Texting and Face-to-Face Speaking in Task-Based Language Teaching

TAKASE Nami

A Thesis Submitted in Partial Fulfillment of

the Requirements for the Degree of

Doctor of Philosophy

Graduate School of

Languages and Cultures,

Nagoya University

September 2022

Acknowledgments

This work would not have been possible without the exceptional support of my supervisor, Professor Ozeki. His expertise and passion for the field of computer-assisted language learning have inspired me to pursue this research topic. The encouragement I received to carry out my research allowed me to complete it, even when I had to interrupt it for work or family reasons.

I am especially indebted to Professor Sugiura for his expertise and for providing me with invaluable support in this research. His suggestions, sincerity, and vision have encouraged me to grow as a researcher. It was a great honor to work under his guidance.

I would also like to express my deepest gratitude to Associate Professor Murao, for her insightful comments and for supporting me throughout this long journey. Your constructive suggestions and persistent guidance have helped me improve my work. Without your support, this dissertation would not have been possible.

I would like to extend my immense gratitude to Adam Jenkins of Shizuoka Institute of Science and Technology for bolstering my first step forward on this research. Your passion for online technology and technical skills enlightened me as I began conducting this research.

I also would like to thank many other faculty members at my former and current workplaces for kindly accepting to be raters for analyzing the data and encouraging me to complete the research. I am deeply grateful that I was able to be part of the faculty.

I would like to express my thanks to my classmates, who gave me input and ideas for improvement. Moreover, to all my students who participated in this research, your cooperation has helped improve future pedagogy in this field. I appreciate your understanding and your interest in learning a foreign language.

Finally, I wish to thank my family for allowing me to continue teaching and conducting research.

ABSTRACT

Texting (text messaging or chatting) is a form of communication that is widely used in our daily lives. The purpose of this study is to investigate whether texting as a communication medium is effective in task-based language teaching compared to face-to-face for novice language learners (CEFR A1-B1). Language learners at an early stage struggle to communicate in a foreign language. The use of online communication may alleviate their problems and expand their opportunities to practice the target language. This study examined the effects of texting and learning English as a second language through tasks on lexical accuracy, syntactic complexity, grammatical accuracy, lexical diversity, and fluency in the subsequent English speaking test and compared them with face-to-face learning.

Chapter 1 explains the inherent feature of texting and the differences between face-to-face communication. Previous research has found linguistic similarities between texting and spoken language (Jonsson, 2016). In addition, texting supports learners with lower phonological abilities, suggesting the use of texting with a broader group of language learners (Payne &Whitney, 2002).

Chapter 2 reviews a series of previous studies on the usefulness of texting for second language acquisition. Findings suggest that self-repairs during texting improve the efficiency of second language acquisition because the information is given visually, and the pace of exchange is slower (Sauro & Smith, 2010). On the other hand, it is still unclear how language learning through texting affects subsequent speaking tests. This research analyzed whether the self-repairs in texting affect subsequent speaking tests on lexical accuracy, syntactic complexity, grammatical accuracy, lexical diversity, and fluency.

Chapter 3 refers to tasks for foreign language learning. Crucial factors in elucidating the relationship between texting and speaking are the task-implementing conditions. According to Skehan (2014), the conditions of the task include repetition and time constraints. These are found to affect the language produced. It has also been pointed out that the task type also affects it. Therefore, in this research, the task type was determined considering the proficiency level of the learners who were to be the research participants.

Chapter 4 states research questions based on task implementation conditions and how self-repairs affect subsequent speaking outcomes. The hypothesis is that the use of texting directs attention to linguistic forms to allow self-repairs in texting (Sauro & Smith, 2010). As texting may follow the language production model of Payne and Whitney (2002), the practice of texting may support other skills, such as speaking. The inherent feature of visual saliency and slow-paced exchange in texting may support the language production process.

ABSTRACT

In Chapter 5, a pilot test was conducted on three task types (interview, narrative, and decision-making) in texting and face-to-face. The purpose of this pilot study was to select the task type to be used for the main studies. The results of the lexical analysis showed that the interview task had the least differences between participants compared to the narrative and decision-making tasks. In addition, it was found that the interview task, unlike the other two task types, did not require specific lexical knowledge to complete the task. Therefore, the interview task was selected for the main studies.

Chapters 6, 7, and 8 report on the three studies (Study 1, 2, 3) conducted to investigate the effect of texting on speaking under task repetition and time constraint conditions. Study 1 investigated the effect of texting under three task repetitions for both face-to-face and texting groups. The results were compared on a one-minute speaking test between pre- and post-test for both modes.

The results showed that the texting group had lower lexical error rates for both content and function words and higher sentence structure scores for syntactic complexity. Further analysis of lexical error types also revealed that the texting group used fewer L1 (native language), repeated, and omitted words. No statistical improvement was found in the face-to-face group. The reason for the improvement in lexical accuracy in the texting group is that it spent about three times as much time on the task compared to the face-to-face group. Therefore, in Study 2, the time between the two modes was kept constant.

For Study 2, the research was conducted under the same procedure as Study 1, except for a constraining practice time of 15 minutes. The results showed that the lexical error rate improved only for the content words in the texting group. There was no repetition or use of L1 in the texting group during the practice session. This may be because texting allows the participants to visually check the language produced, which enables them to notice and correct errors before sending the message.

To test the feasibility of using texting as an effective communication medium for learning new terms, Study 3 was conducted to investigate the learning of lexical phrases in the two modes. The study was conducted under task repetition and 15-minute time constraints for three weeks. The exercise included two tasks: an interview task with a questionnaire of new lexical phrases and a repetition task. Both tasks were performed once a week for three weeks. The face-to-face group performed the tasks orally only, while the texting group performed everything online via texting. The outcome was measured using a cloze test on lexical phrases and a one-minute speaking test before and after the practice.

The results showed that there was no difference in the outcome of the cloze test. High scores were achieved in both modes. In the post-speech test, the texting group showed a decrease in the lexical error rate for both content and function words and a higher score for syntactic complexity and speech rate. On the other hand, the speech and articulation rates

improved in the face-to-face group. The articulation rate, which is considered an indicator of comprehensibility of speech (Suzuki & Kormos, 2019), did not improve in the texting group.

Chapter 9 discusses the results from the three studies conducted for this research. Under certain conditions, texting influenced lexical accuracy, syntactic complexity, and speech rate in the post-speaking test. In contrast, when the same task was performed face-to-face, both speech and articulation rates improved under sufficient timed conditions.

This study found that lexical error types, such as repetitions or the use of L1, were reduced in the texting group. This could be due to the availability of visual information and the slow exchange during texting. In other words, the lexical error types showed that texting helped foreign language learners to recognize and correct errors in lexical usage and syntactic structure, which had a positive effect on their later speaking performance.

In contrast, for the group that spoke face-to-face, both speech and articulation rates improved. Texting had no effect on articulation rates because it does not involve speech, revealing differences in the effects of the two modes on language learning. This research implies that mode differences impact different areas of speaking and that educators may select the mode according to their specific educational purpose. More details are discussed in Chapter 10 on the implication and limitations of this study.

Table of Contents

LIST OF TABLES

TABLES

LIST OF FIGURES

LIST OF APPENDICES

CHAPTER 1 INTRODUCTION

In terms of language teaching, experts believe that the use of computers can offer learning opportunities for high-quality learning (Golonka et al., 2017; Michel & O'Rourke, 2019; Sauro, 2012). This paper aims to examine the use of texting in task-based language teaching and to identify its effects on second language acquisition by comparing the results between texting and face-to-face instruction. The study investigates whether the difference in the mode of communication affects participants' speaking performances.

According to Information and Communications in Japan (2017), the use of smartphones and other Information and Communication Technology (ICT) devices has spread rapidly, showing that over 94.7% of households own mobile phones and 73.0% own personal computers (PCs). In addition, reading and writing text messages, including mail messaging, blogs, and SNS postings, make up over 50% of device usage (Ministry of Internal Affairs and Communications, 2017). Overall, texting increased by more than 7,700% in the past decade, making it the most common method of communication globally (Statistic Brain, 2014). Thus, texting strategy as a global communication method needs to be studied in greater detail, especially regarding its impact on language learning.

This chapter will discuss the features of texting and the language production process of texting and face-to-face speaking.

1.1. What is Texting?

Texting, "text-messaging," "short messaging service," "chat," or "chatting" is an online text-based communication method that has been in use for the last ten years (Peslak & Hunsinger, 2018). It is manually entering text into an electrical device or reading text from it. It was first defined by the Federal Motor Carrier Safety Administration in the United States (US) in 2012 as "manually entering text into or reading text from an electronic device. Texting includes (but is not limited to) short message services, e-mailing, instant messaging, a command or requests to access a Web page, pressing more than a single button to initiate or terminate a call using a mobile telephone, or engaging in any other form of electronic text retrieval or entry, for present or future communication."

Texting can refer to different types of computer-mediated communication. There are two types of text message communication techniques: synchronous computer-mediated communication (SCMC) and asynchronous computer-mediated communication (ASCMC). SCMC involves real-time communication, such as texting, while ASCMC, such as e-mail, involves time delays. In the 1990s and early 2000s, the distinction between SCMC and ACMC was distinct because networking infrastructures were still developmental, and transferring data was costly and time-consuming. However, due to technological progress and infrastructure development, the difference between SCMC and ACMC is now subtle. In

terms of technological progress, the modalities of texting have become more diverse, as verbal or textual language can vary depending on the tool. For example, the exchange of recorded messages, voice chat, can provide a more verbal interaction than text-only chat, blurring the lines between spoken and written language (Yus, 2011). This study is limited to text because bandwidth issues do not interfere with text-only communication, and the study can be implemented with actual students in the classroom.

Texting and face-to-face communication are similar in that they are instant exchanges. The difference is that text is a written mode of communication, while face-to-face exchanges are in speech. Significantly, the written mode also provides visual information that remains after it is produced, whereas information received via speaking diminishes the moment it is articulated. Further differences include that texting may require more production time, and the exchange is slower. In contrast, speaking allows for timely responses, and exchanges are fast. An advantage of texting is that it has no location restrictions; distant interlocutors can communicate instantly. Texting ignores the nonverbal aspect of communication, whereas face-to-face communication requires understanding both verbal and nonverbal messages. As texting and face-to-face communication are immediate, both have the potential for language acquisition. This study examines language acquisition differences between using texting and face-to-face.

Finally, texting input devices vary in type. For instance, the QWERTY keyboard for mobile devices generally has two input types: a physical keypad and a touch screen (Kietry et al., 2015). Additionally, electronic pens are also used as input devices. However, this research is limited to the physical QWERTY keyboards on a computer.

1.2. Language Production Process of Face-to-Face Speaking and Texting

Although the language production process of texting is not yet clear, Payne and Whitney's (2002) model can serve as a basis for understanding the use of texting in the second language process (L2) (which has a processing stage similar to face-to-face speaking production). Payne and Whitney (2002) attempted to show how texting and face-to-face speaking production overlap in some areas by referring to the oral production model (Levelt, 1989). They claim that the production of texting can be predicted and L2 oral proficiency can be developed indirectly through texting. The only difference is the production of overt speech.

According to Levelt (1989), there are three main phases in the process of face-to-face speech production. First, in the conceptualization, formulation, and articulation phase, where ideas are generated, lexical, grammatical, and phonological encoding occurs. For each phase, a phonetic plan is translated into muscle movements. An important phase in the speech process is the formulation phase, in which preverbal messages are encoded and provided with words, structures, and a phonetic plan to be articulated in the final phase.

Payne and Whitney argued that texting and face-to-face speaking production processes are almost equivalent; the only difference is the phonetic planning and the articulation stage. However, because texting is text-based, the writing process is also interwoven, allowing planning, reformulation, and the availability of greater attentional and memory resources before sending the message. Further, time availability eases the cognitive load in texting at the important stage in language production. This frees up attentional resources to enable L2 learners to attend to retrieving vocabulary and attach grammatical features to match the intended message in texting.

One less stage of articulation in texting may be beneficial for some learners with lower phonological buffering capacity. Payne and Whitney (2002) examined the relationship between working memory and speaking performance. In their study, they compared non-word repetition and pre- and post-speech outcomes between two groups who communicated face-to-face and via texting. They found that the group that communicated face-to-face had a higher correlation between non-word repetition and speaking performance outcomes than the group that communicated via texting. This indicates that learners with lower non-repetition scores, indicating lower phonological buffering capacity, were not disadvantaged compared to the face-to-face group. This suggests that the use of texting could be beneficial among a wide range of learners for phonological buffering capacity.

1.3. Constructs of Speaking and Texting

Speaking and texting are both output activities that share similar constructs. Previous researchers claim that the construct of spoken language consists of linguistic features from three main subgroups: phonological, lexico-grammatical, and discourse (Bygate, 2009: Kormos, 2006). Texting and speaking share lexico-grammatical and discourse features, and according to the speaking rubric for the International English Language Testing System (IELTS) and Test of English as a Foreign Language (TOEFL), lexico-grammatical features are described as vocabulary, grammar, and syntactic variation; and discourse features include coherence, pausing, stress, topic development and pragmatic structures (Byagte, 2009: IELTS, 2022; Iimura & Takanami, 2016; TOEFL, 2022). Although the idea of constructs is evolving due to the idea that language is part of a social context, the lexico-grammar and discourse feature plays a crucial role in conveying the message of these two modes. This research seeks to identify how practice in different modes affects these constructs.

The distinction between the two modes lies in the way the information is organized. In written language, complete sentences are formed to organize and express thoughts. In a speech, smaller groups of words in clauses or phrases can convey a single thought (Brown & Lee, 2015). Although texting is a form of written language, it has been found that sentences in texting are shorter than in traditional writing (Jonsson, 2016). Since texting is an instantaneous exchange, it is likely that sentences will be shorter. Single words can act as a response, as found in the spoken exchange. Another clear difference between the two modes

relates to prosody. Since texting is a written exchange, there are no prosodic features. However, in English, stress and rhythm convey subtle messages when speaking (Fraundorf & Watson, 2011). Intonation patterns can also be used to express information for criticism and praise, and pauses are used to convey and emphasize ideas. The absence of prosody in texts illustrates another uniqueness of direct exchange in the written mode.

This study investigates how the different constructs in the two modes affect subsequent speaking performance. For example, in texting, the grouping of words and the prosodic features found in speaking are lacking. However, texting may support the production process given its inherent features, such as the presence of visual text. Face-to-face may also be advantageous as the subsequent test is conducted in speech, not in written form. This research will investigate how the two modes differ in the learner's speech which follows practice.

1.4. Purpose and Objectives of the Study

As texting is now a dominant communication method, investigating the methodology and the effect of this communication medium on foreign language learning will benefit both teachers and learners. Additionally, it may create and expand opportunities to practice and further develop understanding of the target language. The question, however, is whether texting is an effective medium for improving speaking. Thus, this research investigates the differences and impact of texting and face-to-face learning in English as a Foreign Language (EFL) context.

This research also discusses the uniqueness and significance of using texts in language learning compared to face-to-face learning. Texting can assist students in solving the problems they encounter in face-to-face learning and practicing their existing knowledge outside the physical classroom. Further, it provides an additional avenue for learners to practice within the online virtual world, allowing for global interaction with different people. Inherent features of texting may aid language learning as it offers visual text. In addition, it has the potential to draw attention to linguistic aspects and gaps in the target language and its production (Blake, 2000; Hayati et al., 2013; Zhang et al., 2011).

This study investigates whether speaking can be improved by texting in an EFL environment in Japan. It is expected that the theoretical and pedagogical implications will contribute to research on second language acquisition (SLA) and teaching in EFL environments.

1.5. Summary

Methods of communication have evolved due to technological advancement. Consequently, texting is now a widely used form of communication. This research investigates texting as one of the synchronous computer-mediated-communication (SCMC) tools for foreign language learning compared to face-to-face teaching. The difference lies in

the availability of visual text, time, location constraints, and non-verbal aspects of communication. Previous research has not shown how speech is processed during immediate texting exchanges. Payne and Whitney (2002) claimed that texting has one less stage than the model of oral language production presented by Levelt (1989) for L1. In writing texts, there are no articulatory and auditory inputs, which could be helpful for learners with lower phonological abilities. Their findings suggest that texting could be a helpful communication tool to support many learners in an educational setting.

Furthermore, the absence of prosodic features in written exchanges may affect speech production. Therefore, the purpose of this study is to investigate how the use of texts differs from the practice of face-to-face speaking in language acquisition. Finally, examining the effects of this medium of communication may provide learners and teachers with additional ways to compensate for or replace face-to-face language instruction.

Chapter 2 reviews previous studies on texting and foreign language learning and explores its potential to impact speaking performance.

CHAPTER 2 FOREIGN LANGUAGE LEARNING AND TEXTING

Many researchers have been interested in incorporating Computer-Mediated Communication (CMC) in foreign language learning. Studies on the use of texting in language learning concerning linguistic competence and patterns in interactions and discourse markers are well-represented (Abe & Roever, 2019; Blake, 2000; Hayati et al., 2013; Smith, 2003; Zhang et al., 2011). For instance, Hayati et al. (2013) focused on texting for vocabulary learning. Abe and Roever (2019) established sequential structures and interactional practices similar to spoken interaction by adopting the conversational analysis method.

Several studies have highlighted the spontaneity of texting exchanges and therefore suggested that texting could be used as a substitute for spoken language in the language-learning process (Abe & Roever, 2019; Blake, 2000; Hayati et al., 2013). However, there are also studies that classify texting as a pure exchange of written text and find no effect on speaking (Abrams, 2003; Beauvois, 1997). These studies limit their effects to specific domains (Ziegler, 2018). In this section, we review previous research on the potential of texting for second language acquisition, its linguistic features, and how texting behaviors support language acquisition in contrast to face-to-face learning.

2.1. Synchronous Communication in Texting and Second Language Acquisition

Research has shown that inherent features of texting facilitate language learning (Blake, 2000; Payne & Whitney, 2002; Ziegler, 2016). As a form of synchronous communication, texting provides opportunities for instant and comprehensible interaction, which has significant consequences for SLA (Second Language Acquisition) (Blake, 2000; Kitade, 2000; Long, 1981; Swain & Lapkin, 1995). Ziegler (2016) states that the use of technology for language learning is beneficial as it offers not only interactional features but also "additional opportunities to notice target language features" (p.142), such as text-based visual saliency, additional time for interactions, and it is less intimidating than face-to-face interactions. All of which could positively impact individual engagement.

Language learning through face-to-face and computer-mediated interactions has been studied and provides opportunities to test students' language knowledge and solve language problems that are critical for L2 development (Long, 2015; González-Lloret, 2022). The challenge of learning through face-to-face interactions is that learners need to notice the gap between the interlanguage and the target language. In comparison, the use of texts and written exchanges focus on language production and thus helps learners notice their errors (Kitade, 2000; Lai & Zhao, 2006). The early research suggests that the saliency of text brings enhanced noticing to their own language errors compared to face-to-face conversation (Lai & Zhao, 2006).

Another advantage of texting is that it encourages learners to take additional time to process information when producing messages in ongoing interactions. Lai and Zhao (2006)

claimed that this might reduce learners' cognitive burden so they can use their attention to monitor and evaluate their linguistic production.

In addition, texting could encourage greater student participation because it is not as intimidating compared to face-to-face communication. For example, there is less social pressure, more time to respond (González-Lloret & Ortega, 2014). Students who are afraid of speaking face-to-face may therefore find it easier to engage in a virtual world.

2.1.1. Linguistic Features of Texting

Despite the wide range of research on texting, there are different opinions on whether texting exchanges are spoken language or written language and on how it impacts these modes. Recent research on corpus comparisons has examined the properties of texting by comparing them to speech and writing corpora. Based on Biber's (1988) multidimensional approach, Jonsson (2016) has shown that texting has a similar distribution of linguistic features as speech. To this end, L2 text types (genres), including writing, conversation samples, emails from an electronic company, and Internet relay chats, were analyzed for 67 linguistic features. Jonsson successfully analyzed different text types by utilizing the method provided by Biber (1988). By comparing with a balanced set of corpora, Jonson based the findings on the variation feature of mode among different text types. Table 1 describes the results of the salient linguistic features that appear in 1,000 words among the different modes; texting (SCMC, SSCMC), speaking (Speech), e-mails (ACMC), and writing (Jonsson, 2016: p.150). Accordingly, the SCMC and SSCMC are more similar to speech distribution than writing. In sum, Jonsson (2016) concluded that the language used in texting could be regarded as the most oral-like form of writing. Thus, Jonsson (2016) has named texting "conversational writing" (Jonsson, 2016).

Table 1

Frequencies per 1,000 words for the Most Salient Linguistic Features

	Writing	ACMC	Speech	SCMC	SSCMC
first person pronouns	17.0	57.8	52.8	56.9	88.9
second person pronouns	5.0	17.6	23.0	50.4	45.0
direct WH-questions	0.1	2.9	0.8	3.5	3.9
analytic negation	6.4	15.1	13.9	13.1	29.7
demonstrative pronouns	2.3	6.5	10.6	6.6	16.4
indefinite pronouns	0.9	4.6	3.1	11.7	6.0
present tense verbs	64.6	67.6	112.3	147.2	168.5
predicative adjectives	4.8	n.a.	4.9	8.4	15.3
contractions	4.6	16.6	36.1	30.8	55.0
prepositional phrases	117.3	116.9	91.1	47.0	42.0

(Jonsson, 2016: p.150)

An analysis of the language used in texting suggests that a greater part of the language is a standard form, and the message is semantically understandable (Kemp & Bushnell,

2011). Shortis (2007) and Carrington (2004) claim that the spelling in texting is only in "squeeze-text" form, which captures the principle and is capable of predicting and recognizing from the standard language. Research on letter and number homophones, contractions, and emoticons shows that these elements are used less frequently. Further, texting is distinctly unique in its use of abbreviations and nonstandard forms of words. This includes features such as letter and number homophones (4 for "for" and u for "you"), contractions (text for txt), and nonconventional spelling by substituting parts of words, for example, "gr8", "4", "sum1" (Lyddy et al., 2014; Thurlow & Brown, 2003). Although one might assume that this unique form of writing may harm academic skills, research shows no significant difference between the traditional literacy skills of spelling and reading, and texting (Kemp & Bushnell, 2011). In addition, it shows that the majority of text language is in standard form (Thurlow & Brown, 2003).

McSweeney (2017) confirms that students who send more English texting have higher academic skills. Thus, it has been proven that texting does not harm learning. Rather, it leads to higher English literacy scores. Other research with American undergraduates found that texting and literacy ability is related to the style of texting or textese, such as the omission of letters or accent stylization, rather than the use of texting itself. Ouellette and Michaud (2016) found that the use of textese in personal messages has declined and that there is little relationship between texting and literacy skills. They found that most participants used QWERTY keyboards and some form of autocorrect or predictive technology that did not negatively impact the literacy domain. In addition, they claim that spontaneous communication in texting is similar to spoken language in the form of visual text, and it does not hinder academic learning.

2.2. Self-Repairs in Face-to-Face and Texting

Self-repairs are how speakers direct their attention during language production (Segalowitz, 2010; Zuniga & Simard, 2019), and it has been of much interest to many researchers. Self-repair behavior provides information on how language has been processed and acquired (Kormos, 2006; Levelt, 1989). Although there are similarities in the production process of texting and face-to-face speaking (Payne & Whitney, 2002), self-repairs are demonstrated differently in how attention is brought to linguistic elements (Smith, 2008; Sauro & Smith, 2010). In speaking, for example, the information disappears from one moment to the next, while in texting, a certain time passes before the message is transmitted and the visual text is presented.

Face-to-face production consists of multiple processes and requires the allocation of attention at a rapid pace (Levelt, 1989). For language learners, processing information rapidly for communication can sometimes lead to difficulties, errors are often overlooked, or breakdowns can occur. Self-repairs in face-to-face speech may occur before or after the articulation phase in the speech production phase (Payne & Whitney, 2002). It requires

instant processing to detect errors, retrieve linguistic information, and then make corrections while also processing information for what to say next. Self-repairs in face-to-face communication may be articulated, but they can also be covert (Kormos, 2006). For example, when speaking, acts of hesitation, false starts, and pauses can manifest as self-repairs, while covert repairs are difficult to observe as they are carried out mentally (Zuniga and Simard, 2019). Both overt and covert repairs are conducted under multiple parallel processes that may burden the cognitive load, especially for novice language learners.

In contrast, self-repair in texting is perceived as less burdensome because there is one less stage of articulation in the speech process than in speaking. In addition, visual highlighting in texting is beneficial for detecting errors (Smith, 2008). Smith (2008) conducted a study of self-repairs and documented all self-repairs, including those that were not sent and did not appear in texting logs. He found a significant number of self-repairs in texting behaviors and that those repairs acted to correct or rephrase a message before it was sent. Therefore, Smith (2008) found that self-repairs in texting are overt, and although Smith (2008) failed to find statistical significance, self-repairs in texting are reported to impact grammatical-related problems and vocabulary.

Sauro and Smith (2010) have further explored self-repair in texting. Their research was conducted using a video-enhanced chat script which enabled them to capture all behavior during texting. The logs were analyzed, including the three categories: plain text with no evidence of self-repair, deleted text, and text after repair (post-deleted text). They found that texts that contained evidence of repair had higher scores on linguistic complexity and lexical diversity than plain text with no evidence of self-repair. In their study with learners of German using a jigsaw task, the post-deleted text measure scored higher on syntactic complexity, grammatical accuracy in relation to gender, and lexical diversity compared to the plain text or deleted text. The results suggest that the language learners in this study created more complex and sophisticated language through their self-repair behavior when texting. In addition, they found evidence of self-repair in the language production process when texting, which allowed learners to monitor their production as they typed by revising linguistic aspects and repairing what they saw on the screen.

Textual saliency is an advantage for L2 learning and has been reported in several other studies (Lee, 2021; Lee & Révész, 2020). In a study using textual captions for news clips, Lee & Révész (2020) explained that increased salience enhanced by textual captions allowed learners to reflect and apply their declarative knowledge. Lee & Révész (2020) stated that “the learners were able to automatize their explicit knowledge throughout the task” and found textual captions to be effective in drawing attention to the target grammatical form. Repair in the text has a similar benefit. The visual text allows learners to use moment-to-moment online planning time to carefully produce and monitor language use (Sauro & Smith,

2010). It follows that the visual effects of texting are beneficial to second language acquisition.

2.3. Empirical Research on the Use of Texting for Vocabulary Learning

Previous research on self-repairs and texting found that texting is an effective medium for catching errors, using tasks that require the retrieval of existing linguistic knowledge, such as a jigsaw task chosen to induce information exchange (Smith, 2008; Sauro & Smith, 2010). Moreover, previous research has also addressed learning multiword expressions (MWEs) online through the saliency of texting and other affordances such as audio and videos offered in the computer-assisted learning environment.

MWEs, also known as multiword sequences, formulaic sequences, or lexical phrases, are the collection of single words that are stored as single units (Wood, 2010) or a "continuous or discontinuous string of meaningful elements commonly interpreted together as a single unit, in some cases allowing modifications of specific elements" (Christiansen & Arnon, 2017, p.544). Recently, MWEs have gained attention for vocabulary acquisition as they play an important role in language acquisition and process (Boers et al., 2014; Christiansen & Arnon, 2017; Lewis, 2009; Towell et al., 1996). Studies show that MWEs are processed faster than individual words because they are treated as a bundle (Erman, 2007), thus affecting fluency development (Towell et al., 1996). However, as MWEs are often composed of familiar words, L2 learners skip the short words and may not recognize them as units (Carrol & Conklin, 2020). In addition, MWEs are not frequently repeated in a short text, which L2 learners find difficult and time costly (Boers et al., 2014; Lewis, 2009; Towell et al., 1996).

The learning of MWEs has been explored in an online setting to ease the difficulty of acquisition (Boers et al., 2014; Cucchiarini et al., 2020; Stengers et al., 2016). In a series of studies, Boers et al. (2014) presented MWEs with picture cards to link the meaning and word unit to MWE's meaning retention and linguistic form. However, the results concluded that using picture cards for learning the MWE was insufficient to retrieve the form, leading to further investigation of the form by Stenger et al. (2016). Stenger et al. (2016) found that typing practice of MWEs drew attention to linguistic forms. The cloze test revealed that typing is a valid practice for learning new MWEs with no difference in recall compared to the mean-oriented task. Recently, Cucchiarini et al. (2020) conducted a study on idiom learning using texting and investigated the effect of repetition. In their study, German students learning Dutch practiced idioms twice for one group and 12 times for another using the Computer Assisted Language Learning (CALL) program with different types of exercise, including practice through typing. The results of the cloze test showed that even for the group that practiced twice, using the CALL program with typing exercise facilitated the acquisition of certain types of idioms. The input of the new idioms was through reading or listening, but the results imply that it did not affect the cloze test results. The findings from this study

indicated no difference in the type of input, reading or listening, to the learning of idioms. Attention to the orthographic form of the idioms can be possible without verbal sounds. However, the study is limited as results vary depending on the type of idiom. The opaque phrases showed lower scores. Nevertheless, this demonstrates that online practice that included typing supports idiom learning, as some words require less repetition than others. Studies have shown that the meaning and form of the new vocabulary could be learned in the texting environment.

2.4. Limitations of Previous Research and the Current Research

Previous research has found two important pieces of evidence on how behavior in texting can facilitate SLA, self-repairs, and typing. Self-repairs are present while texting (Smith, 2008), and it is beneficial in improving syntactic complexity, grammar, and lexical diversity. For instance, Sauro and Smith (2010) noted that seeing text on the screen allowed learners to detect and modify errors. Practice through typing has also been found to be effective in learning new words, such as MWEs (Cucchiarini et al.,2020). The post-test gap score revealed learning new MWEs is possible without orally practicing them. This suggests that new items can be learned through texting.

However, there are limitations to the research that has been conducted. Task selection deserves further attention (Lin, 2014;Sauro & Smith, 2010) because some tasks require specific knowledge or language to perform, which affects language production. If a research task is too complex, it may hinder language production and affect the study of texting for language learning. The current study attempts to fill this gap by further investigating the effects of texting on task types and conditions.

So far, texting has been shown to be a valid medium for SLA. As the corpus study of spoken, written, and texted data has shown, the linguistic features of texting are similar to speaking (Jonsson, 2016). It could be hypothesized that texting practice affects speaking. However, there is currently no evidence to support this hypothesis, and further research on the relationship between text use and speaking practice is needed. The present study aims to fill this gap by investigating whether language practice via texting affects syntactic complexity, grammar, and lexical diversity of speaking performance under specific task types and conditions. It will also investigate how self-repairs in texting affect speaking performance. In addition, this research will clarify whether learning new terms via texting can be used in speaking tasks.

2.5. Summary

A review of studies of texting in foreign language learning has revealed important implications for SLA. The inherent property of synchronous interaction through written texts allows learners to expand their metalinguistic knowledge and modify their output in real-

time. Therefore, identifying the gap between linguistic output and the target language is crucial in L2 learning.

Self-repair emphasizes a language user's linguistic elements and directs their attention to errors. Self-repair in texting impacts linguistic complexity and lexical diversity (Sauro & Smith, 2010). Additionally, studies show that texting is beneficial for learning new linguistic items such as MWEs (Boers et al., 2014; Stengers et al., 2016; Cucchiarini et al., 2020). Jonsson (2016) found positive results on the linguistic similarity of texts to speech. However, no study has found a relationship between texting and its effect on subsequent speaking performance. Therefore, this study investigates the effects of texts on SLA through careful task implementation (Lin, 2014; Sauro & Smith, 2010). The next chapter will evaluate the variables that influence the learning process in face-to-face task-based learning. This assessment will provide suggestions for task-based learning with texting.

CHAPTER 3 TASK-BASED LANGUAGE TEACHING

Tasks play an important role in pedagogy, as they target and enhance the development of appropriate L2 processes and outcomes. Tasks are “holistic activities” (Samuda & Bygate, 2008, p. 7) that incorporate the knowledge used in context. They can also be applied to texting, a communication tool widely used today. It has been pointed out that using it alone will not lead to successful communication activity for second language learning, as texting is merely a medium (Blake,2000; Carr, 2020). However, through engaging in tasks, learners will practice the language in an authentic context, and the inherent features of the computer medium may allow for highlighting language aspects that can affect the process and outcome of learning via texting.

Many studies have examined the use of face-to-face tasks in learning (Long, 2016; Moore, 2018; Samuda & Bygate, 2008) and found that the implementation conditions that apply to texting tasks are important. In other words, the design and implementation of language tasks need to be carefully considered for face-to-face or texting tasks. This chapter discusses these issues by first defining task-based language instruction. It then examines variables that have been researched in face-to-face task practice that affect language performance, such as task type, task repetition, time pressure, and familiarity. Reliable assessment measures are also considered to understand the effect of texting through task-based learning.

3.1. Definition of Task-Based Language Teaching

Task-Based Language Teaching (TBLT) is an approach to language teaching research that emanated from communicative language teaching in the mid to late 1980s. It has been adopted widely in language classrooms to elicit practical use of the target language (Long, 2016; Moore, 2018; Samuda & Bygate, 2008). TBLT’s definition has been widely used to describe simple activities and curriculum development (Long, 2015). Moreover, the term “task” also has various definitions. Moore (2018), drawing on previous research, defines a task as having five features: 1) there is a work plan; 2) a language-based focus on meaning; 3) a reflection of real-world language use; 4) the promotion of language development; 5) it has a communicative outcome. In assignments, the learner is presented with a work plan that includes instructions, procedures, and assessments. It requires communication to solve authentic problems, such as activities with information gaps. An important characteristic of tasks is that they require meaningful outcomes or goals. The difference between a task and an activity is that tasks involve an end product of information, a key feature of language development (Samuda & Bygate, 2008).

3.2. Features of Task and Implementation

Extensive research on TBLT has revealed that tasks may require several conditions to succeed. Skehan (2014) suggested a framework that illustrates task application's direct feature in actual teaching situations. The framework has two phases; one deals with features of the task itself, such as task types and difficulty. The second point is the condition and performance of the task. Skehan lists the conditions for performing face-to-face tasks (these conditions also apply to text tasks): Repetition, timing and planning, and familiarity with the content. These conditions outline the "preparedness" of a task which may directly affect language production regardless of the mode (Skehan, 2014).

3.3. Task Type

Task type is a term related to different cognitive loads that execute a particular task on discourse features or linguistic function (Foster & Skehan, 1996). Foster and Skehan (1996) conducted a study with three types of tasks (interviews, narration, and decision-making) and found that strategic planning affects performance by trading off performance areas among language complexity, accuracy, and fluency (Foster & Skehan, 1996). Trade-off effects are likely to occur in many parts of language production as learners might have difficulty bringing attention to various aspects of the language simultaneously (Foster & Skehan, 1996; Rahimpour & Safarie, 2011). The trade-off effect is based on theories of information processing (Anderson, 2005) and limited processing capacity. When language users are faced with high task demands, they selectively allocate their limited attentional resources to some parts of the task, thereby neglecting other parts. The nature of the task and other factors, such as task implementation conditions, for example, repetition, time, familiarity with the task, and individual learning style, influence performance (Hu, 2018; Skehan, 2014).

The condition of task type on language production has been explored widely in face-to-face studies (Egusa & Yokoyama, 2004; Foster & Skehan, 1996; Hu, 2018). Studies report that a task with a clear structure will likely promote fluency and accuracy (Hu, 2018), while tasks requiring more attention to language form or the need for cohesiveness may impact certain grammatical forms or syntactic complexity (Hu, 2018). In early research, Foster and Skehan's (1996) study found that among personal, narrative, and decision-making tasks that discriminate between cognitive load, the personal interview task positively affects fluency and accuracy at the expense of complexity. In contrast, narrative and decision tasks tend to promote accuracy and complexity at the expense of fluency. That is, narrative and decision-making tasks are more concerned with argumentation and justification and sequential or conceptual structure to accomplish the task than with the construction of ideas.

In a study comprising English as a Foreign Language (EFL) students, Egusa and Yokoyama (2004) found that an information gap task-induced fluency while decision-making tasks enabled students to produce accurate and complex language. Accuracy improved with

the task with more cognitive load. They believe that the trade-off effect is not as pronounced for EFL learners as it is for English as a Second Language (ESL) learners because of their limited exposure to the language. Even if learners attempted to correct their errors, those corrections might not be correct. Thus, although there are subtle differences in outcomes between EFL and ESL learners, this demonstrates the trade-off effect between task types and their impact on language production.

There is a considerable body of research examining language production under different task types on SCMC (Blake, 2000; Smith, 2003; Yilmaz & Granena, 2010). However, it is not yet known whether there is a trade-off effect in task performance on SCMC. Previously, Blake (2000) performed three types of tasks (jigsaw, information gap, and decision-making) via texting. The study found that while a jigsaw task elicited a lot of negotiation, there was no evidence that other task types were unsuccessful. Language-related episodes (LREs) indicate opportunities for language learning. Yilmaz and Granena (2010) compared the number of LREs in the jigsaw and dictogloss tasks per texting. They found that between jigsaw and dictogloss tasks, the latter had more opportunities for LREs. In addition, they stated that the LREs were qualitatively different between jigsaw and dictogloss tasks and reached a solution in the latter. In addition, Yilmaz and Granena (2010) claimed that comparing this research result to others is problematic as the number of LREs may not directly reflect the performance among different task types. Task conditions and individual learner factors may have impacted the results.

Thus, the task-type studies on SCMC are inconclusive. In other words, how tasks impact language in texting is unclear.

3.4. Repetition

The impact of repetition has been widely researched in face-to-face learning (Ahmadian & Tavakoli, 2014; Fukuta, 2016; Gass et al., 1999), and according to Segalowitz and Segalowitz (1993), there are two ways in which practice may benefit performance. One is the facilitatory effect, which involves speeding up the processes of executing a given task. After the first attempt, learners can focus more on linguistic form than content in the following review task. Another effect is that it becomes easier to convert declarative knowledge into procedural knowledge, which is useful for automation. Kormos (1999) notes that in L2, learners need more time to retrieve linguistic knowledge and to load the conceptualizing and formulating processes of the language production model. Since texting provides additional time, practice can thus reduce this load and provide cognitive space for conceptualizing and formulating a process for qualitative information exchange. Anderson (2005) claims that repetition leads to automatic processing that can close the gap between L1 and L2 learners (Bot, 1992).

There are two types of task repetition: (1) using the same tasks (content repetition) and (2) repeating the same procedure with different content (procedural repetition). The results differ between the types of repetition. However, the findings of previous studies suggest that repetition of the same task plays a facilitative role in accurate oral development (Fukuta, 2016: Hawkes, 2012) and fluent oral production (Bei, 2012; Bozorgian and Kanani, 2017; Lynch & Maclean, 2000; Kim & Tracy-Ventura, 2013). Hawkes (2012) conducted a study with Japanese junior high school students and found more error correction in repeated tasks than in the initial tasks. Bozorgian and Kanani (2017) studied the effects of task repetition on the accuracy and fluency of speaking of intermediate EFL learners. Their study indicated that learners in the experimental group that completed the listening recall task repetition outperformed the learners in the control group in the post-test on speaking. Bozorgian and Kanani (2017) investigated the effect of task repetition between the pre and post-test speaking tests with students from Iran who used story retelling as an intervention. They found that participants' accuracy and fluency improved. In a study using cartoon video clips, Bei (2012) confirmed that repetition affects accuracy and fluency.

However, the study by Jong and Tillman (2018) found that immediate task repetition resulted in reusing words, thus impacting only fluency. Without a clear objective, the benefit of repetition may be limited to fluency.

Further research has been conducted to explore the effect of multiple task conditions and their effect on performance. For example, Ahmadian and Tavakoli (2010) explored the repetition effect of video narration tasks under two conditions (repetition and online planning) on face-to-face learning. They found that this improved accuracy, complexity, and fluency with greater effect than in single-condition groups. They explained that the additional time spent on careful online planning and task repetition provided learners with additional cognitive space to rely on their rule-based system during the conceptualization and formation phases of the speech production process. It also allowed students to test their new linguistic knowledge simultaneously. These results provide important information for understanding online planning time in writing texts. Since online planning is associated with text writing, text writing may predict positive outcomes. Through the effect of task repetition, texting is expected to improve several performance domains.

Some researchers point to the negative effects of repetition, such as limited performance effects (Boers, 2014; Thai & Boers, 2016) and a decline in motivation in some individuals (Qui & Lo, 2016). Using the 4/3/2 activity, in which language learners repeated a task under increasing time pressure, Thai and Boers (2016) found that many errors were repeated across iterations and that accuracy was not observed. Their results suggest that time pressure has a greater impact on accuracy than repetitions do. The same result was found in another study by Boers (2014), compromising accuracy over fluency. This suggests that consistent time for task implementation may be more prudent for improving accuracy (Boers,

2014). The disadvantages of repetition are also mentioned in Qiu and Lo's (2017) study on speaking with EFL learners. While some learners felt more relaxed and confident with repeated tasks, others felt less motivated and showed a decrease in self-repair. For some, repeating the same task can be boring, causing learners to lose interest. However, overall, task repetition is effective in improving speaking performance in face-to-face learning. The effects of task repetition on texting have not yet been reported. The current study aims to investigate this condition.

3.5. Time Constraint

Planning has attracted researchers' attention (Ellis, 2005; Foster & Skehan, 1996; Housen & Kuiken, 2009) as it is necessary for language production, especially second language production when the production process is difficult (Ellis, 2005). Planning is distinguished by the time at which it occurs, before, during, and by the repetition of the task. According to Ellis (2005), planning emphasizes linguistic form, whether planning occurs before the task (pre-task planning), during the task (online planning), or by performing a similar task before the main task (rehearsal). A recent study has shown that different types of planning affect specific performance domains in face-to-face studies.

3.5.1. Pre-Task Planning

Across studies, pre-task planning has been reported as having little impact on accuracy (Mehrang & Rahimpour, 2010; Rahimpour & Safarie, 2011; Ortega, 2005). A study on EFL intermediate learners found that pre-planning a narrative speaking task using cartoon strips did not affect accuracy and fluency but increased complexity (Mehrang & Rahimpour, 2010). Mehrang and Rahimpour (2010) discussed that learners "choose to focus on meaning and plan what they want to say rather than planning grammatical forms" when they are given time to plan before a task. Similar results were also obtained for the effect of pre-task planning on writing. Rahimpour and Safarie (2011) conducted a study with EFL college students on descriptive writing. The group that engaged in pre-task planning had better results in terms of fluency or the number of T-units per text. However, there was no effect on sentence complexity or accuracy in writing. They stated that pre-task planning might impact the content and facilitate organization, which only affects learners' writing clarity. While pre-planning affects fluency, it is not easy to generalize the results as studies were conducted under various timed conditions. The common finding can be concluded that pre-planning does not positively affect accuracy for speaking and writing.

3.5.2. Online Planning

Online planning allows language users more time to check their speech for grammatical accuracy. It also allows learners to look up appropriate lexical items, thus

improving the accuracy of speaking tasks in face-to-face studies. (Ahmadian & Tavakoli, 2014; Saeedi, 2020; Wang, 2014). In Wang's (2014) study, online planning time allowed learners to monitor their language production using video annotation. This improved accuracy and complexity. A study by Ahmadian and Tavakoli (2014) confirmed these findings. They performed a narration task with silent videos with two groups and found that the group with online planning had better accuracy and complexity, but fluency was impaired. They found that careful online planning gave room to speakers' limited attentional capacities and enabled them to formulate accurate and complex speech structures (Ahmadian & Tavakoli, 2014).

Additionally, self-repairs are reported to increase under online planning conditions (Ahmadian & Tavakoli, 2014; Saeedi, 2020). Saeedi (2020) found that online planning correlates with cases of repairs to accuracy. Saeedi (2020) states that additional online time allows learners to reach their explicit L2 knowledge, emphasize language production, and facilitate accuracy through self-repairs. This finding corroborates Ahmadian & Tavakoli's (2014) results. Their study found more error-type repairs in the online planning group than appropriate repairs or different repairs (which focus on error correction concerning language use sophistication).

In summary, the types of planning affect performance differently. For example, while accuracy was not affected under pre-task planning conditions, it was affected by online planning conditions (Ahmadian & Tavakoli, 2014; Saeedi, 2020; Wang, 2014). Online planning also affected complexity (Ahmadian & Tavakoli, 2014; Wang, 2014). Thus, for the face-to-face learning condition, additional time spent on tasks affects accuracy and complexity.

3.6. Task Familiarity

Task familiarity is a condition that impacts learners' performance by affecting the accessibility or organization of ideas based on prior experience (Nazemi & Rezvani, 2019; Wolters & Taylor, 2012). For example, Nazemi and Rezvani (2019) claimed that participants working on a familiar task produce more words, can handle more elaborate tasks in a shorter time, and make fewer errors than an unfamiliar task. This is because a less familiar task requires more attention to its structure.

The problem with task familiarity, however, is that it is difficult to assess the validity of a particular task. The familiarity elements of a task may include its number of abstract and concrete features. According to Skehan (2014), these task features are complex and are influenced by the context of the particular task. In addition, Ahmadian et al. (2015) found that task features differ depending on factors such as proficiency level, participant age, and cultural knowledge, which makes it difficult to determine the difficulty of a task. To address this issue, this study conducts a pilot study to assess the familiarity and difficulty of a task.

3.7. Measuring Task Performance

Recent assessments of task performance in TBLT use a cognitive approach by referring to three dimensions of mastery and the main stages of L2 acquisition. These dimensions include 1) internalization of new L2 items (complexity), 2) modification of L2 knowledge (accuracy), and 3) consolidation and proceduralization of L2 knowledge (fluency) (Housen et al., 2012; Norris & Ortega, 2009; Skehan, 1998). For example, Housen et al. (2012) explained that a new L2 target object is internalized due to complexity as knowledge is developed. It is then modified for more sophisticated and accurate language, eventually aiming to consolidate the L2 knowledge into fluency for higher performance. However, researchers mentioned that "language learning does not progress linearly" (Ellis, 2005), and these three measures have appeared throughout L2 studies to measure language performance (Foster & Skehan, 1996; Norris & Ortega, 2009; Yuan & Ellis, 2003).

This study aims to identify whether texting impacts speaking performance, and this section discusses how the three dimensions (complexity, accuracy, fluency) can be applied to assess its aim.

3.7.1. Syntactic Complexity

A wide range of measurement scales has been adapted to measure syntactic complexities. For example, the average number of words per clause (mean length of clause [MLC]), T-units, and AS-units (Foster et al., 2000; Norris & Ortega, 2009; Hunt, 1965). However, each measure has limitations that must be carefully considered before adoption; for instance, the validity of how the measure corresponds to what it is intended to measure. Further, there is an ongoing debate on whether these measures capture modification at all levels, including task content, proficiency, and language mode, either for writing or speaking, and the accuracy of scores (Foster et al., 2000; Ortega, 2005).

In this study, the IPSyn score (Scarborough, 1990) is used to measure complexity because it is automated and accurate. Although the IPSyn score was developed to measure child language development, it also captures spoken data in the early stages of language acquisition. This automated measure, widely used to measure children's syntactic development, provides information on syntactic forms and grammatical emergence. Language data are analyzed for 56 syntactic structures within four subscales: noun phrases, verb phrases, question/negation, and sentence structure. Each sentence receives zero to two points according to emergence. Altenberg, Roberts, and Scarborough (2018) claim that while the score does not reflect mastery of the language, the automatic score can be processed accurately and provides information about sentence structure and grammar.

3.7.2. Accuracy

The accuracy of a second language performance may refer to how it deviates from a rule (Wolfe-Quintero et al., 1998) or is interpreted as adequacy and acceptability. The present study follows this standard, and accuracy is examined by two evaluators. First, inter-reliability is calculated to highlight discrepancies. Then, all items are discussed to reach a consensus.

Recent research claims that repairs account for fluency (Lambert et al., 2017; Suzuki & Kormos, 2019; Tavakoli & Skehan, 2005). However, repairs can also be an indicator of accuracy as they may directly affect accuracy. In other words, repairs are measures that affect accuracy. Further, repairs have been operationalized as repetition or corrections made by the speaker (e.g., Saito et al., 2019). It is widely understood that L2 speaking follows similar stages as the L1 model (Kormos, 1999; Hartsuiker & Kolk, 2001). The difference being L2 speakers may require more processing capacity and time, which might make it difficult to notice errors or monitor their output (Gilabert, 2007). Despite this difference, repairs indicate errors noticed by speakers and allow for the possibility of correcting them, which may directly affect accuracy.

3.7.3. Fluency

Fluency is defined as ease, eloquence, smooth, and native-like speech (Chambers, 1997). Unlike other dimensions of performance which rely on linguistic knowledge and repertoires, fluency is performed in real-time (Chambers, 1997: Lennon, 1990). There has been much debate about its constructs and identifying reliable ways to measure them (Lennon, 1990; Segalowitz, 2010). Recent research has found that the construct of comprehensibility is strongly related to articulation rate, the mean number of words produced in phonation time (Suzuki & Kormos, 2019). This study encourages comprehensible interaction through tasks and adopts the articulation rate to measure fluency improvement.

Speed is another indicator of fluency. Speech rate refers to the number of words spoken in a given time. Speed can refer to all stages of speech production. Towell et al. (1996) noted that observed progress in speech rate reflects how linguistic knowledge is stored. A change in speed can be interpreted as a shift in knowledge from declarative to procedural. The mode may affect the production procedure; therefore, speech rate is also used in this study to measure fluency.

3.8. Summary

Task-Based Language Teaching (TBLT) has garnered great attention as it promotes language development. Although many TBLT studies are conducted face-to-face, this chapter provided insights for learning via texting. Learning through tasks enables language learning in context and promotes second language acquisition. Recent research has shown that task

characteristics and implementation conditions such as task types, time pressure, planning, task repetition, and familiarity influence language performance (Skehan, 2014). Functional demands on language can vary by task, and these differences affect different domains of language performance. In addition, differences in planning conditions also affect learner performance. Accuracy was affected in all studies only when the conditions included online planning. This finding predicts positive results for texting because they inherently allow for online planning. Task repetition and familiarity improve language performance by compensating for the lack of knowledge from the initial task experience and providing additional processing space for language elements.

Another important issue of TBLT is how performance is measured. Based on previous L2 acquisition research, complexity, accuracy, and fluency measures are adopted to indicate task performance (Housen et al., 2012; Norris & Ortega, 2009; Skehan, 1998).

The next chapter will state this study's research questions and make predictions about texting on speaking performance.

RQ

CHAPTER 4 RESEARCH QUESTIONS AND HYPOTHESIS

The traditional pedagogical approach to developing a particular performance, such as speaking, is based on the assumption that practicing speaking will improve speaking ability. A study by Payne and Whitney (2002) suggests that synchronous exchanges via texting can also indirectly improve L2 speaking. However, there are no studies along the same lines, nor are there any studies that address how texting might affect subsequent speaking performance. Since the texting mechanism has not been sufficiently researched and fully understood, this chapter poses the research question and hypothesizes about the effects of texting on speaking in L2.

4.1. Research Questions

This research investigates the effect of texting on speaking in L2 learning by investigating it under different task conditions. Skehan (2014) mentioned that task implementation conditions, such as task types, time repetition, time pressure, and task familiarity, affect language performance. This research also examines the influence of task variables using texting on speaking performance.

Texting has similar linguistic aspects to spoken language (Jonsson, 2016), and texting follows a similar production process to speech (Payne & Whitney, 2002). In addition, research has shown the presence of covert repair during texting, which displays evidence for syntactic complexity, grammatical accuracy, and lexical diversity of language (Sauro & Smith, 2010; Smith, 2008). Despite this evidence, there has been no research on how these repairs in texting affect subsequent speech performance. The present study aims to fill this gap.

Previous studies examined the effects of texting on speaking and writing (Abrams, 2003; Beauvois, 1998; Payne & Whitney, 2002) and vocabulary acquisition, such as multi-word expressions (Littlemore & Eyckmans, 2008; Cucchiarini et al., 2020; Stengers et al., 2016). While there have been attempts to explore texting for language learning, the findings have been mixed. The differences in these results emphasize the variables of task implementation that this current research seeks to pursue. In addition, a detailed analysis of syntactic complexity, grammar, fluency, complexity, accuracy, and fluency (Norris & Ortega, 2009; Housen et al., 2012; Skehan, 1998) will lead to a better understanding of the effects of texting on speaking.

Additionally, exploring the validity of learning new content is essential. Based on previous studies (Littlemore & Eyckmans, 2008; Cucchiarini et al., 2020; Stengers et al., 2016), this research explores lexical phrase acquisition through repetition practice and using tasks. As mentioned in Chapter 2.3, learning lexical phrases is not easy, as they carry less literal meaning. Consequently, language learners are often confused with the combination of the words such as “look at” or “look for.” The current research investigates how texting

supports learning new lexical phrases and how non-articulated language learning affects speaking.

The following research questions are addressed:

RQ 1) How does task repetition in task-based language learning via texting affect post-speaking tests on lexical use, grammar, and syntactic complexity?

RQ 2) How does time pressure in task-based language learning via texting affect post-speaking tests on lexical use, grammar, and syntactic complexity?

RQ 3) Is there a difference between learning new lexical phrases via texting and face-to-face learning?

4.2. The Hypothesis of the Language Production Process

Texting is believed to ease the language production process as texting allows more time for interaction, and the interactions are visually salient. In addition, texting features allow users to monitor and edit their errors, which is considered beneficial for SLA. Accordingly, this study hypothesizes that learning with texting can influence subsequent speech performance. Speaking is considered a complex task for language learners, in which many phases of speech production are not automated. Since speech behavior consists of several concurrent processes, it becomes difficult to retrieve accurate linguistic information and monitor production (Levelt, 1989). Furthermore, Levelt (1989) explains that language learners notice errors depending on the context and that the level of error detection is unstable. Therefore, it is believed that the use of texts can mitigate speech production difficulties in speaking through the inherent features it possesses for communication.

Recent research found self-repairs in texting that are not apparent in the texting logs, found in computer transcripts (Smith, 2008) but contribute to language production. Sauro and Smith (2010) discovered that covert repairs in texting exhibit syntactic complexity, grammatical accuracy, and lexical diversity. Current research predicts that this behavior will affect subsequent speaking performance. Because task conditions affect speech production, the current study considers each variable and its influence on speaking.

4.3. Task Implementation

This section hypothesizes the effect of task implementation based on Sauro and Smith's (2010) findings on syntactic complexity, grammar, and lexical use.

4.3.1. Hypothesis on the Condition of Task Repetition

The rationale for task repetition is that it reuses certain words and grammatical structures or uses more advanced linguistic elements, making future recall faster and less effortful (Jong & Tillman, 2018; Kim & Tracy-Ventura, 2013). In addition, prior experience

automates some retrieval processes by drawing more attention to the message or formulating new messages (Jong & Tillman, 2018).

Previous research found that task repetition impacts accuracy and fluency (Fukuta, 2016; Hawkes, 2012; Bei, 2012; Bozorgian & Kanani, 2017; Lynch & Maclean, 2000; Kim & Tracy-Ventura, 2013). Therefore, texting is hypothesized to resemble the result of face-to-face practice and improves accuracy and fluency.

4.3.2. Hypothesis on the Condition of Time

The time condition has been discussed under two types, pre-task planning and online planning. Online planning has been reported to affect the accuracy and complexity of face-to-face studies. For example, in a repair study, online planning was found to increase error repairs more than appropriate or different repairs, showing attention to accuracy (Ahmadian & Tavakoli, 2014). Since online planning is a feature of texting, it is reasonable to assume that texting leads to the same results and affects the accuracy and complexity of subsequent speech performance.

4.3.3. Hypothesis on Learning Lexical Phrases

Previous studies show that using the task is expected to elicit target language acquisition, and multiword expressions were successfully acquired (Chapter 2.3). This study hypothesizes that engagement in tasks via texting facilitates lexical phrase learning based on previous studies. Although texting is an exchange in written form, speech production follows the same process as face-to-face learning (Payne & Whitney, 2002). However, the absence of the articulation phase in texting may be a disadvantage for the subsequent speaking task. Since articulation is not practiced during texting, the new words may not be used in the subsequent speaking task. Nevertheless, under the same conditions as face-to-face learning, texting practice can improve overall speaking performance as more time and visual text is available for planning online, which may benefit performance (Ahmadian & Tavakoli, 2010).

4.4. Summary

When speaking in a second language, the high demands of parallel processing, working memory constraints, and online planning limitations make oral tasks more challenging, especially in the earlier stages of language learning (Anderson, 2005; Levelt, 1989). This study investigates how texting can solve this problem and support the developmental process. In addition, this study posed three research questions based on previous research on self-repair in texting (Sauro & Smith, 2014).

RQ 1) How does task repetition of task-based language learning via texting affect post-speaking tests on lexical use, grammar, and syntactic complexity?

RQ 2) How does the time pressure of task-based language learning via texting affect post-speaking tests on lexical use, grammar, and syntactic complexity?

RQ

RQ 3) Is there a difference between learning lexical phrases via texting and face-to-face learning?

This study uses the speaking language production model (Levelt, 1989; Payne & Whitney, 2002) and previous research on TBLT. Subsequently, this study predicts that learning L2 through texting will affect accuracy and complexity through the inherent features of texting, namely visual salience and slow exchange. The language process will be analyzed based on the findings that texting repairs affect syntactic complexity, grammar, and lexical usage. In addition, the use of tasks is expected to support language acquisition regardless of modality.

CHAPTER 5 PILOT STUDY 1: SELECTING THE TASK TYPE

Previous studies on texting have provided positive evidence for language learning (Abrams, 2003; Payne & Whitney, 2002). For example, Sauro and Smith (2010) found that texting repair behavior has important consequences for foreign language acquisition. Since texting has been found to share similar characteristics with speaking (Jonsson, 2016), the study of how texting affects speaking has great potential. The literature on TBLT also suggests that the approach to language learning is maximized when there is a context or purpose in the language task. The current study uses the TBLT approach to investigate how texting affects speaking performance.

TBLT includes several factors that influence language learning. According to Lin (2014), variables such as task type have a crucial influence on language production. Before examining the task condition of texting in conjunction with other variables (task repetition and time pressure) in more detail, this chapter discusses and selects a task type that can be used in texting and face-to-face interactions between learners at levels A1-B2 of the Common European Framework of Reference (CEFR). The chapter reports on a pilot study to select the most appropriate task type for subsequent studies.

5.1. Aims and Scope

To further investigate the effect of texting on speaking, eliminating the influence of task type is crucial. This pilot study explored the language production of different task types and selected the type that provides equal opportunity for the participants to finish the task. The selected task is required for minimum disparities among the participants, as the research's goal is to investigate the effect of the texting intervention. Therefore, a detailed language outcome analysis of the different task types used in both modes is compared.

The pilot study explored the following questions in texting and face-to-face modes:

(1) What vocabulary differences are found for CEFR A1-B1 learners in each task type, that is, interview, narration, and decision-making?

(2) In which task type does prior lexical knowledge necessary for CEFR A1-B1 learners?

The three task types were selected based on Foster and Skehan (1996), according to which the three tasks have different structures that affect cognitive load. The interview tasks involve sharing personal experiences, the narrative task involves describing an event in the correct order, and the decision-making task requires finding solutions to a problem. It is assumed that the differences in task types affect the allocation of attention, which in turn affects speech production. The current study investigates the use of vocabulary among these task types for CEFR A1-B1 learners. As novice language learners have a limited lexical repertoire, measuring lexical items may indicate completing the task. Additionally, this pilot study aims to choose a task type that does not require specific vocabulary knowledge to complete the task. If the task type required specific lexical knowledge, it would be difficult to

determine the effect of the intervention. To compare task type with vocabulary, this pilot study analyzes total word count, word error rate, lexical diversity, and vocabulary level.

The study was conducted in both texting and face-to-face modes to identify factors that may affect one mode over the other.

5.1.1. Participants

Participants in the pilot study were 104 students from two faculties and two departments at a college in Japan. The study was conducted from April to December 2014. The ratio of female to male participants was 3 to 1. All participants were non-native English speakers, and the study was conducted in the context of an English course taught by the researcher. The goal of the course was to improve communication skills and integrate all four skills. All instructions were given in both English and Japanese. Students were informed of the purpose of this study, which had no direct impact on their grades. Informed consent was obtained before the study began (Appendix 1). Participants' language proficiency ranged from 80 to 166 (N = 142; M = 130.66; SD = 15.83) according to the institutional Test of English for International Communication (TOEIC) Bridge Program. This was estimated as A1-B1 level on the CEFR.

Participants were randomly paired and divided into groups so that each group of students participated in both face-to-face and online texting tasks. To avoid the possibility of cheating, participants did not sit next to each other. One group worked on the face-to-face task first, while the other group worked on the texting task. Thereafter, they switched task modes. To avoid mode order and task types affecting results, each group followed the task order shown below in Table 2. For example, when the first task was a face-to-face interview task, the second task was either a texting narration task or a texting decision-making task.

Table 2

Order of Tasks for Each Group

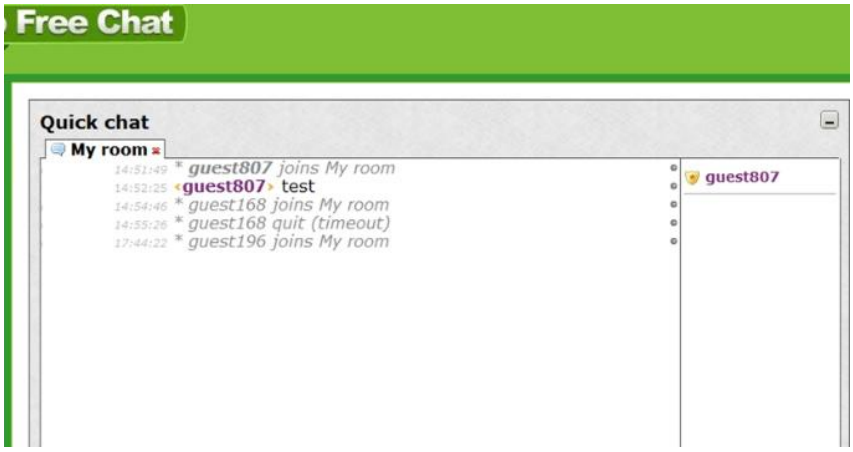
	First Task			Second Task		
group	Mode	Task Type	number of pairs	Mode	Task Type	number of pairs
1	F2F	Interview	11	Texting	Narration	5
				Texting	Decision-Making	6
2	F2F	Narration	11	Texting	Interview	5
				Texting	Decision-Making	6
3	F2F	Decision-Making	8	Texting	Interview	4
				Texting	Narration	4
4	Texting	Interview	8	F2F	Narration	3
				F2F	Decision-Making	5
5	Texting	Narration	8	F2F	Interview	4
				F2F	Decision-Making	4
6	Texting	Decision-Making	6	F2F	Interview	3
				F2F	Narration	3

Note. F2F = face-to-face

The face-to-face tasks were conducted after a one-minute preparation period, while the texting group had no task planning guidance or stipulations. The study took place in the computer lab with CaLabo®EX installed, and participants used the same type of computer. The researcher was able to control the computers remotely if needed and was able to monitor the screens and listen to the participants' exchanges from their seats. In the face-to-face group, each student had a headset, and all exchanges were recorded. The groups performed each task at their own pace and without time constraints. They were asked to speak only English, and their discussions were recorded and transcribed for analysis.

For the texting tasks, an open-source Common Gateway Interface package (http://www.phpfreechat.net/) was used (*Figure 1*), and participants were assigned to different chat rooms so that only designated pairs could join. The texting application allowed only text-based exchanges, and videos, avatars, and audio were not available. In addition, features and applications such as spell check, autocorrect, and translation were blocked to ensure language practice.

Figure 1 *Online Chatrooms for the Texting Group*



5.1.2. Task Types

The types of tasks used were based on those defined by Foster and Skehan (1996), and the content of each task was modified to match the language proficiency and familiarity of the students. In the interview task, students were asked to conduct an interview in which they explored the advantages and disadvantages of using computers and the Internet. Each student took turns asking and answering questions. In the storytelling task, students were given clips from a cartoon and asked to retell the story of the part they were given. In the end, the students completed the cartoon strip together, which represented the whole story. While one student told their part of the story, their partner had to sketch the scene (see Appendix 2 for an example). The decision-making task required partners to decide whom they would save in the event of a fire and in what order they would save them. As in Foster and Skehan's (1996) study, the three tasks in the present study were hypothesized to reflect different levels of attention and to influence the familiarity and predictability of the problem. It is hypothesized that less familiar and predictable tasks will increase cognitive load and impair language performance (Foster & Skehan, 1996).

5.1.3. Data Analysis

All data resulting from student performance on the tasks were collected and used for analysis. The face-to-face conversations were transcribed, and the texting transcripts were retrieved from the server in texting mode. The face-to-face transcripts were first transcribed by a person with experience in L2 instruction. Then, the researcher checked the transcription for errors. Finally, analysis was conducted using the transcription of both modes. Only English words were counted in the text transcripts; symbols, emoticons, and L1 words were

marked. Spelling errors were corrected, and messages generated by the software were deleted.

Vocabulary. Vocabulary errors were first checked for gaps in accuracy. The errors comprised word misuse, unidentifiable words, repetition, the use of L1 words, unnecessary words, and lack of words. Fillers were counted if they filled in any missing words. The vocabulary error rate is a percentage of all the errors by the total number of words. The total number of words includes all the words produced for completing each task. Grammatical errors were not included in the vocabulary errors. To ensure the reliability of the error rate analysis, the obtained data were analyzed by another evaluator. First, 20% of the data was rated for errors by the researcher. Then, the second evaluator (who has a master's degree in teaching foreign languages) assessed the same data. The agreement between the two raters was r = .87. The discrepancies were discussed and solved. Accordingly, the rest of the data were analyzed by the initial evaluator. The vocabulary error rate was calculated by dividing the total number of errors by the number of words and converting them into percentages.

Before analyzing lexical diversity, filler words, repetitions, self-repairs, and L1 usage were extracted. The Measurement of Textual Lexical Diversity (MTLD) was adapted to measure lexical diversity (McCarthy & Jarvis, 2010). To calculate MTLD, the text was first divided into sequences with the same type-token ratio (TTR; set to 0.72). The average length of the sequences was then calculated based on the TTR sets. The higher the MTLD value, the more diverse the vocabulary used.

Vocabulary level was measured to compare the complexity of vocabulary that participants used in the tasks. A lexical analysis was performed using the AntWord profiler program (Anthony, 2014) and Range for Text v.5 (Heatley et al., 2020). The words used in the tasks were then compared with the word lists from The New JACET (The Japan Association of College English Teachers) List of 8,000 Basic Words (2016). These word lists consisted of 8,000 words and were created based on the frequency of words appearing in British National Corpus and the Corpus of Contemporary American English. English textbook-authored entrance exams and proficiency tests in Japan were also used.

Statistical Analysis. The current study was conducted in a natural classroom with a limited sample size. For some measures, the data contained outliers. Therefore, a non-parametric test was considered. To compare results between groups, the Kruskal-Wallis test was chosen, and the Mann-Whitney U test with adjusted Bonferroni analysis was applied for the post hoc test. Independent variables were the interview, narrative, and decision-making tasks, and dependent variables were the linguistic scores from face-to-face and texting modes.

R statistical software was used for the calculations. The letter “n” represents the number of pairs.

5.2. Results and Discussion

5.2.1. The Total Number of Words Produced

The Kruskal-Wallis test was applied to find the differences in the total number of words among the mode and task types. Table 3 illustrates the descriptive statistics of each group. Statistical difference was found among the groups to reject that the total number of words is the same among the groups ($\chi^2 = 58.15$, $df = 5.00$, $p = .00$, $r = .66$). The Mann-Whitney U test was conducted to test pairwise comparison. For the face-to-face group, there was no difference ($p > .05$) among the three task types (interview- narration: $Z = 1.81$, $r = .32$, $p = 1.00$; narration- decision-making $Z = 0.74$, $r = .13$, $p = 1.00$; interview-decision-making $Z = 0.51$, $r = .09$, $p = 1.00$). For the texting group, the interview task was significantly different from the decision-making task ($Z = 3.17$, $r = .54$, $p = .02$) and narration task ($Z = 3.74$, $r = .64$, $p = .00$). The results for the texting group showed that interview task produced fewer words compared to other texting tasks. The increased number of words for narration and decision-making tasks compared to the interview task may reveal that more vocabulary was necessary to carry out the task.

Table 3 showed that the standard deviation score was lower in the interview task than in the decision-making task in both face-to-face (Interview: $SD = 20.70$; Decision-making: $SD = 26.58$) and texting mode (Interview: $SD = 17.00$; Narration: $SD = 39.41$; Decision-making: $SD = 46.81$) (*Figure 2*). The standard deviation results may imply that some learners had difficulty producing words and engaging in decision-making tasks in both modes.

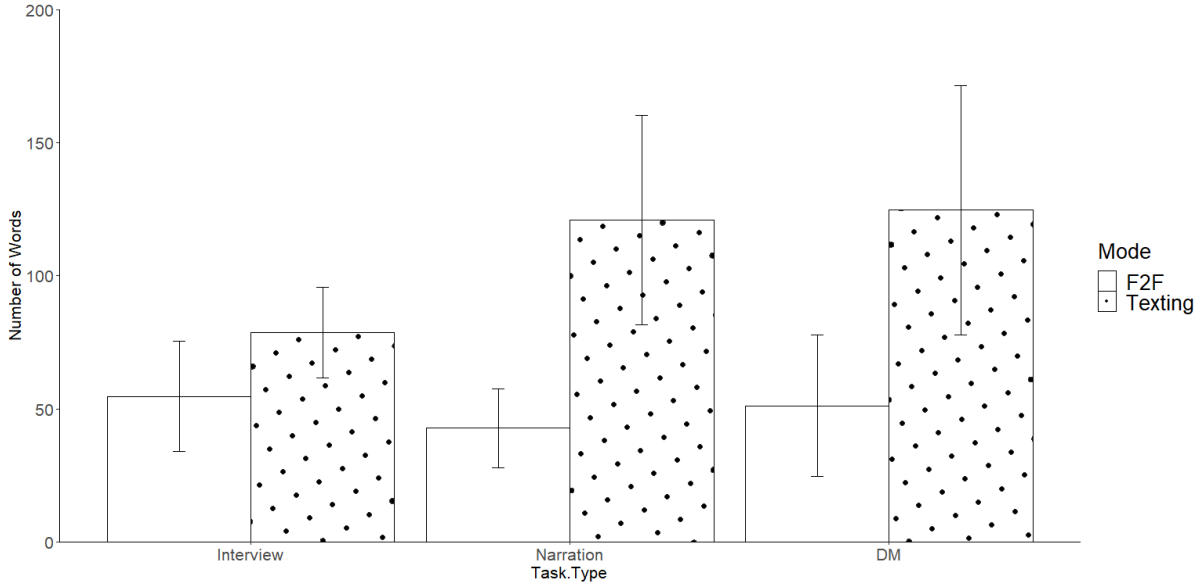
Table 3

Total Number of Words Used for Pilot Study 1

Mode	Task Type	*n*	*M*	*Mdn*	*SD*
F2F	Interview	18	54.61	52.5	20.70
F2F	Narration	14	42.71	41.5	14.95
F2F	Decision-Making	17	51.12	49	26.58
Texting	Interview	17	78.71	74	17.00
Texting	Narration	18	120.94	115	39.41
Texting	Decision-Making	18	124.67	123.5	46.81

Note. F2F = face-to-face

Figure 2 *Histogram for the Total Number of Words Produced for Each Task Type*



Note. F2F = face-to-face, DM= Decision-Making, Error bar indicates standard deviation.

5.2.2. Vocabulary Error Rate

The Kruskal-Wallis test result showed the score for vocabulary error rate is not the same among the six groups ($\chi^2 = 44.09$, $df = 5.00$, $p = .00$, $r = .56$). For the face-to-face group, there was no difference ($p > .05$) among the task types (interview- narration $Z = 0.36$, $r = .06$, $p = 1.00$; narration- decision-making $Z = 1.81$, $r = .33$, $p = 1.00$; interview-decision-making $Z = 2.66$, $r = .45$, $p = .12$). For the texting groups, statistical difference was found that narration task was significantly higher to interview task ($Z = 3.41$, $r = .58$, $p = .01$) and decision-making task ($Z = 3.37$, $r = .57$, $p = .01$). Narration task in texting scored the highest vocabulary error mean ($n = 17$, $M = 11.70$, $Mdn = 11.70$, $SD = 2.43$) in the texting group.

For the standard deviation score, the error rate was the smallest for the interview task face-to-face (Interview: $SD = 3.39$; Narration: $SD = 5.23$; Decision-making: $SD = 5.55$). The texting mode also showed similar results in terms of the standard deviation, with the

interview task scoring lower than the decision-making task (Interview: *SD* = 3.17; Decision-making: *SD* = 3.61) (Table 4, *Figure 3*).

The result of the word error rate score and the standard deviation mean score implies that more participants struggled in the narration and decision-making task compared to the interview task. The significant difference in the error rate for texting showed that the narrative task produced more errors, suggesting that participants may have found it difficult to complete the task.

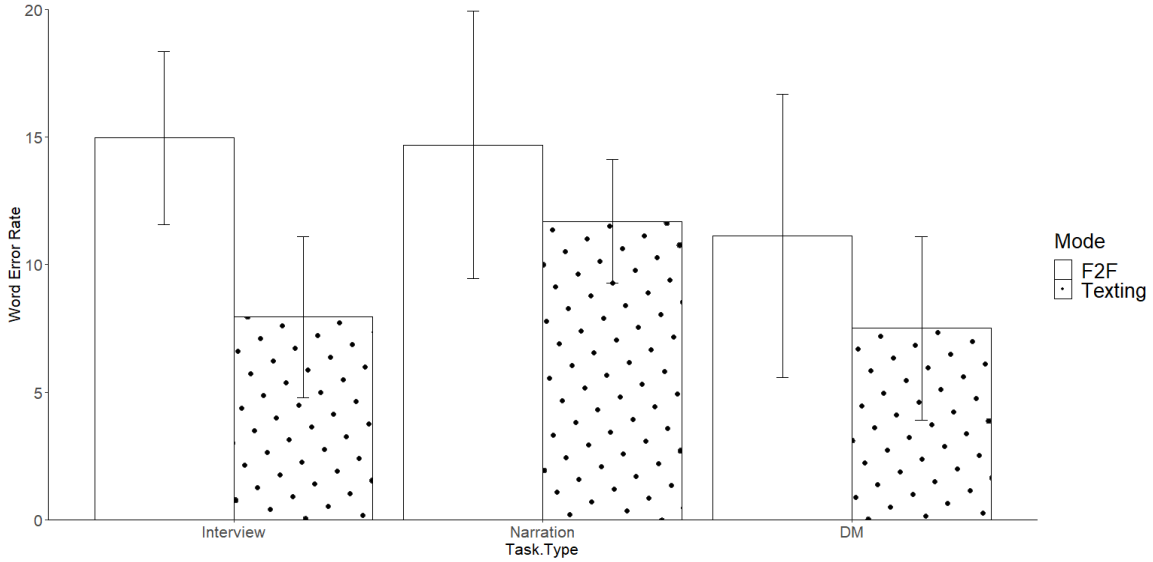
Table 4

Descriptive Statistics for Word Error Rate for Pilot Study 1

Mode	Task Type	*n*	*M*	*Mdn*	*SD*
F2F	Interview	18	14.96	14.55	3.39
F2F	Narration	14	14.69	14.61	5.23
F2F	Decision-Making	17	11.13	11.31	5.55
Texting	Interview	17	7.94	7.36	3.17
Texting	Narration	17	11.70	11.70	2.43
Texting	Decision-Making	18	7.50	7.72	3.61

Note. F2F = face-to-face; *n* = the number of pairs

Figure 3 *Histogram for Word Error Rate for Each Task Type*



Note. F2F = face-to-face, DM = Decision-Making, Error bar indicates standard deviation.

5.2.3. Lexical Diversity

Lexical diversity results are represented in the MTLD score. The Kruskal- Wallis test result found to reject that MTLD scores are the same among the six groups ($\chi^2 = 29.32$, $df = 5.00$, $p = .00$, $r = .42$). There was no difference ($p > .05$) among the task types in face-to-face mode (interview- narration $Z = 0.08$, $r = .01$, $p = 1.00$; narration- decision-making $Z = 0.91$, $r = .16$, $p = 1.00$; interview- decision-making $Z = 1.06$, $r = .18$, $p = 1.00$). However, for the texting groups, a significant difference was found between the interview and decision-making task ($Z = 4.39$, $r = .74$, $p = .00$). The result shows diverse words were necessary to complete the decision-making task ($n = 18$, $M = 34.94$, $Mdn = 33.65$, $SD = 7.39$). The scores' differences may be due to the variety of vocabulary elicited by the task. A decision-making task requires giving a situation to solve a problem. Thus, the related vocabulary for solving the problem is necessary for discussion. Therefore, it can be assumed that various words were necessary to complete the decision-making task for both groups.

The standard deviation also showed differences in scores (Table 5), suggesting that some participants had a lower score than others who may have had difficulty finding the right words. The lack of prior knowledge of the vocabulary for a particular task may have influenced the results. The standard deviation was lowest for the interview task for both modes (interview task F2F: $SD = 6.95$; interview task texting: $SD = 4.28$) (*Figure 4*). Possibly, the interview task was more successful for larger participants.

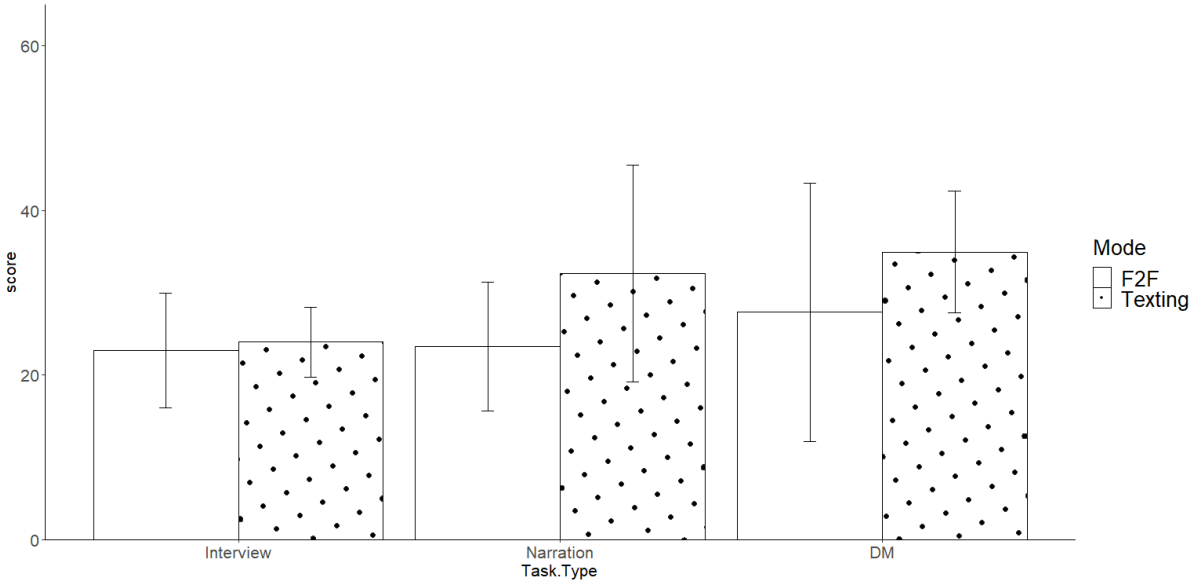
Table 5

Descriptive Statistics for Lexical Diversity for Pilot Study 1

Mode	Task Type	*n*	*M*	*Mdn*	*SD*
F2F	Interview	18	23.01	20.87	6.95
F2F	Narration	14	23.47	21.23	7.78
F2F	Decision-Making	17	27.66	27.12	15.69
Texting	Interview	17	24.00	23.04	4.28
Texting	Narration	17	32.38	31.06	13.15
Texting	Decision-Making	18	34.94	33.65	7.39

Note. **F2F = face-to-face**

Figure 4 *Histogram of Lexical Diversity for Each Task Type*



Note. F2F= face-to-face, DM= Decision-Making, score = MTLD score. Error bar indicates standard deviation.

5.2.4. Level of Vocabulary

Further vocabulary analysis was conducted to find out any differences in vocabulary level usage among the task types. Table 6 shows the percentage of words used in each level of the New JACET List of 8,000 Basic Words (2016). Base Level 1,000 is a list that includes the 1,000 most frequently used words in the English language. Base Level 2,000 includes the second 1,000 most frequently used words. Word level percentages were calculated for the first 1,000 words and words beyond base level 2,000 (Table 10). Proper nouns and first-language vocabulary were not included in any base-level lists.

As shown in *Figure* 5 and *Figure* 6, the levels of vocabulary used in the three tasks followed the same pattern. These results seem reasonable given that the student's proficiency levels were equal across the groups in the two modes and three tasks. The result indicates that the narration task produced higher-level words in both the face-to-face and texting modes. This may be because the task involved narrating a comic-strip scene, which required employing words from the 2,000-8,000 Base Levels. Some examples of the words in the 2,000-8,000 Base Levels lists include "messenger," "dash," and "stripes."

The results for the vocabulary level show that it would be difficult to continue the task if the participants did not have access to 2,000-8,000 level words. A higher percentage of Base level 2,000- 8,000 words in narration tasks compared to other task types could indicate greater difficulty.

There is still concern regarding the decision-making task in which the percentage of base level 2,000 to 8,000 words was the highest after the narration task. In conjunction with

the standard deviation results from linguistic diversity, it could imply a disparity among the participants in completing the decision-making task as it may require a certain vocabulary as a prerequisite to completing this task as well as the narration task.

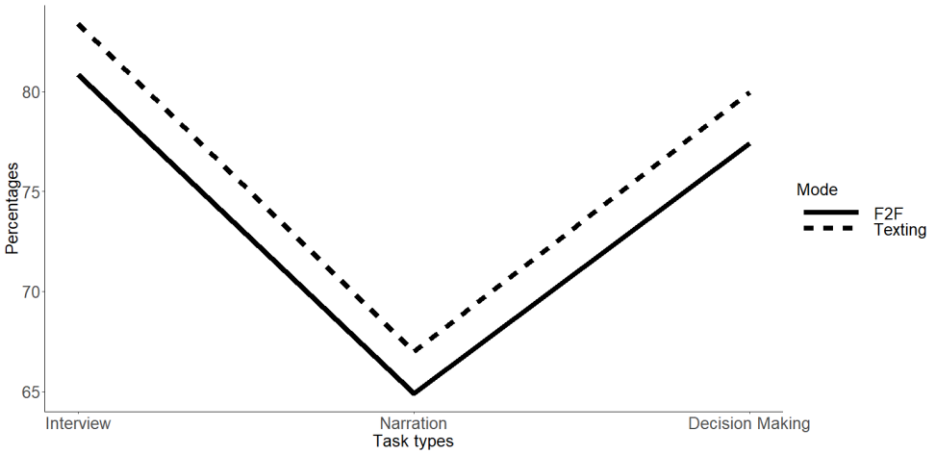
Table 6

Word Level Percentages for Each Task Type

	F2F			Texting		
	Interview	Narration	Decision-Making	Interview	Narration	Decision-Making
Base level 1000	80.85	64.91	77.41	83.39	67.00	79.95
Base level 2000-8000	2.86	7.66	4.56	2.71	9.36	5.64
Words not on the list	16.28	27.43	18.03	13.9	23.64	14.41

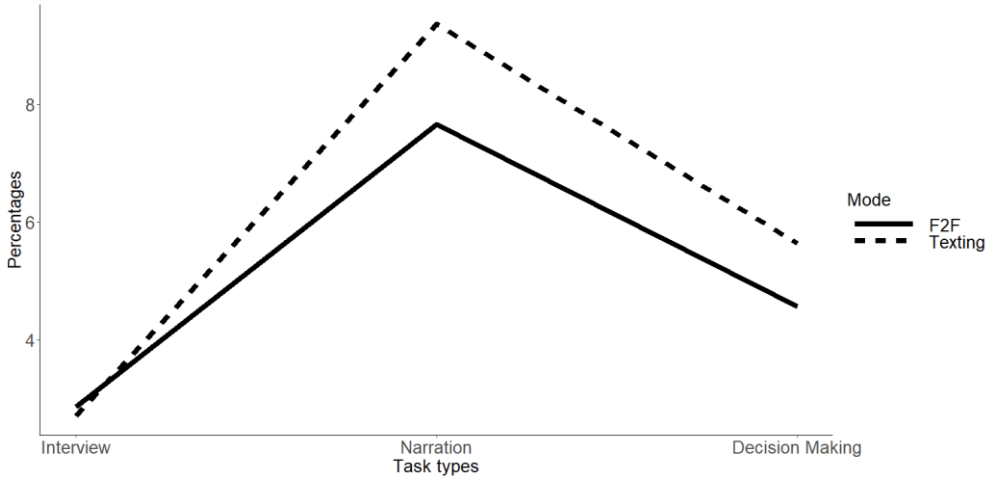
Note. F2F = face-to-face

Figure 5 *Percentages of Base Level 1,000 Words*



Note. F2F = face-to-face

Figure 6 *Percentages of Base Levels 2,000-8,000 Words*



Note. F2F = face-to-face

5.3. Conclusions: Task Type Selection

The purpose of this pilot study was to answer the two proposed research questions and to select a task type for subsequent main studies. To answer the first research question on vocabulary use between task types, statistical analysis showed differences between task types in total word count, lexical error rate, and lexical diversity, indicating that the interview task is different from the narrative and decision tasks. A stepwise comparison in the texting group showed that the narrative and decision-making tasks had higher total word count scores. In addition, the lexical error rate was significantly higher in the narrative task than in the other two tasks in texting, and lexical diversity was higher in the decision-making task than in the interview task in both modes. In terms of standard deviations of lexical diversity, the narrative and decision-making tasks had higher values, indicating inequalities between participants. The result indicated that narration and decision-making tasks might have been challenging in terms of vocabulary for the participant's level.

Further analysis of participants' word-level usage revealed that the narration task required an advanced vocabulary. Referring to the previous studies, linguistics constraints may result in low performance (Crowther et al., 2015; Derwing et al., 2004), and CEFR A1-B1 learners may have difficulty completing the task.

The second research question, namely, what type of task requiring prior knowledge has the least impact on performance, was critical to the selection of the interview task. The interview task allows CEFR A1-B1 students to complete tasks without requiring a specific vocabulary level or other language skills. In addition, the interview task was found to have lower variance in vocabulary error rates and linguistic variety scores; therefore, this task type provides learners of both modes with equal opportunity.

Regarding selecting the task for the main study, the interview task may best suit the purpose considering the difficulty and because it requires minimum prior linguistic knowledge to complete the task. In addition, in terms of familiarity, the interview task is commonly used in the language classroom, and learners do not need higher cognitive skills to complete it. Thus, it is fair to learners who have not engaged with unfamiliar English communication tasks such as narration or decision-making. Accordingly, the interview task may be suitable for CEFR A1-B1 learners.

That being said, the interview task requires careful attention to the degree of freedom it allows in vocabulary, sentence structure, and content (Hu, 2018). Compared to narration or decision-making tasks, less structured tasks allow for greater freedom, leading to disparities in information among individuals.

5.4. Summary

As far as we know, no previous study has examined the influence of task type on texting. Therefore, the aim of this study was to compare speech production between task types in face-to-face and texting and to select the task type to be used in the main studies. In addition, two research questions were pursued: The first examined the differences in vocabulary that participants showed in each task type. The second was to identify the task type that was least influenced by the prior knowledge of CEFR A1-B1 learners.

The research presented in this chapter examined the effect of the task type (interview, narration, and decision-making tasks) on participants' performance in the face-to-face and texting modes.[1] The 104 participants were divided into 6 groups and engaged in both face-to-face and texting tasks. The responses were analyzed for vocabulary use (total number of words, vocabulary error rate, lexical diversity, level of vocabulary). By comparing the results, the interview task was chosen to fit the purpose best as the standard deviation score showed lower disparities among the participants on the analyzed measures. In addition, higher-level (Base Levels 2,000-8,000) vocabulary words such as "messenger," "dash," and "stripes" was necessary to complete the narration task, which may not have matched the proficiency level of all the participants. Furthermore, the narrative and decision-making tasks also require reasoning and language skills that may have overwhelmed students, as these skills correspond to a higher CEFR level. An analysis of vocabulary revealed that the interview task relied least on prior knowledge. Therefore, it was selected as the task type for the main studies.

[1] The linguistic analysis has been modified from the researcher's previous study to investigate the research questions of the pilot study (Takase, 2014).

CHAPTER 6 STUDY 1: TASK REPETITION CONDITION

The pilot study described in the previous chapter suggests that task type affects language production, and the interview task was confirmed to be the best fit for the proficiency level of CEFR A1-B1 learners. Thus, it has been adopted for the main studies. The first main study, Study 1, explores the condition of task repetition, one of the variables that affect performance in task-based learning (Skehan, 2014). This study compares the speaking performance under the same task repetition condition in both face-to-face and texting modes to identify the effect of texting on speaking.

The research presented in Chapter 6 aims to answer the following research question RQ (1) How does task repetition in task-based language learning via texting affect tests of syntactic complexity, grammar, and lexical usage after speaking (Sauro & Smith, 2010)?

6.1. Method

6.1.1. Participants

Study 1 involved 32 first-year students attending an English communication class; the study formed part of the class's content. The students were told about the purpose of this research and that the results do not reflect their grades. All students were non-native English speakers studying at a prefectural university in Japan. The instructions for this study were conducted both in English and Japanese. The participants' language proficiency TOEIC IP scores ranged from 170 to 440 (M = 348.79, SD = 71.25, N = 33), placing them at the A1 to A2 beginners' level in the CEFR.

6.1.2. Task type

Based on the findings from the pilot study, the interview task was adopted. This task type matched the knowledge of the participants. In addition, the participants had already worked on interview tasks in other class assignments, so they were familiar with conducting interviews in their L2.

6.1.3. Procedure

The students took part in an English communication class from October to November 2017, integrating the four skills. Two groups were formed: a face-to-face group (the control group) and a texting group (the experimental group) from two classes taught by the same instructor. Before the intervention, both groups participated in a pre-test interview task. Then, they recorded a one-minute speech on what they did over the summer using the recording software in the university's computer-assisted language learning (CALL) lab, where CaLabo®EX is installed. All the students were using headsets that had microphones for recording. Thereafter, another one-minute speech was recorded as a post-test after the intervention tasks.

In the intervention interview task, students were grouped into pairs and asked their partners what they did during summer vacation. The task was repeated three times with different partners. Students were also asked to switch roles to ensure that each student both asked and answered questions. Although the number of repetitions of a task for optimal efficiency varies across studies (as research objectives differ), a recent study on task repetition suggests three to four repetitions per set (Lambert et al., 2017). In addition, popular activities for improving speaking fluency, such as the 4/3/2 technique, use three repetitions (Boers, 2014; Jong & Tillman, 2018; Lambert et al., 2017); therefore, in the present study, participants were asked to repeat the task three times.

Additionally, different partners were paired each time to "preserve task integrity" (Lambert et al., 2017, p.31). Task repetition may cause boredom and fatigue (Bygate, 1996); thus, working with different partners facilitates different interpretations of the task. In the face-to-face group, learners selected their new partners after each round. Whereas with the texting group, the teacher assigned a different partner per round.

Learners in the face-to-face group recorded each interview round using the same recording software used for the pre-tests and post-tests. The texting group used the chat function in CaLabo®EX. Only the texting feature was used for the task. The participants were familiar with the chat function in the CALL lab because they had used it in other class activities. The computers in the CALL lab could be controlled and monitored from the instructor's console. Spell check, auto-correct, and translation were turned-off, and learners could not search web pages. In such a case, a warning was sent to both the teacher and the user. Learners were instructed to use English only. After the post-test, all learners were asked to report to the teacher what they had learned from each other.

6.1.4. Data Analysis

The sound files for the pre- and post-tests were collected and transcribed for analysis. The transcription was first transcribed by a competent English speaker. It was then reviewed by the researcher for missing information. Only the data set that included full participation in all tasks was used; therefore, data from 32 participants were ultimately analyzed. In addition, the transcribed data and the texting transcripts were checked for spelling and all computer-generated messages were deleted.

For statistical analysis, because the research was conducted in a natural classroom setting with a small sample size, the non-parametric test was conducted. Paired Wilcoxon signed-rank test was applied between the paired pre-and post-test results. Statistical analysis software R and statistical power G*Power 3 software (Faul, Erdfelder, Lang, & Buchner, 2007) were used for calculation. The effect size is referred to as small, $r = .10$; medium, $r = .30$; and large, $r = .50$ (Mizumoto & Takeuchi, 2011).

Vocabulary. Vocabulary was analyzed in terms of lexical diversity and error rate. Vocabulary measurements followed the same procedures as in Pilot Study 1. The MTLD was

used to measure lexical diversity. To investigate linguistic accuracy, the errors were analyzed in detail by counting the errors for both the content words and the function words. The lexical error rate for content words was calculated by dividing the total number of errors for content words by the total number of content words and then converting them to percentages. The lexical error rate for function words was also calculated using the same procedure as for function word errors. For further analysis, the error types for content words and function words were categorized. The errors were manually reviewed by two raters who have experience in teaching English as a foreign language and have conducted research in this area. The second evaluator scored 20% of the data from each group. The agreement between the two evaluators was $r = 0.84$. Errors that included incorrect word choice, unidentifiable words, repetition, use of L1, and missing words for both content and function words were counted as one error. Filler words that filled in missing words were also counted as errors. Grammatical errors such as incorrect plural forms and tense problems were not counted as lexical errors. After the discrepancies were resolved, the first rater independently coded the remaining data. Vocabulary errors for both content and function words were then manually grouped according to the error type determined by the researcher. Error type labeling was checked on 20% of the data by an experienced native English speaker. The agreement was 95%. The discrepancies were discussed and changed. Then the researcher labeled the rest of the data, and the error types were counted manually.

Grammar. Grammatical correctness was assessed by analyzing participants' use of the past tense, as the task topic required participants to talk about what they did during their summer vacation. A Stanford Parts-Of-Speech tagger (Toutanova et al., 2003) was used to detect verbs in the data, and the past tense error rate was calculated by dividing the number of past tense errors by the total number of verbs and then converting them to percentages. Past tense errors were analyzed by two evaluators. Only verbs that contained errors in the past tense were counted. First, the researcher checked the errors. Then, 20 percent of the data from each group was checked by the second evaluator, the same person who checked the lexical errors. The agreement was $r = .98$. After all discrepancies were reviewed and discussed, the first evaluator counted the grammatical errors for the rest of the data.

Syntactic Complexity. The IPSyn (Scarborough, 1990) score measures syntactic complexity by reviewing instances of noun phrases, verb phrases, questions, negation, and sentence structures. Each receives a maximum of two points per utterance. It is recommended that the IPSyn score be used between 50 to 100 per utterance; thus, the randomly selected utterances were added and compared for analysis. To compare the score within groups, the baseline utterance was set to 43, the least number of utterances formed in a single group, including Studies 2 and 3 in this research. Then, random speech data from each participant were collected based on stratified sampling. Repeated words, self-repair words, fillers, and the use of L1 were eliminated before the analysis. Finally, the IPSyn score was

calculated using Computerized Language ANalysis (CLAN) (MacWhinney, 2000). Table 7 shows the number of utterances per group and the ratio of utterances analyzed.

Table 7

Utterances Analyzed per Group for the IPSyn Score (Study 1)

		n	*Mdn*	*SD*	Total Number of Utterances per Group	Ratio of utterances analyzed (%)
F2F	Pre	14	5.00	1.56	70	61.43
F2F	Post	14	5.64	1.84	79	54.43
Texting	Pre	18	5.00	1.80	90	47.78
Texting	Post	18	5.83	2.73	105	40.95

Note. F2F = face-to-face

Fluency. Fluency is a construct that encompasses several dimensions. Articulation rate and speech rate are used to measure fluency. Speech rate is a value that reflects the total number of words produced in one minute. Articulation rate, on the other hand, is measured by the number of words articulated in an utterance. Because speech often occurs in chunks of words, Suzuki and Kormos (2019) found that articulation rate best predicts comprehensibility among various linguistic dimensions. Since this study uses the interview task between two people, articulation rate is also used to measure speech comprehensibility. The articulation rate is calculated as the number of mean words by phonation time. Phonation time refers to the time between one pause to the next. The pauses are set to -25 (dB) (De Jong & Bosker, 2013; 2018; Saito et al., 2019). *Praat* (Boersma & Weenink, 2019) was used to measure pauses. For this analysis, words that contain repetition, repairs, fillers, and L1 words are deleted.

6.2. Results

6.2.1. Lexical Diversity and Lexical Error Rate

Table 8 presents the descriptive statistics for the lexical diversity and error rate of the content and function words produced in the pre-tests, post-tests, and paired t-test results. The paired Wilcoxon signed rank test showed no difference in pre and post-test results on lexical diversity for both modes ($p > .05$). The error rate for both content and function words improved for the texting group on pre- and post-tests (*Figure 7, Figure 8*). The statistical power obtained was higher than .80, indicating strong significance (content word error rate: power = 0.97; function word error rate power = 0.82). For the face-to-face group, the results

showed no significance on lexical error rate for both content and function words ($p > .05$), and mean and median scores gained a higher error rate on the post-test (Table 8).

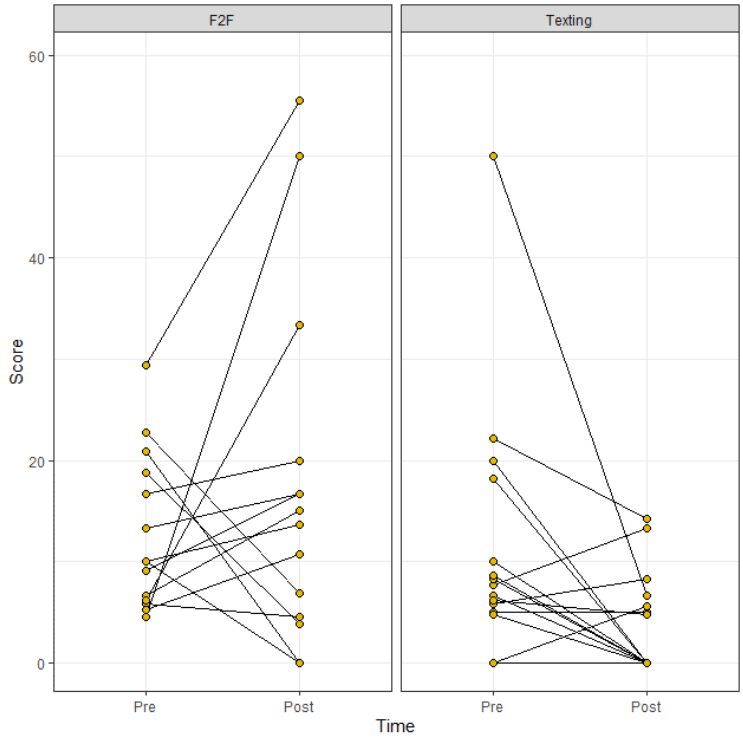
Table 8

Results of Lexical Diversity and Lexical Error Rate for Study 1

		Pre-Test				Post-Test				Pretest-Posttest comparison				
	Mode	*n*	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*		Power (1-β)
Lexical Diversity	F2F	14	27.24	24.86	10.63	14	29.61	27.83	9.82	0.46	0.73	0.14		0.14
	Texting	18	29.27	25.54	13.69	18	27.28	22.29	13.07	0.50	0.68	0.11		0.10
Content Word Error Rate	F2F	14	12.82	10.00	7.71	14	17.63	14.32	17.34	0.43	0.80	0.15		0.27
	Texting	18	10.66	7.18	12.08	18	3.22	0.00	4.78	0.01	2.61	0.44	**	0.97
Function Word Error Rate	F2F	14	12.97	12.25	9.84	14	16.53	15.04	7.09	0.27	1.11	0.21		0.37
	Texting	18	15.86	14.36	12.70	18	7.22	5.97	7.16	0.02	2.26	0.38	**	0.82

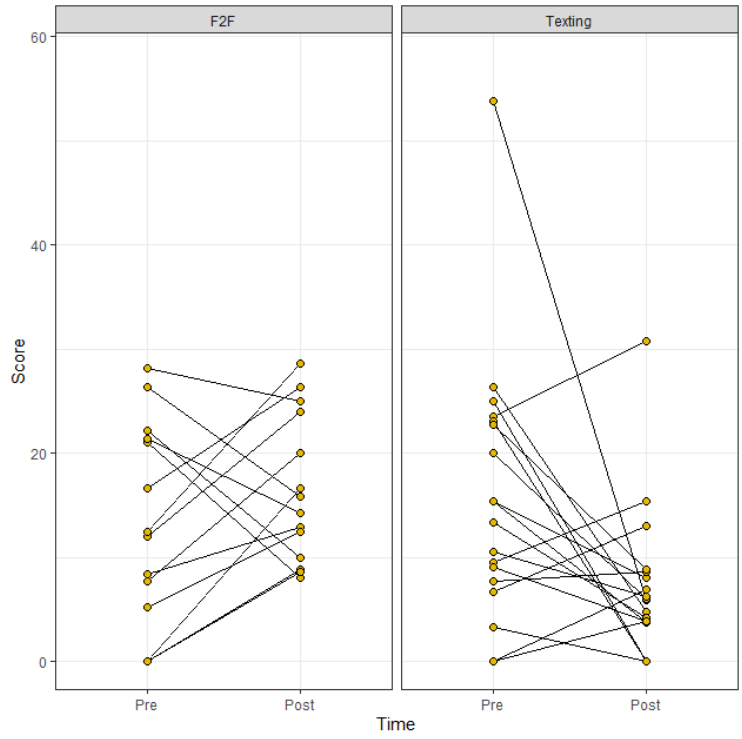
Note. *small r = .10, **medium r = .30, ***large r = .50 (Mizumoto & Takeuchi, 2011)

Figure 7 *Content Words Error Rate for Study 1*



Note: F2F=Face-to-face, Score represents the percentage of errors on content words.

Figure 8 *Function Words Error Rate for Study 1*



Note: F2F=Face-to-face, Score represents the percentage of errors on function words.

Error Types for Content Words and Function Words. Table 9 and Table 10 present the types of content and function word errors that occurred: the numbers represent the accumulated raw occurrence of each error type per data set. Table 9 shows that the use of L1 and repetition were not found in the post-test for the texting group. Under the repeated task condition, the face-to-face practice increased L1 use and repetition in the post-test. It is worth noting that incorrect word choice declined in the post-test for the face-to-face group.

Table 9 *Number of Error Occurrences for Content Words for Study 1*

	Pre-F2F	Post-F2F	Pre-Texting	Post-Texting
Incorrect choice	20	10	8	9
Use of L1	0	11	5	0
Missing (Fillers)	5	4	7	5
Repetition	5	7	4	0
Other (Unidentifiables)	1	1	1	0

Note. F2F = face-to-face.

For function word errors, there was no evidence of repetition or missing function words in the post-test for the texting group (Table 10). For the face-to-face group, the repetition of function words increased nearly three times. In addition, other error types, such as unnecessary words found to increase.

Table 10 *Number of Error Occurrences for Function Words for Study 1*

	Pre-F2F	Post-F2F	Pre-Texting	Post-Texting
Incorrect choice	16	17	16	9
Repetition	11	34	8	0
Unnecessary	8	15	25	14
Missing	7	5	11	0
Other (Unidentifiables)	0	1	0	0

Note. F2F = face-to-face.

6.2.2. Grammar

The Wilcoxon signed rank pair test showed no statistical difference between the two modes on the grammatical error rates in using past-tense verbs ($p > .05$). However, the mean and median score showed a decline for both modes (F2F Pre: $M = 25.08$, $Mdn = 26.14$, F2F Post: $M = 24.35$, $Mdn = 18.33$; Texting Pre: 17.01, $Mdn = 15.48$, Texting Post: 13.85, $Mdn = 10.12$), which shows that both mean in the pre-and post-tests for either mode (Table 11).

Table 11 *Results of Grammatical Error Rate for Study 1*

		Pre-Test				Post-Test				Pretest-Posttest comparison			
	Mode	*n*	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*	Power (1-β)
Grammatical (Past-tense verb) Error Rate	F2F	14	25.08	26.14	18.41	14	24.35	18.33	21.75	0.94	0.07	0.01	0.05
	Texting	18	17.01	15.48	17.6	18	13.85	10.12	14.29	0.45	0.78	0.13	0.14

Note. *small *r* = .10, **medium *r* = .30, ***large *r* = .50 (Mizumoto & Takeuchi, 2011)

6.2.3. Syntactic Complexity

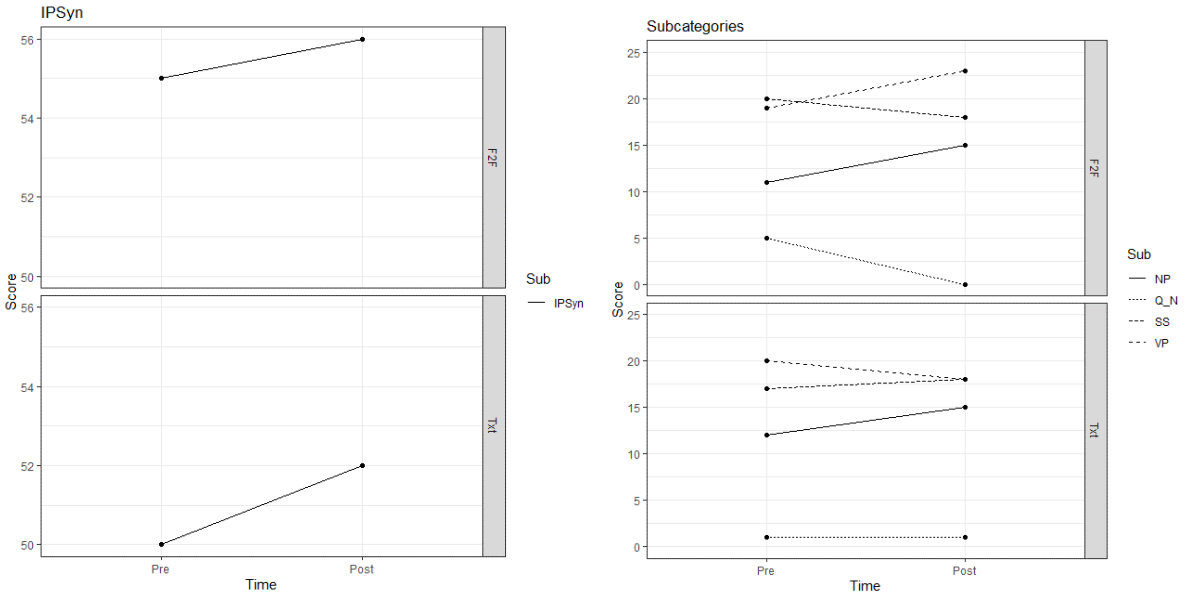
Table 12 shows the IPSyn score and syntactic complexity scores for each subcategory. Both groups increased their overall scores in the post-test. However, the face-to-face group lost points in the sentence structure subcategory (F2F pre = 20, F2F post = 18; Texting pre = 17, Texting post = 18). For verb phrases, scores increased for face-to-face (pre = 19, post = 23) but decreased for texting (pre = 20, post = 18).

Table 12 *Results for the IPSyn Score for Study 1*

		IPSyn Score	Noun phrase	Verb Phrase	Question/ Negation	Sentence Structure
F2F	Pre	55	11	19	5	20
Texting		50	12	20	1	17
F2F	Post	56	15	23	0	18
Texting		52	15	18	1	18

Note. F2F = face-to-face.

Figure 9 *IPSyn Score for Study 1*



Note. F2F = face-to-face.NP = noun phrase, Q_N = question and negation, SS = sentence structure, VP = verb phrase

6.2.4. Fluency

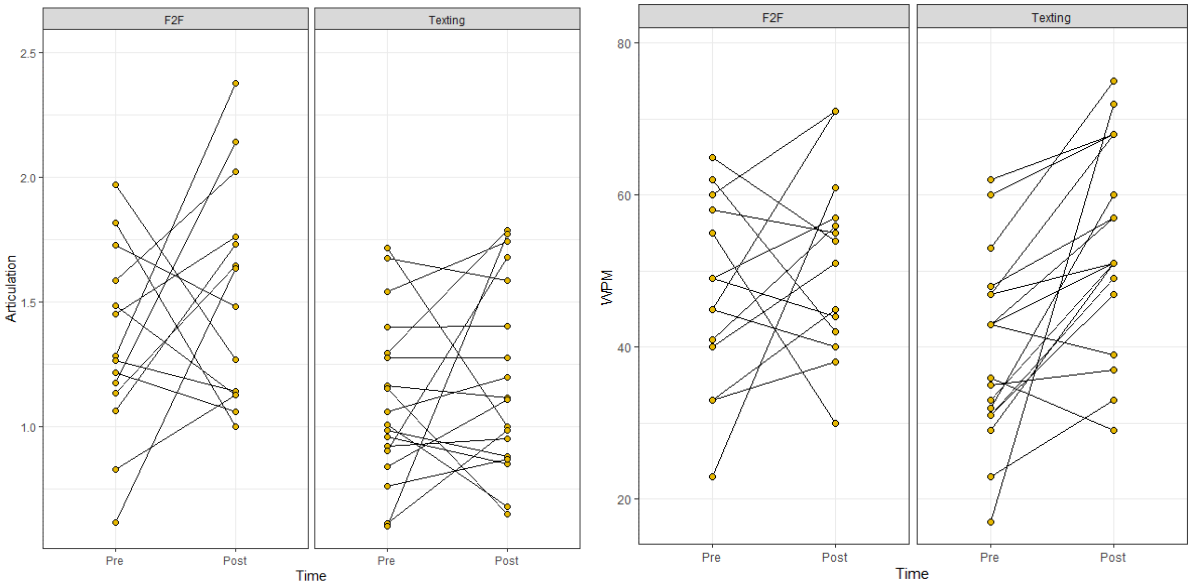
The statistical results showed that the texting group significantly increased the speech rate in the post-test with a high statistical power of .99 (Table 13, *Figure 10*). Face-to-face showed no significant difference ($p > .05$). The result on articulation rate did not show any statistical difference between the two modes ($p > .05$) (Table 13).

Table 13 *Results of the Fluency Measures for Study 1*

		Pre-Test				Post-Test			Pretest-Posttest comparison				
	Mode	*n*	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*		Power (1-β)
Articulation Rate	F2F	14	1.33	1.27	0.37	14	1.54	1.56	0.44	0.24	1.17	0.22		0.54
	Texting	18	1.10	1.03	0.33	18	1.20	1.11	0.38	0.42	0.81	0.14		0.23
Speech Rate	F2F	14	47.00	47.00	12.32	14	51.07	52.50	12.03	0.35	0.94	0.18		0.26
	Texting	18	39.61	39.50	12.11	18	53.50	51.00	13.49	0.00	3.40	0.57	***	0.99

Note. *small r = .10, **medium r = .30, ***large r = .50 (Mizumoto & Takeuchi, 2011)

Figure 10 *Articulation Rate and Speech Rate for Study 1*



Note: F2F = face-to-face, WP*M* = Word per minute (speech rate)

6.3. Discussion

This study aims to answer how task repetition conditions in task-based language learning via texting affect post-speech test scores in syntactic complexity, grammar, and lexical usage (Sauro & Smith, 2010). In summary, post-test results showed improvement in lexical error rates and increased syntactic complexity only in the texting group. In addition, no statistically significant improvements were found between the pre-test and post-test for any of the measures in the face-to-face group; rather, the results showed an increase in lexical errors, for both content and function words, in the post-test for this group. The results suggest that repetition significantly improves error rates in texting. The results regarding error rates and syntactic complexity were further examined by examining the types of errors and rounds of practice between the pre-test and post-test.

6.3.1. Post-test Results

Lexical Error Rate and Error Types. The lexical error rate was measured to determine differences in content and function words used between the pre-test and post-test. Content words are an indicator of information load, and errors in content words can indicate the success of information delivery. Function words have no meaning themselves, but they allow users to link information and then convey it effectively. The difficulty with function words is that they are not easily recognized because they are often processed as units (Healy, 1994).

In the current study, the lexical error rate for both content and function words improved only for the texting group. This may be because texting helps detect lexical errors since learners can visually capture what they have produced on the screen. Previous literature

supports the claim that visual saliency assists in identifying and correcting mistakes (Lee & Revez, 2020). However, when speaking, it loses nuance with time, and students may fail to notice the errors compared to when they are texting.

In addition, certain types of errors occurred more frequently in one mode than in another. For example, content word errors, use of L1 words, and repetition occurred more frequently in the face-to-face mode than in the texting mode. In contrast, most of the function word error types (incorrect word choice, repetition, unnecessary words, omission of words) decreased in the texting mode. This can be explained by the inherent property of texting that the slow pace of texting helps to avoid the use of L1 words and give them enough time to retrieve the correct linguistic information. In contrast, speaking requires quick responses, using repetition or filler words to buy more time. In addition, visual support in texting enables the recognition of errors such as repetition. On the screen, it is obvious when words are repeated. The additional support of visual text can also help to detect small errors, such as function word errors, which are difficult to detect when speaking. The decrease in error types on function demonstrates how texting enabled the detection of those errors. Accordingly, it can be argued that repeated task conditions allowed learners to detect and modify errors in texting. For instance, the decrease in error types, such as using L1 words and repetition, supports this argument.

In contrast, no effect on the lexical error rate was found in the face-to-face group for the three repeated task interventions. Although no statistical difference was found, fewer occurrences of the wrong word were reported for the content error types. This could be due to the practice effect of task repetition.

Lexical Diversity and Grammatical Accuracy. No statistical difference in lexical diversity and grammatical accuracy was found between the pre-test and post-test for either mode. Under the repeated condition of texting use, this result is not consistent with the previous findings of Sauro and Smith (2010), who claimed that texting use affects these areas. Although lexical diversity and grammatical accuracy were not observed, the lexical error rate and accuracy, respectively, improved in the texting group. Several arguments can be made for the differences in the results. One is that the differences are due to the background of the participants and the target language. In Sauro and Smith's (2010) study, the participants were native English speakers learning German. In this study, the participants were native Japanese speakers learning English. The differences between the participants' native language and the target language may affect the lexical repertoire available to increase lexical diversity. Second, novice learners (CEFR A1-A2) in the current study may not be able to recognize and correct their past tenses. The limited knowledge of past tenses may hinder the improvement of grammatical accuracy. Another explanation regarding the results on lexical diversity and grammatical accuracy is that they may be due to the trade-off effect (Foster & Skehan, 1996). Instead of improving lexical diversity and grammatical accuracy,

the current study improved lexical accuracy. This could be because lexical errors, such as repetitions on the screen, can be detected and modified independently of language proficiency compared to selecting another appropriate vocabulary or correcting grammatical errors.

Syntactic Complexity. The total score for syntactic complexity increased for both the texting and face-to-face groups. However, the results for the subcategory showed different results. The texting group saw a decrease in verb phrases but an increase in sentence structure scores. In the face-to-face group, the sentence structure score decreased. This could be due to the fact that learners often communicate in incomplete sentence parts when speaking and can fill in with words when their interlocutors do not understand them. Conversely, the texting group showed a reduction in phrase use and earned points for complete sentences. This suggests that texting likely allowed participants to notice their sentence structure better.

Fluency. The results showed that the speech rate increased significantly for the texting group. However, the articulation rate, an indicator of comprehensibility, did not improve. Thus, since the articulate stage of the language production process is missing, it might be difficult to improve multiple dimensions of speech when texting.

By contrast, the face-to-face group did not improve either fluency measure under the same repetitive task condition. This begs the question of why improvements were not found in any measures for this group.

6.3.2. Practice Rounds

Further attempts were made to observe the behavior under the two different modes. The practice rounds were further analyzed regarding the total number of words used and the time spent on each task.

Total Number of Words. The total number of words describes all the words produced in each round. Table 14 and Figure 11 list the descriptive statistics of the total number of words produced in the practice rounds. Kruskal- Wallis Test was conducted to compare the three repetition rounds, and significant differences were not found ($\chi^2 = 8.64$, $df = 5$, $p = .12$, $r = .22$) among the six practice rounds.

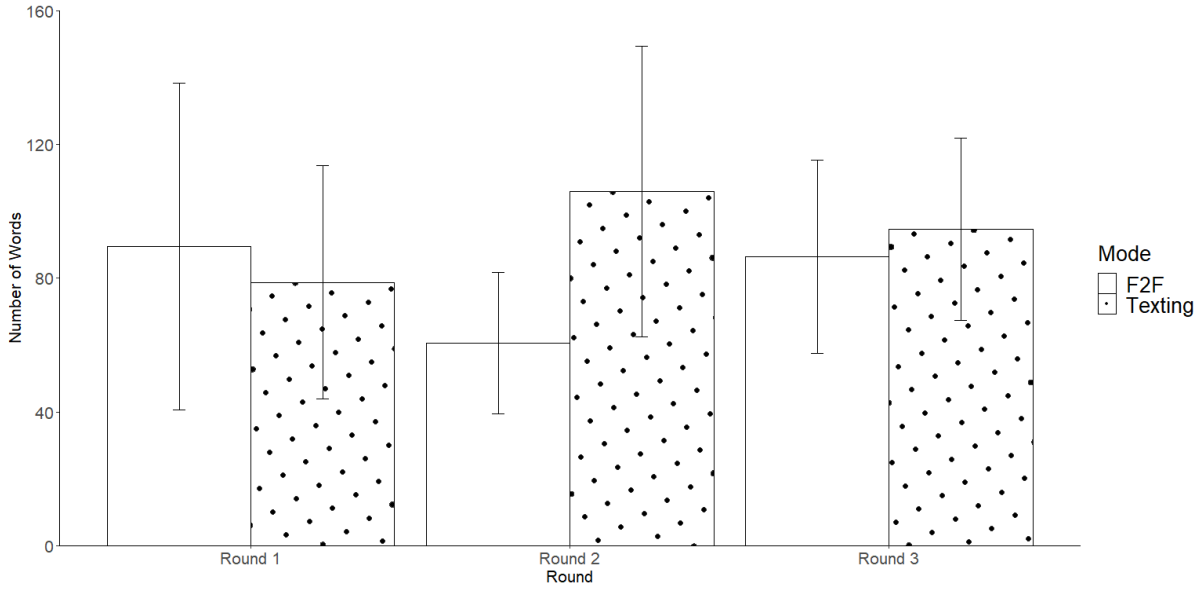
Table 14

Total Number of Words Used in Each Practice Round during Study 1

Mode	Round Number	*n*	*M*	*Mdn*	*SD*
F2F	1	7	89.57	75	48.83
F2F	2	7	60.57	54	21.23
F2F	3	7	86.43	87	28.93
Texting	1	9	78.78	62	34.97
Texting	2	9	105.89	94	43.39
Texting	3	9	94.56	99	27.24

Note. F2F = face-to-face, *n*= refers to the number of pairs

Figure 11 *Total Number of Words Used in Each Practice Round during Study 1*



Note. F2F=Face-to-face. Error bar indicates standard deviation.

Time spent on the task. Table 15 shows the time spent on each round. The results from the Kruskal- Wallis test showed a significant difference among the six groups (χ^2 = 30.52, *df* = 5, *p* = .00, *r* = .63). The post hoc test was applied using the Mann-Whitney test with adjusted Bonferroni to test pairwise comparison. It was found that face-to-face first round was significantly shorter to texting round 1 (*Z* = 3.07, *r* = .77, *p* = . 03) and face-to-face round 2 was significantly shorter to texting round 1 (*Z* = 3.07, *r* = .77, *p* = . 01), texting round 2 (*Z* = 3.34, *r* = .84, *p* = . 01), and texting round 3 (*Z* = 3.02, *r* = .76, *p* = . 04). In addition, face-to-face round 3 was significantly shorter to texting round 1 (*Z* = 3.34, *r* = .83, *p* = . 01), texting round 2 (*Z* = 3.34, *r* = .84, *p* = . 01), and texting round 3 (*Z* = 2.91, *r* = .73, *p* = .05).

The texting group spent two to three times longer on each round (Table 15). This corroborates previous studies' claims that texting takes nearly three times longer than face-to-

face communication (Carr, 2020). *Figure 12* illustrates that the time spent on the task differed between the two modes.

In addition, the time spent on the task may have had a major impact on language production in the post-test. For example, the decrease in lexical errors and increase in syntactic complexity in the texting group could be due to the fact that participants were able to spend more time on each round, which allowed them to process the information at their own pace. In texting, more attention can be paid to the production of content and function words and complex sentences, which affects the lexical error rate and syntactic complexity. This suggests that the time spent on the practice rounds may have affected participants' speech production during texting and calls for the study of speech performance under a fixed time condition.

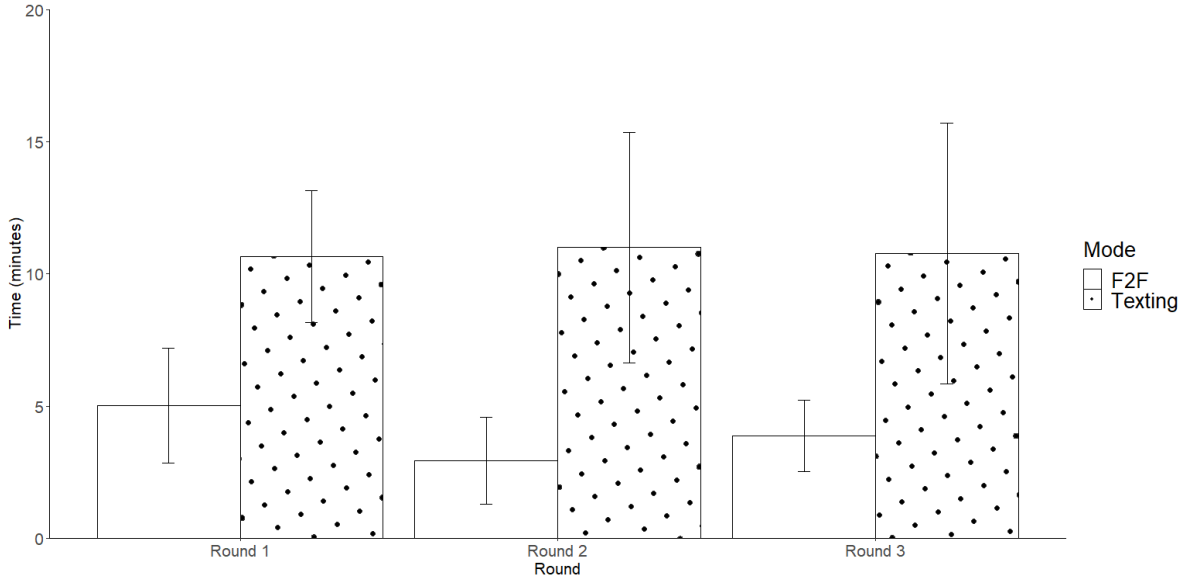
Table 15

Time Spent on Each Practice Round during Study 1

Mode	Round Number	*n*	*Mean*	*Mdn*	*SD*
F2F	1	7	5.01	4.11	2.17
F2F	2	7	2.94	2.41	1.64
F2F	3	7	3.86	4.36	1.35
Texting	1	9	10.67	11.00	2.50
Texting	2	9	11.00	10.00	4.36
Texting	3	9	10.78	10.00	4.94

Note. n= refers to the number of pairs

Figure 12 *Time Spent on Each Practice Round during Study 1*



Note: F2F=Face-to-face. Error bar indicates standard deviation.

6.4. Conclusions

This study examined how repetitive interviewing tasks conducted via texting and face-to-face affected speaking. The study found that repetitive texting tasks affected lexical error rate and syntactic complexity, while repetitive face-to-face tasks did not show statistical improvement. The articulation rate also did not improve in either mode, although the speech rate increased significantly with texting in the post-test. Some error types (e.g., use of L1, repetition, incorrect use of words, overuse of words, and omission of words) decrease in texting but not in face-to-face communication. As mentioned in Sections 2.2 and 2.4, visual support in texting facilitates the detection of common errors, which is beneficial. Furthermore, because visual speech production is an inherent feature of texting, reducing lexical errors shows that it can help learners speak accurately.

The results of the practice rounds showed no difference between the two modes in the total number of words produced. However, it was found that the time spent on the practice rounds differed significantly between the two modes. This may have influenced the results. The texting group required more time per task than the face-to-face group. The reason for the lower lexical error rates, higher syntactic complexity, and higher speech rate observed in the texting group may be due to the time spent on task.

Subsequently, time is another important factor to consider when implementing tasks. Therefore, the next study will investigate language production under time constraints on post-speaking tests of syntactic complexity, grammar, and lexical use.

6.5. Summary

Study 1 examined the effect of texting through task repetition.[2] Pre- and post-test were conducted, and linguistic scores were compared. For the face-to-face task, 14 language learners engaged in the interview activity, while a different group of 18 language learners engaged in the same activity in texting. The results suggest the texting group gained accuracy in content and function word use. In addition, syntactic complexity score showed an increase while demonstrating a higher speech rate. However, the face-to-face group did not show any improvements. The detailed analysis of error types on the post-tests and practice rounds revealed that visual salience on the screen may have enabled the learners to detect the errors and modify them. However, the time on task differed between the two modes, suggesting further examination under the fixed time condition.

[2] This study has been revised from the researcher's former study (Takase, 2019). Detailed analysis of the content and function words, grammar, and syntactic complexity have been added and modified from the researcher's previous study to investigate the research question for this study.

CHAPTER 7 STUDY 2: FIXED TIME CONDITION

Under the repeated task condition, Study 1 improved lexical error rate, syntactic complexity, and speech rate over speaking when practiced in texting. In contrast, the face-to-face group showed no differences. The difference between these two modes could be due to the different amounts of time spent on each practice session. In this chapter, we examine how practicing under the same fixed time condition in both modes affects speaking performance.

Thus, Chapter 7 will investigate the following question:

RQ (2) How does time pressure in task-based language learning via texting affect post-speaking tests on lexical use, grammar, and syntactic complexity (Sauro & Smith, 2010)?

7.1. Method

7.1.1. Participants

The participants in this study were 26 first-year students from a private college in Japan. All students attended an English communication course and learned English as a second language. The majority of the students (24 of 26) passed the pre-second grade of EIKEN (Practical English Proficiency) test, a widely used test in Japan to assess English proficiency. In addition, one student had already passed the second-grade EIKEN test in the previous two years. Therefore, the estimated CEFR level was A1-A2.

7.1.2. Tasks

The task was part of an English communication course in September and October 2020. Two groups were formed from two different classes, and all participating students were gathered in the classroom. Instructions were given in both English and Japanese. The purpose of the examination was explained, and participants were informed that their grades would not be affected. Before the task, both groups recorded a one-minute speech about what they had done during their summer vacation. The speech was recorded using their computers and microphones.

Then the teacher divided the group via face-to-face, and the learners participated in the interview task. All learners were randomly matched by the teacher and were not allowed to exchange partners. In the face-to-face group, learners sat close together, and the task was recorded using a microphone and a computer. In the texting group, participants sat randomly and separately from their partners. The topic of the task corresponded to Study 1: they were asked to interview their partner about what they did during the summer vacation. Learners were asked to act as both interviewers and respondents. However, in this study, learners had only 15 minutes to complete the task in groups face-to-face and via texting. Fifteen minutes

was chosen as the time limit because the practice rounds in Study 1 revealed that most pairs took about 10-15 minutes in each round (Table 16).

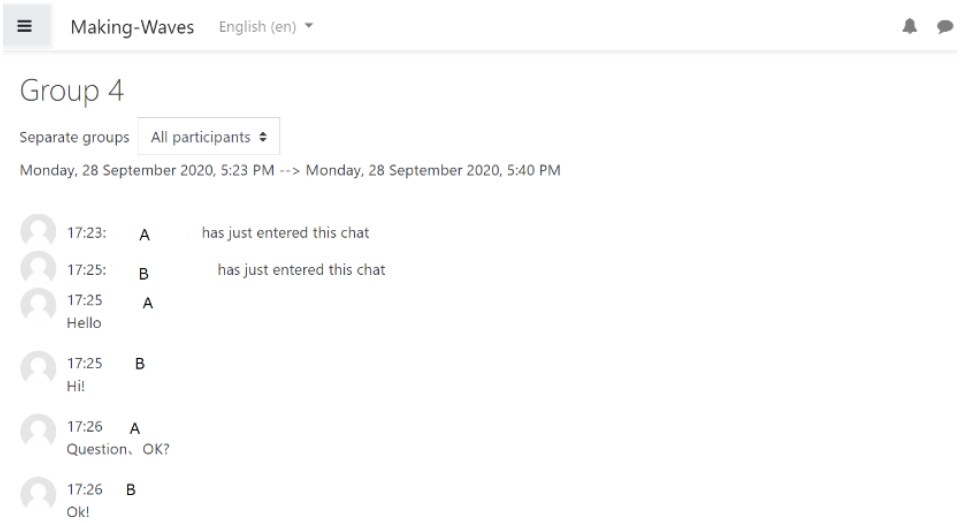
Table 16 *Range of Time Spent on Each Practice Round during Study 1*

	n
less than 9 minutes	10
10~ 15 minutes	13
more than 16 minutes	3

Note. n = number of pairs

The texting group used the chat room module in MOODLE (*Figure 13*) for the interview task. MOODLE was used regularly for sharing class material, so participants were familiar with it. The chat module was limited to texting only. Applications and technical functions such as spell check, autocorrect, and translation software were turned off. After completion of the interview task, a post-test was administered using the same procedures and content as the pre-test. After the post-test, learners reported to the teacher what they had discussed with their partners.

Figure 13 *Chat Module in MOODLE*



7.1.3. Data Analysis

Pre- and post-test audio files, recordings of face-to-face interactions, and text logs were collected. Data were analyzed based on the measurements from Study 1. All data were spell-checked, and computer messages were deleted. Vocabulary and grammar error rates were assessed manually. Vocabulary errors followed the method of Study 1. Vocabulary errors included incorrect word choice, repetition, use of L1, and missing words. Both content and function words were reviewed. For grammatical errors, only errors in the past tense were counted. First, 25% of the data were analyzed for vocabulary and grammar errors by the

researcher. Then, it was reviewed by a native English speaker with teaching experience. The inter-rater reliability for the lexical errors was r = .92, and for the grammatical errors was r =.91. After discussing the discrepancies, the researcher analyzed the remaining data. Part-of-speech (POS) was tagged for analysis using the Stanford Tagger (Toutanova et al., 2003) to calculate the grammatical error rate by the researcher. A native English speaker with teaching experience reviewed 20% of the data. The agreement was 97%. After discussing the discrepancies, the researcher checked the rest of the tagging and counted the error types manually.

Under the condition of repeating the task, the texting group improved some measurements. The face-to-face group, on the other hand, did not. This is due to the difference in the time available per task. Therefore, this Study 2 examines the effect of texting under the fixed time condition in both modes. To compare the differences, the same measures as in Study 1 are used, namely lexical accuracy, lexical diversity, grammatical accuracy, syntactic complexity, and fluency. In other words, lexical accuracy is described by lexical error rate, lexical diversity is measured by MTLD, grammatical accuracy by past tense verbs, syntactic complexity by IPSyn score, and fluency by articulation rate and speech rate.

The IPSyn score was calculated following the same procedure as Study 1. The baseline utterance was set to 43 per group, the minimum number of utterances conducted in this research. Table 17 shows the ratio of utterances chosen for stratified sampling.

Table 17

Utterance Analyzed per Group for the IPSyn Score (Study 2)

		n	*M*	*SD*	Total Number of Utterances per Group	Ratio of utterances analyzed
F2F	Pre	14	3.21	1.70	45	95.56
F2F	Post	14	4.93	2.55	69	62.32
Texting	Pre	12	3.58	1.66	43	100.00
Texting	Post	12	6.67	4.94	80	53.75

Note. F2F = face-to-face

Fluency measures were the same two measures from Study 1, articulation rate and speech rate. Articulation rate is the mean number of words between pauses which has been claimed to reflect comprehensibility (Suzuki & Kormos, 2019). Speech rate is measured by the total number of words in a minute. Repetition, repairs, fillers, and L1 use are deleted for this analysis.

Statistical analysis followed the same procedure as Study 1. Paired Wilcoxon signed-rank test was conducted to compare pre-and post-test scores in both the face-to-face and

texting groups. Statistic software R and statistical power G*Power 3 software (Faul, Erdfelder, Lang, & Buchner, 2007) were used to test the power of significance.

7.2. Results and Discussion

Vocabulary, grammar, and syntactic complexity were analyzed to answer the research question.

7.2.1. Vocabulary

The results for vocabulary measures are summarized in Table 18. The paired Wilcoxon signed rank test showed there was no difference in lexical diversity for both modes between pre- and post-tests (F2F p = .36, Z = .92, r = .17; Texting p = .27, Z = 1.11, r = .23). The statistical power was also low (F2F power = .42, Texting power = .32) indicating no significant difference between the two modes. This followed similar results as Study 1.

The lexical error rate was analyzed for both content and function words. Only the error rate about content words decreased significantly for the texting group (p = .05, Z = 2.04, r = .42), with a statistical power of .88 (*Figure 14*). Although there was no statistical difference in the error rate of content words in the face-to-face mode, the error occurrences tended to increase. For the function word error rate, there was no difference in either mode (F2F p = .26, Z = 1.12, r = .21, power = .32; Texting p = .12, Z = 1.60, r = .33, power = .78) (Table 18). In this study, the fixed time condition may have affected the results. Study 1 reported improvement in both content and function word error rates for the texting group. This study only found a significant difference in the content words. The 15-minute time limit might be insufficient for the texting group to improve both types of words. Conversely, this time, 15 minutes might have been ample time to practice face-to-face. However, no statistical difference in lexical error rate was found between the pre-test and the post-test for the face-to-face group.

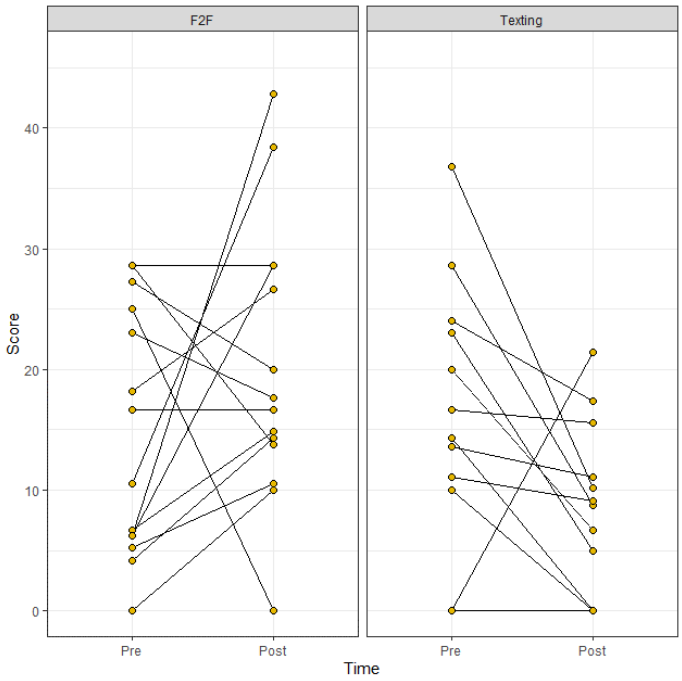
Table 18

Results of Lexical Diversity and Lexical Error Rate for Study 2

		Pre-Test				Post-Test			Pretest-Posttest comparison				Power (1-β)
	Mode	*n*	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*	
Lexical Diversity (MTLD)	F2F	14	26.25	26.07	9.73	14	29.97	28.61	7.86	0.36	0.92	0.17	0.42
	Texting	12	27.85	27.35	6.17	12	30.74	30.32	9.57	0.27	1.11	0.23	0.32
Content Words Error Rate	F2F	14	14.75	13.60	10.29	14	20.20	17.16	11.63	0.22	1.22	0.23	0.53
	Texting	12	16.52	15.48	10.87	12	8.76	8.89	6.99	0.05	2.04	0.42	** 0.88
Function Words Error Rate	F2F	14	13.88	13.03	11.55	14	10.41	9.84	8.23	0.26	1.12	0.21	0.32
	Texting	12	12.50	12.70	6.74	12	8.00	6.67	6.00	0.12	1.60	0.33	0.78

Note. *small r = .10, **medium r = .30, ***large r = .50 (Mizumoto & Takeuchi, 2011)

Figure 14 *Content Word Error Rate for Study 2*



Note. F2F = Face-to-face, Score represents the percentage of content word error.

Table 19 shows the number of occurrences per error type of content words. Again, the face-to-face group increased the number of occurrences of misused words. Additionally, repetition counts remained high for the face-to-face group while the texting group decreased.

Table 19

Number of Error Occurrences for Content Words for Study 2

	Pre-F2F	Post-F2F	Pre-Texting	Post-Texting
Incorrect choice	17	35	17	12
Use of L1	0	2	1	3
Missing (Fillers)	10	6	3	2
Repetition	17	11	10	6

Note. F2F = face-to-face

Table 20 shows the number of error occurrences of function words. Misuse of words and repetition remained high for the face-to-face group in both the pre-test and post-test. However, the texting group was able to decrease its score for incorrect use of words and repetitions. In addition, the number of unnecessary function words increased in the face-to-face group, while it remained the same in the texting group. An example of function words

found across the face-to-face data was the overuse of "so" and "and." These function words may have been used as filler or to keep the conversation going.

Table 20 *Number of Error Occurrences for Function Words for Study 2*

	Pre-F2F	Post-F2F	Pre-Texting	Post-Texting
Incorrect Choice	12	13	11	3
Repetition	8	7	11	4
Unnecessary	9	18	5	6
Missing	9	5	6	6

Note. F2F = Face-to-face.

7.2.2. Grammar

Table 21 shows no statistical difference in the grammatical error rate of past-tense verbs. Similarly, no differences were found in the previous task repetition study (Study 1). In Study 1, the texting group spent more time practicing, but their grammatical error rate score did not improve. In Study 2 (the present study), the face-to-face group received about twice as much time practicing the tasks as in Study 1, yet there was no improvement in grammatical error rate, perhaps because time is not a condition that improves this grammatical feature.

Table 21

Results of the Grammatical Error Rate for Study 2

		Pre-Test				Post-Test			Pre test-Post test comparison			Power (1-β)
	Mode	*n*	*M*	*MD*	*SD*	*n*	*M*	*MD*	*SD*	*p*	*Z*	*r*	
Grammatical (Past-tense verb) Error Rate	F2F	14	23.06	8.33	30.73	14	34.38	25.62	31.7	0.26	1.12	0.21	0.30
	Texting	12	13.88	4.55	20.4	12	10.29	5.51	14.99	0.53	0.7	0.14	0.11

Note. *small r = .10, **medium r = .30, ***large r = .50 (Mizumoto & Takeuchi, 2011)

7.2.3. Syntactic Complexity

For syntactic complexity, the IPSyn score showed improvement in both groups (Table 22). Scores for the sentence structure subcategory improved in the face-to-face group, while they remained the same in the texting group. For both groups, the scores for noun phrases increased, while the scores for verb phrases were lower in the post-tests (*Figure 15)*.

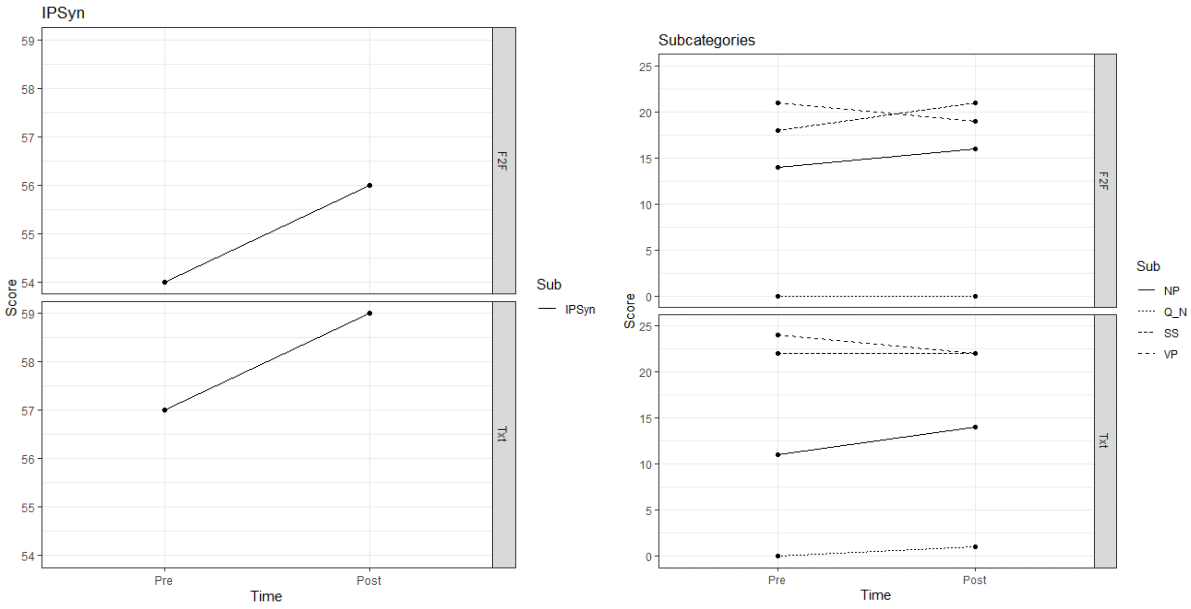
Table 22

Results of the IPSyn Score for Study 2

		IPSyn Score	Noun phrase	Verb Phrase	Question/ Negation	Sentence Structure
F2F	Pre	54	14	21	0	18
Texting		57	11	24	0	22
F2F	Post	56	16	19	0	21
Texting		59	14	22	1	22

Note. F2F = face-to-face.NP = noun phrase, Q_N = question and negation, SS = sentence structure, VP = verb phrase

Figure 15 *IPSyn Score for Study 2*



7.2.4. Fluency

The results of fluency measures for articulation and speech rate are shown in Table 23. There was no difference in articulation rate between the pre-and post-scores of both modes. The 15-minute task condition, again, did not affect the articulation rate for both groups. For speech rate, in this study, the face-to-face group showed an increase, while the texting group tended to decline *(Figure 16)*. The time constraints of this study may have affected the results as they had less time than in Study 1.

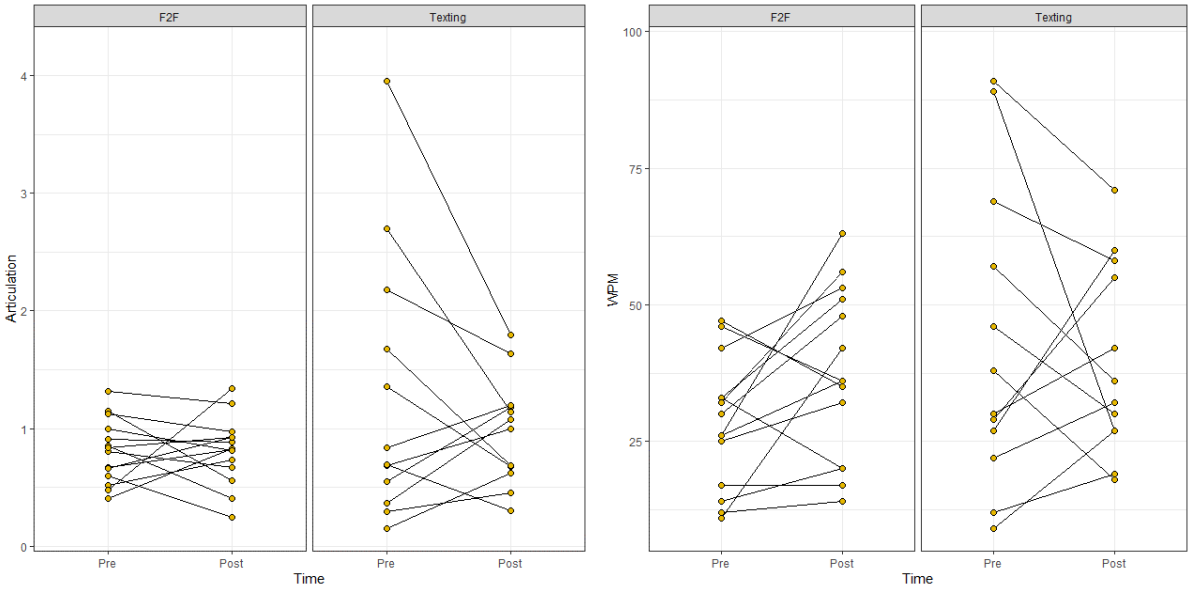
Table 23

Results of the Fluency Measures for Study 2

			Pre-Test				Post-Test			Pretest-Posttest comparison			Power (1-β)
	Mode	*n*	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*	
Articulation Rate	F2F	14	0.81	0.82	0.27	14	0.81	0.83	0.29	0.09	0.12	0.02	0.05
	Texting	12	1.29	0.77	1.15	12	0.98	1.04	0.45	0.38	0.88	0.18	0.05
Speech Rate	F2F	14	28.14	28	11.86	14	37.36	36	15.62	0.07	1.78	0.34	0.81
	Texting	12	43.25	34	27.8	12	39.58	34	17.46	0.70	0.39	0.08	0.08

Note. *small r = .10, **medium r = .30, ***large r = .50 (Mizumoto & Takeuchi, 2011)

Figure 16 *Articulation Rate and Speech Rate for Study 2*



Note. F2F = face-to-face, WPM = words per minute (speech rate)

7.2.5. Practice Session

The practice session was analyzed for further study of content and function words. Content word lexical errors, including incorrect word choice, use of L1, omissions, and repetitions, occurred more frequently in the face-to-face group than in the texting group (Table 24). In contrast, there were no cases where the use of L1 and repetition occurred for the texting group. As the language produced on the screen is visually salient, it can be assumed that the participants were aware of avoiding those errors. This result best explains why the error rate improved only for the texting group.

Table 24

Number of Errors for Content Words in the Practice Session (Study 2)

	F2F Practice Session	Texting Practice Session
Incorrect choice	18	8
Use of L1	49	0
Missing (Fillers)	17	2
Repetition	40	0

Note. F2F = face-to-face.

The number of function word errors, especially repetitions, was also more frequent in the face-to-face group than in the texting group (Table 25). Repetitions and unnecessary function words were not found in the texting logs of the practice session. However, the 15-

minute time constraint did not allow the texting group to significantly improve the error rate for the function words on the post-speaking test.

Table 25

Number of Errors for Function Words in the Practice Session (Study 2)

	F2F Practice Session	Texting Practice Session
Incorrect choice	7	10
Repetition	22	0
Unnecessary	1	0
Missing	3	3

Note. F2F = face-to-face.

The time constraint condition yielded results different from Study 1. The total number of words produced in the 15-minute practice has been presented in Table 26. The Mann-Whitney U test was applied to the two groups. A significant difference was observed in the total number of words between the face-to-face and texting groups ($U = 0$, $Z = 3.00$, $p = .00$, $r = .83$). There were twice as many words for the face-to-face group compared to the texting group (Table 26). In other words, the 15-minute task condition gave the face-to-face group to produce a mass amount of language. In contrast, the texting group may have focused on the accuracy or complexity of the language.

Table 26

Total Number of Words used in the Practice Session (Study 2)

	Mode	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*		Power (1-β)
Total number of words	F2F	7	202.14	197	68.33					
	Texting	6	90.17	88.5	16.31	0.00	3.00	0.83	***	0.95

Note. *small $r = .10$, **medium $r = .30$, ***large $r = .50$ (Mizumoto & Takeuchi, 2011)

7.2.6. Speaking Performance Under Fixed Time

In relation to the research question, the texting group improved the content word error rate under the specified time. In this section, the effect of time on different domains of linguistic features for the face-to-face and texting modes is discussed.

Speech rate. The results show that time impacted both modes greatly, especially as more words were produced in the face-to-face group than in the texting group during the practice session. For the texting group, 15 minutes were insufficient to improve the speech rate, unlike in Study 1, in which participants had more time, given that typing takes longer than speaking.

Notably, the articulation rate did not improve for both modes. Short practice, as three times repetition or 15-minute practice, may not be sufficient for improvement.

Lexical Error. Even on a short 15-minute task, the texting group improved lexical errors on content words. The results from the practice session confirm the improvement of the content error rate. In the texting practice session, the use of L1 and repetition were not found for the content words, and incorrect word use or omission errors were reduced. The unique mechanism of speech production in typing affects subsequent speaking in a short 15-minute task. Not only is typing slower than speaking, but because speech production is monitored visually, participants may spend a lot of time deleting and correcting messages before sending them. This can be explained by the different types of errors. Thus, it is evident that the visual highlighting of messages allowed learners to identify the errors under limited time pressure and correct them before sending the message. Overall, the slow pace of the conversation and the visual highlighting of the texting affect monitoring processes that impact subsequent speaking performance. These factors reduce cognitive load by allocating attentional resources to other areas of the speech production process, such as lexical or grammatical retrieval. Sauro and Smith (2010) have pointed out that the repair process in texting allows learners to notice, modify, and test language during a conversation. The analysis of error rates and types in this study supports this claim.

In contrast, the 15-minute exercise for the face-to-face group showed no positive changes in error rates. The results show prominent differences in the effect between the modes. In face-to-face communication, subtle errors may be ignored or unnoticed. Instead, learners may try to overcome them by switching to a different method of communication, such as gestures. By contrast, the limited communication method using only text may help them focus their attention on linguistic aspects, which may facilitate language learning.

Grammatical Error Rate and Lexical Diversity. This study also did not find improvements in grammatical error rates or lexical diversity in either mode. The grammar item was the past tense, in which the tense needed to share one's summer vacation. The grammatical error rate did not show any statical difference between the two modes. This may be because novice learners struggled to encode their messages after the conceptualization stage, as it is not automatized (Lambert et al., 2021; Levelt, 1989). Because lexical retrieval and syntactic formulation in the encoding process require conscious attention in lower-level L2 speakers, grammar is confounded with other areas of language development. Moreover, this study was conducted in a class that focused on communication. Therefore, maintaining the flow of conversation may have been the focus, resulting in improved accuracy in content words and gaining scores in syntactic complexity.

Lexical diversity score did not show improvement for this study as well. As discussed in Study 1, trade-offs may have occurred at the expense of improving syntactic complexity for both modes and lexical accuracy on content words for texting.

7.3. Conclusions

Results showed that even the short 15-minute time allotted to complete the interview task improved linguistic features in speaking for both modes. The texting group saw greater improvement in lexical error rate for content words. Further investigation revealed that the face-to-face group experienced an increase in lexical errors in word choice, L1 use, and repetition in the practice session. These types of errors may have influenced the error rate in the post-test.

In contrast, in the texting group, there was no repetition or L1 use during the practice session. This could be due to the fact that learners had the opportunity to visually see and correct their output before sending a message.

7.4. Summary

Study 2 examined how 15-minute practice constraints affected the linguistic aspects of the subsequent speaking test. A face-to-face group was formed with 14 language learners, and the texting group had 12 learners. Both groups took one minute pre- and post- speaking tests, before and after the 15-minute practice intervention. The pre- and post-tests were examined according to the same measures as Study 1: lexical usage, grammar, and syntax complexity. The results suggest that the inherent features of texting facilitate error detection for content words. The error types, no repetition or L1 use, showed how they supported this.

Studies 1 and 2 examined the effect of texting by determining the condition that affects speech production. Under both task conditions (repetition and time constraints), texting improved lexical accuracy for the subsequent speaking task. Regardless of the time allotted for the task, texting drew attention to errors, particularly screen errors. Examining task repetition and time eliminated these effects on speech production. However, it is not yet known whether texting is an appropriate method for learning new content. Completing tasks in Study 1 and 2 replied on the existing language knowledge, and it is not clear if texting allows for learning new items. Therefore, the effectiveness of texting in learning new content needs to be investigated. The next chapter will address the validity of using texting to learn new lexical phrases.

CHAPTER 8 STUDY 3: LEARNING LEXICAL PHRASES

Studies 1 and 2 found that completing interview tasks via texting improved speaking performance in terms of lexical accuracy and speech rate under different conditions. The drawback of the previous two studies is that the interview task only used existing language knowledge to complete the task. It is still unclear whether texting can be a platform for learning new terms for improving subsequent speaking performance. To investigate the effectiveness of using texting in language learning, it is inevitable to examine whether learning new content through texting is possible. Study 3 aims to investigate the relationship between texting and learning new content through the acquisition of new lexical phrases.

Previous research has shown that learning fixed expressions or lexical phrases is challenging for most language learners (Lewis, 2009). Since lexical phrases are commonly used among native speakers, learning the target language is challenging for EFL students who have little familiarity with the language. Since the meaning of lexical phrases is quite unclear and combining words can confuse learners, this study uses texting to focus attention on combining lexical phrases and to investigate whether the non-articulated form of learning enables learners to articulate and use the target words in the subsequent speaking test. The learning practice consists of target lexical phrase repetition practice and the interview tasks, and the results are compared with those of the face-to-face groups.

8.1. Research Questions

The following research questions:

RQ3: Is there a difference between the texting and F2F modes when learning new lexical phrases?

First, to answer this question, three sub-questions will be investigated to find the relationship between learning new lexical phrases and speaking.

(1) Is there a difference in cloze test results between a texting group and a F2F group?
(2) Will learners be able to use lexical phrases naturally in oral speech?
(3) How does task practice affect texting and F2F modes regarding lexical use, grammar, and syntactic complexity when learning new lexical phrases?

The hypothesis is based on former multiword expression (MWE) studies of each sub-research question that there may be no difference in the cloze test results and target lexical phrases in the post-test speaking test. Alternatively, texting may outweigh the face-to-face results. According to Cucchiarini et al. (2020), typing out words emphasizes single words, as short-function words are often skipped when reading (Carrol & Conklin, 2020). Accordingly, texting can lead to equivalent or better cloze test results than face-to-face conversation. Moreover, the speech production system follows similar phases in texting as in face-to-face conversation. Therefore, there may be no difference in the use of lexical phrases in the post-

test speech test. However, since the articulation phase is absent in texting, pronunciation of lexical phrases is a problem. Since lexical phrases are a combination of familiar words, a pilot study was conducted to test whether components of the expressions exceed the students' knowledge.

Referring to the results from Studies 1 and 2, task implementing conditions, such as repetition and time constraints, affect language production. The number of repetitions and time is predicted from the results of Studies 1 and 2.

8.2. Method

8.2.1. Participants

Thirty-eight undergraduate students from a college in Japan participated in this study from October to December 2018. All participants were Japanese non-English native speakers aged 19 to 22 years. One student did not complete the pre-test and post-test but participated in the practice tasks. Thus, the data used to analyze the pre-test and post-test were 37. In this study, LexTALE (Lexical Test for Advanced Learners of English) was used to predict the participants' general English proficiency (Lemhöfer & Broersma, 2012). LexTALE correlates with other English proficiency tests (e.g., the Oxford placement test) that reflect CEFR levels (Lemhöfer & Broersma, 2012). The LexTALE scores indicated that participants were lower than the B2 level in CEFR (M = 57.22, SD = 6.16, min = 46.25, max = 72.5, N = 37). All the data for this study was collected during an English class with students' consent, and they were informed that the results would not affect their grades. Task instructions were conducted in both English and Japanese.

This study followed a quasi-experimental design and used pre-tests and post-tests. Two classes were asked to participate in the study. Both groups completed an lexical phrase vocabulary quiz before and after the three-week interventions. The first class acted as the control group, with 16 students completing only the face-to-face tasks. The second class acted as the experimental group, with 21 students solving all tasks by texting only. For the experimental group, typing speed was measured using a website (https://typing-speed-test.aoeu.eu/) before the pre-test, and an exit survey was conducted after the post-test. The average typing speed for the texting group was 125.67 characters per minute (SD = 50.86, min = 76, max = 304, n = 21).

8.2.2. Procedure and Tasks

Two types of tasks (repetition and interview tasks) were used for learning lexical phrases in this study. The repetition task was adopted to ensure the target items were practiced and to present their translation into the students' native language. For the face-to-face group, the repetition task was practiced orally by repeating the teacher three times. In the texting group, the target lexical phrases were practiced silently by typing them, and Google

Forms was used to introduce new lexical phrases (Figure 17). The students were asked to practice typing the target lexical phrases three times on the designated activity page. All texting group activities were presented in the language management system MOODLE. Students used the same school computers and keyboards.

After the repetition exercises, both the control and experimental groups conducted a face-to-face interview and a texting interview, respectively. The results were compared with those of Studies 1 and 2. For the pre-test and post-test, students recorded their speeches individually using the microphone. The teacher divided all learners into two groups, and they interviewed each other using the questions on the worksheet that contained the target lexical phrases. The face-to-face group sat next to each other and recorded the exchange with a microphone. The texting group sat randomly. The duration of the interview task was 15 minutes. The interview questions inquired about their favorite city or what they would like to do there. Some of the interview questions were:

- If you could choose to go anywhere in the world, what do you look for in a place to live?
- Is it easy to get around by public transportation there?

Figure 17 *Web Page for Learning Lexical Phrases for the Texting Group*

All the participants took the pre-test in the first week. Then, through the second to fourth weeks, the learners practiced new sets of lexical phrases each week through repetition and interview tasks. Finally, learners were asked to take the post-test in the fifth week.

8.2.3. Pre- and Post-Test

Two tests were used as pre-tests and post-tests: the cloze test and the one-minute speaking test. The cloze had a time limit of 15 minutes and consisted of 15 sentences containing the targeted lexical phrases. Learners were asked to fill in the blanks to match the sentences to the correct translation. Learners were not allowed to use any external resources. Application software such as a translator or autocorrect was turned off. The test was administered online in a classroom with the teacher watching the screen. The topic of the one-minute speech was, "If you could live in your favorite city, where would you live? Explain your reasons." This corresponded to B1 level in the CEFR descriptor (Council of Europe, 2020). Before the recording, all learners had one minute to prepare their speech.

8.2.4. Target Lexical Phrases

Before the current study, two pilot studies were conducted to develop and ensure the reliability of the cloze test. For the first pilot test, a cloze test draft was developed by referencing the lexical phrases from a class textbook to match participants' proficiency (Bolhke, 2016). The first draft consisted of 34 items (Appendix 3). Since an lexical phrase are multiple words retrieved as a whole from memory (Wood, 2010), it often relies on the conventional norm (Erman & Warren, 2000). Thus, to be proven, it needs to be assessed by several language users, not only the test developer. To reflect the norm of a native speaker, the naturalness and authenticity of the lexical phrases used, the list was consulted by a group of native or highly experienced students of applied linguistics and native-speaking colleagues with master's degrees in the field. The list was then narrowed down to 25 items (Appendix 4). From May to June 2018, 15 lower intermediate to beginner students completed the 25-question cloze. Students were informed that their scores would not affect their grades and that only student-consented data were used for analysis. Each item was calculated as one point, and the maximum score was 25. The average score of the 25 cloze tests was 13.8 out of 25 (n = 15, SD = 4.18). However, there was a decline in the difficulty score (percentage correct) between items 15 and 16 (Table 27). Therefore, the researcher decided to use only the 15 items for study 3 (Appendix 5) based on the difficulty score.

Table 27

Pilot Study 2 Results of 25 Items

	FSs	Acceptable answers	% Correct	di
1	(go) camping		14.29%	0.50
2	set (in)		14.29%	0.50
3	(group) of		14.29%	0.25
4	attracted (to)		20.00%	0.25
5	(with) ease		26.67%	0.75
6	check (out)		28.57%	0.50
7	(in) awe of		30.00%	0.50
8	something (else)		35.71%	0.50
9	engage (in)	with	40.00%	0.25
10	find (out)		42.86%	0.50
11	turn (out)		42.86%	0.25
12	get (around)		50.00%	0.00
13	hang (out)		53.33%	1.00
14	search (for)		57.14%	0.75
15	interact (with)		60.00%	1.00
16	(in) the shape		73.33%	0.50
17	(on) the left	to	80.00%	0.50
18	come (from)		80.00%	0.50
19	different (from)		86.67%	0.50
20	(pay) attention		86.67%	0.00
21	turn (on)	on, in	93.33%	0.00
22	(at) the bottom	towards, near	93.33%	0.00
23	look (for)		93.33%	0.00
24	grew (up)		93.33%	0.25
25	(on) the wall	against	100.00	0.00

The second pilot study was conducted to measure the reliability of the completed cloze. A separate group of 21 lower intermediate to beginning students was presented with the cloze text twice, three weeks apart. Again, a high correlation (r = .81) indicates a reliable measure for assessment (Table 28).

Table 28

Correlation of the Finalized Cloze Test during Three-week Interval

	n	*M*	*SD*	*r*
Week 1	21	5.48	1.47	
Week 5	21	5.19	1.86	0.81

Note. F2F = face-to-face.

8.2.5. Linguistic Analysis

For the cloze tests, the difficulty index (percentage correct) per item was calculated by the ratio of the number of students with correct answers to the total number of students per group.

The recorded one-minute speech and the data from the practice tasks were collected and transcribed. For the texting group, the chat transcripts were downloaded, and all messages were deleted from the software before analysis.

The measures used for the linguistic analysis of the pre-test and post-test follow Studies 1 and 2: lexical error rate (content and function word), lexical diversity, grammatical error rate, syntactic complexity, and fluency (articulation rate and speech rate).

The method of linguistic analysis is the same used in Studies 1 and 2. Two evaluators coded the lexical and grammatical error rates. The evaluators were the researcher and a native English speaker with teaching experience. Lexical errors included incorrect word choice, unidentifiable words, repetition, use of L1, and missing words. Missing words that were replaced with filler words were also counted. To label the error types, the researcher first labeled 20% of the data, after which it was checked by a native English speaker with teaching experience. The agreement was 91%. The differences were discussed and changes made. Finally, the researcher labeled the rest of the data and counted the error types manually.

Grammatical errors only included verb errors. For the lexical and grammatical error rate, the researcher first coded 20% of the transcribed data. Then, it was reviewed by a second coder. The correlation between the two evaluators was r =.93 for the lexical error rate and r = .84 for the grammatical error rate.

Syntactic analysis was performed using IPSyn scoring, and utterances were extracted using the stratified sampling method. The 43 utterances were analyzed per group, and the IPSyn score was calculated (Table 29).

Table 29

Ratio of Utterances Analyzed per Group for the IPSyn Score (Study 3)

		n	*M*	*SD*	Total Number of Utterances per Group	Ratio of utterances analyzed
F2F	Pre	16	3.75	1.20	60	71.67
F2F	Post	16	4.31	1.40	69	62.32
Texting	Pre	21	3.76	1.00	79	54.43
Texting	Post	21	5.10	1.77	107	40.19

Note. F2F = face-to-face.

The measurement of fluency follows the other two studies and records articulation and speech rate. Repeated words, repairs, filler words, and L1 words were omitted prior to analysis. The Wilcoxon Signed Rank Paired Test was performed for the pre- and post-tests. Pearson correlation was calculated to determine the relationship between typing rate and the number of words spoken in the post-test for the texting group.

8.3. Results and Discussion

Cloze Test. Cloze tests on 15 lexical phrases were conducted to answer the first sub-research question. Wilcoxon signed rank paired test results indicated significant differences with a large effect size between pre-and post-test for both modes (F2F: $Z = 3.50$, $p = .00$, $r = .62$. power = .99; Texting: $Z = 4.02$, $p = .00$, $r = .62$, power = .99). Power analysis showed strong power of .99 for both modes. The effect size for both modes had the same result ($r = .62$), indicating a strong effect of the practice. This indicates mode difference was not found.

Table 30

Results of Cloze Test on Lexical Phrases (15 items)

		Pre-Test				Post-Test			Pretest-Posttest comparison				
Mode	n	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*		Power (1-β)
F2F	16	5.31	4.00	3.26	16	10.94	11.00	2.57	0.00	3.50	0.62	***	0.99
Texting	21	5.24	5.00	1.26	21	11.00	12.00	3.49	0.00	4.02	0.62	***	0.99

Note. *small $r = .10$, **medium $r = .30$, ***large $r = .50$ (Mizumoto & Takeuchi, 2011)

Concerning the second research question: "Will learners be able to use the target lexical phrases naturally in oral speech?" A small number of usages occurred in both groups. The target lexical phrases appeared twice in one speech of the face-to-face group and four in four speeches of the texting group in the post-test. Although there were only a small number of target lexical phrases in natural speech, both groups showed use. Problems in pronouncing

the lexical phrases were anticipated for the texting group, but there was no evidence for this. The reason why pronunciation was not a problem in texting could be that the lexical phrases were likely to be a combination of single words whose pronunciation the learners already knew. If the target lexical phrases had consisted of unknown words, the result might have been different.

Lexical Measures. To answer the third research question: "How does task practice affect texting and F2F on lexical use, grammar, and syntactic complexity when learning the target lexical phrases?" The transcribed speech was examined for lexical use, grammar, syntactic complexity, and fluency. The Wilcoxon signed rank paired test was applied for analysis. Table 31 shows the results of the pre-and post-test of lexical diversity and lexical error rate. The results show that the texting group improved significantly on the error rate of content words ($Z = 2.26$, $p = .02$, $r = .35$, power = .99) and function words ($Z = 2.65$, $p = .01$, $r = .41$, power = .99). Face-to-face group also showed improvement; however, it did not show statistical difference. The repeated task condition and the opportunity to spend three weeks on the task allowed sufficient practice time for texting, which improved the accuracy of both content and function words. The lexical diversity result is consistent with those of Studies 1 and 2, showing no difference between the two modes. This means that the improvement of lexical diversity in novice learners cannot be achieved quickly.

Table 31

Results of Lexical Diversity and Error Rate for Study 3

		Pre-Test			Post -Test				Pretest-Posttest comparison				
	Mode	n	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*		Power (1-β)
Lexical Diversity (MTLD)	F2F	16	27.64	26.07	10.76	16	29.32	28.08	7.53	0.67	0.43	0.08		0.18
	Texting	21	26.73	26.94	9.23	21	28.55	29.46	9.45	0.20	1.27	0.20		0.25
Content Words Error Rate	F2F	16	14.4	8.12	14.72	16	13.72	11.15	11.16	0.94	0.08	0.01		0.07
	Texting	21	16.04	12.5	9.24	21	9.25	8.33	7.04	0.02	2.26	0.35	**	0.99
Function Words Error Rate	F2F	16	13.63	13.56	10.55	16	12.53	9.84	7.72	0.67	0.43	0.08		0.12
	Texting	21	14.22	13.04	9.68	21	7.48	5.71	6.21	0.01	2.65	0.41	**	0.99

Note. *small r = .10, **medium r = .30, ***large r = .50 (Mizumoto & Takeuchi, 2011)

Table 32 shows the number of raw occurrences of error types of content words. Incorrect word choice errors decreased for both modes. However, the texting group showed a greater decrease in raw errors: Face-to-face group decreased by seven raw errors, while the texting group had a decrease of 17. This could be a reason for the improvement in the content word error rate. For the face-to-face group, there was a notable difference in the number of word omissions or filler words, meaning that the speakers could not recall the words from their repertoire. Face-to-face practice alone may increase the error rate, which could be difficult for novice learners to overcome.

Table 32

Number of Error Occurrences of Content Words for Study 3

	Pre-F2F	Post-F2F	Pre-Texting	Post-Texting
Incorrect choice	27	20	38	21
Use of L1	0	2	1	0
Missing (Fillers)	8	22	15	17
Repetition	1	6	1	3

Table 33 shows the number of raw occurrences of error types of function words. For the face-to-face groups, the use of unnecessary words increased. The frequency of errors on unnecessary words almost doubled in the post-test for the face-to-face group. By contrast, the texting group decreased the errors for unnecessary words, repetitions, and missing words. The decrease here may have influenced the error rate for texting to improve in Study 3.

The results of content and function word error rates are partially consistent with those of Studies 1 and 2 in terms of no L1 use of content words and repetition of function words in the post-test for the texting group. However, comparing the studies might be difficult because few errors were found in the pre-test for these items in Study 3.

Under the current task condition of repetition and fixed time, more lexical errors were reduced in texting, such as incorrect use of content words or omission of function words in the post-test. In the face-to-face group, errors involving incorrect word choice, omission of function words, and addition of unnecessary words increased. The less frequent errors in texting suggest that it can provide an advantage for visual attention, allowing students to easily identify and correct errors on screen.

Table 33

Number of Error Occurrences of Function Words for Study 3

	Pre-F2F	Post-F2F	Pre-Texting	Post-Texting
Incorrect choice	6	10	10	19
Repetition	0	1	2	0
Unnecessary	12	25	26	24
Missing	26	23	22	11

Note. F2F = face-to-face.

Grammar. The grammatical error rate of verbs was analyzed. There were no statistical differences between pre-and post-tests for both modes (Table 34). Consistent with Studies 1 and 2, the results showed no improvement in grammatical errors for verbs. This result deserves attention because the mean score shows that all scores were zero. This implies that verb grammatical mistakes were rarely found in both pre-and post-test data. As the topic of the post-test was to describe their favorite cities and the reasons, the necessary verbs were mostly present tense. The difficulty of the topic may have affected this result. It can be assumed that the use of such a topic in an interview task does not improve grammatical accuracy in the tenses.

Table 34

Results of the Grammatical Error Rate for Study 3

		Pre-Test			Post-Test				Pretest-Posttest comparison			
Mode	n	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*	Power (1-β)
F2F	16	4.12	0	5.18	16	4.95	0	6.91	0.72	0.36	0.06	0.13
Texting	21	3.71	0	5.68	21	3.48	0	7.19	0.72	0.40	0.06	0.07

Note. *small r = .10, **medium r = .30, ***large r = .50 (Mizumoto & Takeuchi, 2011)

Syntactic Complexity. For syntactic complexity, the IPSyn score showed different results between the face-to-face and texting modes. The texting group's score improved from 65 to 69, but the face-to-face group declined (Table 35). In the texting group, the score for the sentence structure subcategory increased by 3 points, from 24 to 27, whereas in the face-to-face group, it increased by only one point, from 22 to 23. In the subcategory question and negation, the texting group's score remained stable, while in the face-to-face group, it dropped by 5 points from 6 to 1. For noun and verb phrases, both groups recorded an increase in noun phrases, while the face-to-face group recorded a decrease in verb phrases (*Figure 18*). Overall, the IPSyn score showed that practice via texting supports the production of complete sentences, while face-to-face practice was not as successful.

The ISPyn score is consistent with the results of studies 1 and 2; the texting group improved or maintained scores between the pre-test and post-test. However, for the face-to-face group, the results were negative. As the results show, the use of texting enables the achievement of scores on sentence structure with correct parts of speech and word order.

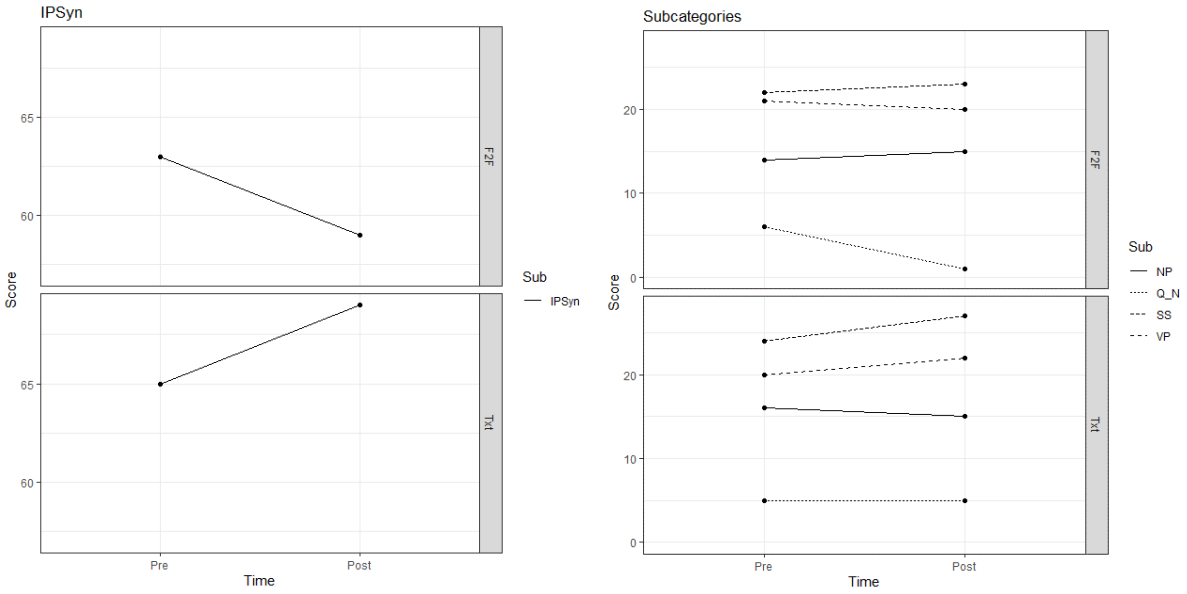
Table 35

Results of the IPSyn Score for Study 3

		IPSyn Score	Noun phrase	Verb Phrase	Question /Negation	Sentence Structure
F2F	Pre	63	14	21	6	22
Texting		65	16	20	5	24
F2F	Post	59	15	20	1	23
Texting		69	15	22	5	27

Note. F2F = face-to-face.

Figure 18 *IPSyn Score for Study 3*



Note. F2F = Face-to-face.

Fluency. The Wilcoxon signed rank paired test result for articulation rate, and speech rate is described in Table 36. Both modes improved significantly on speech rate in the post-test (F2F p = .00, Z = 2.82, r = .50, power = .99; Texting p = .00, Z = 3.89, r = .60, power = .99). However, only the face-to-face group showed significant improvement for the articulation rate (p = .02, Z = 2.42, r = .43, power = .96). There was no impact on articulation rate for the texting groups.

The improvement in articulation rate is observed only in study 3; it was not observed in Studies 1 and 2. This result may be explained by the mode effect. As mentioned in Section

1.3, the construct of speech consists of small groups of words or units. A small group of words, such as sentences and phrases, can form a meaning that allows messages to be conveyed. When speaking face-to-face, small units are formed consciously or unconsciously, which can lead to an improved rate of articulation over the course of practice. However, when writing texts, these small units are not visible on the screen. Whether the clauses and phrases are typed in a small group of words or not, the message's meaning does not affect how the message is sent if it is typed before hitting the "send" button. Thus, small word groups are unlikely to be the focus of attention, and text practice had no effect on articulation rate and showed no effect in all three studies. The difference between modes affected different areas of fluency.

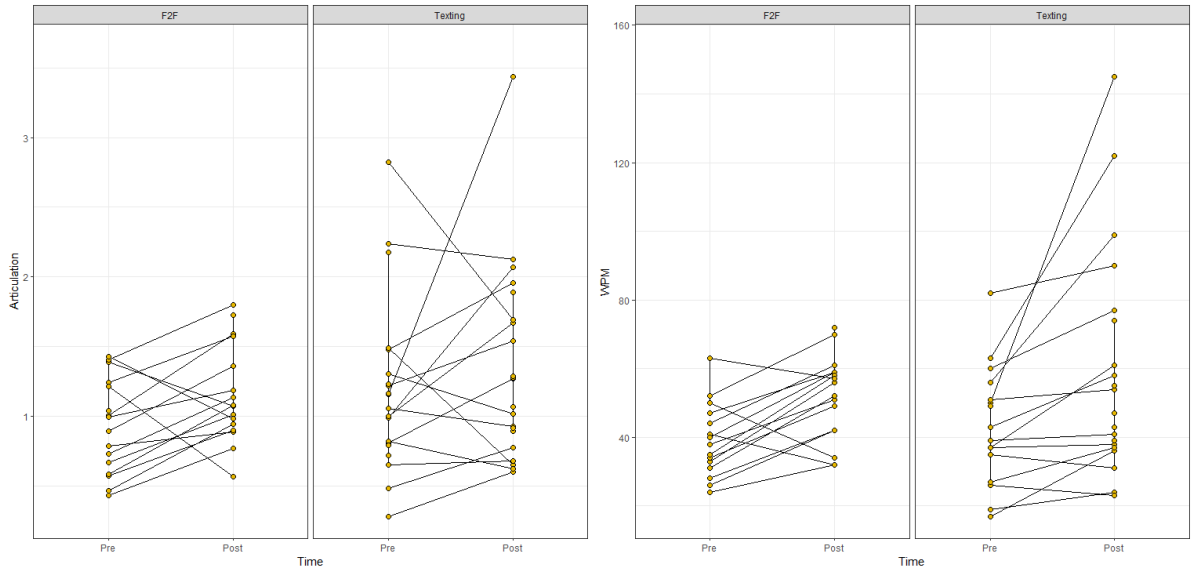
Table 36

Pre- and Post-test Results on Fluency Measures for Study 3

		Pre Test				Post Test			Pretest-Posttest comparison				
	Mode	*n*	*M*	*Mdn*	*SD*	*n*	*M*	*Mdn*	*SD*	*p*	*Z*	*r*		Power (1-β)
Articulation Rate	F2F	16	0.93	0.94	0.34	16	1.16	1.08	0.35	0.02	2.42	0.43	**	0.96
	Texting	21	1.17	1.06	0.62	21	1.35	1.27	0.69	0.18	1.34	0.21		0.40
Speech Rate	F2F	16	39.75	39	10.74	16	51.62	54	12.41	0.00	2.82	0.50	***	0.99
	Texting	21	40.24	37	16.12	21	59.48	54	31.95	0.00	3.89	0.60	***	0.99

Note. *small r = .10, **medium r = .30, ***large r = .50 (Mizumoto & Takeuchi, 2011)

Figure 19 *Articulation Rate and Speech Rate for Study 3*



Note. F2F = face-to-face, WP*M* = words per minute (speech rate)

Typing Speed. Typing speed (*M* = 125.67, *Mdn* =107, *SD* = 50.86, *n*= 21) and the number of spoken words per minute in the post-test of the texting group (*M* = 59.48, *Mdn* = 54, *SD* = 31.95, *n* = 21) were analyzed to find the relationship of texting to speaking. Pearson's correlation test showed no correlation between the two scores (r = .23, t(19) = 1.04, p = .31), indicating that the words produced in texting were not affected by the individual's typing speed. The results suggest that learners with high typing scores may not be able to type more messages in L2. Another factor, language proficiency, may have a direct effect on language production. However, the effects of typing ability may be different for speakers with higher language proficiency who are able to continue the conversation.

8.3.1. Learning Lexical Phrases and Texting

To answer the first sub-research question, is there a difference in cloze test results between the two modes? The results showed that learning lexical phrases via texting was equivalent to face-to-face learning. Both groups saw a large effect size on pre- and post-cloze test results. In addition, there were no significant differences in the effect size. In other words, the cloze test results showed no difference in vocabulary learning between the two modes.

To answer the second subquestion, will learners be able to use the lexical phrases learned in texting for speaking? The lexical phrases used in the post-test were limited. The data showed that there were few lexical phrases in the speaking tests. The interview tasks were designed to facilitate the acquisition of lexical phrases in a natural context, but this intervention lasted only three weeks, which may not have been sufficient. In addition,

speakers self-select their topics in the interview tasks, and those with unfamiliar words may have been avoided.

Samples 1 and 2 are dialogue excerpts of the face-to-face and texting groups that incorporate lexical phrases in their speech. Both samples demonstrated "hang out," which they learned by asking and answering the interview question, "Where do you often hang out with your friends?"

> **Sample 1:** Student 1 from the face-to-face group:
> Pre-test*: ...I think Tokyo has many things. For example, very favorite shops and the Tokyo tower. I want to go on sightseeing...*
> Post-test: *...I think Tokyo has many newest shops. I want to hang out with friends. My best friend lives in Tokyo now, so I want to hang out with her….*
> **Sample 2:** Student 1 from the texting group:
> Pre-test: *... there are a lot of stores and amusement, and I want to go shopping in Tokyo.*
> Post-test: *I want to shop to search for good items and hang out with my friends in some amusement parks.*

8.3.2. Texting and Speaking Improvement

To answer the third subquestion, how does practice affect texting and face-to-face mode? Improvement was found in both modes on the post-test. For the texting group, lexical accuracy on content and function words, syntactic complexity, and speech rate improved on the post-speech test.

The results on texting are consistent with those of Studies 1 and 2 in terms of the relationship between lexical error rates and error types. The task condition in Study 3 allowed sufficient time and repetition practice for students to notice both content and functional word errors in texting. The face-to-group study showed no significant difference between pre-test and post-test scores in this regard. Errors occurred, for example, in the omission of content words and the overuse of function words. Texting decreased the errors, especially in the incorrect choice of content words and the omission of words. This suggests that practicing by texting under the same conditions reduces lexical errors and improves speech accuracy. Linguistic information on the screen facilitates the detection and subsequent correction of errors in texting.

The improvement in the syntactic structure of the language could be due to the visual aid of texting. IPSyn scores increased with texting, while scores decreased with face-to-face. The result in this subcategory indicates an increase in the scores for sentence structure and verb phrase use, suggesting that texting supports the correct organization of words and the use of verbs. The ability to visually see speech production on the screen allows these errors to be noticed, and texting can support language practice.

8.3.3. Face-to-Face and Fluency

In the face-to-face group, scores for articulation rate and speech rate improved significantly. The results for fluency showed that both modes improved the number of words produced per minute. This shows that the speech rate can be increased in both modes. However, the articulation rate improved only in face-to-face speech. Since the articulation rate affects the comprehensibility of speech (Suzuki & Kormos, 2019), this result indicates the limitations of practicing in texting mode. When practicing face-to-face, accuracy was sacrificed over fluency compared to texting. Given the absence of the articulation phase in texting, the result indicates partial acceptance of the hypothesis that texting has little effect on fluency.

This research has clearly shown that conversations in texting and face-to-face groups affect different areas of speech production. The use of texting allows learners to notice errors and correct them, improving the accuracy of speech. In contrast, during face-to-face practice, both measures of fluency, articulation and speech rate, improved significantly. This research has shown that different modes can improve different areas of speech performance.

8.3.4. Learner Feedback

Texting for language learning is not common; therefore, a final survey was conducted for the texting group in Japanese. It examined the difficulty of the interview task, the cloze texts, and the online typing activity, as well as general feedback on online language practice (Table 37). A similar distribution of students found it difficult to participate in the interview tasks and cloze tests. However, most students felt that online activities helped them to acquire better English speakers. This might suggest that online exercises can be an alternative to the face-to-face practice environment. The answer to the follow-up question, why do you prefer one mode over the other, describes how learners perceived the task-based activities via texting. Some feedback from students who preferred texting included:

- I feel more confident speaking after a texting activity.
- I can practice writing and reading at the same time.
- I have more time to write appropriately. Unfortunately, I feel more pressured when face-to-face and cannot continue the conversation.
- I like doing texting activities because I can practice typing simultaneously.
- I have more time to think.
- I am not too fond of communicating orally.

Learners who could not decide said:

- Texting is similar to speaking.
- I can understand the words and sentences in texting. In face-to-face, you can practice pronunciation and accent.
- There are both strong and weak points.
- Those who preferred face-to-face stated:
- I'm not used to typing in English.
- The pressure of face-to-face helps me remember words.
- I don't have good writing skills.
- I like the pressure of having to reply quickly.
- I think there are differences. Maybe I was uncomfortable waiting for a reply because of the typing speed.
- I'm not good at spelling.
- I think speaking is important.

Table 37

Exit Survey

Questions	Easy	Fairly Easily	Neutral	Fairly Difficult	Difficult	Total
1. Was the online interview task difficult?	5	7	13	5	7	37
2. Was the cloze-test difficult?	5	9	13	3	7	37
	Strongly Agree	Agree	Undecided	Disagree	Strongly Agree	Total
3. Do you think the online activities help you become a better English speaker?	4	24	6	3	0	37
	Online Texting	Cannot say	Face-to-face			Total
4. Which mode do you feel you can produce and practice English?	11	11	15			37

8.4. Conclusions

In this study, no differences were found in the learning of lexical phrases according to mode. The cloze test results were the same between the two modes. In addition, after three weeks of practice, lexical accuracy and syntactic structure of speech improved on the post-speech test, which was not the case for face-to-face practice. However, both groups improved speech rates, and the face-to-face group showed a significant improvement in articulation rate.

The results suggest that texting is a valid communication method for acquiring new lexical phrases and improving speaking. Moreover, under certain conditions, lexical accuracy and syntactic complexity can improve significantly compared to face-to-face practice. This is especially true for learners with basic proficiency (CEFR A1-B1) who need more time to process and retrieve linguistic information for output. The use of texting allows for visual support and slow exchange. Additionally, face-to-face practice has a significant impact on improving articulation rate. Overall, Study 3 found that texting can replace face-to-face interactions for learning lexical phrases and that practicing improves different areas of language constructs depending on the mode.

8.5. Summary

Study 3 examined the effect of texting on lexical phrase learning under task repetition and time pressure for three weeks.[3] The cloze test was developed to ensure an examination of the effect of the intervention. The study was conducted by forming two groups, face-to-face and texting, with 16 and 21 language learners, respectively. The intervention included repetition practice of the target lexical phrases and an interview task that lasted three weeks. Cloze test, and pre- and post-speaking tests were conducted to measure the effect of the intervention. The result of the cloze test showed no difference in lexical phrase learning between face-to-face and texting modes. In the post- speaking test, there was small evidence of target lexical phrases used naturally in speech. In addition, the comparison of the scores between pre- and post-test indicates that the texting group improved lexical error rates for content and function words, syntactic complexity, and speech rate. By contrast, the face-to-face group improved the area of fluency, speech rate, and articulation rate. The results support the claim that texting improved accuracy and use of complex speech compared to face-to-face, while the face-to-face group improved fluency scores.

[3] This study has been revised to investigate further the word usage of content and function words, grammar, syntactic complexity, and fluency from the researcher's former study (Takase, 2020).

CHAPTER 9 DISCUSSION

Previous studies on the use of texting for speech have yielded mixed results. This is because speech production is influenced by the nature of the task and the conditions of task implementation (Lin, 2015; Skehan, 2014). Based on research on repair in texting (Sauro & Smith, 2010), the current study examined the effects of texting under different task conditions in the subsequent speaking test. In texting, repair behavior affects linguistic complexity and lexical diversity (Sauro & Smith, 2010), but there are few studies examining these effects to subsequent speech. Using the results of the three studies conducted in this research, this section discusses how texting affects subsequent speaking under different task conditions. Accordingly, the following research questions are raised:

Research Question 1: How does task repetition in task-based language learning through texting affect post-speaking tests on syntactic complexity, grammar, and lexical use?

Research Question 2: How does time pressure in task-based language learning through texting affect post-speaking tests on syntactic complexity, grammar, and lexical use?

Research Question 3: Is there a difference between learning lexical phrases through texting and face-to-face interactions?

9.1. Use of Texting Under Repeated and Timed Condition

The results of the three studies conducted in this research illustrate that the use of texting in task processing improves the areas of lexical accuracy, syntactic complexity, and speech rate. This generally supports Sauro and Smith's (2010) findings about the function of repairing when texting. Repeated practice of tasks in texting improved the subsequent speaking test as learners showed fewer salient errors such as L1 use and repetition (Study 1). The error types describe two important findings about texting use. The use of texting draws attention to lexical errors and allows for modification, and even though texting lacks the articulation phase, it allows for improvement in speech performance. Study 1 showed a significant decrease in lexical error rates for content and function words. The reduced errors in L1 use and repetition in the post-test showed that the error that stood out on the screen was not found or reduced in the post-speech test. The results of Study 2 also supported this claim by analyzing the practice session. Repetition and L1 use were not found when texting was used during practice. Study 3 claimed a similar result. The results found in the three studies under different task conditions support the use of texting to highlight linguistic features and errors that allow for correction by focusing learners' attention on form, thus improving lexical accuracy and syntactic complexity. Although texting is text-only, it has been shown that instantaneous exchange improves speaking. The use of texting is nonetheless inferior to practicing face-to-face or more effective than face-to-face practice in improving lexical accuracy and syntactic complexity. In contrast, face-to-face showed no significant difference in lexical error rates for both content and function words.

Although lexical accuracy improves after text practice, task implementation under time pressure conditions encountered limitations. For example, a 15-minute texting task in Study 2 showed improvements in content word lexical error rates but no difference between the pre-test and post-test for function words. The time limit may have directed attentional resources to content words, as the interaction in the task was more focused on message delivery. However, despite the time limit, the texting group showed a decrease in error types such as incorrect word choice and L1 use for content words and repetition of function words in the tests. Again, this could be because these errors are obvious on the screen. In Study 2, it was found that 15 minutes was not enough to improve function words because they were difficult to recognize. However, even under time pressure, learners can use text to identify errors in content words.

Although the current research generally corroborates with Sauro and Smith's (2010) findings of the function of texting in language learning, the studies conducted in this research did not find improvements in lexical diversity and grammatical error rates in either the face-to-face or texting groups. The discussion was addressed in Section 6.3. This difference can be explained from two perspectives: the occurrence of trade-offs between lexical diversity and lexical accuracy and the different measures used for lexical diversity. First, the improvement in lexical accuracy (not lexical diversity) can be explained by the different L1 backgrounds of the participants. The participants in Sauro and Smith's study were all native English speakers learning German, whereas the current study focused on Japanese learners learning English. Some believe that second language acquisition is influenced by similarity to the L1 (Allen, 2019; Ortega, 2009; Ringbom & Jarvis, 2009). In other words, orthographically similar languages facilitate the acquisition of lexical items (Grosjean & li, 2013; Marian et al., 2012), so German ESL learners may have a lower cognitive load than Japanese ESL learners because of orthographic similarities. The Japanese learners in the current study may have focused on accuracy rather than variety because their vocabulary is limited to expansion.

Second, the measures used for lexical diversity may have impacted the difference as well. In Sauro and Smith's study, the Index of Guiraud was employed for lexical diversity, while the current research used the Measure of Textual Lexical Diversity (MTLD). The Index of Guiraud is affected by the length of the text (Kojima, 2012). In addition, the length of practice and task type may have allowed learners to produce more language. For instance, the Sauro and Smith study utilized the jigsaw task for approximately 40 minutes, while the current research analyzed one-minute speech. More time on the speaking test may have been an advantage for improving lexical diversity.

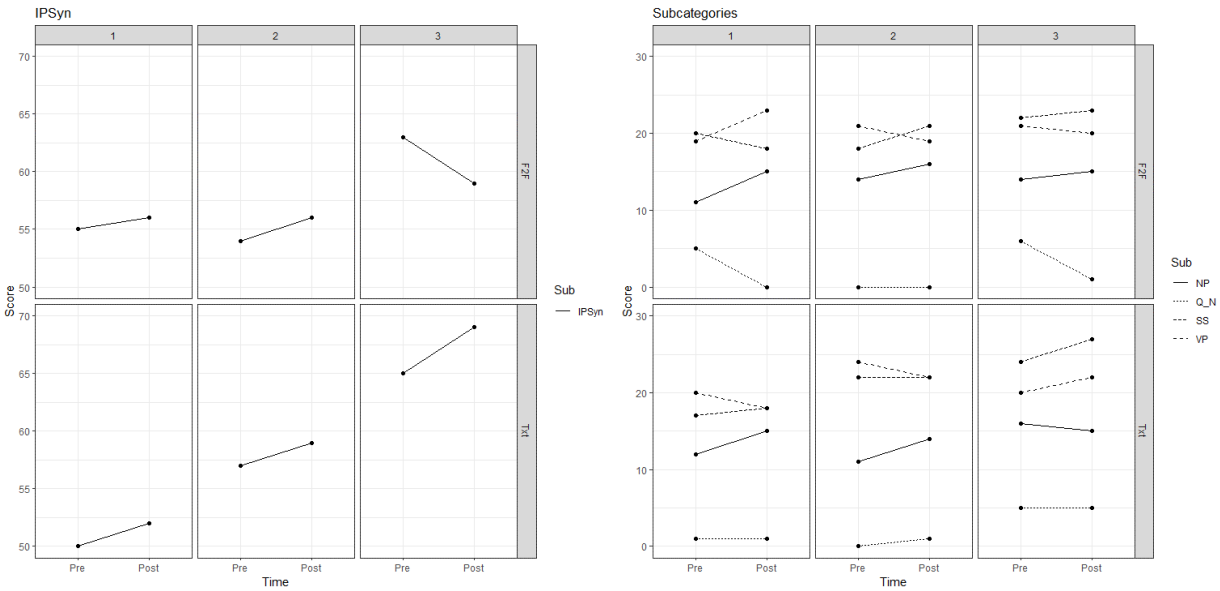
Concerning syntactic complexity, Sauro and Smith's study and the current research both found positive results. Sauro and Smith counted the ratio of clauses to c-units and occurrences of grammatical gender for syntactic complexity. This study corroborated these results using the IPSyn score and found that texting groups either maintained or improved

between pre-and post-speaking tests. *Figure 20* shows how texting improved its score among different task conditions. The subcategory showed that the texting group scored high on sentence structure. Texting showed a lower score for verb phrases and a higher score for sentence structure, indicating word occurrence and syntax.

For the face-to-face group, IPSyn score and subcategory score results were inconsistent across task conditions, and scores did not always improve. While Sauro and Smith (2010) counted only grammatical gender for sentence complexity, the IPSyn score reflects a more detailed analysis of grammar emergence and sentence complexity. Therefore, the inconsistent results for the face-to-face groups under the task conditions suggest little possibility of improvement.

As for the fluency measure, the texting group showed no improvement in articulation rate. Nevertheless, the speech rate improved significantly under the repeated tasks and repetition conditions and under the temporal conditions at three-week intervals. Although the articulation phase is skipped in texting, there was evidence that texting improved fluency measures. The task conditions with sufficient time for practice allowed for improvement in speech fluency. For the face-to-face group, more practice time affected both articulation and speech rate. Further discussion is provided in the next section on face-to-face and fluent speaking.

Figure 20 *IPSyn Score across all Three Studies*



Note. The IPSyn score results reflect how practice in each mode emphasizes sentence structure and facilitates sentence production. This is especially true of the texting mode.

9.2. Face-to-Face Practice and Speaking Improvement

While the current study aims to examine the effects of learning via texting, the comparative analysis highlights the characteristics of the face-to-face group. Under conditions of repetition or time pressure, the face-to-face group showed no significant statistical improvement in lexical error rate, lexical diversity, grammatical error rate, and syntactic complexity after practicing the task. In particular, face-to-face practice showed no statistical improvement in lexical error rate for content and function words, in contrast to texting practice. Moreover, an increase in lexical errors, such as repetitions and unnecessary function words, was observed under several conditions. However, when the task was repeated, a positive outcome was observed on error types for incorrect word choice with content words; a decrease in the occurrences was found in Studies 1 and 3. It could be inferred that in face-to-face learning, repetition may lead to a reduction in content word choice errors.

In many cases, as discussed in section 4.2, since speaking requires the attention of multiple areas simultaneously, many errors may be difficult to detect in face-to-face practice. This statistical result on lexical error rate was consistent across task conditions, suggesting that the amount of practice performed in these studies was insufficient to show improvement. Improving accuracy in face-to-face requires time and practice. By contrast, texting is more effective as improvements in lexical accuracy were found across all the task conditions conducted in this research.

When syntactic complexity was assessed for the face-to-face group, the results were inconsistent across the study. The IPSyn score increased after repetition and 15-minute practice but decreased during the 3-week repetition training for lexical phrases. This suggests that it is difficult for face-to-face modes to improve the same linguistic skills under the same conditions as texting.

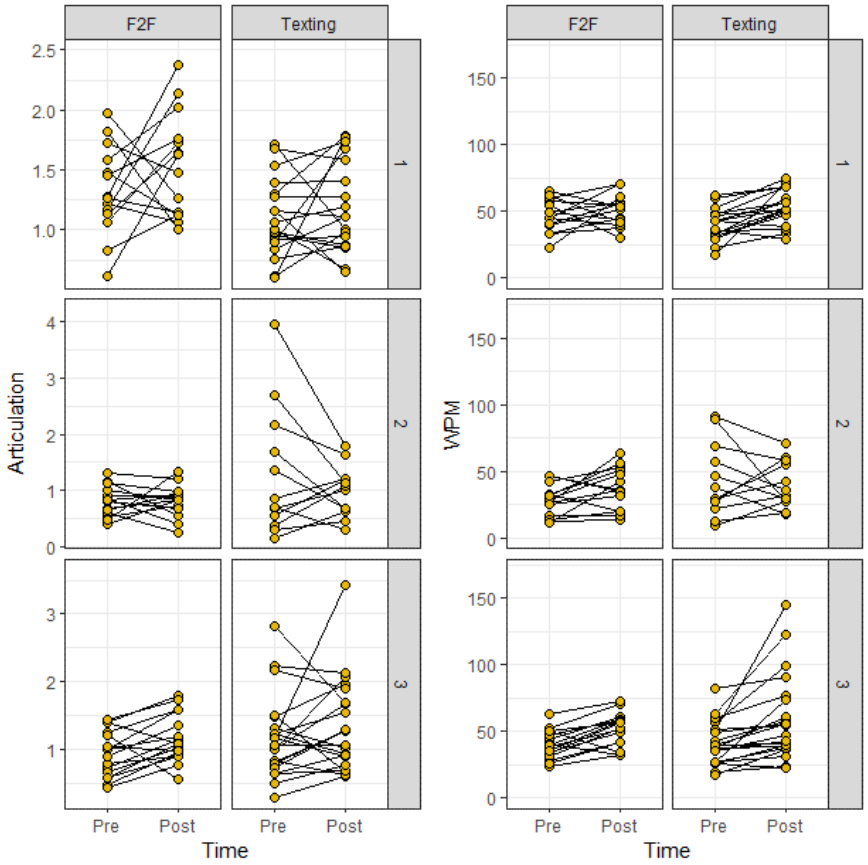
That being said, the significant advantage of the face-to-face mode is that, under the same time and practice conditions, a face-to-face practice allowed improvements in both articulation and speech rates (*Figure 21*). This improvement was not found in texting. The impact was significant when practiced for 15 minutes during lexical phrase learning. Both modes showed an increase in speech rate; however, only the face-to-face group improved articulation, predicting improved comprehensibility when speaking. This implies two things. First, oral practice is necessary to improve articulation since there is no articulation in texting. Speech behavior consists of several constructs, such as lexico-grammatical and discursive features, in addition to articulation mentioned in section 1.3. The feature that directly affects intelligibility is articulation rate. Articulation rate is a small unit of words between pauses that have been shown to be a strong indicator of comprehensible output (Suzuki & Kormos, 2019). Texting may lack this feature because pauses are not visible when texting. In texting, pauses and repairs can be inserted without many constraints, which is an advantage for

improving accuracy, but may not be for improving fluency. This study found how the face-to-face speaking mode affects fluency constructs more than the texting mode and that different modes affect different areas of speech constructs.

The finding of this study can be confirmed from another perspective. The results of Study 1 through 3 have shown that articulation speed or speech rate can be achieved without explicit instructions. Note that none of the studies were explicitly instructed to focus on articulation or speed. However, the face-to-face group did improve its score on articulation rate. This rules out the possibility that only the mode difference affected the improvement.

Second, despite the fact that mode difference affects fluency improvement, task conditions are also an essential construct. Fluency constructs such as articulation rate may not improve without sufficient time and repetition. Under the 15-minute timed condition, fluency did not improve for both face-to-face and texting. However, after the 3-week practice of learning lexical phrases, both articulation rate and speech rate improved for the face-to-face group. The result shows that both mode and task conditions affect task performance.

Figure 21 *Articulation Rate and Speech Rate across all Three Studies*



Note. F2F = face-to-face, WP*M* = words per minute (speech rate)

The face-to-face practice showed a tendency towards improving fluency measures over lexical accuracy or syntactic complexity. In other words, lexical accuracy and syntactic complexity may have been sacrificed to improve fluency. These results provide evidence of how improvements differ between the two modes.

9.3. Using Texting to Learn Lexical Phrases

Lexical phrases were chosen as the target items to test the effectiveness of using texting for learning. The second research question was, "Is there a difference in learning lexical phrases between texting and face-to-face interactions?" Cloze test results revealed no difference between the face-to-face and texting groups. Additionally, lexical phrases were used naturally in post-test speaking. This suggests that texting creates opportunities for lexical phrase learning, not merely retrieving stored information. Thus, texting facilitates lexical phrase learning for similar outcomes as face-to-face.

A limitation of this study, however, is that it may be assumed that the list of lexical phrases was selected from a combination of known words. Had the words been unknown, the

lack of articulation practice in texting could have proved a disadvantage for natural use in speech.

9.4. Summary

This chapter summarizes the results and discusses the main findings from this research. Mode differentially affected the results in terms of lexical error rate, syntactic complexity, and fluency under three different task conditions: repetition, 15-minute time constraints, and repetition and time constraints for learning lexical phrases. The findings of this research were that while texting improved lexical accuracy and syntactic complexity, face-to-face practice improved articulation rate. In addition, there was no mode difference in terms of learning lexical phrases. The effect of texting on lexical accuracy has been proven by the detailed analysis of lexical error types showing a decrease in L1 use for content words and repetition for function words. The error types showed how the salient errors on the screen promote awareness to help achieve lexical accuracy.

The results confirm those of Sauro and Smith (2010) with a few differences. One is that the current research did not show improvement in lexical diversity, but instead improved lexical accuracy. This may be due to the differences in participants' native language engaging in the research and the measures applied.

In the process of exploring the effects of texting, the effects of face-to-face learning were also identified. When equal time is given and repeated, only face-to-face mode improves both constructs of fluency, speech rate, and articulation rate. The mode of practice determines the areas of language improvement, but task conditions also affect the outcome.

CHAPTER 10 CONCLUSION

Based on the results, this chapter presents the theoretical implications and limitations of this study, as well as suggestions for further research. The effect of texting on speaking was investigated in three empirical studies conducted with EFL learners at CEFR proficiency levels A1-B1. The result shows that there are differences in the area where speaking improves depending on the mode. It was found that both modes provide an improvement in speaking.

10.1. Theoretical Implications

The hypothesis, based on production models and previous texting research (Levelt, 1989; Payne & Whitney, 2002), suggests positive outcomes for post-speech outcomes when texting is used as a medium for task-based learning. The results of this study showed that use of texting improved speaking, which implies that texting follows a similar process as speaking. The contributions of this study are not limited to the effects of texting but how mode differs from face-to-face practice by affecting various constructs of speaking; that is, texting may lead one to focus more on lexical accuracy or syntactic complexity when speaking and that face-to-face practice improves measures of fluency.

The inherent features of visual salience and typing behavior in texting emphasize linguistic forms (Sauro & Smith, 2010, Smith, 2008). Thus, one clear difference between face-to-face interactions and texting is self-repair behavior. Although texting repairs are not always seen in transcripts (Smith, 2008), they do influence speech production. The post-speech test supports the claim that texting draws attention to correct forms, which was justified by the results on lexical error rate and syntactic complexity. The analysis of the lexical error types in both content and function words showed that repetition or L1 use of error types reduced or was not found in the texting groups, both in the practice session and in the post-speech test. The noticeable errors on the screen enabled learners to notice and correct them.

Payne and Whitney (2002) predicted, based on Levelt's (1989) language production model, that texting develops fluency. This study was able to identify that speech rate, a construct of fluency, particularly improved in texting practice. Face-to-face practice, by contrast, allows practicing articulating the message, which helps to improve both articulation rate and speech rate. The results illustrate how mode impacts different areas of fluency constructs.

Comparative studies have shown the effects of the different modes on subsequent speech performance. Language practice in texting is an advantage for novice language learners in speaking. Therefore, this study contributes to the existing literature on speaking practice and how texting supports it differently from face-to-face practice.

10.2. Pedagogical Implications

An educational implication of this study is that there is a difference in areas of improvement when face-to-face and texting practices are compared. The findings can help educators in selecting the mode for pedagogical purposes. This study found that one of the benefits of using texting is that it draws attention to vocabulary and sentence structure, based on the outcome of lexical errors and syntactic complexity. The speed of the conversation is naturally slower, allowing them to assess and correct their performance in real-time. In addition to the visual information of the text on the screen, the extra time in texting can support language learning, especially for beginners, as it is less intimidating and responses can be delayed. Compared to face-to-face conversation, interlocutors can have more time to formulate messages and responses, leading to greater attention to correct forms. By contrast, in face-to-face speech production, multiple pieces of information must be processed simultaneously in a short period of time. For novice L2 learners, this can be overwhelming.

Speaking is not just about lexical or syntactic accuracy and complexity. The comprehensibility of speech is an important construct as well. For this improvement, face-to-face may be more effective compared to texting. The result of the articulation rate across the studies provide evidence for this. For improvement in fluency, especially for articulation rate, this study has found face-to-face to be more effective compared to texting.

Overall, this study provides evidence that speaking can be improved through both face-to-face and texting practice. It was found that speaking can be influenced by the way learners practice, this is, mode. It can therefore be suggested that educators consider the differences in pedagogy.

10.3. Limitations and Suggestions for Future Research

This is a small-scale study and is, therefore, subject to some limitations. The first limitation is the sample size and the participants in this study. The purpose of this study was to investigate the use of texting in a natural learning environment. To this end, all studies conducted in this study were held during regular classroom hours. Moreover, although this study design was ecologically valid, the number of participants and the task conditions were limited to the class size and the lesson plan. In addition, the language level of the participants was limited to CEFR A1- B1, and they may have had difficulty holding a spontaneous conversation. Therefore, exchanges between participants were not always smooth and made it difficult to make accurate corrections. Had the participants been native speakers or proficient language users, the impact of the face-to-face and texting exercise might have been stronger. Gurzynski-Weiss and Baralt (2014) note that the feedback provider can influence language performance outcomes. In their study of corrective feedback, the interlocutor had a research and teaching background. This produced different results than a similar study in which the interlocutor was a native speaker. Thus, interlocutors with different backgrounds may be able to provide expert testimony on errors. In other words, the role of the interlocutor may

influence learners' language performance, which could affect researchers' understanding of the texting effect.

Another limitation is the assumption of the interview task and the selection of the topic for the interview. As mentioned in Section 5.3.3, the interview task may lead to variations in the difficulty of the spoken content. Because the task depends on a topic that the speaker selects, some learners may avoid complex, unfamiliar topics or specific topics that require a certain vocabulary. The interview task in Study 3 was a question about one's favorite place to live. This interview task may have been straightforward because grammatical errors related to verbs were rarely found. Other highly structured tasks or advanced topics that correspond to the learners' language level (e.g., narrative or decision-making) might have yielded different results. For example, Ziegler (2018) conducted a study using pictorial narrative tasks via texting and allowed three minutes of planning time. The results showed that learners improved their scores on the Guiraud index. This result provides insight into how task types and topics can influence outcomes.

For further research, given the limitations discussed above, it would be worthwhile to examine various conditions and the extent of texting in other languages. Texting highlights linguistic forms. Ziegler's (2016) meta-analysis on the use of SCMC for language learning found that the use of SCMC had a greater delayed effect on grammar compared to the practice of face-to-face interaction. This research can be continued in the long term to investigate the effects of retrieving information in different modes. In addition, it is worthwhile to investigate how the results differ in learning with other target languages than English. The current research selected English as a foreign language for the focus of the research. The specific language selected may lead to a different result. Future studies may investigate the use of texting in other languages and generalize the effect on language learning.

In addition, tasks and lessons should be explored to make better use of the tool because texting can be useful both inside and outside the classroom. Texting can democratize the different language levels and cultural backgrounds of learners. Telecollaboration pedagogy is widely documented and provides a platform for successful ideas (Akiyama & Cunningham, 2017). Creative use of texting can facilitate learning and sharing ideas for authentic learning.

10.4. Conclusion

The current study sheds new light on the evidence that texting affects the subsequent speaking performance of novice learners, who achieve different results from face-to-face practice. In particular, lexical accuracy and syntactic complexity in speaking are improved by texting. Results showed that error types such as L1 use and repetition were reduced during texting, which affected subsequent speaking performance. Another important discovery was that face-to-face showed improvements in measures of fluency, while texting had limitations.

CONCLUSION

When practicing face-to-face, it was only this mode that improved articulation rate when practiced under the same conditions as texting. This suggests that mode affects different areas of speech performance and that both are important for language learning.

Texting is an important aspect of communication today, and its integration into language learning benefits foreign language acquisition. Moreover, the study proved that texting, although a written mode, nevertheless helps to improve areas of speaking, especially where the face-to-face practice is less effective.

APPENDIX

APPENDICES

Appendix 1 *Participant Consent Form*

Research topic: Texting and Face-to-face Speaking in Task-Based Language Teaching

1. I confirm that I have read and understand the letter dated [insert date] explaining the purpose, method, procedures, benefits and risks for the above research project and I have had the opportunity to ask questions about the project. ☐

2. I understand that my participation is voluntary and that I am free to withdraw at any time without giving any reason and without there being any negative consequences. In addition, should I not wish to answer any particular question or questions, I am free to decline. ☐

3. I understand that my responses will be kept strictly confidential. I give permission for members of the research team to have access to my anonymized responses. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in the report or reports that result from the research. ☐

4. I agree to be recorded for each session. ☐

5. I agree for the data collected from me to be used in future research. ☐

6. I agree to delete any materials and personal information obtained from taking part in the activity and will not disclose it to public or to the third party. ☐

7. I agree to take part in the above research project ☐

APPENDIX

_____	_____	_____
Name of Participant	Signature	Date
_____	_____	_____
Name of person taking consent	Signature	Date
_____	_____	_____
Research Representative	Signature	Date

Retrieved from https://www.sampleforms.com/participant-consent-form-sample.html

APPENDIX

Participant Withdrawal Form

To the Research Representative, _____

Research topic: Texting and Face-to-face Speaking in Task-Based Language Teaching

I hereby request to withdraw from the study, although I have agreed to participate in the study above.

_____ _____ _____

Name of Participant Signature Date of Withdrawal

【Confirmation of the research in charge】 I confirmed that the participant above withdrew from the study.

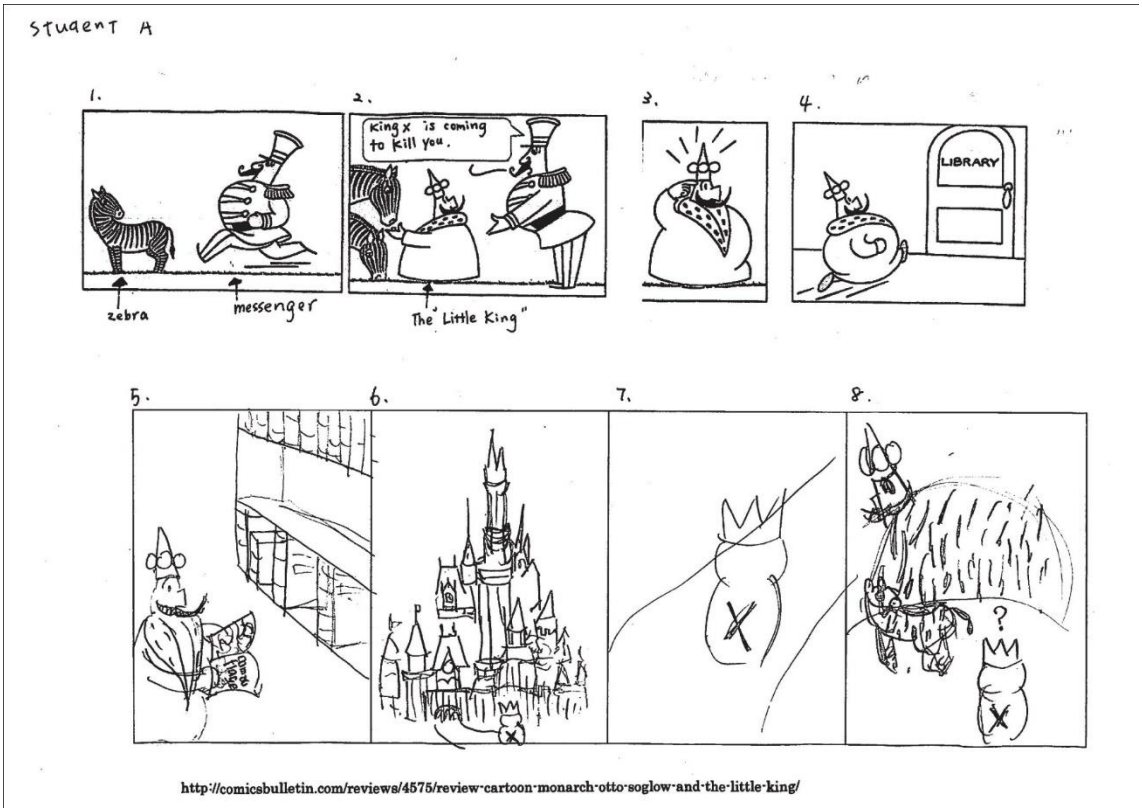
_____ _____ _____

Research Representative Signature Date of Withdrawal

APPENDIX

Appendix 2 *Worksheet for the Narration Task*

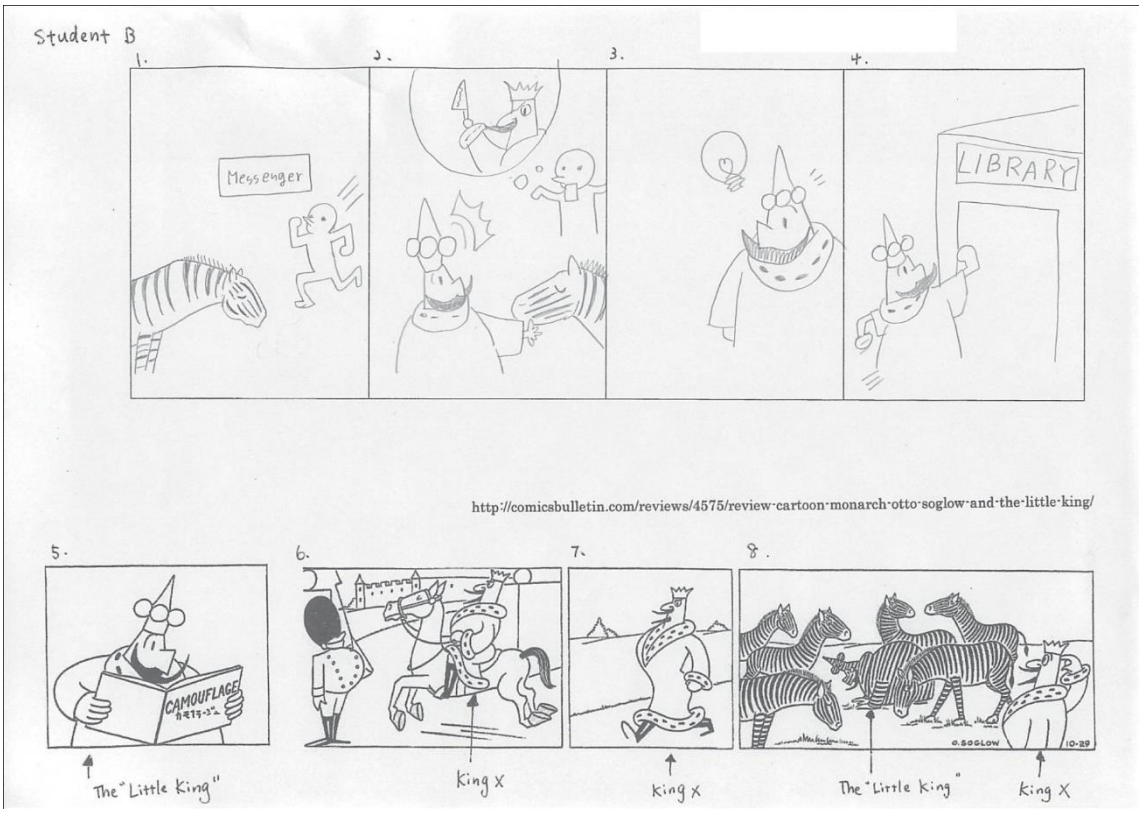
Student A



Reprinted with permission from The Library of American Comics.

APPENDIX

Student B



Reprinted with permission from The Library of American Comics.

APPENDIX

Appendix 3 *The First Draft of Cloze Test*

1. This band is not so famous, but they've been ( ) a while. They've done a lot of albums.
2. When she turned ( ) the car radio, the music was playing.
3. Why I take the piano ( ) the road and in the air.
4. I'm ( ) awe and have so much respect for the music.
5. This sofa does not ( ) with the rest of our furniture in the living room.
6. I don't like the color. Can we look at something ( ), please?
7. A stop sign in Brazil is ( ) the shape of an octagon.
8. There is a unicorn ( ) the left.
9. The family name is ( ) the bottom.
10. When I ( ) down in the barber's chair, there was a Chicago flag on the box that barber kept all his tools in.
11. There was a Chicago flag ( ) the wall behind me.
12. A design that is distinct is different ( ) others.
13. If you engage ( ) something, you show interest in it.
14. When you ( ) attention to something, you concentrate on it.
15. Setting is an element in a story where and when the story ( ) place.
16. This book is about a ( ) of people stuck on an island together.
17. He goes to Egypt looking ( ) treasure.
18. Historical fiction has a fictional story and is set ( ) the past.
19. I grew ( ) in Tokyo.
20. Although we have a good transportation system in our city, I think the best way to get ( ) is by bus.
21. If you can choose to live anywhere in the world, what do you ( ) for in a city?
22. The mall is a good place to hang ( ) with friends.

APPENDIX

23. Since it's raining today, let's ( ) out and do nothing.

24. Can you get to the airport ( ) ease?

25. People come ( ) different part of Asia to live in Tokyo.

26. I found ( ) that living in Hamamatsu is very convenient.

27. I want to check ( ) the new restaurant.

28. I often meet ( ) with my friends at the coffee shop.

29. While you are in New York, take ( ) some live theater.

30. I like to ( ) camping with my friends.

31. There is a real excitement searching ( ) fossils.

32. The book turned ( ) to be extremely exciting.

33. Many people are attracted ( ) another country because they like the way they do things.

34. When you are in a foreign country, try to interact ( ) others.

APPENDIX

Appendix 4 *Second Draft of Cloze Test*

1. When she turned ( ) the car radio, the music was playing.
2. I'm ( ) awe and have so much respect for the music.
3. I don't like the color. Can we look at something ( ), please?
4. A stop sign in Brazil is ( ) the shape of an octagon.
5. There is a unicorn ( ) the left.
6. The family name is ( ) the bottom.
7. There was a Chicago flag ( ) the wall behind me.
8. A design that is distinct is different ( ) others.
9. If you engage ( ) something, you show interest in it.
10. When you ( ) attention to something, you concentrate on it.
11. This book is about a ( ) of people stuck on an island together.
12. He goes to Egypt looking ( ) treasure.
13. Historical fiction has a fictional story and is set ( ) the past.
14. I grew ( ) in Tokyo.
15. Although we have a good transportation system in our city, I think the best way to get ( ) is by bus.
16. The mall is a good place to hang ( ) with friends.
17. Can you get to the airport ( ) ease?
18. People come ( ) from different part of Asia to live in Tokyo.
19. I found ( ) that living in Hamamatsu is very convenient.
20. I want to check ( ) the new restaurant.
21. I like to ( ) camping with my friends.
22. There is a real excitement searching ( ) fossils.
23. The book turned ( ) to be extremely exciting.

APPENDIX

24. Many people are attracted ( ) another country because they like the way they do things.

25. When you are in a foreign country, try to interact ( ) others.

APPENDIX

Appendix 5 *Final Cloze Test with 15 Target Items*

1. I want to check ( ) the new restaurant. 新しいレストランに行ってみたいな。
2. The mall is a good place to hang ( ) with friends.　モールは友達とぶらぶらするに良い場所だ。
3. Although we have a good transportation system in our city, I think the best way to get ( ) is by bus.　私たちのまちではよい交通システムがあるが、あちらこちらに動き回るのにはバスが最適だろう。
4. If you engage ( ) something, you show interest in it. 何かに従事するとうことは、それに興味を示しているということである。
5. I like to ( ) camping with my friends.　友達とキャンプすることが好きだだ。
6. I don't like the color. Can we look at something ( ), please?　私はその色が好きではないな。他のを見せて頂ける？
7. I'm ( ) awe and have so much respect for the music.　私は音楽に対して畏敬の念を抱いている。
8. Historical fiction has a fictional story and is set ( ) the past. 歴史小説とはフィクションで、過去を題材にしている。
9. This book is about a ( ) of people stuck on an island together.この本は、ある一組の団体が一緒に島に取り残される話である。
10. I found ( ) that living in Hamamatsu is very convenient.　浜松に住むことは大変便利だと気付いた。
11. Can you get to the airport (　　) ease? 空港まで簡単に行けますか。
12. There is a real excitement searching ( ) fossils.　化石を探すことは本当にワクワクする。
13. The book turned ( ) to be extremely exciting.この本は結果的にとても面白かった。
14. Many people are attracted (　　) another country because they like the way they do things. 多くの人が他の国に惹かれるのは他のやり方に魅力を感じるからだ。
15. When you are in a foreign country, try to interact (　　) others. 外国に行ったら、他の人と関わるようにしよう。

REFERENCES

Abe, M., & Roever, C. (2019). Interactional competence in L2 text-chat interactions: First-idea proffering in task openings. *Journal of Pragmatics*, *144*, 1–14. https://doi.org/10.1016/j.pragma.2019.03.001

Abrams, Z. I. (2003). The Effect of Synchronous and Asynchronous CMC on Oral Performance in German. *The Modern Language Journal*, *87*(2), 157–167. https://doi.org/10.1111/1540-4781.00184

Adams, R., & Nik Mohd Alwi, N. A. (2014). Prior knowledge and second language task production in text chat. *Technology-Mediated TBLT*, 51–78. https://doi.org/10.1075/tblt.6.03ada

Ahmadian, M. J., & Tavakoli, M. (2010). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research, 15(1),* 35–59. https://doi.org/10.1177/1362168810383329

Ahmadian, M. J., & Tavakoli, M. (2014). Investigating What Second Language Learners Do and Monitor under Careful Online Planning Conditions. *Canadian Modern Language Review*, *70*(1), 50–75. https://doi.org/10.3138/cmlr.1769

Ahmadian, M. J., Tavakoli, M., & Vahid Dastjerdi, H. (2015). The combined effects of online planning and task structure on complexity, accuracy and fluency of L2 speech. *The Language Learning Journal*, *43*(1), 41–56. https://doi.org/10.1080/09571736.2012.681795

Akiyama, Y., & Cunningham, D. J. (2017). Synthesizing the Practice of SCMC-based Telecollaboration: A Scoping Review. *CALICO Journal*, *35*(1), 49–76. https://doi.org/10.1558/cj.33156

Allen, D. (2019). The prevalence and frequency of Japanese-English cognates: Recommendations for future research in applied linguistics. *International Review of Applied Linguistics in Language Teaching*, *57*(3), 355–376. https://doi.org/10.1515/iral-2017-0028

REFERENCES

Altenberg, E. P., Roberts, J. A., & Scarborough, H. S. (2018). Young Children's Structure Production: A Revision of the Index of Productive Syntax. *Language, Speech, and Hearing Services in Schools*, *49*(4), 995–1008. https://doi.org/10.1044/2018_lshss-17-0092

Anderson, J. R. (2005). *Cognitive psychology and its implications*. Macmillan.

Anthony, L. (2022). *Laurence Anthony's AntConc*. [Computer Software, Version 4.0.11]. Tokyo, Japan. Laurenceanthony.net. https://www.laurenceanthony.net/software/antconc/

Anthony, L. (2004). *Laurence Anthony's AntWordProfiler*. [Computer Software, Version 1.4.1]. Tokyo, Japan. Www.laurenceanthony.net. https://www.laurenceanthony.net/software/antwordprofiler/

Beauvois, M. (1997). Write to speak: The effects of electronic communication on the oral achievement of fourth semester French students. In J. A Muyskens (Eds), *New Ways of Learning and Teaching: Issues in Language Program Direction*, (pp.93-115). Heinle and Heinle publishers. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/1aba059a-c044-4554-9f7a-94e125fda5b7/content

Beauvois, M. (1998). Conversations in Slow Motion: Computer-mediated Communication in the Foreign Language Classroom. *Canadian Modern Language Review*, *54*(2), 198–217. https://doi.org/10.3138/cmlr.54.2.198

Bei, G.X. (2012). Effects of Immediate Repetition in L2 Speaking Tasks: A Focused Study. *English Language Teaching*, *6*(1). https://doi.org/10.5539/elt.v6n1p11

Biber, D. (1988). *Variation across speech and writing.* Cambridge University Press.

Blake, R. (2000). Computer mediated communication: a window on l2 Spanish interlanguage. *Language Learning and Technology*, *4*(1), 111–125. https://scholarspace.manoa.hawaii.edu/server/api/core/bitstreams/973bed3f-8657-45ef-8717-bbd758105b45/content

REFERENCES

Boers, F. (2014). A Reappraisal of the 4/3/2 Activity. *RELC Journal*, *45*(3), 221–235. https://doi.org/10.1177/0033688214546964

Boers, F., Lindstromberg, S., & Webb, S. (2014). Further Evidence of the Comparative Memorability of Alliterative Expressions in Second Language Learning. *RELC Journal*, *45*(1), 85–99. https://doi.org/10.1177/0033688214522714

Boersma, P., & Weenink, D. (2019). *Praat: doing Phonetics by Computer*. Www.fon.hum.uva.nl. https://www.fon.hum.uva.nl/praat/

Bohlke, D. (2016). *Keynote 2.* National Geographic Cengage Learning K.K.

Bosker, H. R., Pinget, A.-F., Quené, H., Sanders, T., & De Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, *30*(2), 159–175. https://doi.org/10.1177/0265532212455394

Bot, K. D. (1992). A bilingual production model: Levelt's "speaking" model adapted. *Applied Linguistics*, *13*, 1–24. https://doi.org/10.1093/applin/13.1.1

Bozorgian, H., & Kanani, S. M. (2017). Title: Task Repetition on Accuracy and Fluency: EFL Learners Speaking Skill. *International Journal of English Language and Literature Studies*, *6*(2), 42–53. https://doi.org/10.18488/journal.23.2017.62.42.53

Brown, H. D., & Lee, H. (2015). *Teaching by principles: An interactive approach to language pedagogy* (Fourth edition). Pearson Education.

Brunetti,I., & Soglow, O.(2012). *Cartoon Monarch: Otto Soglow and the Little King.* California: Idea & Design Works Llc.

Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. *Challenge and change in language teaching*, 136–146.

Bygate, M. (2009). Teaching and Testing Speaking. In M. H. Long., & C. J. Doughty (Eds), *The Handbook of Language Teaching* (pp. 412-440). Wiley and Blackwell publications.

REFERENCES

Carr, C. T. (2020). CMC Is Dead, Long Live CMC!: Situating Computer-Mediated Communication Scholarship Beyond the Digital Age. *Journal of Computer-Mediated Communication,* 25(1), 9–22. https://doi.org/10.1093/jcmc/zmz018

Carrington, V. (2004). Texts and literacies of the Shin Jinrui. *British Journal of Sociology of Education*, *25*(2), 215–228. https://doi.org/10.1080/0142569042000205109

Carrol, G., & Conklin, K. (2020). Is All Formulaic Language Created Equal? Unpacking the Processing Advantage for Different Types of Formulaic Sequences. *Language and Speech*, *63*(1), 002383091882323. https://doi.org/10.1177/0023830918823230

Chamberland, C., Saint-Aubin, J., & Légère, M. A. (2013). The impact of text repetition on content and function words during reading: Further evidence from eye movements. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*, 67(2), 94. https://psycNet.apa.org/buy/2012-15519-001

Chambers, F. (1997). What do we mean by fluency? *System*, *25*(4), 535–544. https://doi.org/10.1016/s0346-251x(97)00046-8

Christiansen, M. H., & Arnon, I. (2017). More Than Words: The Role of Multiword Sequences in Language Learning and Use. *Topics in Cognitive Science*, *9*(3), 542–551. https://doi.org/10.1111/tops.12274

Council of Europe. (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume*. Council of Europe Publishing, Strasbourg. https://www.coe.int/lang-cefr.

Crowther, D., Trofimovich, P., Isaacs, T., & Saito, K. (2015). Does a Speaking Task Affect Second Language Comprehensibility? *The Modern Language Journal*, *99*(1), 80–95. https://doi.org/10.1111/modl.12185

Cucchiarini, C., Hubers, F., & Strik, H. (2020). Learning L2 idioms in a CALL environment: the role of practice intensity, modality, and idiom properties. *Computer Assisted Language Learning*, 35(4), 1–29. https://doi.org/10.1080/09588221.2020.1752734

De Jong, N. H., & Bosker, H.R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In R. Eklund (Ed.), *Proceedings of the 6th Workshop on*

Disfluency in Spontaneous Speech (DiSS) (pp. 17-20). http://hdl.handle.net/11858/00-001M-0000-0015-0FB8-8

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second Language Fluency: Judgments on Different Tasks. *Language Learning*, *54*(4), 655–679. https://doi.org/10.1111/j.1467-9922.2004.00282.x

Egusa, C., & Yokoyama, Y. (2004). The Effects of Task Types on Second Language Speech Production among Japanese University Students : Fluency, Accuracy, Complexity, and Trade-off Effects. *ARELE: Annual Review of English Language Education in Japan*, *15*, 129–138. https://doi.org/10.20581/arele.15.0_129

Ellis, R. (2005). *Planning and Task Performance in a Second Language* (Vol. 11). John Benjamins Publishing.

Erman, B. (2007). Cognitive processes as evidence of the idiom principle. *International Journal of Corpus Linguistics*, *12*(1), 25–53. https://doi.org/10.1075/ijcl.12.1.04erm

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text - Interdisciplinary Journal for the Study of Discourse*, *20*(1). https://doi.org/10.1515/text.1.2000.20.1.29

ETS. (2014). *TOEFL iBT TEST Independent Speaking Rubrics.* https://www.ets.org/s/toefl/pdf/toefl_speaking_rubrics.pdf

Faerch, C., & Kasper, G. (1984). Two ways of defining communication Strategies. *Language Learning*, *34*(1), 45–63. https://doi.org/10.1111/j.1467-1770.1984.tb00995.x

Faul, F., Erdfelder, E., Lang, A.-G. & Buchner, A. (2007). G*Power 3: A Flexible Statistical Power Analysis Program for the Social, Behavioral, and Biomedical Sciences. *Behavior Research Methods*, *39*, 175–191. https://www.psychologie.hhu.de/arbeitsgruppen/allgemeine-psychologie-und-arbeitspsychologie/gpower

Favreau, M., & Segalowitz, N. S. (1983). Automatic and controlled processes in the first- and second-language reading of fluent bilinguals. *Memory & Cognition*, *11*(6), 565–574. https://doi.org/10.3758/bf03198281

REFERENCES

Federal Motor Carrier Safety Administration (FMCSA). (2012). *No Texting Rule Fact Sheet: Federal Motor Carrier Safety Regulations*. US Department of Transportation. Retrieved from https://www.fmcsa.dot.gov/sites/fmcsa.dot.gov/files/docs/No_Texting_Rule_Fact_Sheet.pdf

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: a unit for all reasons. *Applied Linguistics*, *21*(3), 354–375. https://doi.org/10.1093/applin/21.3.354

Foster, P., & Skehan, P. (1996). The Influence of Planning and Task Type on Second Language Performance. *Studies in Second Language Acquisition*, *18*(3), 299–323. https://doi.org/10.1017/s0272263100015047

Foster, P., & Wigglesworth, G. (2016). Capturing Accuracy in Second Language Performance: The Case for a Weighted Clause Ratio. *Annual Review of Applied Linguistics*, *36*, 98–116. https://doi.org/10.1017/s0267190515000082

Fraundorf, S. H., & Watson, D. G. (2011). The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language*, *65*(2), 161–175. https://doi.org/10.1016/j.jml.2011.03.004

Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, *20*(3), 321–340. https://doi.org/10.1177/1362168815570142

García-Pastor, M. D. (2018). Linguistic Accuracy in Spanish Adolescent Learners' EFL Writings. *Miscelánea: A Journal of English and American Studies*, *57*, 79–99. https://doi.org/10.26754/ojs_misc/mj.20186323

Gass, S., Mackey, A., Alvarez-Torres, M. J., & Fernández-García, M. (1999). The Effects of Task Repetition on Linguistic Output. *Language Learning*, *49*(4), 549–581. https://doi.org/10.1111/0023-8333.00102

Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *IRAL - International Review of Applied Linguistics in Language Teaching*, *45*(3). https://doi.org/10.1515/iral.2007.010

REFERENCES

Golonka, E. M., Tare, M., & Bonilla, C. (2017). Peer interaction in text chat: Qualitative analysis of chat transcripts. *Language Learning & Technology*, 21(2), 157–178. https://dx.doi.org/10125/44616

González -Lloret, M. (2022). Conversation Analysis of Computer-Mediated Communication. *CALICO Journal*, 28(2), 308-325. https://www.jstor.org/stable/calicojournal.28.2.308

González-Lloret, M., & Ortega, L. (2014). Towards technology-mediated TBLT: An introduction. In M. González-Lloret & L. Ortega (Eds.), *Technology-mediated TBLT: Researching technology and tasks* (pp. 1–22). Amsterdam, the Netherlands: John Benjamins.

Gregori-Signes, C., & Clavel-Arroitia, B. (2015). Analysing Lexical Density and Lexical Diversity in University Students' Written Discourse. *Procedia - Social and Behavioral Sciences*, *198*, 546–556. https://doi.org/10.1016/j.sbspro.2015.07.477

Gurzynski-Weiss, L., & Baralt, M. (2014). Exploring learner perception and use of task-based interactional feedback in ftf and cmc modes. *Studies in Second Language Acquisition*, *36*(1), 1–37. https://doi.org/10.1017/s0272263113000363

Halliday, M. A. K. (1985). *Spoken and written language*. Oxford University Press, USA.

Hartsuiker, R. J., & Kolk, H. H. J. (2001). Error Monitoring in Speech Production: A Computational Test of the Perceptual Loop Theory. *Cognitive Psychology*, *42*(2), 113–157. https://doi.org/10.1006/cogp.2000.0744

Hawkes, M. L. (2012). Using task repetition to direct learner attention and focus on form. *ELT Journal*, *66*(3), 327–336. https://doi.org/10.1093/elt/ccr059

Hayati, A., Jalilifar, A., & Mashhadi, A. (2013). Using Short Message Service (SMS) to teach English idioms to EFL students. *British Journal of Educational Technology*, *44*(1), 66–81. https://doi.org/10.1111/j.1467-8535.2011.01260.x

Healy, A. F. (1994). Letter detection: A window to unitization and other cognitive processes in reading text. *Psychonomic Bulletin & Review*, *1*(3), 333–344. https://doi.org/10.3758/bf03213975

REFERENCES

Hirotani, M., & Lyddon, P. A. (2013). The Development of L2 Japanese Self-Introductions in an Asynchronous Computer-Mediated Language Exchange. *Foreign Language Annals*, *46*(3), 469–490. https://doi.org/10.1111/flan.12044

Housen, A., & Kuiken, F. (2009). Complexity, Accuracy, and Fluency in Second Language Acquisition. *Applied Linguistics*, *30*(4), 461–473. https://doi.org/10.1093/applin/amp048

Housen, A., Kuiken, F., & Vedder, I. (2012). *Dimensions of L2 Performance and Proficiency: Complexity, Accuracy and Fluency in SLA. Language Learning & Language Teaching*. Volume 32. John Benjamins Publishing Company. https://eric.ed.gov/?id=ED539539

Hu, X. (2018). Effects of task type, task-type repetition, and performance criteria on L2 oral production. *Learning language through task repetition*, 143–169. https://doi.org/10.1075/tblt.11.06hu

Hunt, K. W. (1965). *Grammatical Structures Written at Three Grade Levels. NCTE Research Report No. 3.* Illinois, US: National Council of Teachers of English. http://eric.ed.gov/?id=ED113735

IELTS. (2017). *Speaking: Band Descriptors (public version) Band Fluency and coherence Lexical resource Grammatical range and accuracy Pronunciation*. https://www.ielts.org/-/media/pdfs/speaking-band-descriptors.ashx?la=en

Iimura, H., & Takanami, S. (2016). The Effects of Mimicking Practice on Reading Aloud and Speaking Performance. *ARELE: Annual Review of English Language Education in Japan*, *27*, 293–308. https://doi.org/10.20581/arele.27.0_293

JACET Committee of Basic Words Revision (2016). *The New JACET list of 8000 basic words: JACET 8000*. Tokyo, Japan: JACET

Jong, N., & Tillman, P. (2018). Grammatical structures and oral fluency in immediate task repetition: Trigrams across repeated performances. In M. Bygate (Eds), *Learning Language through Task Repetition* (Vol. 11, pp. 43–73). John Benjamins Publishing Company.

REFERENCES

Jonsson, E. (2016). Conversational Writing. In *www.peterlang.com*. PETER LANG. https://www.peterlang.com/document/1067683

Kemp, N., & Bushnell, C. (2011). Children's text messaging: abbreviations, input methods and links with literacy. *Journal of Computer Assisted Learning*, *27*(1), 18–27. https://doi.org/10.1111/j.1365-2729.2010.00400.x

Kietrys, D. M., Gerg, M. J., Dropkin, J., & Gold, J. E. (2015). Mobile input device type, texting style and screen size influence upper extremity and trapezius muscle activity, and cervical posture while texting. *Applied Ergonomics*, *50*, 98–104. https://doi.org/10.1016/j.apergo.2015.03.003

Kim, Y., & Tracy-Ventura, N. (2013). The role of task repetition in L2 performance development: What needs to be repeated during task-based interaction? *System*, *41*(3), 829–840. https://doi.org/10.1016/j.system.2013.08.005

Kitade, K. (2000). L2 learners' discourse and SLA theories in CMC: Collaborative interaction in Internet chat. *Computer Assisted Language Learning*, 13(2), 143–166. https://doi.org/10.1076/0958-8221(200004)13:2;1-d;ft143

Kojima M. (2013). Current trends and issues of lexical diversity studies in productive texts of English learners. *Bulletin of Gifu City Women's College*, 62, 29–38. https://irdb.nii.ac.jp/01405/0002511423

Kormos, J. (1999). Monitoring and Self-Repair in L2. *Language Learning*, *49*(2), 303–342. https://doi.org/10.1111/0023-8333.00090

Kormos, J. (2006). *Speech Production and Second Language Acquisition*. Routledge. https://doi.org/10.4324/9780203763964

Lai, C., & Zhao, Y. (2006). Noticing and text-based chat. *Language Learning and Technology*, *10*(3), 102–120. http://llt.msu.edu/vol10num3/laizhao/

Lambert, C., Aubrey, S., & Leeming, P. (2021). Task Preparation and Second Language Speech Production. *TESOL Quarterly*, *55*(2). https://doi.org/10.1002/tesq.598

REFERENCES

Lambert, C., Philp, J., & Nakamura, S. (2017). Learner-generated content and engagement in second language task performance. *Language Teaching Research*, *21*(6), 665–680. https://doi.org/10.1177/1362168816683559

Lee, B. J. (2021). The effects of proficiency and textual enhancement technique on noticing. *System*, *96*, 102407. https://doi.org/10.1016/j.system.2020.102407

Lee, M., & Révész, A. (2020). Promoting grammatical development through captions and textual enhancement in multimodal input-based tasks. *Studies in Second Language Acquisition*, 1–27. https://doi.org/10.1017/s0272263120000108

Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid Lexical Test for Advanced Learners of English. *Behavior Research Methods*, *44*(2), 325–343. https://doi.org/10.3758/s13428-011-0146-0

Lennon, P. (1990). Investigating Fluency in EFL: A Quantitative Approach*. *Language Learning*, *40*(3), 387–417. https://doi.org/10.1111/j.1467-1770.1990.tb00669.x

Levelt, W. J. (1989). *Speaking: From intention to articulation.* MIT press.

Lewis, M. (2009). *The idiom principle in L2 English: Assessing elusive formulaic sequences as indicators of idiomaticity, fluency, and proficiency* (Doctoral dissertation, Engelska institutionen). http://urn.kb.se/resolve?urn=urn:nbn:se:su:diva-7317 </div

Lin, H. (2014). Computer-mediated communication (CMC) in L2 oral proficiency development: A meta-analysis. *ReCALL*, *27*(3), 261–287. https://doi.org/10.1017/s095834401400041x

Littlemore, H. S., & Eyckmans, J. (2008). Cognitive Linguistic Approaches to Teaching Vocabulary and Phraseology. In F. Boers & S. Lindstromberg (Eds.), *Applications of Cognitive Linguistics [ACL]*. Mouton de Gruyter. https://doi.org/10.1515/9783110199161

Loewen, S., & Erlam, R. (2006). Corrective feedback in the chatroom: An experimental study. *Computer Assisted Language Learning*, *19*(1), 1–14. https://doi.org/10.1080/09588220600803311

REFERENCES

Long, M. (1981). Input, interaction and second-language acquisition ‘in H. Wintiz (eds.): *Native Language and Foreign Language Acquisition.* 379(1), 259–278..Annals of the New York Academy of Sciences.

Long, M. (2015). *Second language acquisition and task-based language teaching*. John Wiley & Sons.

Long, M. H. (2016). In Defense of Tasks and TBLT: Nonissues and Real Issues. *Annual Review of Applied Linguistics*, *36*, 5–33. https://doi.org/10.1017/s0267190515000057

Lyddy, F., Farina, F., Hanney, J., Farrell, L., & Kelly O’Neill, N. (2014). An Analysis of Language in University Students’ Text Messages. *Journal of Computer-Mediated Communication*, *19*(3), 546–561. https://doi.org/10.1111/jcc4.12045

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, *4*(3), 221–250. https://doi.org/10.1177/136216880000400303

MacWhinney, B. (2000). The CHILDES project: Tools for analyzing talk. *Child Language Teaching and Therapy*, *8*. https://doi.org/10.1177/026565909200800211

Matsumura. (2012). *Tasuku wo katsuyousita Eigo jyugyou dezain* [Tasks for English Classes]. Tokyo, Japan: Taishukan.

McCarthy, P. M., & Jarvis, S. (2010). MTLD, vocd-D, and HD-D: A validation study of sophisticated approaches to lexical diversity assessment. *Behavior Research Methods*, *42*(2), 381–392. https://doi.org/10.3758/brm.42.2.381

McSweeney, M. (2017). I Text English to Everyone: Links between Second-Language Texting and Academic Proficiency. *Languages*, *2*(3), 7. https://doi.org/10.3390/languages2030007

Mehrang, F., & Rahimpour, M. (2010). The impact of task structure and planning conditions on oral performance of EFL learners. *Procedia - Social and Behavioral Sciences*, *2*(2), 3678–3686. https://doi.org/10.1016/j.sbspro.2010.03.572

REFERENCES

Meredith, J., & Stokoe, E. (2014). Repair: Comparing Facebook "chat" with spoken interaction. *Discourse & Communication, 8*(2), 181–207. https://doi.org/10.1177/1750481313510815

Michel, M., & O'Rourke, B. (2019). What drives alignment during text chat with a peer vs. a tutor? Insights from cued interviews and eye-tracking. *System, 83,* 50–63. https://doi.org/10.1016/j.system.2019.02.009

Michel, M., & Smith, B. (2017). Measuring lexical alignment during L2 chat interaction. In S. M. Gass., P. Spinner., & J, Beheney (1st Eds), *Salience in Second Language Acquisition* (pp. 244-268). Routledge.

Ministry of Internal Affairs and Communications. (2017). *Information and Communications Japan*. Soumu.go.jp. https://www.soumu.go.jp/english/

Mizumoto, A. & Takeuchi, O. (2011). Introduction to Effect Size and Statistical Power Analysis: For Proper Use of Statistical Tests. *Proceedings of the 2010 Sectional Meeting of The Japan Association for Language Education and Technology, 2010,* 47–73. https://www.mizumot.com/method/mizumoto-takeuchi.pdf

Moore, P. J. (2018). Task-Based Language Teaching (TBLT). *The TESOL Encyclopedia of English Language Teaching*, 1–7. https://doi.org/10.1002/9781118784235.eelt0175

Nairne, J. S. (2002). The myth of the encoding-retrieval match. *Memory, 10*(5-6), 389–395. https://doi.org/10.1080/09658210244000216

Nakatani, Y. (2010). Identifying Strategies That Facilitate EFL Learners' Oral Communication: A Classroom Study Using Multiple Data Collection Procedures. *The Modern Language Journal*, *94*(1), 116–136. https://doi.org/10.1111/j.1540-4781.2009.00987.x

Nation, I. S. P., Heatley, A., & Coxhead, A. (2002). *Range: A program for the analysis of vocabulary in texts.* [Computer software]. Wellington, NZ: LALS, Victoria University of Wellington, New Zealand. Retrievd from http://www. victoria. ac. nz/lals/staff/paul-nation/nation. Aspx

REFERENCES

Navidinia, H., Mobarak, M., & Malekzadeh, F. (2019). Investigating the Effect of Noticing on EFL Students' Speaking Accuracy. *International Journal of Instruction*, *12*(1), 83–98. https://eric.ed.gov/?id=EJ1202216

Nazemi, M., & Rezvani, E. (2019). Effects of Task Familiarity and Task Repetition on Iranian EFL Learners' Engagement in L2 Oral Performance. *Contemporary Research in Education and English Language Teaching*, *1*(4), 45–56. https://doi.org/10.33094/26410230.2019.14.45.56

Norris, J. M., & Ortega, L. (2009). Towards an Organic Approach to Investigating CAF in Instructed SLA: The Case of Complexity. *Applied Linguistics*, *30*(4), 555–578. https://doi.org/10.1093/applin/amp044

Ortega, L. (2005). 3. What do learners plan?: Learner-driven attention to form during pre-task planning. *Planning and Task Performance in a Second Language*, 77–109. https://www.jbe-platform.com/content/books/9789027294647-lllt.11.07ort

Ortega, L. (2009). *Understanding Second Language Acquisition.* Hodder Education. Oxford, UK, 2(66), 136.

Ouellette, G., & Michaud, M. (2016). Generation text: Relations among undergraduates' use of text messaging, textese, and language and literacy skills. *Canadian Journal of Behavioural Science / Revue Canadienne Des Sciences Du Comportement*, *48*(3), 217–221. https://doi.org/10.1037/cbs0000046

Payne, J. S., & Whitney, P. J. (2002). Developing L2 Oral Proficiency through Synchronous CMC: Output, Working Memory, and Interlanguage Development. *CALICO Journal*, *20*(1), 7–32. https://www.jstor.org/stable/24149607

Peslak, A., & Hunsinger, D. S. (2018). The evolution of text messaging: An expanded review of influencing variables over time. *Issues in Information Systems*, 19(3), 10. https://doi.org/10.48009/3_iis_2018_61-70

Pilotti, M., Maxwell, K., & Chodorow, M. (2006). Does the Effect of Familiarity on Proofreading Change With Encoding Task and Time? *The Journal of General Psychology*, *133*(3), 287–299. https://doi.org/10.3200/genp.133.3.287-299

REFERENCES

Pinet, S., & Nozari, N. (2021). The role of visual feedback in detecting and correcting typing errors: A signal detection approach. *Journal of Memory and Language*, *117*, 104193. https://doi.org/10.1016/j.jml.2020.104193

Plonsky, L., & Oswald, F. L. (2014). How Big Is "Big"? Interpreting Effect Sizes in L2 Research. *Language Learning*, *64*(4), 878–912. https://doi.org/10.1111/lang.12079

Qiu, X., & Lo, Y. Y. (2017). Content familiarity, task repetition and Chinese EFL learners' engagement in second language use. *Language Teaching Research*, *21*(6), 681–698. https://doi.org/10.1177/1362168816684368

Rahimpour, M., & Safarie, M. (2011). The effects of on-line and pre-task planning on descriptive writing of Iranian EFL learners. *International Journal of English Linguistics*, 1(2), 274. https://doi:10.5539/ijel.v1n2p274

Ringbom, H., & Jarvis, S. (2009). The Importance of Cross-Linguistic Similarity in Foreign Language Learning. *The Handbook of Language Teaching*, 106–118. https://doi.org/10.1002/9781444315783.ch7

Rodrigues, H. L., Gallo, D. A., & Geraci, L. (2002). Processing approaches to cognition: The impetus from the levels-of-processing framework. *Memory*, *10*(5-6), 319–332. https://doi.org/10.1080/09658210224000144

Saeedi, M. (2020). Task Condition and L2 oral performance: Investigating the combined effects of online Planning and immediacy. *International Journal of Foreign Language Teaching and Research*, 8(32), 35-48.

Saito, K., Sun, H., & Tierney, A. (2019). Explicit and implicit aptitude effects on second language speech learning: Scrutinizing segmental and suprasegmental sensitivity and performance via behavioural and neurophysiological measures. *Bilingualism: Language and Cognition*, *22*(5), 1123–1140. https://doi.org/10.1017/s1366728918000895

Salomone, A. M., & Marsal, F. (1997). How to Avoid Language Breakdown? Circumlocution! *Foreign Language Annals*, *30*(4), 473–484. https://doi.org/10.1111/j.1944-9720.1997.tb00854.x

REFERENCES

Sauro, S. (2012). L2 performance in text-chat and spoken discourse. *System*, *40*(3), 335–348. https://doi.org/10.1016/j.system.2012.08.001

Sauro, S., & Smith, B. (2010). Investigating L2 Performance in Text Chat. *Applied Linguistics*, *31*(4), 554–577. https://doi.org/10.1093/applin/amq007

Scarborough, H. S. (1990). Very Early Language Deficits in Dyslexic Children. *Child Development*, *61*(6), 1728–1743. https://doi.org/10.2307/1130834

Segalowitz, N. (2010). *The cognitive bases of second language fluency*. Routledge.

Segalowitz, N. S., & Segalowitz, S. J. (1993). Skilled performance, practice, and the differentiation of speed-up from automatization effects: Evidence from second language word recognition. *Applied Psycholinguistics*, *14*(3), 369–385. https://doi.org/10.1017/s0142716400010845

Shintani, N. (2015). The effects of computer-mediated synchronous and asynchronous direct corrective feedback on writing: a case study. *Computer Assisted Language Learning*, *29*(3), 517–538. https://doi.org/10.1080/09588221.2014.993400

Shortis, T. (2007). Gr8 txtpectations. *English Drama Media*, 21-26. http://nl.ijs.si/janes/wp-content/uploads/2014/09/shortis07.pdf

Skehan, P. (1998). *A cognitive approach to language learning*. Oxford University Press.

Skehan, P. (2014). *Processing Perspectives on Task Performance* (Vol. 5). John Benjamins Publishing Company.

Skehan, P., & Foster, P. (2007). Complexity, accuracy, fluency and lexis in task-based performance: A meta-analysis of the Ealing research. *Complexity, accuracy, and fluency in second language use, learning, and teaching*, 207-226.

Smith, B. (2003). The use of communication strategies in computer-mediated communication. *System*, *31*(1), 29–53. https://doi.org/10.1016/s0346-251x(02)00072-6

REFERENCES

Smith, B. (2008). Methodological hurdles in capturing CMC data: The case of the missing self-repair. *Language Learning & Technology*, 12(1), 85-103. http://llt.msu.edu/vol12num1/smith/

Smith, B. (2012). Eye tracking as a measure of noticing: a study of explicit recasts in scmc. *Language Learning and Technology*, *16*, 53–81. http://llt.msu.edu/issues/october2012/smith.pdf

Smith, B., & Renaud, C. (2013). Using eye tracking as a measure of foreign language learners' noticing of recasts during computer-mediated writing conferences. In K. McDonough & A. Mackey (Eds), *Second Language Interaction in Diverse Educational Contexts* (pp. 147–166). John Benjamins Publishing Company.

Statistic Brain Research Institute. (2014). *Text message statistics*. Statistic Brain. https://www.statisticbrain.com/?s=text+message+statistics&post_type=post

Stengers, H., Deconinck, J., Boers, F., & Eyckmans, J. (2016). Does copying idioms promote their recall? *Computer Assisted Language Learning*, *29*(2), 289–301. https://doi.org/10.1080/09588221.2014.937723

Suzuki, S., & Kormos, J. (2019). Linguistic dimensions of comprehensibility and perceived fluency: an investigation of complexity, accuracy, and fluency in second language argumentative speech. *Studies in Second Language Acquisition*, *42*(1), 1–25. https://doi.org/10.1017/s0272263119000421

Swain, M., & Lapkin, S. (1995). Problems in Output and the Cognitive Processes They Generate: A Step Towards Second Language Learning. *Applied Linguistics*, *16*(3), 371–391. https://doi.org/10.1093/applin/16.3.371

Tajgozari, M. (2019). The effect of watching captioned TV series on speaking accuracy of Iranian advanced EFL learners. *Global Journal of Foreign Language Teaching*, *9*(3), 157–166. https://doi.org/10.18844/gjflt.v9i3.4221

Takase, N. (2014). Accuracy of Output across Online Text-Based Chat rooms and Face-to-Face Tasks on Foreign Language Learning. *LET Journal of Central Japan, Vol26,* p.53-62.

REFERENCES

Takase, N. (2019). Developing Speaking Fluency through Tasks in Chat Rooms. *PanSIG Journal*, *2018*, 250–257.

Takase, N. (2020). Learning Formulaic Sequences Through Text Messaging Tasks. *PanSIG Journal*, *2019*, 331–338.

Tavakoli & Skehan. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Eds), *Planning and task performance in a second language* (pp. 239–277). John Benjamins Publishing.

Thai, C., & Boers, F. (2016). Repeating a Monologue Under Increasing Time Pressure: Effects on Fluency, Complexity, and Accuracy. *TESOL Quarterly*, *50*(2), 369–393. https://doi.org/10.1002/tesq.232

Thurlow, C., & Brown, A. (2003). Generation Txt? The sociolinguistics of young people's text-messaging. *Discourse analysis online*, 1(1), 30. http://www.shu.ac.uk/daol/articles/v1/n1/a3/thurlow2002003-paper.html

Toutanova, K., Klein, D., Manning, C.D., & Singer., Y. (2003). Feature-Rich Part-of-Speech Tagging with a Cyclic Dependency Network. In *Proceedings of the 2003 Human Language Technology Conference of the North American Chapter of the Association for Computational Linguistics*, 252–259. https://aclanthology.org/N03-1033.pdf

Towell, R., Hawkins, R., & Bazergui, N. (1996). The Development of Fluency in Advanced Learners of French. *Applied Linguistics*, *17*(1), 84–119. https://doi.org/10.1093/applin/17.1.84

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, *80*(5), 352–373. https://doi.org/10.1037/h0020071

Tzelgov, J., Henik, A., Sneg, R., & Baruch, O. (1996). Unintentional word reading via the phonological route: The Stroop effect with cross-script homophones. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *22*(2), 336–349. https://doi.org/10.1037/0278-7393.22.2.336

REFERENCES

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it. In Norbert, S (Eds), *Formulaic Sequences: Acquisition, Processing, and Use*, 9, 153.

Wang, Z. (2014). On-line time pressure manipulations. *Processing perspectives on task performance*, 5, 27-61.

Wolfe-Quintero, K., Inagaki, S., & Kim, H. Y. (1998). *Second language development in writing: Measures of fluency, accuracy, & complexity* (No. 17). University of Hawaii Press.

Wolters, C. A., & Taylor, D. J. (2012). A Self-regulated Learning Perspective on Student Engagement. *Handbook of Research on Student Engagement*, 635–651. https://doi.org/10.1007/978-1-4614-2018-7_30

Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence and classroom applications.* A&C Black.

Yilmaz, Y., & Granena, G. (2010). The effects of task type in Synchronous Computer-Mediated Communication. *ReCALL*, *22*, 20–38. https://doi.org/10.1017/S0958344009990176

Yuan, F. (2003). The Effects of Pre-Task Planning and On-Line Planning on Fluency, Complexity and Accuracy in L2 Monologic Oral Production. *Applied Linguistics*, *24*(1), 1–27. https://doi.org/10.1093/applin/24.1.1

Yus, F. (2011). *Cyberpragmatics: Internet-mediated Communication In Context*. John Benjamins Publishing Company. https://doi.org/10.1075/pbns.213

Zhang, H., Song, W., & Burston, J. (2011). Reexamining the Effectiveness of Vocabulary Learning via Mobile Phones. *Turkish Online Journal of Educational Technology - TOJET*, *10*(3), 203–214. https://eric.ed.gov/?id=EJ944968

Ziegler, N. (2016). Taking Technology to Task: Technology-Mediated TBLT, Performance, and Production. *Annual Review of Applied Linguistics*, *36*, 136–163. https://doi.org/10.1017/s0267190516000039

Ziegler, N. (2018). Pre-task planning in L2 text-chat: Examining learners' process and performance. *Language Learning and Technology*, *22*(3), 193–2013.

https://scholarspace.manoa.hawaii.edu/items/f64da4ec-360a-4134-8e11-ad9250e6ea94

Zuniga, M., & Simard, D. (2019). Factors Influencing L2 Self-repair Behavior: The Role of L2 Proficiency, Attentional Control and L1 Self-repair Behavior. *Journal of Psycholinguistic Research*, *48*(1), 43–59. https://doi.org/10.1007/s10936-018-9587-2