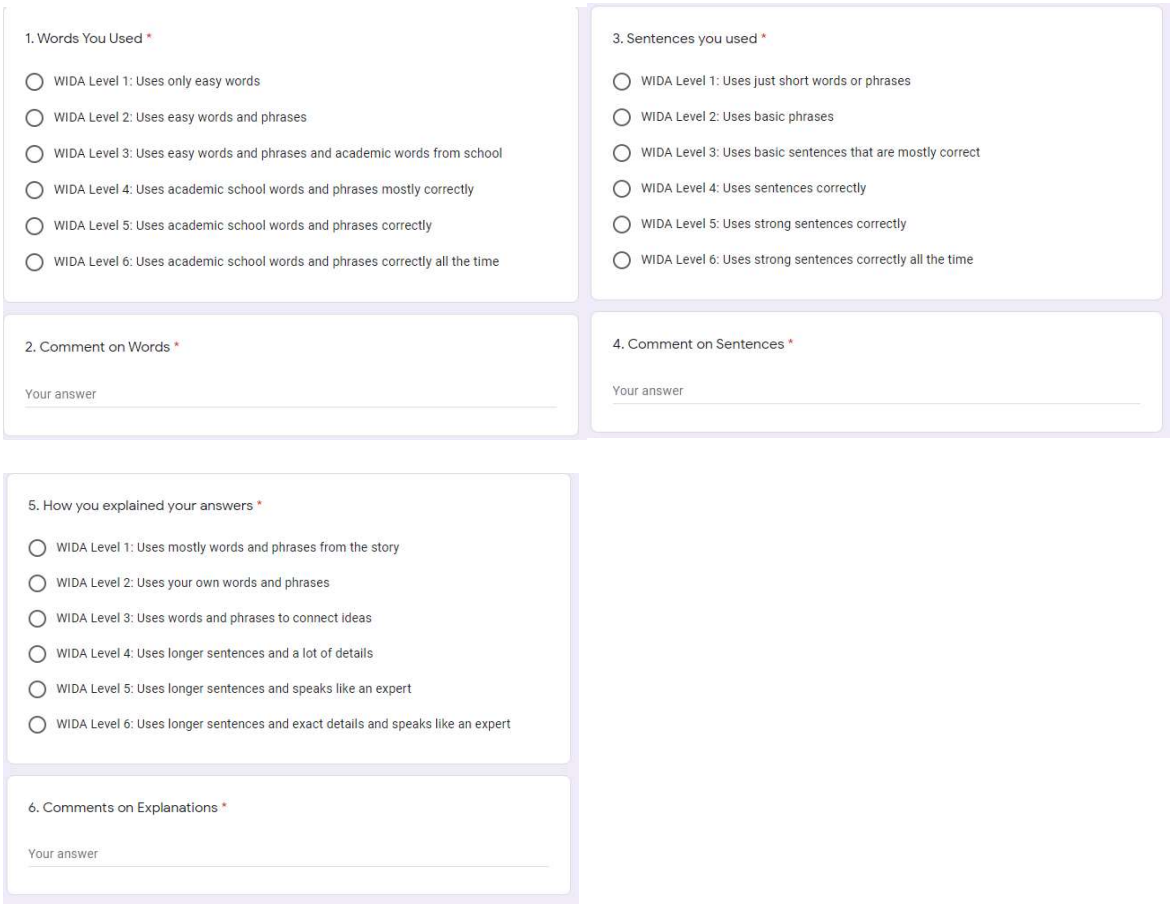
Date _____ Signature of the Child _____

Parental Consent obtained by: Jennifer François
Name
Jennifer Francois
Signature

APPENDIX D: WIDA SPEAKING INTERPRETIVE RUBRIC

WIDA Speaking Interpretive Rubric Grades 1-12			
	Discourse Level Linguistic Complexity	**Sentence Level Language Forms**	**Word/Phrase Level Vocabulary Usage**
Level 6 Reaching	Response is fully comprehensible, fluent, and appropriate to purpose, situation and audience; comparable to the speech of English proficient students meeting college- and career-readiness standards; characterized by:		
	--sustained, connected oral language characterized by confidence, coherence, and precision in the expression of ideas tailored to purpose, situation, and audience --Clear evidence of consistency in conveying an appropriate perspective and register	--a full range of oral phrase and sentence patterns and grammatical structures matched to content area topics --controlled, skilled use of oral language to convey meaning, including for effect	--consistent usage of just the right word or expression in just the right context related to content area topics --facility with precise vocabulary usage in general, specific, or technical language
Level 5 Bridging	Response is comprehensible, fluent, and generally related to purpose; generally comparable to the speech of English proficient peers; characterized by:		
	--sustained, connected oral language that shows appropriate and coherent expression of ideas related to purpose, situation, and audience --clear evidence of conveying an appropriate perspective and register	--a broad range of oral phrase and sentence patterns and grammatical structures matched to the content area topic --controlled, fluid use of oral language to convey meaning, including for effect	--usage of technical and abstract content-area words and expressions as appropriate --usage of words and expressions with precise meaning related to content area topics as appropriate --vocabulary usage that fulfills the speaking purpose
Level 4 Expanding	Response is generally comprehensible, fluent, and related to purpose; characterized by:		
	--connected oral language that supports the expression of expanded or related ideas through emerging coherence, detail, and clarity --some evidence of conveying an appropriate perspective and register	--a range of oral phrase and sentence patterns and grammatical structures characteristic of the content area --generally controlled and fluid use of oral language to convey meaning	--usage of specific and some technical content-area words and expressions as appropriate --usage of words and expressions with multiple meanings or common idioms across content areas as appropriate --vocabulary usage that generally fulfills the speaking purpose
Level 3 Developing	Response is generally comprehensible (though comprehensibility and fluency ay from time to time be compromised in more complex speech); characterized by:		
	--oral language that shows the development on connected language in the expression of an expanded idea or multiple related ideas --evidence of a developing sense of perspective and register	--developing range of oral phrase and sentence patterns and grammatical structures common to content areas --developing control in use of oral language to convey meaning	--usage of some specific content words and expressions as appropriate --usage of words or expressions used frequently in content areas, as appropriate --vocabulary usage that attempts to fulfill the speaking purpose.
Level 2 Emerging	Response is generally comprehensible (though comprehensibility and fluency ay often be compromised in more complex speech characterized by:		
	--oral language that shows emerging expression of ideas; some attempt at connecting ideas may at times be evident --some amount of language that may be repeated from the prompt	--chunks of language, repetitive oral phrase patterns, and formulaic grammatical structures used in social and instructional situations or across content areas --variable control in use of oral language to convey meaning	--usage of general content words and expressions --usage of social and instructional words and expressions across content areas --possible usage of general vocabulary where more specific language is needed
Level 1 Entering	Response is generally comprehensible (though comprehensibility and fluency may be significantly compromised in language beyond words, oral phrases, or memorized chunks); characterized by:		
	--words, oral phrases, or memorized chunks of oral language used to represent ideas --varying amounts of language that may be repeated from the prompt	--words, chunks of language, or simple phrasal patterns associated with common social and instructional situations --occasional control in use of oral language to convey meaning	--usage of highest frequency general content-related words --usage of everyday social and instructional words and expressions

APPENDIX E: STUDENT FRIENDLY WIDA RUBRIC



1. Words You Used *

- WIDA Level 1: Uses only easy words
- WIDA Level 2: Uses easy words and phrases
- WIDA Level 3: Uses easy words and phrases and academic words from school
- WIDA Level 4: Uses academic school words and phrases mostly correctly
- WIDA Level 5: Uses academic school words and phrases correctly
- WIDA Level 6: Uses academic school words and phrases correctly all the time

3. Sentences you used *

- WIDA Level 1: Uses just short words or phrases
- WIDA Level 2: Uses basic phrases
- WIDA Level 3: Uses basic sentences that are mostly correct
- WIDA Level 4: Uses sentences correctly
- WIDA Level 5: Uses strong sentences correctly
- WIDA Level 6: Uses strong sentences correctly all the time

2. Comment on Words *

Your answer

4. Comment on Sentences *

Your answer

5. How you explained your answers *

- WIDA Level 1: Uses mostly words and phrases from the story
- WIDA Level 2: Uses your own words and phrases
- WIDA Level 3: Uses words and phrases to connect ideas
- WIDA Level 4: Uses longer sentences and a lot of details
- WIDA Level 5: Uses longer sentences and speaks like an expert
- WIDA Level 6: Uses longer sentences and exact details and speaks like an expert

6. Comments on Explanations *

Your answer

APPENDIX F: SPEAKING TASK TRANSCRIPTS

Task One

00:00
metamorphosis and biology means the
00:03
process of transformation from an
00:05
immature form to an adult form in two or
00:07
more distinct stages good examples are
00:09
insects life for most insects begins as
00:13
a larva or nymph then progresses to the
00:15
pupa stage and ends as an adult there
00:18
are two main types of metamorphosis and
00:21
insects incomplete metamorphosis and
00:24
complete metamorphosis insects change
00:28
how they look and what they can do when
00:30
they grow some insects with incomplete
00:33
metamorphosis have three different life
00:35
stages these insects start as eggs which
00:38
are sometimes so small you cannot see
00:40
them when the egg hatches a larva or
00:43
nymph comes out nymphs are just baby
00:45
insects most of the time the nymph looks
00:48

similar to the adult but is smaller may
00:51
have different coloration and does not
00:52
have wings the nymph goes through stage
00:55
called instars
00:56
shedding its skin at each stage finally
01:00
it changes into a mature adult with
01:02
wings
01:03
some insect nymphs are aquatic which
01:05
means they live in water these names
01:08
usually have gills and look very
01:09
different from the adults they will turn
01:11
into nymphs that live in water are
01:14
called naiads dragonflies are an example
01:17
of incomplete metamorphosis when all
01:20
insects grow they change how they look
01:22
insects that have complete metamorphosis
01:25
have four different life stages these
01:28
insects start as eggs which are very
01:30
small the eggs hatch and a larva comes
01:33
out the larva looks like a worm and eats
01:36

and eats so that it can grow much bigger
01:38
when the larva has grown it changes into
01:40
a pupa the pupae usually cannot move or
01:44
eat the pupa is a special time when the
01:47
insect is changing into an adult that
01:49
will look very different from the larva
01:51
or the pupa the pupae are inside cocoons
01:55
when the pupa opens the adult insect
01:58
comes out many insects have a life cycle
02:02
so as a review remember that in insects
02:07
that have an incomplete metamorphosis
02:09
there are only three stages the egg
02:12
stage the
02:13
larvae stage in the adult stage and in
02:16
insects that undergo a complete
02:18
metamorphosis like a monarch butterfly
02:20
go through four stages the egg stage the
02:24
larva stage the pupa stage and the adult
02:27
stage on behalf of Layne and I thank you
02:31
for watching
02:39

(ACrameThirtyTwo, 2011, November 14)

Task Two

00:00
(printing machine)
00:08
(rock music)
00:09
- [Voiceover] I bet you thought rocks are just rocks, right?
00:11
(record scratching, music stops)
00:13
Nope. There are three major types of rocks:
00:15
sedimentary,
00:17
igneous,
00:18
and metamorphic.
00:19
But the coolest thing about rocks is that each one
00:22
has the ability to change into the other kind.
00:26
- Huh?
00:27
- How is that possible?
00:29
(rock music)
00:36
(record scratching, music stops)
00:37
- [Voiceover] Sedimentary, metamorphic, and igneous rocks
00:41
change into each other in a process we call the rock cycle.
00:45
(rock music)
00:47
(music stops)
00:48
No, not that kind of rock.

00:50
This kind of rock.
00:52
(rock music)
00:55
(music stops)
00:56
Yeah, that's more like it.
00:58
The first type of rock we'll talk about is sedimentary.
01:01
On the surface of the Earth,
01:03
wind and water break down rock into tiny pieces.
01:06
Those pieces might collect in a riverbed, on a flood plain,
01:10
be swept into sand dunes, or collect on the ground.
01:14
Over time, layers of these rock fragments build up
01:17
and start to weigh down on one another.
01:19
Eventually they get fused together
01:21
to form sedimentary rocks.
01:23
The cool thing is that, if you look closely,
01:25
you can still see pieces of the original rocks
01:28
or sediment that were bound together.
01:31
- [Voiceover] Let's do a demo.
01:33
For our rocks, we're gonna use jelly beans.
01:35
Each flavor of jelly bean represents
01:37
a rock or a mineral that has been broken down

01:39
by wind and water through a process called erosion.
01:42
We put our jelly beans in this bowl,
01:44
and add some honey and corn starch,
01:46
they're the bonding agents to hold our pieces together,
01:49
kind of like glue for rocks.
01:51
A little time and pressure has turned
01:53
our jelly bean pieces of sediment
01:55
into a brand-new rock.
01:57
- [Voiceover] So what happens if you apply
01:58
both heat and pressure?
02:00
It becomes a metamorphic rock.
02:02
Metamorphic rock may form by friction
02:04
of the Earth's shifting crust,
02:06
pressure deep within the Earth,
02:07
or even radioactive decay.
02:10
The heat and pressure cause the rock structure to change
02:13
so it takes on a new form.
02:15
Even though it's changed, you can often
02:16
still see structures of its original components.
02:20
- [Voiceover] Let's take our sedimentary jelly bean rock

02:22
and turn it into a metamorphic one with heat and pressure.
02:26
To add pressure, we'll put this heavy pot on top.
02:29
For heat, we'll stick it in the oven for about 30 minutes.
02:32
After it's cooled, you can see how our jelly bean rock
02:35
has formed a more solid unit.
02:37
However, you can still see the individual pieces of candy,
02:40
but the structure has fundamentally changed.
02:44
- [Voiceover] The third type of rock
02:46
in the rock cycle is igneous.
02:48
When rocks get super-heated deep within the Earth,
02:51
they melt and form a liquid called magma.
02:54
If magma rises to the surface or moves up
02:56
in the Earth's crust, it begins to cool.
02:59
Igneous rocks have a uniform structure throughout,
03:02
but will have different properties depending on whether
03:04
they cooled on the Earth's surface or within the crust.
03:09
- [Voiceover] To turn our jelly bean metamorphic rock
03:11
into an igneous rock, we're gonna melt it
03:13
in this pot of boiling water.
03:15
When our rock is cooled,

03:16
you can see how all the different pieces
03:18
combined to make an igneous rock,
03:20
with uniform structure throughout.
03:22
Pretty cool, huh?
03:24
- [Voiceover] But this is only part of the story.
03:26
We showed you one path for the rock cycle,
03:28
but really any rock can go from one type to another.
03:32
For example, igneous rocks can turn
03:35
into either metamorphic or sedimentary.
03:37
And metamorphic rocks don't have to become igneous rocks,
03:41
they can be broken down again and become sedimentary.
03:44
Or, the sedimentary rocks can get pushed deep
03:46
within the Earth to form igneous.
03:49
See? All of the rock types are connected,
03:51
making a cycle that never ends.
03:55
The end!
03:57
(rock music)
04:00
♫ - We are the rocks of the world
04:02
♫ Whoa, ho ho, rocking
04:05
♫ We are rocking so much

04:07
♫ Until the night
04:12
♫ Rocking baby, whoo oh
04:15
♫ Rock it to the beat
04:16
♫ Rockin' baby
04:18
♫ Whoo, whoo
04:20
♫ Rocking until the sun comes up
(MITK12, 2012)