LANGUAGE IN MULTIMODAL WRITING PROCESSES AND PERFORMANCE: DEVELOPING MULTIMODAL WRITING TASKS FOR L2 LEARNERS

By

Jung Min Lim

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Second Language Studies—Doctor of Philosophy

2020

ABSTRACT

LANGUAGE IN MULTIMODAL WRITING PROCESSES AND PERFORMANCE: DEVELOPING MULTIMODAL WRITING TASKS FOR L2 LEARNERS

By

Jung Min Lim

In this sequential mixed methods research project, I first investigated learners' needs of multimodal writing in the higher education setting and then examined adult L2 writers' multimodal writing performances and processes, and their task perceptions. For the first study, a qualitative needs analysis, I conducted individual semi-structured interviews with 7 instructors of undergraduate courses in different disciplines to explore how they implement and perceive multimodal course assignments. Additionally, I collected 161 course syllabi data from which I identified 104 multimodal tasks. Triangulating two data sources, I found three themes that emerged from the two data sources: (1) goals and instruction of multimodal writing: disciplinary versus creative expression; (2) linguistic mode in multimodal texts: written and spoken words; and (3) tasks of multimodal writing: individual versus collaborative work. Based on the needs analysis, I developed a timed multimodal writing task that is to construct a narrated slide presentation. I utilized this task as one of the instruments in the subsequent phase.

In the second study, adopting a convergent parallel mixed methods design, I investigated L2 learners' multimodal writing performances and processes and their perceptions toward the multimodal writing task. Thirty-one adult Korean learners of English intermediate- to high-proficiency individually completed a multimodal writing task (i.e., a timed argumentative narrated presentation task) and a monomodal writing task (i.e., a timed argumentative essay task) while their on-screen writing behaviors were screen recorded. After the multimodal writing task, each writer completed a stimulated recall interview in their first language (Korean) watching a

video of one's own writing processes. They all completed task perception and background questionnaires. Writing process data—on-screen writing processes and stimulated recalls— were qualitatively analyzed using an inductive approach. In terms of task performances, five experienced academic English instructors evaluated the multimodal writing performances in terms of the overall quality, visualization quality, and language; and three of them also rated the monomodal performances using an analytic rubric.

Findings from the performance data revealed that multimodal text quality is strongly associated with language performance, but another dimension of nonlinguistic performance also contributes to the overall text quality. More specifically, the multimodal performance data fit a regression model that explains 83% of the variance of multimodal text quality with language scores ($ß$ = .62) and visualization scores ($ß$ = .45). Furthermore, the language scores of participants' multimodal writing performance showed significant positive correlations to all subscores and total score of monomodal writing performances with moderate to strong effect sizes; however, none of the scores of monomodal writing task performances were correlated with the visualization score of the multimodal writing task performances.

From the writing process data, I found that L2 writers spent a smaller amount of time and effort on constructing visual texts than language, especially in the middle of the text construction processes. When focusing on language, they spent considerable time on selecting and upgrading words for scripts and evaluating information they found from the Internet and their own texts-constructed-so-far. Results from the task perception data showed that the multimodal task was was perceived to be more complex, difficult and interesting than the monomodal task. I discuss implications for L2 writing research and pedagogy focusing on how to understand multimodal tasks as language tasks for learners whose goal is to improve language.

To my parents

ACKNOWLEDGMENTS

This project would not have been possible without the support of many amazing people around me. First, I would like to express my sincere gratitude to Dr. Charlene Polio for her supportive supervision and encouragement, which was a multimodal ensemble of verbal, nonverbal, and culinary modes. She is the best mentor one can ask for, and her great passion for research is a legacy that I wish to keep for years to come.

I am also deeply grateful to my dissertation committee members. Dr. Peter De Costa has always offered me valuable suggestions, allowing me to employ research methods and frameworks appropriate for qualitative and mixed-methods research. Thanks to Dr. Paula Winke, I could build a greater understanding of what it means to write with a professional tone, as well as the ability to design a well-structured study. Dr. Koen Van Gorp has endeavored to equip me with the knowledge and skills needed for the development of pedagogically meaningful tasks, which served as a standard of my multimodal task building. I firmly believe their guidance was essential for the completion of my project.

I was fortunate to have wonderful colleagues in the Second Language Studies program. We have shared a lot of memorable events at many different places. Just to name a few, we have spent much time working hard in the office, trying different cocktails and beers through cocktail nights and brewery tours, and supporting each other's presentations at local and international conferences. We have successfully completed several collaborative research projects over coffee and donuts. Through these unforgettable, joyful experiences, I learned that academic and leisure hours can overlap, and will miss this special crowd full of passion for second language research.

Most importantly, I would like to express my thanks to my family. My parents and sister have always shown their unwavering love and trust in me. Hyung-Jo, who is an older academic child of Dr. Charlene Polio, has been the most supportive husband with a great sense of humor and wisdom that he inherited from my parents-in-law. All the love and trust they have shared with me made this journey more meaningful and grateful.

Finally, I would like to thank the instructors and second language learners who participated in my study. The conversation I had with them was a great opportunity for me to learn from their unique insights, leading me to uncover new implications. This project was funded by the National Federation of Modern Language Teachers' Association, the TESOL Association, and the College of Arts and Letters at Michigan State University.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1.

INTRODUCTION

Researchers have recently been debating the inclusion of multimodal composition in second language (L2) classrooms (e.g., see the dialogue and a special issue of the *Journal of Second Language Writing* in 2017 and in 2020; the special issue of *TESOL Quarterly* in 2015), but little is known about how to construct multimodal writing tasks for language development (Polio, 2019). In the dialogue of the *Journal of Second Language Writing* (2017), for example, some researchers acknowledged the multimodal nature of contemporary writing practice and underscored the importance of multimodal composition for L2 writing (Belcher, 2017; Warschauer, 2017; Yi, 2017). At the same time, others have expressed considerable concerns about the integration of multimodal writing into the L2 writing classroom (e.g., Manchón, 2017; Qu, 2017) based on the following assumptions: (1) multimodal writing is not academic and (2) a focus on multimodal writing results in less language learning. These premises, however, merit empirical research because there is no evidence from L2 research that supports such claims.

According to previous L1 research, these assumptions about multimodal writing may not be accurate. In terms of the first concern that multimodal texts are not academic, if we follow the definition of genre by Swales (1990, p. 58) that "a genre comprises a class of communicative events, the members of which share some set of communicative purposes", multimodal texts such as academic posters and presentation slides can be counted as academic genres. Given this extended notion of academic genres, researchers have conducted genre analyses on the multimodal texts (e.g., academic posters in D'Angelo, 2010, 2016; presentation slides in Rowley-Jolivet, 2002, 2012) and revealed patterns of nonlinguistic texts for particular genres and disciplines. These multimodal genres can be challenging to novice writers because linguistic

patterns in these genres are different from other genre conventions. According to Rowley-Jolivet (2012), for example, texts on presentation slides, which is one of the academic multimodal genres, demonstrated higher lexical diversity, more frequent nominalization, and fewer pronouns than transcribed verbal presentations. These differential linguistic characteristics of multimodal texts can be related to the fact that some meanings can be conveyed better in one mode than others (Bezemer & Kress, 2008; Jewitt, 2008; Kress & Van Leeuwen, 1996; Unsworth, 2007); for example, images can show spatial relations better than words whereas language is more powerful in making categorical distinctions (Unsworth, 2007). Because genre-specific lexicogrammatical features of academic texts are not directly transferrable to the multimodal academic texts, it is necessary to identify multimodal writing tasks that students need and how linguistic and visual resources are interrelated in the multimodal texts (Early, Kendrick, & Potts, 2015).

The second concern regarding language development is perhaps more central to the concerns and based on an assumption that language may no longer be the goal of language classes if multimodal writing tasks are introduced in the course content. This assumption is due to the fact that researchers who proposed the multimodal turn in L2 writing advocated the strong version of multimodality that emphasizes the importance of teaching nonlinguistic modes as equivalent tools of communication as language (Grapin, 2019; Kress, 2000; Van Leeuwen, 2015). These researchers have grounded their proposals in the findings of previous L1 research that was conducted in K-12 classrooms (Dalton, 2012; Edwards-Groves, 2011; Howell, Butler, & Reinking, 2017; Smith, Pacheco, & de Almeida, 2017; Unsworth, 2006) and first-year or creative writing classes (e.g., Archer, 2010; Depalma & Alexander, 2015; Fraiberg, 2010; Vankooten & Berkley, 2016) where the learning goal is to advance literacy skills while students use their first

language. In these settings, the strong version could be persuasive; however, this position lacks the consideration of the L2 speaker's more limited linguistic knowledge and adult L2 writers in this context articulate language development as a primary goal (e.g., Zhou, Busch, & Cumming, 2014). Many L2 students have experience using nonlinguistic modes such as visual modes (e.g., visualizing idea in a flowchart; making graphs for comparison; and using typographical cues for emphasis) but may lack linguistic knowledge that can be retrieved and produced in real time. The lack of contextualization of ideas from L1 research may have led to L2 researchers' increased concern that students cannot achieve language development though multimodal writing practice. Nevertheless, as Grapin (2019) noted, a weak version of multimodality sets language as the main goal of learning with other modes are regarded as compensatory, which may be appealing to L2 instructors for its emphasis on language.

Even though the weak version of multimodal writing may be adopted, how a curriculum including multimodal tasks can facilitate writing development is questioned. According to Polio (2017), writing development can be defined as change over time in a wide range of written text (e.g., linguistic features, genre knowledge, writing process, and strategy use); and the target of progress could differ according to learning purposes and contexts. Given that L2 learners at some point need to produce multimodal genres (Chun, Smith, & Kern, 2016; Elola & Oskoz, 2017), the definition of writing skill can be expanded to include writing for multimodal genres and using linguistic and nonlinguistic knowledge for these genres. This change in the operationalization of writing subsequently changes the target of progress; learning to write a multimodal text can be a part of the learning goal for EAP writing classes. What is more relevant to language learners, however, will remain in the linguistic development. Thus, understanding to

what extent and how language plays a role in multimodal texts is crucial for further discussion of multimodal writing tasks in the context of L2 writing instruction.

Furthermore, multimodal writing tasks can help the processes of writing alphabetic texts according to compositionists' theory. Flower and Hayes (1984) explained that a mental writing plan is intrinsically multimodal, and this plan is later translated into an alphabetical text. They claimed that the distance between the mode of the writing plan and the mode of production contributes to writing difficulty. Contextualizing this original claim into contemporary writing, Palmeri (2012) proposed that composition teachers design planning activities to align the mode of mental representation and the mode of production; these activities eliciting linguistic and/or nonlinguistic output can ease the process of translating a multimodal writing plan into a prose. For example, if a writing task is to describe a past experience, a writer may have visual, spatial, and olfactory images; in this case, the writer may better benefit from a writing activity that help visual shaping than a traditional outlining activity. Palmeri suggested that providing activities to think and write multimodal to writers is not only helpful for rhetorical development but also timely because writing has become multimodal itself. This cognitive account for multimodal writing, however, has not been discussed in L2 writing research.

There have been only a few studies that explore the issue of multimodality in L2 writing to date. L2 research on multimodal writing has primarily focused on a descriptive analysis of what writers do when they are asked to create a product that includes language and other modes (e.g., Cimasko & Shin, 2017; Smith et al., 2017). Two studies observed the instructional effect of using multimodal writing tasks on language gains (e.g., Dzekoe, 2017; Vandommele et al., 2017). Although the research was noteworthy, in these studies the authors did not offer sufficient justification as to why they used particular types of multimodal writing tasks. In many cases, a

multimodal project was included in a writing class (e.g., first year writing for undergraduate students and EAP writing course for nonnative speakers of English) where researchers identified the course goal as the exploration of various academic genres However, some researchers examined multimodal writing tasks such as reproduction of persuasive writing into a short movie, which may not be a common academic genre and may not raise students' awareness of multimodal academic genres. To better integrate multimodal writing into an existing syllabus, researchers must first examine what types of multimodal writing are targeted in academic contexts.

I have thus far discussed a controversy over the value of multimodal L2 writing pedagogy, which is attributable to researchers' concerns that L2 writers cannot practice academic writing with multimodal tasks. This issue might be able to be resolved to some extent by addressing the following questions: How useful are multimodal writing tasks to L2 students? How can educators design multimodal writing tasks that pertain to L2 course language objectives? How can applied linguists design multimodal writing tasks that facilitate language development? To answer these questions, in my dissertation study, I conducted a needs analysis that offers implications for curriculum and multimodal writing task development (Study 1).

In Study 2, based on the needs and the compositionists' account of the multimodal processes of writing, I designed a pedagogical multimodal writing task and investigated the relationships between students' performances and processes for a traditional writing task and the multimodal writing task. In addition, I examined how students perceived the difficulty and complexity of the multimodal writing task compared to a traditional writing task. Integrating the two sequential studies, I present empirical evidence as to how (ir)relevant a multimodal writing

task is to language tasks and offer insights into how to incorporate the new aspect of writing for greater pedagogic values.

CHAPTER 2.

REVIEW OF THE LITERATURE

In this chapter, I first briefly review different approaches that have explained the mechanisms of multimodal writing: the cognitive process of writing (e.g., Leijten, Van Waes, Schriver, & Hayes, 2013), social semiotics (e.g., Bezemer & Kress, 2008; Cimasko & Shin, 2017; Lemke, 1998; Pacheco & Smith, 2015), systemic functional linguistics (e.g., Alyousef, 2016; Anderson, 2008; Daly & Unsworth, 2011; Hagan, 2007; O'Halloran, 2004), and genre studies (e.g., Archer, 2010; D'Angelo, 2016; Rowley-Jolivet, 2002, 2012). This review begins with the theoretical background to L1 multimodal writing research. Next, I review previous empirical studies that have explored L2 learners' multimodal writing and identify research gaps.

In this dissertation, the term *multimodal writing* indicates writers' use of nonlinguistic resources along with written words to achieve a goal of constructing messages as opposed to *multimodal communication* or *multimodal literacy* that does not require linguistic mode of communication for meaning construction (e.g., dance performance, music, visual arts). Writers' *multimodal texts* refer to the outcome of their cognitive processes of multimodal writing that incorporate writing and design schemas for a given task. I limit the scope of the inquiry to *multimodal writing* and *multimodal texts* given the purpose of the current dissertation project that seeks ways to understand and develop multimodal writing tasks for language learners.

Theoretical Background

Writing is inherently multimodal. It combines written words as well as nonlinguistic texts constructed in other modes such as figures, tables and typefaces. For example, an APA style paper, one of the most traditional academic styles, involves many visual choices in making tables and figures comprehensible and using boldface and italicized typefaces to indicate different

levels of headings. As writing on computer for a greater audience has become common, the multimodality of writing has been expanding to incorporate videos, sounds and social networks. However, because creating formal prose has been discussed as the dominant mode of writing in previous research, little has been discussed in L2 research regarding how other resources in nonlinguistic modes contribute to meaning construction.

Most of the multimodal writing research to date has been conducted using functional approaches to language, for example, systemic functional linguistics and social semiotics (Halliday, 1978, 1985; Kress & Van Leeuwen, 1996; Van Leeuwen, 2005). In these approaches, each of the modes within the multimodal texts has distinct contributions to meaning making. Important questions in this line of research are why the writer particularly chose one option to another and how readers would interpret the writer's choice of resources; for example, what would be the intention behind using an arrow as a bullet point instead of other symbols? Would readers interpret that as a causal relationship or a simple listing? Given this focus, researchers sought to investigate the underexplored aspects of authentic multimodal texts such as writers' purposeful choices of different modes in particular forms (e.g., Archer, 2006; Bezemer & Kress, 2008; Liu & O'Halloran, 2009; Pacheco & Smith, 2015; Smith & Dalton, 2016; Unsworth, 2006, 2007). In other words, research on the multimodal text from the social semiotics and systemic functional linguistics has provided possible interpretation of why writers would use and combine linguistic and nonlinguistic resources in their texts.

Another line of research has been centered in genre analysis. In multimodal genre studies, similar to textual and discourse analyses of earlier genre research, researchers focused on outlining regular semiotic choices, or patterns, in a discourse community (Bateman, 2008; D'Angelo, 2010, 2016; Tardy, 2005). Researchers focused on different aspects of writing, from

the lexicon to metadiscourse; at the same time, many studies utilized the notion of multimodality that is defined from the perspectives of social semiotics and systemic functional linguistics.

While most of the studies on multimodal writing have adopted either social semiotics or systemic functional linguistics approaches, some compositionists explained multimodal writing as writers' cognitive processes (Flower & Hayes, 1984; Hayes & Flower, 1980; Leijten et al., 2013; Palmeri, 2012). I revisit the original ideas in writing process research that have explained the translation of a multimodal writing plan to prose (Linda Flower & Hayes, 1984; Hayes & Flower, 1980) and introduce the current cognitive model of mulitmodal writing (Leijten et al., 2013; Palmeri, 2012). I begin my review with the cognitive accounts of multimodal writing that has received the least attention despite its relevance to the current discussion of the cognitive task-based language teaching.

The cognitive process of writing. The lack of empirical research on the cognitive processes of building a multimodal text might have arisen from the assumption that the ultimate outcome of a composing behavior is formal prose (Flower & Hayes, 1980, 1984; Hayes & Flower, 1980). This strong assumption may have circumvented researchers from exploring the processes of multimodal writing in a way that aligns with how compositionists have analyzed the production of alphabetic texts. However, in the earlier papers from cognitivists' and compositionists' perspectives, for example, the work of Hayes and Flower, there has been a discussion on the multimodal representation of meanings, which is highly relevant to the current issue of multimodal writing. Building on their own seminal model of the cognitive model of writing, Flower and Hayes (1984) proposed the Multiple Representation Thesis, with which they attempted to illustrate the ways writers compose a formal prose text from thoughts, or meanings, stored and accessible in multimodal forms. Their argument was summarized as follows (p. 122):

> As writers compose they create multiple internal and external representations of meaning. Some of these representations, such as an imagistic one, will be better at expressing certain kinds of meaning than prose would be, and some will be more difficult to translate into prose than others. Much of the work of writing is the creation and the translation of these alternative mental representations of meaning.

In this excerpt, meaning indicates the current writing plan that a writer has been creating, considering the writing purpose and ideas, in their working memory. The types of mental representations include automatic and conscious *procedural knowledge*, *non-verbal imagery* (e.g., auditory, kinesthetic, and visual representations), *declarative knowledge* (e.g., semantic representations, gists, episodic representations), and *verbal images* (e.g., keywords and chunks). They explain that a writer activates different mental representations in optimal modes, and a writing plan, which is a composite of information in multiple modes, is later translated into language by the mental translator. A novel argument in this thesis is that the mode(s) of mental representation is critical to the difficulty of writing a formal essay. More specifically, Flower and Hayes proposed a schematic representation of the distance between the modes of a writing plan in a writer's mind and formal prose in written words. Because a writer translates a writing plan composed of different shapes to a formal alphabetic prose, the writer needs to consider linguistic choices and prose constraints (see Figure 1 retrieved from Flower and Hayes, 1984). For example, non-verbal imagery includes auditory and kinesthetic information that is more challenging to put on words than explaining an abstract concept. Given that in the 1970s and 1980s most compositionists focused on the translation from writing plan to language, this pluralistic approach toward a multimodal writing plan can be contextualized into the current multimodal writing and provide meaningful suggestions.

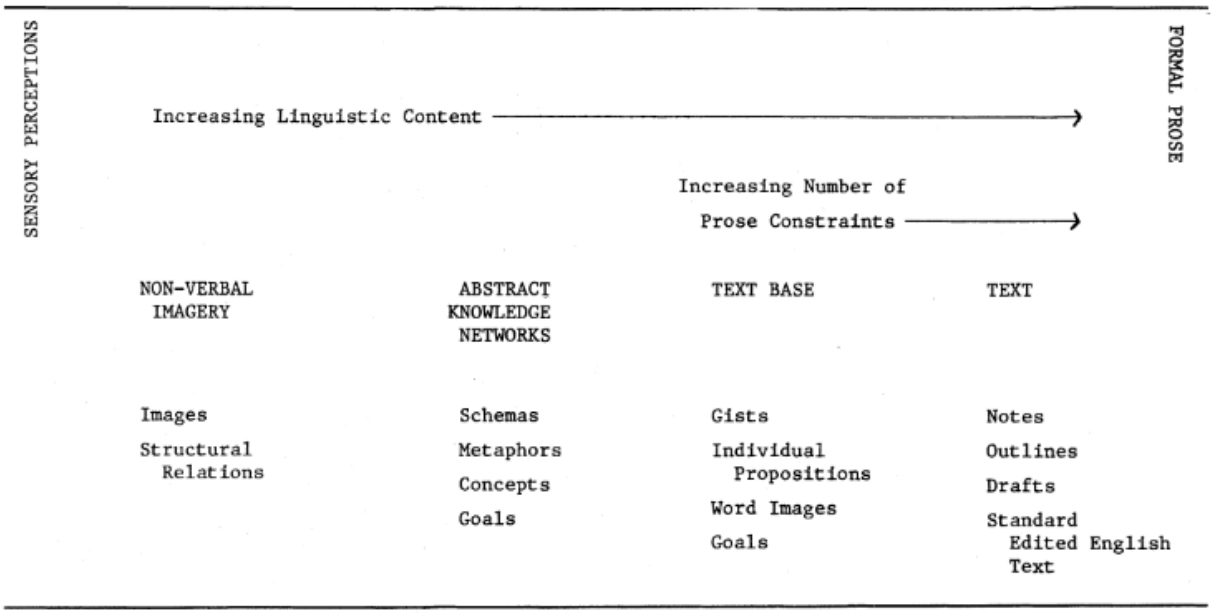

Figure 1. Some alternative representations of meaning generated during planning (Flower & Hayes, 1984: 131, Copyright © 1984, © SAGE Publications)

The Multiple Representation Thesis merits further attention for the current discussion of multimodal writing. Theoretically, this thesis can be used to explain the human mind during the translation of mental representations to multimodal texts (Palmeri, 2012). If some meaning is best represented in the visual mode in the human mind a multimodal text including a visual presentation along with explanation in the linguistic mode can be more effective than a monomodal text of written words. Coupled with a task-based approach, this thesis can inform educators how to sequence writing tasks; the (non)alignment of a writing plan in mind and a written outcome can be used as a scale for sequencing. For example, if a prompt elicits the cognitive process of making a writing plan that is likely to be realized in a visual mode, tasks can be sequenced from the visualization of an idea and proceed to the description of the visuals that learners produced. In this way, linguistic elements and prose constraints can be introduced later in one's own writing processes.

In addition to the Multiple Representation Thesis, Palmeri (2012) highlighted the *problem finding* in Flower and Hayes' (1980) *The Cognition of Discovery*, which means a writer's act of identifying a problem to be solved, as a potential framework to explain multimodal writing. Flower and Hayes originally explained that the ability to find and formulate a problem is a critical component to the general creativity that is demanded in cognitive activities in different academic disciplines. Interpreting this notion of problem finding in the context of multimodal writing, Palmeri argued that the skills for problem finding are transferrable from one mode to another mode if students are taught metalanguage that is common across modes (e.g., words and images). His proposal was to work with other disciplines where problem finding is critical to students. This proposal that is grounded in L1 research, however, warrants a careful examination before being implemented in the L2 classroom because it directly challenges extensive research suggesting that linguistic activities are governed by a language-specific domain in the human mind. For adult L2 writers, if the skills for problem findings is transferrable as Palmeri hypothesized, it is questionable whether L2 instruction needs to focus on these skills. Students may bring their rhetorical problem finding and solving skills into the class while they have difficulties in translating these mental representations into visible texts in their second language. In other words, adult L2 writers' knowledge in nonlinguistic modes might be mature and full-fledged but their language to express their ideas is limited.

Despite these earlier and recent proposals relevant to multimodal writing, there is only one empirical study that investigated multimodal writing from the perspective of the cognitive writing process. Extending Hayes' (2012) earlier cognitive model of the writing process, Leijten et al. (2013) proposed a revised model that reflects the mental steps a professional communication designer demonstrated during authentic proposal writing. Adopting an

ethnographic method, they observed a focal participant who had expertise in the multimodal writing task (i.e., business proposal writing). They collected interview data and keystroke logs from the beginning to the end of proposal writing which spanned eight and a half hours in total. The participant worked on the proposal over five sessions, of which the first was the longest and producing the largest amount of output. He began by inserting headings and notes to be used in later writing, which they counted as *writing schema*, on a template, which is one of the *task-related-sources*. He tended to recycle chunks from previous proposals, which are also considered as task-related-sources. In the third session, he looked for the optimal visualization for the project; in session four, he used Excel software to create visuals for the budget section in the proposal. The last session was primarily to review and revise the proposal more coherent and consistent.

Based on their case study, the authors updated the Hayes' model of writing process at three levels—control, process, and resource; what is new in this model are as follows (see Figure 2). First, the new model added design schemas at the control level. For the visualization of the proposal, a writer sought for physical and mental representations of fitting visuals to a purpose. Even though they included this schema in the theoretical model, and further acknowledged the importance of visual designing process, no separate processor for the visuals was added due to the lack of their understanding about this new component. Second, at the process level, they kept the four systems that Hayes (2012) elaborated where a *proposer* generates ideas in non-verbal forms; a *translator* recodes the non-verbal ideas into language; a *transcriber* recodes these verbal ideas into written texts; and an *evaluator* oversees the writing process. A *searcher*, the new addition, is a processor that "looks for information in external sources as one of the basic writing processes (p. 325)", which is widely used in different writing genres. Another notable change is

on the task environment at the resource level. They renamed task components to fit multimodal writing: *text-created-so-far* to *text-and-graphics-created-so-far*, *task materials* to *task-related sources*, and *transcribing technology* to *production technology* (find these components in Figure 2 from Leijten et al., 2013). Thirty years after the Multiple Representation Thesis (Flower & Hayes, 1984) which held a strong assumption that writing is equivalent to a creation of formal prose, this case study shows somewhat expanded definition of writing. It shows that real life writing tasks, which require a writer to construct a multimodal text that integrates their writing and design schemas with one's own idea, warrants further exploration, especially with regard to the writer's cognitive multimodal writing processes.

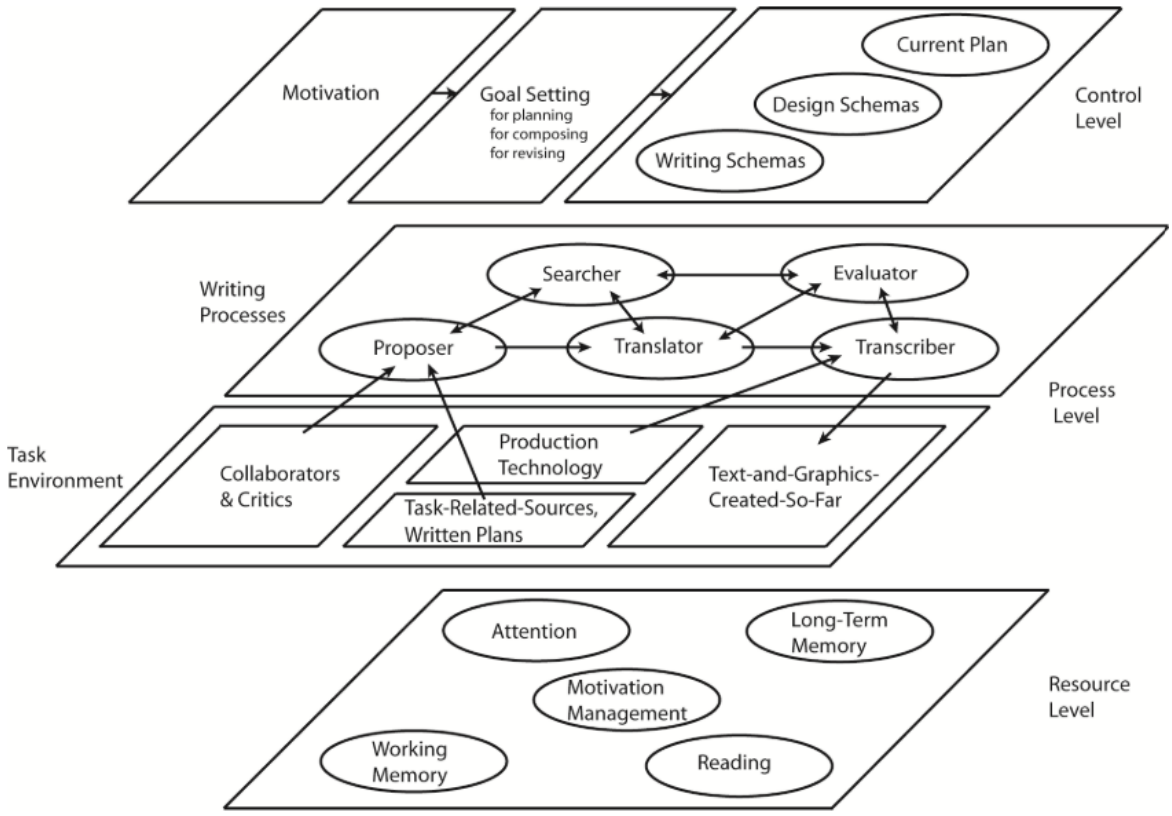

Figure 2. Model of composing elaborated to encompass activities of skilled professional communicators (Leijten et al., 2013: 324, licensed under Creative Commons Attribution-Noncommercial-No Derivative Works 3.0)

Considering these changes in the definition of writing process, I suggest an alternative schematic representation of the Multiple Representation Thesis that updates the original model proposed in 1984 and reflects the new writing model in Leijten et al.'s study (2013) in Figure 3. What is changed from the original schematic representation is the addition of *multimodal texts* to indicate that formal prose is not necessarily the only mode of written outcome. A multimodal text embraces visual elements and written words and each mode of representation can align with multimodal writing plans. In addition, writing plans for multimodal writing can have timestamps different from the ones for essay writing. In the original thesis, writing plans were discussed to "make it easy to mix images, sounds, and schemas in the same pot, and they allow the writer to delay decisions that are better made later in the writing process" (Flower & Hayes, 1984: 145); however, for multimodal texts, writing plans may not have to be hold to the final prose writing but can be written in multiple stages and modes as shown in Leijten, et al. (2013). Not every non-text plan is recoded into a text form but transcribed to the best option for a writing goal.

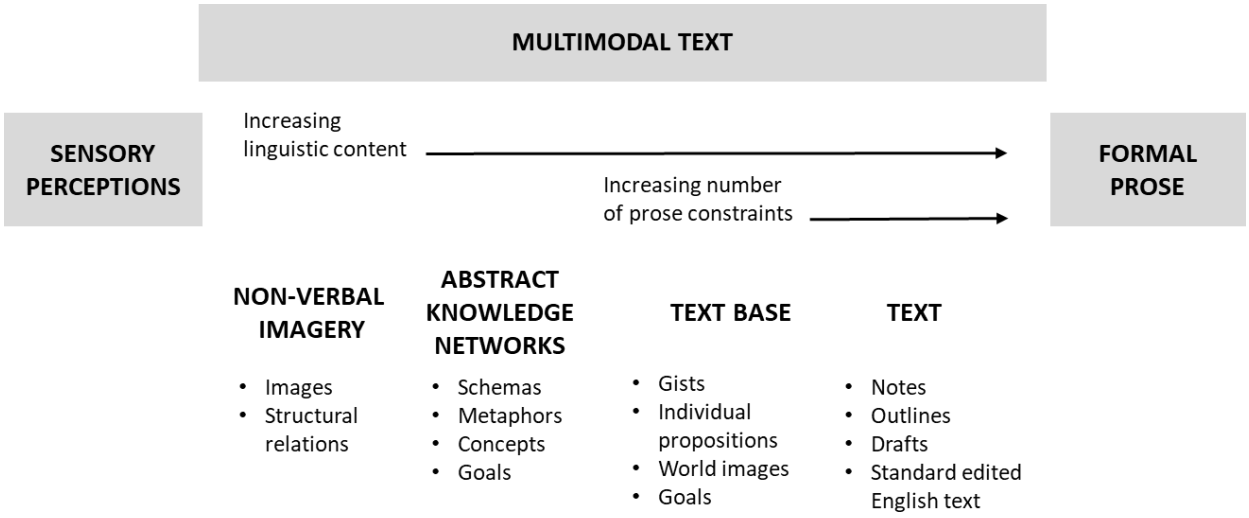

Figure 3. Schematic representation of multimodal plans and multimodal text

While the cognitive theories of writing were not developed for explaining the cognitive processes of constructing multimodal texts, these provide explanations for how human mind

works during and for multimodal writing. With respect to the Multiple Representation Thesis, it could be argued that the cognitive complexity of a multimodal writing task is related to the number and diversity of mental representations of a writing plan and the alignment of modes between a writing plan and written text. The latest version of the cognitive model of writing reflects the multimodal characteristics of authentic writing tasks and explicates multimodal writing processes. If multimodal writing tasks are used for pedagogical purposes, these theories can be used to manipulate the demands for linguistic, nonlinguistic, and intermodal choices.

Social semiotics and systemic functional linguistics. Halliday's view of language, as proposed in his books *Language as Social Semiotics* (1978) and *An Introduction to Functional Grammar* (1985), triggered substantial changes in research. He established a strong argument to move research focuses from traditional structures (e.g., sentence and grammar) to functions (e.g., discourse and semiotic resources). Simply put, social semiotics research focuses on how people use semiotic resources (e.g., spoken words, written words, pictures, movements, gestures, and sounds), which encompass all possible means of communication (*meaning-making*, in their own term) such as linguistic and cultural resources, and how these uses change through social interaction. Researchers have shed light on individual writers' agency and identity in constructing a multimodal text by looking into how they choose and utilize different semiotic resources (e.g., Cimasko & Shin, 2017; Jiang, 2018; Nelson, 2006; Tardy, 2005). This focus contrasts to traditional semiotics, which views semiotic systems as fixed rules of meaning and signs.

On the other hand, systemic functional linguistics (SFL) examines how language, which is one of the semiotic systems, interrelates with other semiotic resources in meaning making to fulfill one of the three metafunctions: *ideational (field)*, *interpersonal (tenor)*, and *textual (mode)*

metafuctions. The *ideational* function is to show experience and ideas, which is related to one's world view; the *interpersonal* function is to engage in society through language, for example taking turns and understanding others' feelings; and the *textual* function is to organize the text for easier and communication. Thus, instead of focusing on syntactic structures and/or thematic roles within sentences, SFL brings attention to discourse-level features of language. In addition, SFL linguists assume that a particular linguistic feature is chosen over other alternative options when circumstances meet the conditions for the feature being selected. In this sense, the language that a person produces is a system that is selected to carry a specific function in a particular context. SFL thus underscores the importance of the context where language is used.

Van Leewen, Kress, and Jewitt, to name a few, integrated these two interrelated concepts (i.e., social semiotics and SFL) and developed a research tradition that targets the metafunctions of different semiotic resources. While traditional SFL researchers focused on how language communicates social functions, researchers who looked into multimodal texts analyzed the metafunctions of semiotic resources beyond language (e.g., graphics, layout, and gestures). Kress and Van Leeuwen (1996) defined the metafunctions of semiotics as follows (pp. 40–41):

- The ideational metafunction: Any semiotic system has to be able to represent, in a referential or pseudo-referential sense, aspects of the experimental world outside its particular system of signs.
- The interpersonal metafunction: Any semiotic system has to be able to project the relations between the producer of a sign or complex sign, and the receiver/reproducer of that sign; to project a particular social relation between the producer, the viewer and the object represented.

- The textual metafunction: Any semiotic system has to have the capacity to form texts, complexes of signs which cohere both internally and with the context in and for which they were produced.

As can be seen in each definition, Kress and Van Leeuwen explained that any semiotic resources can perform such metafunctions. Based on this operationalization of systemic functional social semiotics, researchers developed two close but separable research traditions. One weighed more on describing the functions of different modes (i.e., *multimodal discourse analysis*) while the other focused on how writers choose different social semiotics (i.e., *social semiotic multimodal analysis*).

With *multimodal discourse analysis*, researchers placed a focus on "the metafunctional systems underlying semiotic resources and the integration of system choices in multimodal phenomena" (Jewitt, 2014b: 35). Researchers focused on the relations between images and written words (Hagan, 2007; Liu & O'Halloran, 2009; Martinec & Salway, 2005; Unsworth, 2007). For example, Martinec and Salway (2005) proposed a system that is generalizable to a broad range of genres. They identified *status* and *logico-semantics* of the two modes to interpret intermodal relations. The status of the relation could be either equal and unequal. For equal status, image and text could independently or complimentarily contribute to the meaning. In terms of logico-semantic relations, they suggested that two relationships are possible. First, information in one mode could expand the meaning by the other mode, for example elaborating, extending, or enhancing meanings. Second, one mode could project the meaning presented by the other mode. This projection could be observable in comic strips and diagrams that appear in textbooks and academic publications. An example for projection, more specifically locution, was a Venn diagram and a separate prose that explained the content presented in the Venn diagram.

Because diagrams displayed not only images but also some texts within the graphic, Matinec and Salaway elaborated that in some cases the texts within a diagram need to be further analyzed as a category of expansion. This generic model of image-text relation influenced the following studies: e.g., Alyousef, 2016; Chandler, Unsworth, & O'Brien, 2012; Daly & Unsworth, 2011; Martinec, 2013; Unsworth, 2006, 2007).

Focusing on pedagogical implications for multiliteracy education, Unsworth (2006, 2007) provided schemes that explain the interaction between language and image in the ideational meaning. His 2007 study specifically analyzed textbooks and websites for K-12 school sciences, focusing on the logico-semantics relation in the original scheme of Martinec and Salaway. With the examples from school textbooks, he revealed that in many cases images and written words expand meanings from one mode to another mode (e.g., concurrence and complementarity); however, he could not find cases where content presented in one mode is enhanced by the other mode (i.e., enhancement). While Martinec and Salway (2005) and Unsworth (2007) shared similar systems, other researchers used different terms to demonstrate the multimodal relationships. For instance, Liu and O'Halloran (2009) placed more emphasis on cohesion than the ideation metafunctions and suggested a different system that identified four types of intermodal relations: comparative, additive, consequential, and temporal. Hagan's (2007) system was also distant from the Martinec and Salaway's scheme. She combined works of systemic functional linguistics, particularly on cohesion (Halliday & Hasan, 1976), and the psychological accounts of interpreting graphics (Arnheims, 2004). I summarized the schemes used in these multimodal discourse analyses in Table 1.

In summary, previous studies using multimodal discourse analysis documented the relations between words and images in multimodal texts and advanced the systems as some

added different emphasis on other concepts than Halliday's metafunctions. For example, Hagan (2007) emphasized that her representation of the intermodal relation reflects psychological accounts for the interpretation of visual information. Liu and O'Halloran (2009) highlighted that their approach better accounts for 'discourse', while earlier studies by Martinez and Salaway (2005), and Unsworth (2006, 2007) focused on 'grammar'. Taken together, these differences can contribute to provide multiple ways to interpret intermodal relations.

Table 1.

Schemes for Analyzing Image-text Relations in Previous Research

	Martinec and Salway (2005)	Unsworth (2006)	Unsworth (2007)	Liu and O'Halloran (2009)	Hagan (2007)
Focus	Generalized system of image-text relations (status and meaning)	Ideational meaning	Ideational meaning	Cohesion (logical relation)	Perceptual tie (structure) and cohesive tie (content)
Data	Samples from various genres (e.g., advertisement, drawings in newspaper, comic strips)	Samples from various genres (e.g., online advertisements, online teaching material)	School science textbook and website	Samples from various genres (e.g., drawings in newspaper, instruction sheet)	Samples from 109 professionals and 21 students (e.g., cover page, advertisements, assignments)
Scheme	*Status* 1) Equal a) independent b) complementary 2) Unequal a) image subordinate to text b) text subordinate to image *Logico-semantics* 1) Expansion a) elaboration • exposition • exemplification b) extension • enhancement temporal • spatial • causal 2) Projection a) locution (wording) b) idea (meaning)	1) Concurrence a) redundancy b) exposition c) instantiation d) homospatiality • image instantiates text • text instantiates image 2) Complementarity a) augmentation • image extends text • text extends image b) divergence 3) Connection a) projection • verbal • mental b) conjunction • casual • temporal • spatial	1) Expansion a) concurrence • clarification • exposition • exemplification • homospatiality b) complementarity • augmentation • divergence c) enhancement • manner • condition • spatial • temporal • causal 2) Projection a) verbal b) mental • perception • cognition	1) Comparative a) Generality b) Abstraction 2) Additive 3) Consequential a) Consequence (cause) b) Contingency (purpose) 4) Temporal (Successive)	1) Typographic interplay • shared location • blended content 2) Interplay in Parallel • similar location, shape; alignment; overlap • exophoric tie 3) Interplay in Sequence • Contrast breadcrumb; similar location, shape; alignment; overlap • referencing image; substitution tie; repetition tie; collocation tie; referencing tie 4) Interweaving • similar location, shape; alignment; overlap • overlap collocation; collocation tie

Another line of research originated from systemic functional linguistics and social semiotics is *social semiotic multimodal analysis* that focuses on the agency of sign makers (i.e., writers) and why writers select particular semiotic resources in a given context. A selection and use of semiotic resource is determined by writers' purposes, their understanding of audience, as well as the potentials and limitations of each of the semiotic resources. In previous studies using social semiotic multimodal analysis, researchers aimed to describe the contextual factors that affected writers' choices in modes, or semiotic resources, thus placed more focus on the contextualized act of writing than the multimodal discourse analysis, which primarily focused on the relationships represented on a page, did. Earlier studies devoted to show that writing is becoming multimodal across genres (Bezemer & Kress, 2008; Lemke, 1990, 1998).

Lemke (1998) reported how linguistic and nonlinguistic modes were collectively used in scientific articles published in prestigious science journals. He counted the number of graphics (e.g., figures, tables, charts, graphs, and other visual presentations) and page counts of articles in two issues of two journals *Science* and *Physical Review Letters,* where each article was about 3 pages long, and one issue of *Bull NY Acad Med.* The *Science* included six graphics on average; the *Physical Review Letters* included about 3.8 graphics and 8.5 equations per article. *Bull NY Acad Med* published longer papers (15 page-long) and included about 16.2 graphics and 17 equations. With these numbers, Lemke argued that technical scientific academic writing should be viewed as multimodal ensembles. He suggested that meaning presented by different modes can play *presentational, orientational,* and *organizational* functions, which correspond to *ideational, interpersonal,* and *textual* metafunctions in Halliday's terms. He provided examples regarding how tables and figures were presented in papers and how writers used linguistic texts to direct readers to different places in the multimodal text. For example, a writer used "see Table

X" in a journal article to relocate readers' attention from text to graphics which only happens in multimodal genres. In addition, Bezemer and Kress (2008) demonstrated that learning resources have become multimodal, and questioned how these changes in representation would affect learning. The ideas of *transduction* and *transformation* in this paper was influential to subsequent research that looked into the writer's identity development while creating multimodal texts (e.g. Cimasko & Shin, 2017). *Transduction* indicates "the move of semiotic material from one mode to another (p. 175)" while *transformation* means within-mode changes. For example, according to these definitions, a written instruction that includes an image of a compass and some description in language in imperative (e.g., First, put the point on the dot.) about how to use a drafting compass to draw a circle is a transduction of an act of drawing a circle with a compass. As a result of transduction, which reduces the event or act of using a compass into a static picture and some written words, this multimodal text loses some information (e.g., the person who uses a compass) while adding new information (e.g., command function if the sentence is in an imperative mood).

Based on these earlier studies that explicated the characteristics of multimodal texts, some empirical research adopting this social semiotics approach explored the outcome of using multimodal writing projects in literacy education (e.g., Smith & Dalton, 2016; Tan & Guo, 2009). For example, Smith and Dalton (2016) invited two college freshmen students who had participated an AP literature and composition class where they were required to engage in three multimodal composition tasks. When they were 12^{th} grade students, they had written a reflection post about a novel on a webpage, produced presentation slides reporting their literacy analysis on the novel, and created an audio letter to the main characters. One year later, they were asked to make short videos that project their identities as multimodal designers and to reflect on one of

the multimodal products they had created before. Smith and Dalton revealed that the participants used different resources to author their stories in creative ways; both participants reported this reflection project was helpful in representing their identities and reflecting their composing processes. Tan and Guo (2009) applied the notion of critical multimedia literacy (Lemke, 2006), which was defined as an analytic technique to demonstrate how images and texts are arranged to reinforce or undermine each other, to English curriculum in Singapore as an attempt to pilot how new literacies can be integrated into the current system. They focused on the development of activities and lessons to incorporate this critical literacy and reported the challenges the focal teachers had faced in due course.

Placing an explicit focus on adolescent's (grades 5-12) digital multimodal composition, Smith (2014) reported a systematic review of literature and outlined six themes that were salient in the previous research. She found 79 studies in journals, book chapters, and conference papers from 1999 to 2012. In terms of research design, she reported that studies, except for one study that was quasi-experimental, were conducted within a qualitative approach. Among the themes, she revealed that digital video (48.7%) was the most frequent type of multimodal product that students created. This video project covered a wide range of genres from public service announcements to digital stories that remix different sources such as music, pictures and voice. One of her findings echoed the importance of overt instruction. She observed that teachers in the previous studies had placed explicit focus on the technological skills and metalanguage, which are important and helpful for students in interpreting and constructing multimodal texts. Because modes other than language also have their particular grammar, using metalanguage and giving explicit instructions on how to produce multimodal writing can be equivalent of form-focused instruction in language classrooms.

Most studies to date rely on the definition of multimodal texts from Kress and Van Leeuwen (1996) and Jewitt (2008) who developed this line of research explicating multimodal texts from the SFL and social semiotics. While these approaches have initiated research inquiries about multimodal texts and provided thick descriptions, researchers to date have not attempted to provide regular genre-specific features of multimodal writing partially due to their fundamental emphasis on the contextualized interpretation of phenomena, not discussing regularity or conventions across writers. Now I turn to the next approach that shed lights on the multimodal genres in academic contexts where researchers paid attention to the regular patterns in multimodal texts.

Academic genre studies. Given that the awareness of multimodality is increasing, academic genre studies have begun to unveil the conventions of different modes in specific discourse communities (e.g., Archer, 2012; D'Angelo, 2010, 2016; Li & Lodge, 2017; Mogull & Stanfield, 2015; Morell, 2015; Rowley-Jolivet, 2002, 2012). Similar to the original genre studies following Swales'(1990) rhetorical analyses, these studies demonstrated the typical roles and affordances of visual and linguistic modes in academic discourse communities. To note, even though these genre studies set out to document regular patterns of discourse in a particular community, researchers have adopted the use of analytic frameworks from multimodal discourse analysis and social semiotic multimodal analysis (with exception of Mogull & Stanfield, 2015). In addition, many studies in academic and technical multimodal communication focused on the presentation genres where visual aids for presentations such as power point slides were investigated along with speaking and gestures (e.g., Mogull & Stanfield, 2015; Morell, 2015; Rowley-Jolivet, 2012). In this section, I discuss the previous studies on multimodal texts that involved written texts because the scope of the current disseration project is limited to the

writing of multimodal texts in a second language; however, a caveat should be noted that the conventions of multimodal genres have been changing fast (Reid, Snead, Pettiway, & Simoneaux, 2016); it could be problematic to interpret the findings as generalizable genre conventions.

Mogull and Stanfield (2015) analyzed journal articles published in *Science* in 2014, resulting in 264 articles, in terms of inscriptions (i.e., modes). Their inscription types were as follows: diagrams, equations, graphs, instrument output, photographs, and table. They revealed that tables were not frequently used while graphs and diagrams appeared often. Notably, the types of graphs and diagrams were becoming more divergent, which merits more pedagogical guidance for researchers. Interestingly, this descriptive study collected data from the identical journal that Lemke (1998) used and focused on the specific types of graphics researchers used to further discuss pedagogical implications for technical writing; however, it did not follow the analytic framework of social semiotics.

Several studies focusing on academic multimodal genres adopted multimodal discourse analysis. Investigating international conference presentations, Morell (2015) and Rowley-Jolivet (2002, 2012) included presentation slides as a subgenre. Morell proposed an evaluation model for 20-minute academic presentations in social science and technical science. She collected four nonnative researchers' presentations at an intensive workshop on academic English, which she interpreted as model presentations. Then she coded how each of the speakers used spoken, written, nonverbal (e.g., graphs, tables, bar charts, images or videos) modes, body language and combined different modes. When examining the combination of modes, instead of closely looking at each intermodal relations as multimodal discourse analysis would do, she evaluated the overall balance of the different modes. Results showed that the presentations that the four

participants gave at the workshop, the presenters used key words and condensed structures (e.g., bullet-pointed lists) to display texts on slides. She also revealed that written information on slides, alphabetically and non-verbally, was repeated in the speaking simultaneously or consecutively; however, she did not discuss the intermodal relation on the visual information on each slide.

Rowley-Jolivet (2002, 2012) investigated academic research presentations which were recorded at conferences in science. In her 2002 paper, she analyzed 90 presentations that contained 2048 non-verbal visuals (e.g., tables, graphs, and diagrams) in terms of shared visual lexicogrammar, which is some rules that researchers have to follow in constructing the visuals which she categorized into four types: scriptural (linguistic), graphical, figurative and numerical. The distinction between graphical and figurative visuals was on the possibility of being interpreted in multiple ways. Rowley-Jolivet (2012) shifted focus to the language between visuals and what presenters verbally presented. Interestingly, this study combined textual analysis using corpus and the manual analysis of metafunctions. In terms of the corpus analysis, she built and analyzed parallel corpus of texts on slides and transcribed spoken commentaries; and revealed that slides had higher lexical diversity and nominalization because researchers eliminated function words (e.g., pronouns, determiners, and auxiliary verbs). For the metafunctions, she demonstrated that in conference presentation genre, the role of slides is to communicate the ideational function while the roles of verbal comments are to communicate textual and interpersonal functions. Based on the findings, she recommended ESP training courses to focus on the transition from the dense information on slides to the verbal commentaries.

While Rowley-Jolivet and Morell investigated research paper presentations with slides, D'Angelo (2010, 2016) analyzed how visual and textual information on academic posters were presented, integrating Kress and Van Leeuwen's (2006) grammar of visual design and Hyland's (2005) metadiscourse. Hyland defined metadiscourse as "the cover term for the self-reflective expressions used to negotiate interactional meanings in a text, assisting the writer (or speaker) to express a viewpoint and engage with readers as members of a particularly community." (Hyland, 2005: 37) and explicated that non-verbal expressions such as printing, genre and media, and punctuation can express metadiscourse. Even though non-verbal metadiscourse signals were proposed, researchers have so far only focused on the textual element in communicating metadiscourse. In this sense, D'Angelo's works on revisiting the idea of textual metadiscourse and expanding the scope to visual modes can be seen as original attempts. She conducted a mixed-methods study where she built and analyzed a multimodal corpus of 120 posters that consisted of 40 posters from Law, Clinical Psychology, and High Energy Particle Physics and conducted an online survey for experienced and novice researchers as well as interviews with twelve researchers to understand the use of academic posters. To examine metadiscourse realized through textual and visual modes, she manually annotated textual metadiscourse markers based on Hyland's list of metadiscourse markers, and interactive visual components using a qualitative analysis software and reported frequencies and examples for each resource. For the interactive visual metadiscourse, she developed a coding scheme consisted of five categories of interactive resources: 1) *information value* to organize the layout of information (e.g., left-right, top-bottom, center-margin); 2) *framing* to divide text sections (e.g., frame lines, color contrasts, empty spaces); 3) *connective elements* to connect ideas and parts of visual and textual discourse (e.g., vectors, repetition of colors, alignment) 4) *graphic elements* to clarify and organized data for the

viewer (e.g., taxonomies, flowcharts, networks, tables, figures), and 5) *fonts* to enhance legibility and highlight important parts of the words (e.g., type, size, color). D'Angelo revealed differences between posters in the three disciplines. While posters of clinical psychologists contained many words and less textual interactive resources, those of lawyers had small number of running words and many interactive resources. Posters in hard science were found to be more succinct than those in psychology and include few textual interactive resources. In terms of visual information, she revealed that the three disciplines used similar amount of interactive resources in total; however, the three disciplines demonstrated differential preference to interactive metadiscourse resources. For example, in hard sciences, researchers put more graphic elements; psychologists preferred to use fonts for interactive purposes; and lawyers used framing resources more often than others.

Similarly, Li and Lodge (2017) reported a corpus-based quantitative analysis of university lecture slides in social sciences and engineering. They computed syntactic and lexical complexity of written words and manually coded visual type and visual-text relations to quantify linguistic and multimodal features of each PowerPoint slide. Lecture slides of social sciences courses were composed of more complex structures and showed lexical variation than those of engineering. However, engineering slides contained more sophisticated lexical items, numerical and visual elements. In terms of the relationship between linguistic and nonlinguistic modes on slides, they found that lecture slides in both disciplines most frequently employed concurrence relations to repeat, explain or provide examples. Even though multimodal genre analyses have not found much discipline-specific features of visual information, they expanded genre research to cover multimodal texts and exemplified coherent schemes for both images and texts. A more

comprehensive study could include intermodal analysis as well as linguistic patterns for different sections.

In summary, multimodal academic genre studies demonstrated functional analyses on sample texts in different genres and attempted to provide general tendency in the use of multiple modes (e.g., verbal presentation of empirical studies with visual supports in Rowley-Jolivet, 2012; slides in Morell, 2015 and Rowley-Jolivet, 2002; academic posters in D'Angelo, 2010, 2016; and journal articles in Mogull & Stanfield, 2015). Identifying metafunctions and metadiscourse realized through different modes were the primary interests in many studies, which led to some interim conclusion that each of the modes are typically used to exhibit different metafunctions (e.g. visuals for ideational functions and verbal commentaries for organizational and interpersonal metafunctions in Rowley-Jolivet, 2012). After reporting the regular uses of different modes, researchers tended to suggest implications for novice presenters or writers; however, they did not explain how these conventions can be taught or provided to the novice researchers.

Summary and implications for L2 research. Less than two decades ago researchers began to investigate multimodal composition and texts. The current review revealed that the underlying dominant theoretical background has been systemic functional social semiotics. Following this tradition, researchers have focused on identifying the roles of nonverbal texts as well as linguistic texts in communicating meanings (i.e., social semiotic multimodal analysis and multimodal discourse analysis). Some researchers set the primary goal as defining the regular patterns in using multiple modes in a multimodal genre (i.e., multimodal genre studies). Only one empirical study was rooted in the cognitive approach to writing. I summarized different approaches to multimodal writing in Table 2. I outlined different perspectives toward multimodal

writing; however, a caveat should be noted that these approaches are not mutually exclusive. For instance, as discussed for D'Angelo's and Rowley-Jolivet's academic genre studies, researchers pulled coding schemes from the notions of metafunctions of systemic functional linguistics. With this review on multimodal writing from different theoretical orientations, I observed the following implications for L2 research: (1) the potential contribution of the cognitive writing model to explaining multimodal academic writing; (2) the importance of discipline-specific approaches; (3) and the role of language in multimodal texts.

Table 2.

Summary of Approaches to Multimodal Writing

Approach	Focus	Theoretical background	L1 studies	L2 studies
Social semiotic multimodal analysis	Situated choice of resources	Social semiotics (Halliday, 1978)	Bezemer & Kress (2008); Lemke (1998)	Cimasko & Shin (2017); Nelson (2006); Smith, Pacheco, & Rossato De Almeida (2017)
Multimodal discourse analysis	Metafunction system of available resources	Systemic functional grammar (Halliday, 1985)	Daly & Unsworth (2011); Hagan (2007); O'Halloran (2004)	Alyousef (2016); Anderson, Stewart, & Kachorsky (2017)
Multimodal genre analysis	Genre-specific grammar for language and visuals	Genre analysis (Swales, 1990)	Archer (2010) D'Angelo (2010, 2016);Li & Lodge (2017); Rowley-Jolivet, (2002, 2012)	Molle & Prior (2008); Tardy (2005)
Multimodal cognitive writing process	Writers' processes of composing a multimodal text	Cognitive model of writing (Hayes & Flower, 1980)	Leijten, Van Waes, Schriver, and Hayes (2013)	–

First, given that many pedagogical decisions for L2 writing consider the cognitive process of writing, it would be helpful to update the model of writing as the current multimodal writing practices as attempted in Leijten et al. (2013). To date, cognitive accounts for multimodal composition have not exercised much impact in this domain because of the emphasis on the social semiotics in discussing multimodal research and the focal dominance of linguistic texts in composition research. For example, Jewitt (2014) from the social semiotics approach explicitly stated that "multimodality is distinct from cognitive psychological approaches that focus more explicitly on the internal, notions of mind, and cognitive process (p. 31)". In composition literature, Leijten et al.'s (2013) study demonstrated that the cognitive processes of composing multimodal texts is realized through translating a multimodal writing plan reflecting a writer's visual schema as well as writing schema. In addition, based on the Multiple Representation Thesis, distance between the modes of writing and representation can explain the amount of effort a writer invests in translatinge a writing plan into a written text. This alignment of internal and external representations can be connected to the task complexity research in L2 writing studies, which I come back in the last section of the literature review for further explanation.

Second, multimodal texts also display discipline-specific features. Across different approaches looking at multimodal writing, researchers emphasized the social context such as target readers and the conventional practice of a discourse community. Even in a study conducted from a cognitive approach to writing, Leijten et al. (2013) used the ethnographic method to explain the contextual variables influencing the focal participant's writing processes and weaved those social factors into control and process levels in the writing. D'Angelo (2016) who focused on a genre analysis compared the use of metadiscourse signs in academic posters of three disciplines. These studies explained that researching multimodal texts in a specific context

could inform pedagogical decisions as to which writing skills made experienced writers. On the other hand, in a K-12 literacy study, Smith (2014) revealed that digital videos (e.g., documentaries, digital stories) were the most popular type of multimodal writing practice. Because the pedagogical goal for the K-12 students is to practice different technological tools and develop multiliteracies (The New London Group, 1996), digital video projects could be a plausible choice in this context. However, they would not be the most relevant type of multimodal composition to EAP students, whose goal is presumably to become familiar with academic tasks such as final papers and paper presentations and obtain the basic skills to successfully complete such tasks. Because any pedagogical choice for multimodal task necessitates a careful interpretation of language use domain and students' goals (Long, 2005, 2016), a needs analysis would be the first step for any meaningful discussion of multimodal writing in an EAP context.

Lastly, the role of language in multimodal writing should not be underestimated especially for the L2 writing context. In literacy education perspectives, researchers have discussed that language classes need to focus on metalanguage development (Archer, 2006, 2010; Unsworth, 2006) and technology (Hundley & Holbrook, 2013; Walsh, 2010); in content courses, English language learners' nonlinguistic communication skills can help achieve course goals (e.g., Science subject of elementary school students in Grapin & Llosa, 2020; Lee, Llosa, Grapin, Haas, & Goggins, 2019). However, L2 writers in tertiary level EAP courses aim to learn how to use language for academic tasks that they will encounter once they exit language courses. For example, novice writers, even native speakers of English who are new to the context, may encounter difficulties in producing clear and purposeful graphics for a compelling proposal. There could be some academic conventions to presenting graphics along with texts; as revealed

in some multimodal academic genre studies (Morell, 2015; Rowley-Jolivet, 2002, 2012), presentation slides and posters tend to contain more low-frequency words in simpler syntax, and less metadiscourse markers than regular papers. Learning such genre conventions becomes more complicated when novice writers are nonnative speakers of English. While L1 writers can focus on the expansion of genre schema in visual and linguistic modes, L2 writers have to expand linguistic knowledge at the same time. Thus, a better understanding of language in multimodal texts can provide supports to L2 writers and further inform some generic multimodal writing tasks for EAP learners.

Previous L2 Research on Multimodal Writing

In L2 writing context, only a few studies have tried to broaden the learning goals to cover multimodal writing. Kress (2000) and Van Leeuwen (2015), for example, provided a theoretical foundation for multimodal texts, and called for more attention to multimodal writing. Kress (2000) emphasized that each mode has specialized functions and students need to learn how to exploit various modes to communicate different functions effectively. Van Leeuwen's (2015) response to a special issue on multimodality suggested that future research should develop assessment criteria for multimodal literacy and to take a closer look at the visual literacies for different school subjects. Specific to L2 writing, Elola and Oskoz (2017) noted that writing genres have changed in contemporary digital settings, and recommended studying multimodal genres for teaching and assessment.

While some conceptual papers have introduced interesting arguments, there have been only a few number of studies that investigated multimodal writing. Grapin (2019) explained that this lack of empirical research on multimodality is related to the operationalization of the term *mode*; in L2 studies, mode has been regarded as the channel of communication (e.g., spoken and

written) while in other content areas it indicates various social semiotics covering verbal and nonverbal resources. In his definitions of multimodality in weak and strong versions, the weak version assumes language as a privileged mode and students may stop use other modes than language when they achieve proficiency; he found that this weak version is the predominant position in ESL education for K-12. On the other hand, the strong version of multimodality emphasizes strategic use of multiple modes regardless of language proficiency and all modes are valued based on their affordances and norms in each discipline. He argued that the strong version of multimodality should be encouraged because the strong version is more helpful for students to participate in content classes. In summary, previous conceptual papers claimed a complete reconceptualization of the goal of learning writing in general. These papers, however, may have not been persuasive to researchers whose primary concern is tertiary EAP because none of the papers clearly argued why semiotic resources other than language should be taught to adult EAP writers who might have developed some basic skills to use semiotic resources from previous, possibly L1, learning experiences.

In terms of empirical research, only a few studies have looked into L2 writers' multimodal composition (see Tables 4 and 5 for L2 studies on multimodal writing at K-12 and tertiary levels). I found following themes from this review. First, despite the fact that in the field of L2 learning is devoted to the language development, previous studies on multimodal writing, to date, have not set development as a primary enquiry. Following a systemic functional social semiotics approach, many researchers focused on how individual writers engaged in the process of choosing and using semiotic resources to construct meaning in a particular context. Among L2 studies, two studies attempted to address language development (Dzekoe, 2017; Vandommele et al., 2017). Dzekoe (2017) adopted a multiple case study design to explore the effects of using

digital poster projects in EAP class on self-revision behaviors and reported that the multimodal practice helped students revise contents and improved overall text quality. Vandommele et al. (2017) compared changes in linguistic complexity, accuracy, and fluency in L2 writing of the three conditions: task-based instruction, out-of-classroom digital project, and non-intervention. The two intervention groups' goal was to make a website that could help new immigrants to their city. While students in task-based instruction group were provided with 18 tasks and scaffolding activities, students in out-of-classroom condition met youth workers and free-lancer artists who helped them learn website designing skills. This study revealed that the multimodal writing project, regardless of the contexts, resulted in more gains than non-intervention condition; and the out-of-classroom project which gave more autonomy to students led to higher language learning gains than the in-class project.

Second, among the eleven studies summarized in Tables 4 and 5, nine studies adopted a case study design to describe multimodal writing processes, particularly as a practice of building one's identity. In these studies, researchers showed more interest in how this alternative writing tasks could empower marginalized students (Anderson et al., 2017; Pyo, 2016) and how students' identities developed over the course of producing multimodal texts (Jiang, 2018; Smith et al., 2017; Tardy, 2005). Building on previous descriptive and qualitative studies, future research can expand to quantify such processes to show overall tendency in writing processes. Gánem-Gutiérrez and Gilmore (2018), for example, examined how much time L2 writers spent on different writing processes (e.g., text construction, revising, pausing, rereading, using external resources) when completing a timed argumentative writing task by analyzing screen capture videos of their on-screen writing behaviors with eye-gaze traces. In addition, it could be possible that the focus of systemic functional social semiotics in previous multimodality research restricts

research methods. Only two studies used other more quantitative approach stated their theoretical framework as TBLT (Vandommele, Van den Branden, Van Gorp, & De Maeyer, 2017) and Noticing (Dzekoe, 2017).

Third, researchers, in both L1 and L2 studies, have not justified why they chose the tasks they used for multimodal writing. The New London Group's proposal on multiliteracies has been reflected in many K-12 studies, which could explain the studies summarized in Table 3. Participants in these studies were in the instructional contexts where both basic literacy and linguistic skills had to grow. However, in many studies that were conducted in tertiary education, researchers identified the research context being language courses (see Table 4). It is questionable whether these tasks are carefully adopted ones for the target students. While teachers can provide students opportunities to explore different modes to compose multimodal texts, it is problematic that there was no justification on why somewhat creative tasks were implemented. In fact, this line of research did not follow TBLT where learners' needs are analyzed, and tasks are sequenced to align with the mental processes of writing. From the perspectives of task-based language teaching, ignorance of multimodal writing tasks in the real world limits quality pedagogical practice.

Lastly, there has been little attention to the evaluation of multimodal texts. Multimodal writing skill may not be the goal of writing course in immediate future, but knowing how it is relevant, or irrelevant, to language proficiency can help instructors and material developers design pedagogical tasks. Learners need to know what they are expected to produce and/or how they are expected to perform when using nonlinguistic modes along with linguistic mode. Much of the work on the assessment of multimodal writing has focused on generating guiding principles (e.g., Hung, Chiu, & Yeh, 2013) or discussing general challenges faced by educators

(e.g., Yi & Choi, 2015; Yi, King, & Safriani, 2017). There have been no large-scale surveys nor empirical studies designed to understand multimodal writing performances. The absence of a basis of interpreting multimodal writing performance thus has generated teachers' and students' reluctance to incorporate the authentic modes of writing. Furthermore, the lack of common understandings of multimodal writing performance has been a challenge for researchers. For L2 writing literature, rubrics have served for systematic analyses of learners' linguistic development (e.g., Connor-Linton & Polio, 2014; Jacobs, Zinkgraf, Wormuth, Hartfiel, & Hughey, 1981). When a new construct is introduced, researchers have developed a new rubric to account for the construct (e.g., authorial voice in Zhao, 2012; integrated writing ability in Chan, Inoue, & Taylor, 2015; Cumming, Kantor, & Powers, 2002). Without any empirically tested rubrics of multimodal writing, challenges would remain for both research and practice.

In conclusion, L2 researchers have recently begun investigating multimodal writing. Many methodological practices that L1 multimodal researchers made directly influenced L2 research, which include the dominance of systemic functional social semiotics as theoretical framework and the lack of justification of choosing a particular type of multimodal writing task. Researchers have tried to implement a strong version of multimodality, which may have triggered some L2 researchers' backlash against introducing multimodal writing practice to writing classroom. In addition, most of the studies to date are disjointed with the current discussion in L2 writing research for adult learners such as TBLT and EAP genre studies. Taken together, multimodal writing research needs to be contextualized in the adult EAP writing contexts. Central to contextualizing, one of the immediate issues is the task-based needs assessment of multimodal writing in EAP context.

Table 3.

Previous L2 Research in Secondary School

Study	Context	Framework	Focus	Tasks	Methods (Participant and Data)
Anderson, Stewart, & Kachorsky (2017)	Secondary school (Persuasive writing unit, age 14-15)	Interpersonal metafunction (rhetorical force, authorial stance)	Students' renegotiation of positioning through designing multimodal text	Video of a persuasive argument (Modes: texts, image, sound)	Case study 17 multimodal texts by 3 academically marginalized students Open coding, axial coding, presentation of exemplar cases
Smith, Pacheco, & Rossato De Almeida (2017)	Secondary school (8th grade)	Translanguaging; Social semiotics	Multimodal codemeshing	A multimodal video project about ones' hero (Modes: images, text, songs, voice)	Comparative case study 3 eighth grade bilingual students Screen capture and video observations, student design interviews, multimodal products Open coding, timescapes
Vandommele, Van den Branden, Van Gorp, & De Maeyer (2017)	Secondary school in Belgium (age 14-15)	Task-based language teaching	Effects of a collaborative multimodal writing on different settings on writing development of novice learners	Design a website that should include photo-comic, video-based interview, etc. (Modes: text, image and video)	Experimental study 84 novice learners of Dutch in-class (n=26); out-of-school project (n=26); control group (n=32) pre- and post- test performance on traditional writing tasks (one narrative, one persuasive) multi-level modeling three fixed effects (pre/post, condition, interaction)
Pyo (2016)	ESL class in multicultural service program for youths	Multiliteracies pedagogy	Student's engagement with multimodal project; authorial identity	Presentation after reading a book about immigrants' life at students' choice of presentation (Modes: image and text)	Case study One participant (out of a bigger study) observation, six semistructured interviews, project output (slides) 9-page slides: 3 pages with written words, 4 pages including words and images, 1 page with image only Inductive analysis

Table 4.

Previous L2 Research in Tertiary Level

Study	Context	Framework	Focus	Tasks	Methods (Participant and Data)
Jiang (2018)	College (Chinese EFL for non-English majors)	Identity; Investments	Processes of writers' investment change a digital multimodal composing program	Five video projects on five textbook topics (Modes: image, voice, caption)	Multiple case study 3 focal undergraduate students (22 in total) Observation, interview, and student-authored multimodal texts (selectively transcribed) Qualitative inductive analysis Recursive cross-case analysis
Cimasko & Shin (2017)	College (English 101)	Sociosemiotic ethnography; Resemiotization; Recontextualiza-tion	L2 writer's authorial decisions and contextual factors in multimodal designing	Reproduction of argumentative essays students wrote into animated video or slide (Modes: characters' action, text, image, voice)	Ethnographic case study One college ESL writer Her argumentative paper, video transcript, multimedia video, interview transcripts, observation notes
Dzekoe (2017)	College (ESL)	Noticing; Intersemiotic complementarity	Effect of computer-based multimodal composing activities on students' revision	Online multimodal posters (Modes: image and text)	Case study with embedded quantitative data 22 advanced-low proficiency ESL students surveys, students' revision history, online posters, reflections, listening activities, stimulated recall interviews, final written drafts, writing scores Intersemiotic analysis of visual and linguistic elements

Table 4 (*cont'd*)

Study	Context	Framework	Focus	Tasks	Methods (Participant and Data)
Alyousef (2016)	International students in undergraduate Business program	Multimodal discourse analysis; Theme system	Thematic progression and composition of information value	Business marketing plan reports (Modes: text, graphs, tables)	Case study 3 international undergraduate students in marketing classes and 2 tutors text analysis of multimodal marketing plans thematic progression patterns
Molle & Prior (2008)	EAP course for graduate students	Genre; Sociocultural approach (multimodal and semiotic approach)	Genre and needs of EAP students (graduate)	Authentic writing tasks students performed in their content courses. (Modes: image, equation, notes, table and text)	Needs analysis International graduate students Native instructors in the students' disciplines student texts, class observations ethnographic methods
Nelson (2006)	College (first year writing)	Synaesthesia (transformation and transduction)	Multimodal authorship	Design experiments with students (Modes: image and text)	Case study 5 writers in UC Berkeley Students' written journals, in-class interaction recordings, interviews, digital essay-related artifacts
Tardy (2005)	EAP course for graduate students	Habitus and identity; Genre	Identity development (disciplinarity and individuality) observed in slides	Presentation slides participants made for their own academic purposes (Modes: text, figures and tables, and style of slides)	Case study 4 international graduate students 20-month period (12 slides in total) Genre analysis

Developing multimodal writing tasks for L2 learners. TBLT researchers, in particular, have emphasized the necessity of conducting a needs analysis, including a systematic analysis of what students need to learn to perform adequate functions in the discourse community, which informs what types of tasks should consist of course content (Long, 2005, 2016). For a valid needs analysis, it is stressed that multiple sources (e.g., literature, leaners, domain experts, and applied linguists) and methods (e.g., interviews, surveys, and observations) must be incorporated because the interaction between sources and methods can triangulate data (González-Lloret, 2014; Long, 2005; Serafini, Lake, & Long, 2015; Van Avermaet & Gysen, 2006); however, Serafini et al. (2015) reported that less than a half of previous needs analysis studies used such interaction. Problematizing the lack of source and method interaction in many previous needs analysis studies, Serafini et al. (2015) provided an detailed example of utilizing this interaction in a needs anlaysis. The study aim was to build ESP courses for international post-docs and other professional researchers. For sources, they invited current international post-docs, researchers, domain experts, graduate students in applied linguistics taking TBLT seminar, and an expert applied linguistics; in terms of methods, they first conducted semi-structured interviews with some of the participants who were the insiders of the target language use domain and used the preliminary findings of the target tasks to construct surveys that were sent to a large number of insdiers. By recycling questions to different respondent groups, they identified that international students tended to be unaware of some detrimental influences of their lack of language skills on the work effectivness of which many domain experts (i.e., principal investigators) were aware. These critical functional deficiencies were then recommended to be considered in making training materials. Because this paper aimed to provide detailed description of needs analysis, the

authors did not provide how these findings were represented in the course content. In addition, this study was limited in describing language associated with the language use domain.

Other studies have also shown the multiplicity of sources and methods to increase the validity of needs. Chaudron et al. (2005) conducted a needs analysis to construct a course for learners of Korean at a tertiary level. To identify the task types, they interviewed a subset of Korean learners and collected survey data from all students enrolled in Korean courses at an institution. Using two methods to the same population, they were able to identify generalizable needs. Based on the findings, they preceded to the next step where they collect the language samples in target tasks. On the other hand, Malicka et al. (2017) incorporated two groups of sources (novice and experts to language use domain) and used semi-structured interviews. Their goal was to build a needs-based TBLT syllabus for future hotel receptionists. They particularly focused on the task sequencing thus needed domain novices (i.e., students in internships) to triangulate the task difficulty identified by a range of sources. Even though these studies did not include multiple interactions for triangulation, they provided good examples of why these interactions were useful for the particular contexts and aims of the studies.

While previous studies proposing needs-based syllabi focused on face-to-face speaking events in language use domain, González-Lloret (2014) placed a focus on the fact that current communication events often require technology. She asserted that the content of needs analysis for learners who are going to engage in technology-mediated contexts should cover language and technology and inform pedagogical language and technology tasks. Technology-mediated TBLT (González-Lloret & Ortega, 2014) grants comparable amount of attention to technology as language because performing adequately in current language use domains, which often engage technology, demands both language and technology skills. Except for this addition of

technology, the basic ideas of TBLT, including the necessity of needs analysis and language use analysis, remain intact in the technology-mediated TBLT framework. Even though multimodal writing does not necessary require digital literacy, multimodal writing is discussed mostly in the context of computer-mediated settings. This strong emphasis on technology in multimodal writing is also found from previous survey studies in other disciplines such as composition and communication (Anderson et al., 2006; Lutkewitte, 2010; Reid et al., 2016).

Anderson et al. (2006) conducted an online survey study in 2005 to investigate the teaching practices of multimodal writing in college composition classes. Their survey instrument included questions about the access to software, hardware, and supports for learning technologies for multimodal writing. Lutkewitte's (2010) dissertation project partially replicated Anderson et al.'s (2006) survey and focused on the teaching practice of multimodal writing in first-year composition courses. Even though these studies did not adopt the technology-mediated TBLT framework, researchers also considered technology as an inseparable component to contemporary multimodal writing. While Anderson et al. (2006) and Lutkewitte (2010) focused on the analysis of current teaching practice in a writing class, Reid et al. (2016) focused on the multimodal writing in different majors and conducted a survey study. They investigated the types of multimodal writing professors use across discipline in a large public university. They reported that science faculty used more multimodal writing for their own writing than the humanities and social science faculty; however, humanities and social science faculty gave more multimodal assignments to undergraduate students than science faculty. For undergraduate students, professors across disciplines indicated that presentations with visual/multimedia component is the most frequent text type and technical/academic multimodal writing was also frequent. They also thematically analyzed some open-ended questions and reported that professors agreed on the

prevalence of the multimodal writing while the conventions and genres of multimodal writing in academic contexts are in flux.

Investigating writing processes for task development. A needs analysis can inform what the learning goals and content should be included in curriculum, but researchers have expressed the challenges of translating the needs to the pedagogical tasks (e.g., Ellis, 2017; Malicka, Guerrero, & Norris, 2017), particularly for the sequencing of pedagogical tasks (Malicka et al., 2017). For task sequencing, cognitive task complexity dimension, which is one of the three components of Robinson's (2005) Triadic Componential Framework, has been utilized as criteria. Robinson (2005) proposed that the cognitive complexity of a task can be manipulated by the variables that affect cognitive and conceptual demands (i.e., resource-directing) and the ones influence the procedural and performative demands (i.e., resource-dispersing). Increased demands in resource-directing variables, such as more elements and spatial reasoning, let learners challenge different linguistic features, in turn help learners produce more accurate and complex language. Given this cognition hypothesis, Chaudron et al. (2005) determined that here-and-now variable (i.e., close and easy vs. far and hard directions) and the number of elements (i.e. number of purchase decisions) could determine task complexity. Malicka et al. (2017) asked domain novice and expert participants about the difficulty of the tasks they found from needs analysis to identify what features make tasks more complex. Based on Robinson's cognition hypothesis, they demonstrated a sample pedagogical unit about dealing with overbooking situation where they exemplified how two complexity variables (i.e., the number of elements and reasoning demands) can result in three pedagogical tasks. For example, the simplest task elicited learners to describe available hotel rooms to customers; the most complex task was to interpret a range of available hotels and the complaining customers' profiles

and talk to the customers about the overbooking situation. With these findings, Malicka et al. proposed that sequencing tasks from simple to complex tasks helps learners to ultimately practice language at the target-like situations. These studies, however, did not provide any evidence as to whether this principle was helpful in language development. It is possible that giving a task with challenging problems are resource-directing, thus facilitate language development.

According to Ellis (2017), many factors come into play in the cognitive complexity of output tasks. He noted that it is hard to determine whether the resource-dispersing and resource-directing variables work as they were anticipated to affect the cognitive process of production. Instead of relying on the putative variables that Robinson proposed, he recommended to have a theory that could explain why such variables interactively contribute to the cognitive complexity. As Ellis pointed out, previous studies on task complexity and sequencing focused speaking tasks and Levelt's (1989) model of speaking has been used to explain how task complexity variables affect task performance and language development (Skehan, 2016). In writing research, the model of cognitive writing processes (Linda Flower & Hayes, 1981; Hayes, 2012; Hayes & Flower, 1980) has explained how writing tasks affect writing processes and influence writing performance (e.g., Gánem-Gutiérrez & Gilmore, 2018; Johnson, Mercado, & Acevedo, 2012; Sasaki, 2000; Yoon, 2019). Kellogg's model of working memory in written composition (Kellogg, 1996; Kellogg, Whiteford, Turner, Cahill, & Mertens, 2013) also informed many writing studies that specifically focused on the task demands on the cognitive capacity with less concerns with knowledge in long-term memory (e.g., Ellis & Yuan, 2004; Johnson, 2017; López-Serrano, Roca de Larios, & Manchón, 2019; Révész, Michel, & Lee, 2019). This model explains that a writing system is composed of six basic processes: planning, translating, programming,

executing, reading, and editing. Each of the processes demands certain dimensions of working memory (i.e., spatial, central executive and verbal dimensions). This model explains that, because working memory has only limited capacity to hold and process writing plans, writers manage the attentional resources optimal for each writing process. However, it does not consider long-term memory that stores language and writing schemas and task environments, which are included in Hayes (2012) and Leijten et al. (2013).

Johnson's (2017) meta-analysis of the L2 writing studies that investigated the effects of manipulating cognitive task complexity variables on the linguistic features of essays, for example, is based upon the Kellogg's model of writing. He revealed that researchers preferred the number of elements and reasoning demands to manipulate resource-directing variables and planning time and topic familiarity to operating resource-dispersing components. For example, the positive effect of increased reasoning demands on lexical complexity was attributed to writers' attentional resources were directed to a translating process while such resources were not available to monitoring system. While this model with an exclusive focus on the working memory capacity has well addressed task complexity research and writing research, it has less capability to explain diverse undergoing cognitive processes from formulating writing plans with different schemas to translating these to output.

Gánem-Gutiérrez and Gilmore (2018) adopted Hayes's model for their study on L2 writers' processes when completing a timed argumentative writing task with an access to the Internet. They collected L2 writers' writing behaviors on computer screen and their eye gazes with an eye tracker, and manually coded writing processes (i.e., text construction, revising, rereading, use of external sources, and pausing) to identify duration and frequency. Additionally, they segmented each participant's video into five equal intervals to examine how writers'

processes change throughout task execution. They revealed that text construction and revising were dominant writing processes, but after three fifths writers spent more time on rereading and using external sources. In addition to its contribution to theory building and relevance to the current cognitive model of writing, methodologically this study shows a great example of utilizing screen capture videos of writing as to investigate online processes of writing without interruption.

To summarize, previous work on L2 writers' cognitive processes during writing tasks has employed different writing models that fit the scope of the writing processes researchers wanted to discuss. Because of the close relationship between the cognition hypothesis and the working memory model, Kelloggs' model has been popular in TBLT studies. However, Hayes and Flower's (1980) model—and its updated versions (Hayes, 1996, 2012; Leijten et al., 2013)—has provided a theoretical basis for many studies that painted a fuller picture of writing processes, including the writing plans, long-term memory and working memory (e.g., Gánem-Gutiérrez & Gilmore, 2018; Sasaki, 2000). Particularly for multimodal writing processes, the most relevant model of writing is Leijten et al.'s (2013) model of writing, an extended version of Hayes's (2012) latest model on empirical data. However, there has not been any attempt to examine multimodal writing in light of the cognitive writing processes. Because a fuller conceptualization of what processes undergo during writing processes, this study investigated what writing processes L2 writers demonstrate when completing a multimodal writing task that elicits linguistic and nonlinguistic modes of communication.

CHAPTER 3.

STUDY 1: A NEEDS ANALYSIS

Methods

Study 1 identifies multimodal writing tasks that international students may perform in their degree-pursuing undergraduate programs in the US and to explore design components of multimodal writing tasks. I adopted a qualitative approach and triangulated two data sources (i.e., instructor interviews and syllabi) and previous literature. Triangulating data from the two different sources concurrently helped increase the validity of the study. This needs analysis should be useful not only to L2 writers but also to L1 writers who need to learn academic English genres.

Study context.The Present Study

In the past two decades, researchers have delved into the functions of modes in different multimodal texts with different primary goals. Most studies focused on identifying the functions that different modes exercise in communication, which contributed to the understanding of new multimodal genres. When it comes to pedagogical implications, researchers have coherently demonstrated the importance of addressing nonlinguistic modes of communication in class. While this suggestion is timely and draws attention to the real-world writing, a crucial question is whether it applies to L2 learners at tertiary level whose goal is to develop language (Zhou, Busch, & Cumming, 2014; Manchón, 2017; Polio, 2019) and who would have developed skills for using other semiotic resources. Furthermore, despite the rising popularity of multimodal writing in L2 research, no empirical research has yet explored how multimodal writing can be integrated into current instructional practices.

The goal of the project is to examine the relevance of multimodal writing tasks to language learning and answer the following questions: To what extent does language contribute to multimodal task performance? How much time do students spend on language when doing a multimodal task? Do they care little about language when doing a multimodal task? I explored what multimodal writing assignments are used in undergraduate courses, devised a timed multimodal writing task for L2 writers and investigate how learner perform and perceive the non-traditional multimodal writing task. Furthermore, based on the latest cognitive model of writing by Leijten et al. (2013), I investigated L2 writers' processes of multimodal writing task execution by using on-screen writing behavior and stimulated recall data. Given that L2 writers' primary goal is on language, I shed light on students' processes and production of language while completing a multimodal task. The following research questions guided the current project:

1. What are undergraduate students' needs for multimodal writing?
2. How do L2 students perform a multimodal writing task compared to a monomodal writing task?
    2.1. To what extent is language related to the quality of L2 multimodal texts?
    2.2. What writing processes are demonstrated during L2 writers' multimodal task performance? Do L2 writers attend to language throughout the task?
3. How do L2 writers perceive multimodal writing tasks compared to a traditional task?

This dissertation project is an exploratory sequential mixed methods (Creswell & Creswell, 2018; Polio & Friedman, 2017) with two studies in the context of higher education design (QUAL → QUAN + QUAL). Study 1 addresses the first research question on the needs

of multimodal writing tasks in EAP classes[1]. Based on the findings of Study 1, I developed a timed multimodal writing task and investigated how students respond to the developed task in Study 2. The goal of Study 2 is to answer the second and third research question regarding college EAP students' multimodal writing task performance and their perceptions toward the multimodal task. I adopted a convergent parallel design (Creswell & Creswell, 2018; Polio & Friedman, 2017) in which qualitative and quantitative data collection occurred concurrently. Figure 4 summarizes the overall design of the project.

The context of Study 1 was U.S. higher education setting whereas participants of Study 2 were L2 writers attending Korean universities. Despite their geographical and contextual differences, such as the use of English as a foreign language and a second language, the two studies targeted users of English as academic purposes. All Korean participants of Study 2 indicated their experience in taking English-medium courses in which they participated in and performed academic tasks in English. By changing the site of study, Study 2 was able to offer implications that multimodal writing tasks, that had been believed to be less relevant to language learning than monomodal language tasks, can be useful for EAP learners across different language contexts. Therefore, by conducting two sequential studies, I aimed to achieve the goal of the current dissertation project which is to examine how multimodal writing tasks can be implemented in EAP classes in higher education. I present the two studies separately in the following two chapters (Chapter 3 for Study 1; Chapter 4 for Study 2) and discuss integrated findings in the final chapter.

[1] This needs analysis is published in the *Journal of Second Language Writing* (Lim & Polio, 2020).

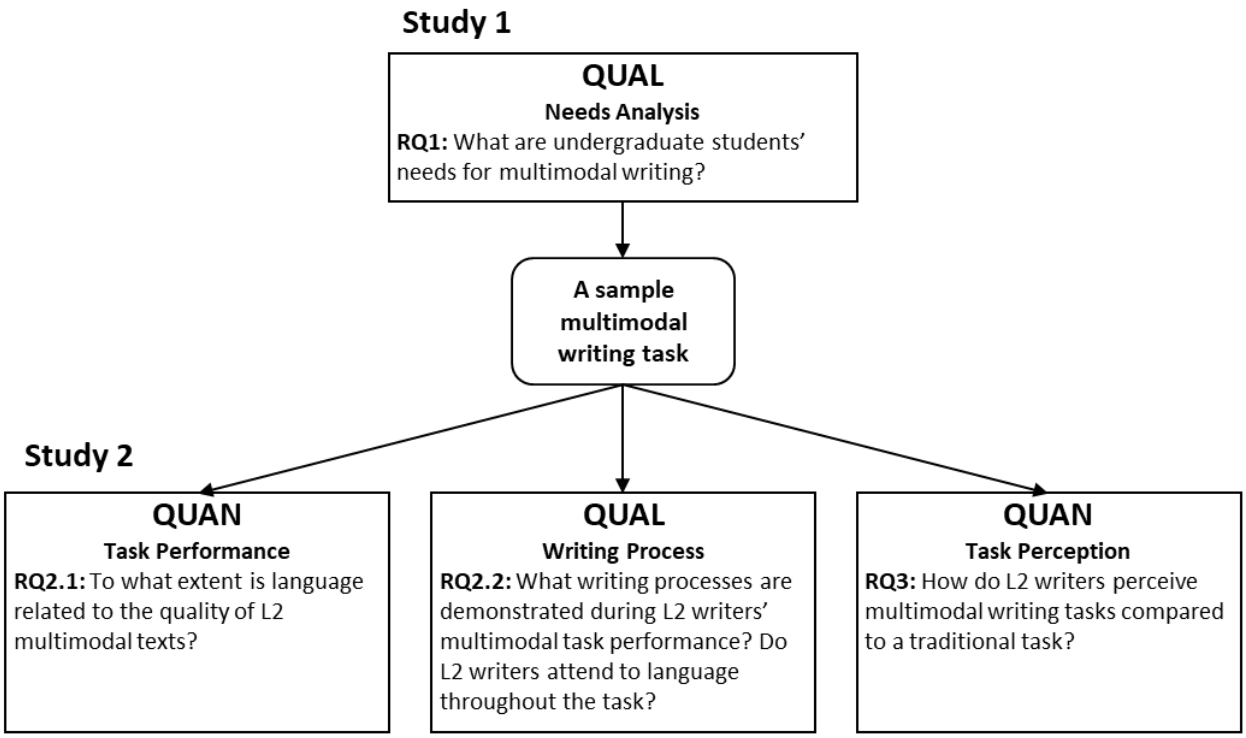

Figure 4. Overview of the research design

Study 1 was conducted at a US public university in Midwest where 3,862 international students (10.2% of undergraduate population) were enrolled in undergraduate programs during Fall 2018. In terms of their majors, according to the institutional report in 2016, about a quarter of the students identified their major as Business and about 20% were in Engineering majors. Other popular majors were in Nature Science and Social Science; and students in Education and Humanities consisted about 6% of the population.

Participants. Instructors and faculty members who taught undergraduate courses across disciplines were recruited for an individual semi-structured interview on undergraduate course requirements and assignments. I interviewed seven professors who taught undergraduate courses in the following disciplines: Education (n=3), Engineering (n=2), Business (n=1), and First-Year-Writing (n=1). All participants had experience having international students whose first languages are other than English; the instructor of the first-year-writing program had more extensive experience in designing and teaching courses for English language learners.

Semi-structured interviews. The interview protocol was designed to elicit instructors' descriptions of their course syllabi and major assignments (e.g., instruction and grading criteria); and to address their thoughts on the similarities and differences of multimodal tasks and formal writing tasks (see Appendix A for the interview protocol). Each interview lasted about 45 minutes, and all interviews were voice-recorded and fully transcribed. Instructors provided sample syllabi and course materials, which were not included in the syllabi data to avoid overlaps.

Course syllabi. I collected and analyzed 161 undergraduate-level course syllabi from four disciplines: Education (n=25; Teacher Education), Science (n=11; Chemistry, Physics), Engineering (n=25; Mechanical Engineering, Computer Engineering), Social Science (n=41;

Economics, Political Science, Psychology), and Humanities (*n*=59; Philosophy, Writing). Most of the syllabi included in the current study were publicly available on the department websites in the four colleges at the moment of data collection; the business school declined to share their syllabi. Because of the convenience sampling, our findings may not representative of the entire university.

Data analysis and triangulation. All audio-recorded interviews were fully transcribed; and all materials including artifacts provided by the instructors and the separate dataset of course syllabi were imported to qualitative analysis software MAXQDA. I conducted a thematic analysis (Braun & Clarke, 2006) on the interview data without a pre-existing coding scheme and thus looked to see what themes emerged.

Initial analysis on the syllabi data was to identify multimodal writing tasks from the description. Based on the operationalization of multimodal writing for the current study, I coded for any assignments that explicitly stated the inclusion of multiple modes including written English word. Exclusion criteria included (1) only one mode such as texts (e.g., two-page double-spaced essays on previous experience) or computer language (e.g., code for computer program); and (2) multiple modes but no English text (e.g., an excluded assignment in Chemistry was to draw a picture and a formula of a chemical compound). As a result, I identified 104 multimodal writing tasks from Education (*n* = 38), Humanities (*n*= 42), Science (*n* = 7), and Social Sciences (*n* = 17). None of the tasks in the Engineering syllabi data were described in enough detail to determine if they were multimodal tasks.

The second round of coding was to triangulate the themes that emerged from the interview data. After identifying multimodal tasks and finding themes from the interview data, I

re-examined the 104 multimodal tasks. Appling the three themes identified from the interview data, I investigated whether these issues were relevant and applicable to the task classification.

Results

Three themes characterized multimodal writing tasks in academic contexts and could be considered in developing multimodal tasks for research or pedagogic purposes: (1) Goals and instruction of multimodal writing: disciplinary versus creative expressions; (2) Linguistic mode in multimodal texts; and (3) Tasks of multimodal writing: individual versus collaborative work.

Goals and instruction of multimodal writing. It was commonly indicated by the interviewers that the main goal of multimodal writing is to communicate an intended meaning to a target audience. However, I identified two different functions of multimodal tasks. One is for students to understand and meet the audience's expectations of academic genre conventions (i.e., disciplinary expression), and the other is to have students experience various modes and mediums for creative production (i.e., creative expression). Unlike multimodal tasks for disciplinary expressions, those for creative expressions were found not to have clear expectations or conventions to follow.

I found that tasks for *disciplinary expressions* were mostly structured with explicitly stated preferred styles and components prevalent in a specific discipline. Examples of such assignments included papers containing the presentation of data, PowerPoint slides for an in-class oral presentation, and lab reports. For these tasks, students were expected to demonstrate their ability to follow sets of established conventions and rules. The conventions were not always explicitly written, but occasionally listed as requirements. For example, for a technical report for a senior-year Engineering course, an instructor used two class sessions to illustrate what he expected students to do for the report and to provide feedback on their report drafts. In a similar

course, another Engineering instructor asked students to research target multimodal products (e.g., posters) and follow some genre conventions:

> *I don't actually tell them how to present. I give them a number of websites that talk about preparing posters... I leave these [posters] up and it's like go look at them. What works what doesn't. Critique it, think about it, critique it, critique it amongst your group and then use that information to inform your questions. The posters have gotten better over the years I think because of that something they can learn from doing*. [Engineering Instructor 1]

Another way to focus on disciplinary practice was to provide a detailed description of the components that students should include in their final outcome. For a lab note assignment in Science, for example, students had to include "coversheet, data, formulae, and graphs based on the data." Instructors often provided templates for the students. From the syllabi, I found 74 tasks were designed with specific disciplinary conventions (see Table 5). In-class presentations based on course readings and students' own papers (n = 31) and data analysis papers (n =13) were common types of assignments across the disciplines.

Table 5.

Multimodal Writing Tasks in Two Approaches

	Education	Humanities	Science	Social Science	Total
Creative expressions	8	22			30
Essay to visual representation	2	10			12
In-class presentation	3	3			6
Video*		4			4
Online discussion posts	1	2			3
Reflection project	1	1			2
Portfolio		1			1
Paper (data analysis)	1				1
Journals/lab notes/field notes		1			1
Disciplinary expressions	30	20	7	17	74
In-class presentation	12	12		7	31
Mini lesson and lesson plan	15				15
Paper (data analysis)	2	3	2	6	13
Journals/lab notes/field notes		1	5	2	8
Professional webpage		2			2
Video resume		1			1
Online discussion posts**	1				1
Others (map, art description for an exhibition, diagram)		1		2	3

* Video assignments focusing on creative expressions include documentary, resemiotization tasks, and a promotional video.

** Unlike other three online discussion assignments, one assignment specified the structure of the post.

Thirty assignments, on the other hand, focused on students' achievement of rhetorical goals, with little attention to disciplinary practice. These tasks inducing the writer's purposeful choices of nonlinguistic and linguistic modes enable students to express their ideas in creative methods. Tasks for creative expressions were found from the syllabi of Humanities (n = 22) and Education courses (n = 8) (see Table 5). It should be noted here that one type of multimodal

writing does not necessarily have one exclusive function of disciplinary expressions over creative expressions, or vice versa. For example, an in-class presentation with slides was one of the common multimodal tasks (n = 37); for 31 presentation assignments, students were given a particular format to follow (e.g., a conventional academic presentation), while they were given medium options for six presentation tasks (e.g., a skit, a video, a poster, presentation slides).

In a first-year writing class, through a "digital remix project," which is coded as "essay to visual representation", students transform linguistic texts they previously wrote into multimodal texts such as "a video, a photo essay, a poem, a web page, a painting, a poster, a collage." Through this resemiotization process, as illustrated in the following interview excerpt from its instructor, students are expected to raise their awareness of the affordances of different modes and use linguistic and nonlinguistic resources strategically to achieve their rhetorical goals:

> *We'll talk about the ways that different forms operate and how they have other things in. Are you using that? So one of the questions might be, are you fully using the tools of this new genre... And then there's always the understanding. Is it clearly understandable? Is the music too loud? Did you do your words too fast on the screen so nobody could read them? Are there parts of it people don't understand because they don't come from your culture?* [Writing Instructor 1]

Multimodal tasks for creative expressions allow students to explore different modes, but they can be perceived as overly challenging for students without an explicit provision of new authoring tools and resources for them (Cimasko & Shin, 2017). In this regard, an instructor of Education-major courses indicated that pre-service student teachers needed more assistance and preparation for multimodal task performance than she had thought earlier:

> *So we saw something I think we need to work on that course actually is the video crafting part because we think the students, they are part of a particular generation and we think that they come in already knowing how to use technology. And actually a lot of our students don't know a lot. Some of our students don't know how to use Google Docs.* [Education Instructor 3]

There were, however, some instructors who considered multimodal writing tasks to be easier than formal writing tasks. For example, an instructor from Education mentioned that a good essay requires "another level of skill set" that is beyond what is needed for effective multimodal performance such as creativity and abstract thinking. Another instructor said that students will be able to perform well on a multimodal task as long as they comply with its guidelines; thus, poor multimodal performance can be interpreted as a lack of investment.

Linguistic mode in multimodal texts. The second theme mainly involves how linguistic mode is used and interplay with other nonlinguistic modes. I found that the multimodal writing tasks required either written words or a mixture of written and spoken words (e.g., a written script for an in-class presentation with slides). For the assignments that required written words in the final product, it was anticipated that authors create a visual presentation of data analysis (n = 27; e.g. graphs, diagrams, tables) (see Table 6). While the visual mode was somewhat dominantly used with written words, other modes were also available in some assignments such as building professional webpages and interactive maps for which writers could utilize spatial and aural modes. Except for three assignments out of 32, multimodal writing assignments that did not elicit spoken words focused on promoting disciplinarity. More specifically, students were given some specific guidelines for visualizations that include layout and formatting. While

academic papers were found to be the most popular multimodal text type in the context of this study, they still had a heavy reliance on linguistic resources to convey information.

Table 6.

Linguistic and Nonlinguistic Resources Anticipated from the Multimodal Writing Tasks

	Creative expressions	Disciplinary expressions	Total
With written words	3	29	32
Data presentation (graphs, diagrams, tables)	2	25	27
Webpage		2	2
Medium of the writers' choice	1		1
Art pieces at an exhibition		1	1
Interactive map		1	1
With written and spoken words	27	45	72
Visual aids for presentation (e.g., slides)		42	42
Medium of the writers' choice*	21		21
Video	5	2	7
Poster	1	1	2

**Note.* Six of them should be shared online, thus limit some non-digital resources (e.g., interactive gestural and spatial modes).

The other category of multimodal writing assignments required written and spoken words as well as nonlinguistic modes including visual, aural, spatial, and gestural modes. In terms of the linguistic mode, spoken words tends to be a more dominant method in meaning making than written words (e.g., spoken narration and written caption in a digital story); however, it should be noted that the spoken words are expected to, and sometimes required to, be rehearsed in written words. For example, a script is either read naturally for recordings (e.g., digital storytelling) and practiced for an in-class academic paper presentation with slides. During the interview, an instructor of Business explained how a nonnative speaker of English was assisted by his group members to write a script for presentation in the classroom, and this collaborative preparation

made his presentation qualitatively better than his earlier presentations. Furthermore, students in an Education course were required to turn in the script as shown in the following excerpt:

> *We are looking at sort of the images and the script and how those things interplay together and really like that stuff is particularly on the rubric for that multimodal project because they have to do video they have to do the sound over.* [Education Instructor 3]

When an assignment required written and spoken words coupled with visual aids for presentation, it focused on the expression of disciplinary voices (e.g., academic presentation based on data analysis); on the other hand, when an assignment allowed a student to choose other media to supplement written and spoken words, it tended to promote the student's creative expressions (e.g., digital story, performance, skit). It was also observed that assignments using posters (n = 2) and videos (n = 7) as media could be designed to serve both functions. For example, a resemiotization assignment focusing on fostering creative expressions limited the final medium to poster presentation. Another poster I found was a part of academic poster presentation. In this case, as D'Angelo's (2010, 2016) showed, students are expected to use nonlinguistic resources with linguistic resources (e.g., words woven into visualization, juxtaposition of graphics and linguistic texts, and font styles for information hierarchy). The following quote by an Engineering instructor specifies grading criteria for poster presentations. That is, for the successful competition of academic posters, students should be able to use graphics coherently in terms of relevance, layout, and color schemes, and make them easily readable, aiming to achieve the interplay between these different modes for successful communication:

People don't want to sit there and read a whole article. Better when it's bulleted because it's just easier… they've got some pictures here that you know their pictures are relevant. They're one of the things that I would really emphasize you know with the graphic is to make it easy to read... I say keep everything across the board consistent to make it easy for your reader because otherwise you get lost in that looking after code every time. [Engineering Instructor 1]

When it comes to the task sequence, we found that oftentimes multimodal writing is accompanied by other writing tasks (see Table 7). Sixty multimodal tasks were stand-alone; seventeen tasks had pre-writing tasks to inform the multimodal writing process such as discussion posts, essays, papers and presentation. For 32 multimodal tasks, we found post-writing tasks, mostly in the form of reflection essays. Given that prose writing might be a subsequent step following the production of multimodal tasks in undergraduate courses, this might be an important task to integrate into ESL courses.

Table 7.

Pre- and Post-Tasks for Multimodal Writing Tasks from the Syllabi Data

Pre-task (*n*=17)		Post-task (*n*=32)	
Discussion posts	2	Essay	2
Essays	3	Paper	7
Paper	7	Presentation	2
Presentation	1	Reflection	21
Proposal	4		

Tasks of multimodal writing: Individual versus collaborative work. The majority of multimodal tasks in academic settings have been designed to involve individual writing performance, as evidenced by the course syllabi. Of the total of 104 tasks described in the syllabi, 19 tasks were group work, and six tasks were individual performance followed by the

initial stage of collaborative writing; the remaining tasks (n = 80) were identified as individual tasks (see Table 8). I found that many assignments focusing on creative expressions involved individual performance (n = 23), which might indicate that such multimodal assignments were designed to encourage individual writers to make authorial choices on mediums and modes. All of the resemiotization task (i.e., essay to visual representation), for example, were described as individual work.

Table 8.

Authors of Multimodal Tasks Focusing on Creative and Disciplinary Expressions

	Creative expressions	Disciplinary expressions	Total
Individual	23	57	80
Group	7	12	19
Group and individual work*		5	5

* *Note*. Individual reports based on problem-solving activities in groups where writers build outlines and notes together (e.g., lab reports).

It is worth noting that 17 out of 24 collaborative assignments had their focus on disciplinary expressions. An instructor in Engineering reported that such collaborative assignments were expected to help students "get groomed toward producing this [professional] level of expectation." The instructor further presented the rationale for developing collaborative multimodal writing projects:

> *So the way that it's divided is set on a project. There are actually seven different types of roles that they could take. So I have engineers. They are doing buildings, they're doing foundations and transportation engineers that do parking areas. And they all work together on the same big project but they're only going to report on their specific part of it*. [Engineering Instructor 2]

As shown in the excerpt, a course for senior-year students would mimic an authentic project that requires collaborative problem-solving and that each student is responsible for reporting the part they were in charge of. While each member composes their own segment, they all collaborate to make a coherent technical report eventually.

CHAPTER 4.

STUDY 2: MULTIMODAL TASK IMPLEMENTATION

Methods

I adopted a convergent parallel mixed methods design (Creswell & Creswell, 2018; Polio & Friedman, 2017) to provide a holistic view of how L2 writers respond to a multimodal writing task in terms of multimodal text quality (RQ2.1), cognitive processes (RQ2.2), and perceptions (RQ3). Figure 5 provides a summary of Study 2 research design.

Participants. A total of thirty-one adult learners of English for academic purposes in a Korean university participated in this research (Age M = 22.84, SD =2.52). To be eligible for participation, they were asked to bring a copy of TOEFL or IELTS test transcripts showing above or equal to B2 level in Common European Framework of Reference (e.g., iBT TOEFL 72 and IELTS 5.5). I set this requirement to ensure that students have experiences in taking timed-writing tests and be able to produce some written responses. All of them had experience taking English-medium courses (1–2 courses, n = 2; 3–4 courses, n = 3; 5–6 courses, n =6; above 6, n = 20) and 17 of them had study abroad experience (less than a year, n = 9; 1-2 years, n = 3; more than three years, n =5). Participants reported that they had experience completing academic writing without visual components (e.g., argumentative essay), academic writing with visual components (e.g., academic posters), presentations with slides for their courses. More details per participant are available in Appendix A.

Not all writers' data were used for the RQ2. For RQ2.1, I used data from 29 participants because of technical issues in the screen capture videos (P17, P31). Investigating multimodal writing processes, I analyzed 12 participants' stimulated recall and writing behavioral videos. All responses from 31 participants were explored for RQ3 on writers' task perceptions.

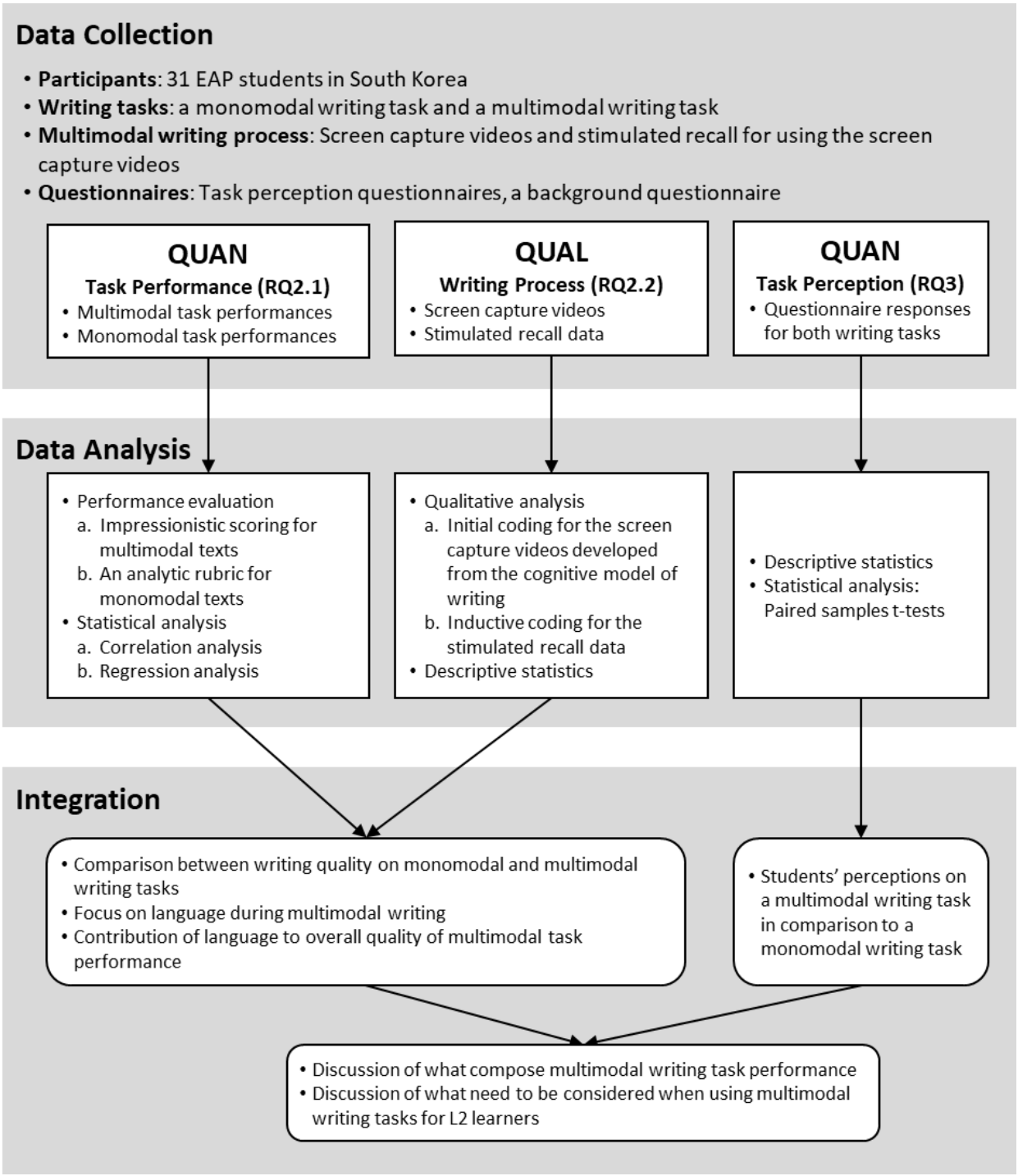

Figure 5. The convergent parallel mixed methods design of the Study 2

Instruments. Participants completed two timed writing tasks: a monomodal writing task and a multimodal writing task that I developed as a result of Study 1. While performing the tasks, their on-screen writing processes were screen-recorded; after the multimodal writing task, they watched their videos for stimulated recall interviews.

Monomodal writing task. For the traditional task, I used one of the prompt from Yoon and Polio (2017) that elicits the argumentative genre and gave participants twenty minutes to complete the task. The task was delivered on Microsoft Word without spell-checking function and they had no access to the Internet. Following is the full instruction for the monomodal writing task:

> *Many college students now carry and use their laptops while taking a class. What do you think about students' use of laptops in the class? Write an argument supporting your opinion. You should spend about* ***20 minutes*** *on this task.*

Multimodal writing task. In this task, students composed a three-minute slide presentation with voice recording (i.e., narrated slide presentation) on Microsoft PowerPoint as an authoring platform. Similar to the monomodal writing task, this multimodal writing task was a timed task (50 minutes); and the prompt focused on a topic of technology and elicited argumentation. The prompt was as follows: *Does technology make us more alone? Plan and make a 3-minute video showing your position using examples*. Participants were able to use any external resources; in addition, given that the time constraint limits participants' time for data search, I provided three links to webpages that include information relevant to the topic. Figure 6 shows the task prompt delivered on PowerPoint platform. As I met each participant individually, I explained each participant important functions of the authoring tool (e.g., how to record their voice on PowerPoint) and responded to questions before beginning the task.

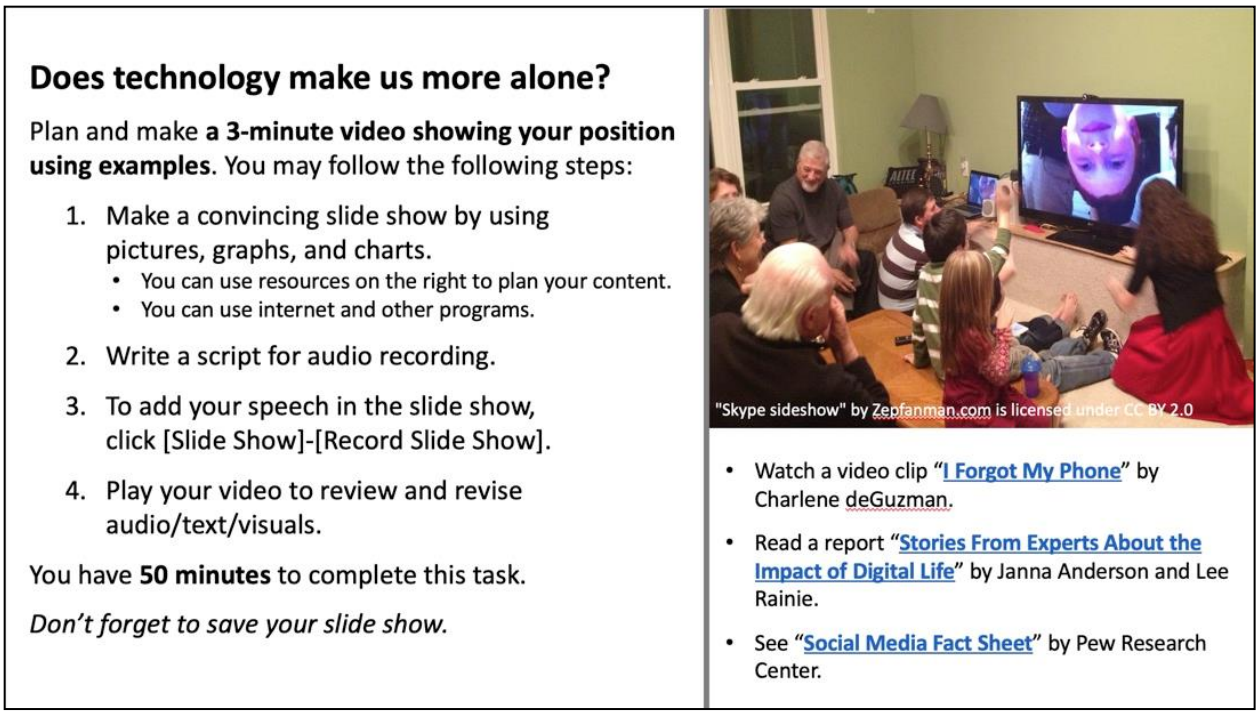

Figure 6. Multimodal writing task for Study 2

The multimodal writing task allows writers to freely use online resources; however, while completing the monomodal writing task, participants were not allowed to use any online resources. I kept the traditional format for the monomodal writing task in order to interpret students' monomodal writing proficiency that is measured through other conventional writing tests (i.e., independent timed writing tasks). One caveat to this methodological decision is that such may limit the comparability of the two writing tasks.

On-screen writing behaviors. While completing the multimodal writing task, participants' writing behaviors on screen were recorded using a screen recording software *Snagit*[2]. The length of videos on average were 52 minutes and 57 seconds with a standard deviation of 5 minutes and 43 seconds. All videos files were saved to be used as a prompt for

[2] Downloadable from: https://www.techsmith.com/screen-capture.html.

immediate stimulated recall interviews and a separate data source for an investigation of writing behaviors. Twenty-nine participants' video data were available for further analysis.

Stimulated recall interviews. Because think-aloud protocol can increase the cognitive burden to the writers, especially when they are completing a task eliciting multiple modes, hence can impact performance, I adopted a retrospective interview for verbalization. After completing the multimodal writing task and a short task perception questionnaire for the task, each participant reported their thoughts when completing the task. To help retrieving their thoughts at the time of writing, I replayed the screen capture of the multimodal writing. For practicality, videos were speeded up to three times of the original speed. Participants were asked to stop when they wanted to explain. At pauses, I prompted participants to verbalize their thoughts. Verbal protocol for the stimulated recall was as follows:

> *What we're going to do now is watch the video. We are interested in what you were thinking at the time you were responding to the task. We can see what you were doing by looking at the video, but we don't know what you were thinking. So what I'd like you to do is tell me what you were thinking, what was in your mind at that time while you were composing this text. You can pause the video any time that you want. So if you want to tell me something about what you were thinking, you can click pause. If I have a question about what you were thinking, then I will push pause and ask you to talk about that part of the video. Your verbalization will be audio-recorded with the video play so that researchers can find what your verbal report referred to.*

Participants used Korean, their first language, for the stimulated recall interviews. All interviews were recorded using a screen recording software. The analyzed stimulated recall data were videos of a voice overlay on the original writing video; and this data showed participants'

retrospective verbalization as they watched their writing behaviors. On average, a videoclip lasted 23 minutes and 40 seconds with a standard deviation of 4 minutes and 59 seconds.

Task perception questionnaires. Following previous studies that argued for the use of explicit measures of learners' task perceptions to compare the cognitive demands of tasks (Sasayama, 2016; Yoon, 2019), this study asked participants to answer six questions on the tasks immediately after completing each task. The six nine-point Likert scale questions were derived from Yoon (2019) which looked into the task manipulation and genre effects on learners' perceptions and performance on timed writing tasks. Each item measured task complexity, difficulty, anxiety, confidence, interest, and motivation. Marking one for the task difficulty item, for example, indicates the task was not difficult at all; marking nine for the item indicates that the task was extremely difficulty. The full items are as follow:

This task required no mental effort at all.	1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9	This task required extreme mental effort.
This task was not difficult at all.	1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9	This task was extremely difficult.
I felt really relaxed doing this task.	1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9	I felt frustrated doing this task.
I didn't do well on this task.	1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9	I did well on this task.
This task was not interesting at all.	1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9	This task was very interesting.
I don't want to do more tasks like this.	1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9	I want to do more tasks like this.

Background questionnaires. The background questionnaire was composed of two parts: (1) academic background and multimodal writing experience; and (2) self-assessment of communication in L2. The first part consisted of five open-ended questions that ask participant' majors, years in the college, previous experiences in taking courses in English. The second part includes thirty 9-point Likert scale questions that ask the participants' evaluation on their own communication skills. Twenty-five questions that address reading (k = 9), listening (k = 5), writing (k = 5), speaking (k = 3), and grammar (k = 3) were from Mills and Moulton (2017) which adapted ACTFL's can-do statements to measure the language goal achievement. While they used 5-point Likert scale, I changed the scale to 9 because the greater number of options can reveal the distance between categories and improve reliability (Adelson & McCoach, 2010; Wakita, Ueshima, & Noguchi, 2012). I added five questions that address multimodal communication (see Appendix C for the background questionnaire). I used Qualtrics software for data collection platform. In this dissertation, I did not include any analysis of the self-assessment data.

Data collection procedure. I met each participant individually for the experiment. After collecting participants' consent, students began one writing task and completed a task perception questionnaire. After a 10-minute break, students began the other writing task and completed a task perception questionnaire. Then they completed a background questionnaire. To avoid an order effect, 16 participants completed the multimodal task first and the other 15 participants started with the monomodal task. After the questionnaire for the multimodal task, there was a stimulated recall interview. In total, participation lasted about 2.5 hours. All participants received monetary compensation ($40) after the completion of all tasks. In addition to Table 9 that lists the order of research activities, I report a full counterbalancing schedule in Appendix B.

Table 9.

Counterbalanced Data Collection Procedures

Group A		Group B	
Tasks	Time (min)	Tasks	Time (min)
Consent form		Consent form	
1. Narrated presentation task	50	1. Essay writing task	20
2. Task survey	5	2. Task survey	5
3. Stimulated recall	30	3. Stimulated recall	30
4. Break	10	4. Break	10
5. Essay writing task	20	5. Narrated presentation task	50
6. Task survey	5	6. Task survey	5
7. Background questionnaire	15	7. Background questionnaire	15
Compensation		Compensation	

Evaluating performance. To answer the second research question on the interpretation of multimodal writing performance, both monomodal and multimodal tasks were scored by raters who had extensive English teaching and test rating experiences.

Monomodal task performance. Three native speakers of English who have extensive experiences in teaching English for academic purposes evaluated the monomodal texts using an analytic scoring rubric by Connor-Linton and Polio (2014). The analytic rubric consists of five categories, including content, organization, vocabulary, language, and mechanics (See Appendix D for the rubric). Maximum possible scores of the categories are 20 point, except for the mechanics score at half weight (10 point). There was a rater training and norming session using other argumentative essays. Raters individually scored the essays which were presented in random order. Interrater reliability was acceptable (Cronbach's α for total score = .85, content = .86, organization = .80, vocabulary = .78, language = .86 and mechanics = .76) and the mean

score of each text was used for further analysis.

Multimodal task performances. To evaluate multimodal texts, I invited two additional raters with EAP teaching experience, so a total of five native speakers of English provided scores for the multimodal texts. They were asked to provide scores ranging from 1 to 9 for the overall quality, visualization, and verbal delivery and language. Other categories that are considered in Hung et al. (2013) were not considered because participants did not use additional sounds nor gestures for meaning construction. Likewise, the order of multimodal texts was randomized. Even though there was no explicit attempt to norm the grading criteria, interrater reliability was acceptable for the two scores on specific focuses Cronbach's α for visualization = .76, verbal delivery and language = .84) and the overall quality of the multimodal texts (Cronbach's α overall quality = .74).

Table 10.

Descriptive Statistics of Monomodal and Multimodal Task Scores (n = 29)

Category (Total possible score)	*Mean*	*SD*	*Min.*	*Max.*
Monomodal task				
Content (20)	14.84	2.22	8.67	19.00
Organization (20)	14.90	2.06	9.33	19.33
Vocabulary (20)	14.77	1.79	10.67	18.00
Language (20)	14.33	1.87	11.00	18.67
Mechanics (10)	8.23	0.80	6.00	9.50
Total (90)	67.06	8.35	47.33	83.67
Multimodal task				
Visualization (9)	6.55	1.25	2.60	8.90
Verbal delivery and language (9)	6.51	1.31	3.60	8.70
Overall quality (9)	6.58	1.16	3.40	8.40

Unfortunately, two files of the multimodal texts had damaged sounds (P17 and P31) with which raters could not evaluate on the language or overall quality. Thus, further analyses for the first and second research questions were limited to twenty-nine participants' performance. The descriptive statistics of monomodal and multimodal text qualities are reported in Table 10.

Coding writing process data. I used a pro version of MAXQDA, a qualitative data analysis software, for all data analyses of writing process data (i.e., videos of writing behaviors on screen and videos of stimulated recalls). I imported all video data into the analysis software that is equipped with a function to directly code on the video files and a replay of video in sync with transcript. I provided specific examples of the functions I used in the following sections. For both data analyses, I took an inductive approach and developed a coding scheme that can explain the data in reference to the cognitive model of writing (Leijten et al., 2013).

To be more specific, after reviewing the collected data, I found it necessary to develop separate coding schemes for the on-screen writing behavior data and the stimulated recall data. Video of writing behaviors provided information regarding how much time was spent on observable behaviors, but it could be only speculated in connection with particular writing processes. For *Proposer*, plans are non-verbal at the time and exist only in the participants' minds; for *Translator*, such ideas are formulated in the mind, but not appeared on screen; for *Evaluator*, the overall changes may give some clues about what the student is doing, but only their retrospective interview can provide valid data about what they have done. On the other hand, stimulated recall data cannot provide temporal information but they reveal what participants were thinking and can be linked to the process- and control-levels of the cognitive model of writing. Triangulating two data sources provided quantifiable data in terms of

observable behaviors that provide the magnitude of each observable process and qualitative data that speak to the cognitive processes.

Writing behaviors on screen. First, I segmented each video into writing process episodes, each of which represent a segment reflecting a writer's switch to a different writing process as defined and used in Gánem-Gutiérrez and Gilmore (2018). After reviewing all episodes in conjunction with the latest cognitive model of writing (Leijten et al, 2013), I developed a coding scheme consisted of seven categories describing observable processes (see Table 11). The observed writing behaviors cannot be one-to-one match with writing processors of the cognitive model of writing, but certain processors are likely to control writing behaviors. In Table 11, I marked the writing processors and task environments that are hypothesized to operate the observable writing behaviors. Mostly, the writing behavior data were generated from *Searcher*'s interaction with *Task-related sources, written plans* and from the interaction of *Transcriber* and *Translator* to *Production technology* and *Text-and-graphics-created-so-far*.

Figure 7 shows a screenshot of how I used the analysis software for this coding. Each small square, a *coded segment*, indicates one writing process and its color represents a code. After coding, I exported time information for the codes from the MAXQDA using *Code Coverage* function. This data provides exact starting and ending time for coded segments. I divided each participant's video into five equal intervals (period) to investigate how writing behaviors change as multimodal texts develop. For example, a 50-minute video was cut into five 10-minute videos; a 40-minute video was divided into five 8-minute videos. Then I converted time spent for a behavior into a percentage of a total time within each period. This conversion was to control the fact that each participant spent different amount of time on task. For example,

two minutes for an 8-minute interval (25%) is more than 2 minutes for a 10-minute interval (20%). I reported descriptive statistics with graphs for results.

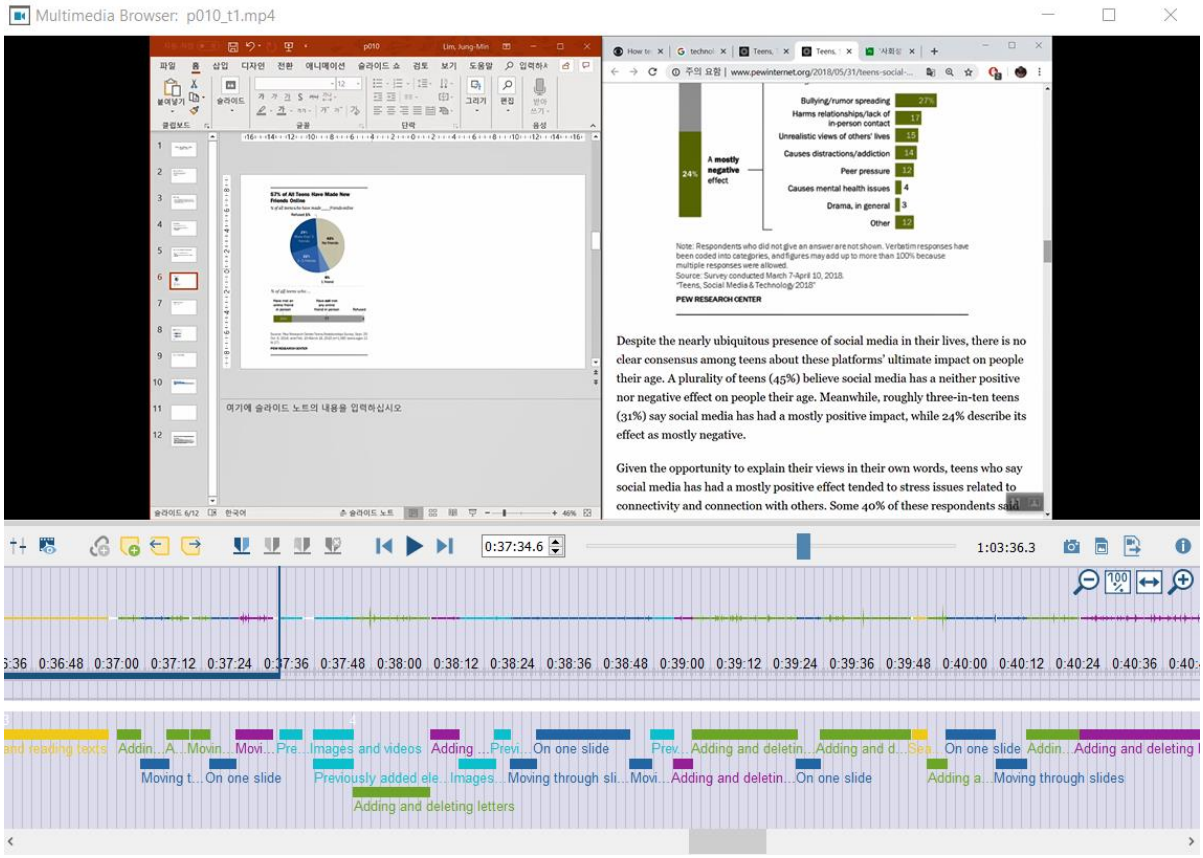

Figure 7. A screen capture of the writing behavior coding

Stimulated recall data. First, I fully transcribed the video-recorded recall verbatim and imported both video data and transcripts with time information to MAXQDA. Like the writing behavior data, I segmented transcripts into the shortest units that contains one idea (i.e., Plakan's (2009) idea units). A coding scheme was developed to explain participants' writing processes while completing the timed multimodal writing task. It contains seven categories that are hypothetically influenced by control- and process-level cognitive activities. Three categories provided evidence for current plan, writing schemas and design schemas; four categories showed how *Evaluator* and *Searcher* interact with task environments. In Table 12, I listed the codes

emerged from the data and how each of them corresponds to the cognitive model of writing. For example, the following translated excerpt from a participant contained two idea units as described in the parentheses:

> *What I was thinking is, when I do tasks like these, I make structures.* [explaining writing schema] *Because just listing my three reasons wouldn't be really logically convincing. So I wanted to structure this text discuss to whom technology influence, or how different kinds of technology affect people.* [explaining current plan]

After coding the stimulated recall data, I found that this elicitation method with the video stimuli prompted them to verbalize their plans at the time of writing. I thus further explored the coded segments explaining writers' current plans and divided them into two categories: explaining current plans for contents and explaining current plans for writing processes. When explaining their contents, they explained what they planned to include, which was sometimes realized in the texts or changed to another plan as they proceed. Their current plans for writing processes involved their metacognitive strategies to complete the timed multimodal writing task.

Inter-rater reliability. I recruited a second coder who is a native speaker of Korean and holds a Ph.D. in applied linguistics with a research focus on second language writing. The second coder coded a subset of the writing behavior data and the stimulated recall data. First, I trained the second coder with one participant's stimulated recall data. Using the coding scheme, the coder analyzed five participants' stimulated recall data (42% of the participants) and the interrater reliability was acceptable (kappa = .72, percentage agreement = 75.31%). Next, I delivered another norming session for writing behavior data coding. The second coder analyzed three participants' writing behavioral videos using the coding scheme, and the interrater reliability was high (kappa = .89, percentage agreement = 90.42%). I coded the rest of the data.

Table 11.

Coding Scheme for Writing Behavioral Data and its Relevant Writing Processes in the Cognitive Model of Writing

	Process level	
	Writing processes	Task environment
Reading text-constructed-so-far e.g., *playing slideshow; moving through slides; staying on a slide; playing recorded narration*	Evaluator	Task-related sources, written plans; Text-and-graphics-created-so-far
Searching internet e.g., *dictionary, reading texts, still images, clip arts, data summary*	Searcher	Task-related sources, written plans
Accessing provided resources e.g., *task prompt slide and links to provided webpages*	Searcher	Task-related sources, written plans
Editing words on the slides e.g., *adding and ordering slides; adding and deleting letters (in L1 and L2) on a slide; moving to another slide change words in text-constructed-so-far; pasting others' texts to a slide*	Translator, Transcriber	Production technology; Text-and-graphics-created-so-far
Editing words on the slide notes e.g., *adding and deleting letters on a slide note (in L1 and L2); moving to another slide note change words in text-constructed-so-far; pasting others' texts to slide notes*	Translator, Transcriber	Production technology; Text-and-graphics-created-so-far
Speaking e.g., *reading out loud; speaking without script*	Translator, Transcriber	Production technology; Text-and-graphics-created-so-far
Editing visual elements e.g., *animation, design functions, previously added elements, video link, tables, images, fonts, objects*	Translator, Transcriber	Production technology; Text-and-graphics-created-so-far

Table 12.

Coding Scheme for Stimulated Recall Data and its Relevance to the Cognitive Model of Writing

	Control level	Process level	
		Writing processes	Task environment
Explaining current plan for process and for content e.g., *I had 15 minutes left so I had to start recording; My opinion was that...*	Current plan		
Explaining writing schemas e.g., *Because presentation should have a thank-you page...*	Writing schemas		
Explaining design schemas e.g., *My criteria is to choose real pictures, not a pictogram. Using pictograms gives an impression that the content is not reliable. Who uses them for serious business meetings?*	Design schemas		
Evaluating source texts e.g., *I thought, great, this page had objective information to use*.		Evaluator	Task-related sources, written plans
Evaluating own text e.g., *I figured this was off-topic and panicked*.		Evaluator	Text-and-graphics-created-so-far
Searching for language e.g., *This word [assume] is not a perfect match to what I think. I tried to think about another option...*		Searcher	Task-related sources, written plans
Describing technology difficulties e.g., *It took much longer than I thought because I don't make slides that often.*		Searcher	Production technology

Statistical analysis. For RQs 2.1. and 3, I used inferential statistics with bootstrapping method in order to estimate more accurate confidence interval thus provide more generalizable findings (LaFlair, Egbert, & Plonsky, 2015; Larson-Hall & Herrington, 2009). I thus report bias corrected confidence intervals for the correlation coefficients and mean differences that were based on the resampling for 10,000 times with Simple method of sampling as recommended in LaFlair et al. (2015).

To answer research question 2.1., I first inspected the correlations of overall quality scores of the multimodal texts to the two scores for visualization and verbal delivery and language. After checking the correlations, I fitted three models to predict the multimodal text quality by the two other scores from the two raters of the monomodal writing task and the interaction between visualization and language scores. All predictor variables were centered around the mean. The statistical assumptions for multicollinearity (VIF < 2.50) and normality ($1.5 < DW < 2.5$) were checked following statistical guidelines (Allison, 1999; Field, 2013; Jeon, 2015) . Furthermore, to investigate the association between multimodal text quality and monomodal text quality, I conducted Pearson's correlation analysis. Lastly, I conducted correlations analyses using Spearman's *rho* for the process data to examine the association between multimodal writing performance and processes.

For RQ3, participants' responses to the task perception questionnaire were first inspected in terms of the descriptive statistics. Due to the repeated measure, I used a series of paired samples t-tests to see whether the difference between perceptions on the two writing tasks were statistically significant. To address the problem of multiple comparisons, the alpha level of all inferential statistic results was set with the Bonferroni adjustment at $\alpha = .0083$ (.05/6).

Results

Multimodal text quality predicted by language use and visualization scores. The three scores for the multimodal texts showed positive correlations to each other as shown in the scatter plots (see Figure 8). The strongest relationship was found from the correlation between the overall quality and language use scores at medium to large effect size (r = .83, BCa 95% CI [.67, .92], p < .01, see Figure 8(b)) followed by the correlation between the overall quality and visualization score (r = .73, BCa 95% CI [.27, .92], p < .01, see Figure 8(a)). The two different effect sizes showed that the raters' impressionistic scoring on the multimodal quality was more strongly related to writers' ability to use language than their ability to design attractive visual aids. In addition, the correlation between language use scores and visualization scores (r = .47, BCa 95% CI [-.05, .79], p = .01) was weaker than the others.

While the correlations among the three scores of the multimodal task performances yielded positive correlations, the correlations of visualization scores to the other two scores indicated that some writers may have unbalanced competence in the command of elements in visual mode and in verbal mode. In Figure 9, I provided examples of such cases. Both examples are limited to the introduction section given the space limit. First, P13, whose writing proficiency was advanced (total score = 83.67 out of 90), showed great performance in terms of verbal delivery and language but scored low for visualization (visualization score = 4.60, verbal delivery and language = 8.60, overall quality = 7.60). This writer only used one slide with three pictures that were copied and pasted from web search for the introduction section; the images on the slides were consistent with the examples that were provided in the narration. From the script that the writer constructed, we can infer that the writer had good commands of grammatical language and diverse lexical items.

On the other hand, P27 with intermediate writing proficiency (total score = 63.50) scored higher on the visualization (7.60) than language and verbal delivery (5.40) and the overall quality (6.60). She used six slides for the introduction section, and total of 16 slides for the task. As the script shows, she used simple syntactic structures and ungrammatical forms (e.g., errors in verb-noun agreement and interrogative sentences). However, she used the visual elements effectively to show transitions, using text boxes with animation and slide changes and referred to the images when expressing her thoughts in verbal mode. Both writers used three pictures that they borrowed from web search, but P27 applied another layer to the images to better communicate her message to audience. This effort might have improved the quality of the multimodal text thus she received higher score for the overall quality than language and verbal delivery.

As these examples show, some writers showed a wide gap between the ability to express ideas in linguistic mode and that in nonlinguistic mode, namely visualization ability for this multimodal task. For P13, the multimodal task that elicits nonverbal ability had negative influence on the the performance score; on the other hand, P27 was able to elevate her score by utilizing the additional mode that is not usually available in the traditional writing task. The overall holistic quality scores for both writers were around the mean of the language use and visualization scores, which could indicate that multimodal performance scores were concluded as composites of language use and visualization ability by raters. From the three correlation coefficients and the examples, it could be inferred that a writer's ability to use language and design visual aids have independent contributions to the ability to compose a multimodal text.

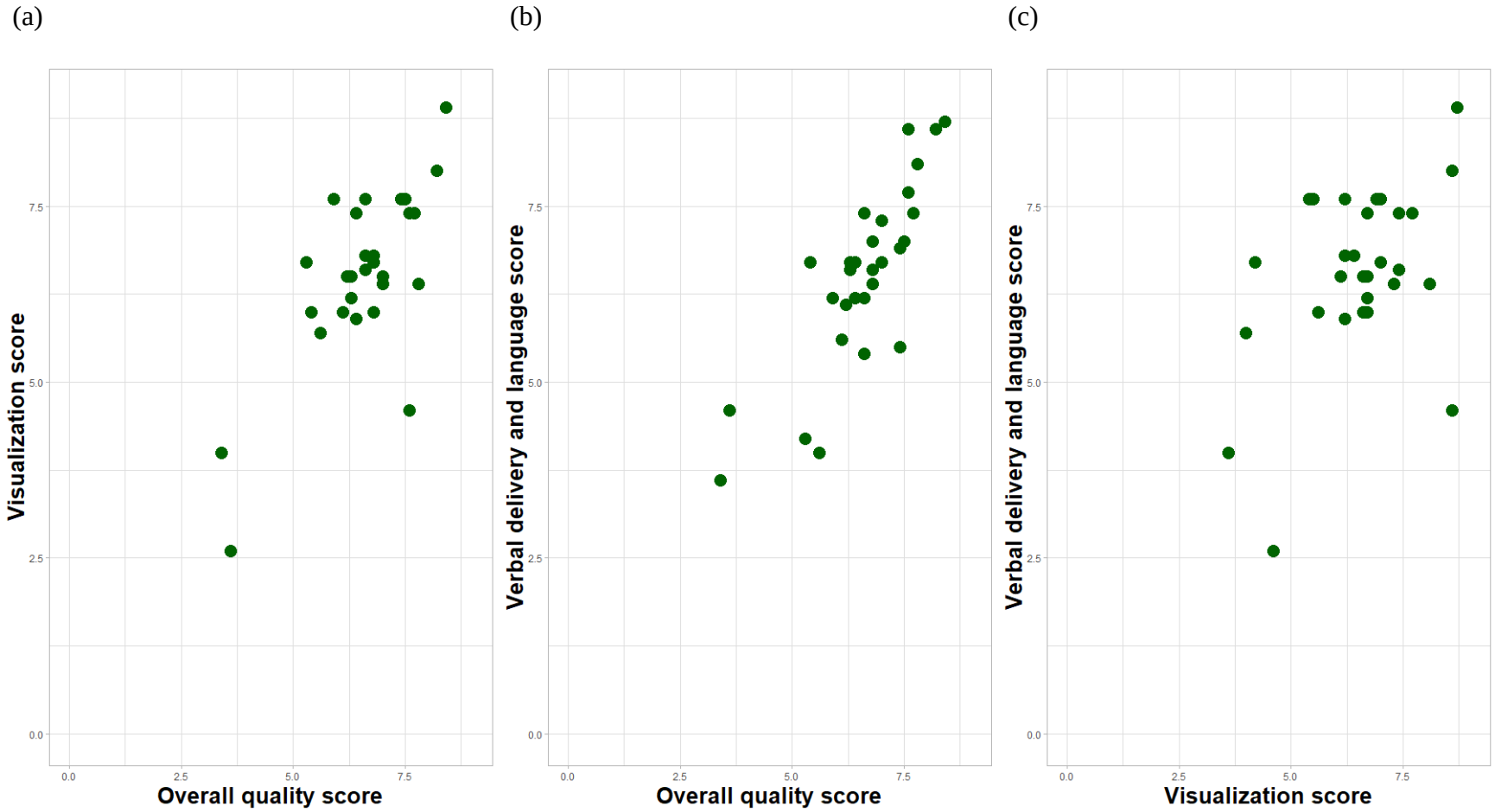

Figure 8. Scatterplots of the three impressionistic scores for the multimodal tasks

Slide	Script
P13 	The development of technology seems to be at a quick, nonstop rate these days. Electronic devices with even better speculations, such as smart phones are being introduced to the public each day. Smart phones, one of the main results of technology development, is mainly used for efficient technology usage and communication with other smartphone users. Thus, it would be appropriate to say that smart phones, along with technology development, is playing a significant role in communication and relationship among people, especially through social network service, widely known as SNS. Now, let's take a moment for you to think about how SNS and the advanced technology has influenced your relationships with other people.
P27 	Nowadays, the advance of technology changes a lot of things. And, people get many benefits from it, as you know, wifi, SNS, listening to music easily, taking photography. [Textbox appears] But, some people believes technology makes us more alone. Yes, I agree with them partly.

Figure 9. Introduction section of two writers' multimodal texts

Figure 9. (cont'd)

Slide	Script
P27 (cont'd) 	As you see, in 2005, the people in the picture look and talk each others. However, in 2011, they just look at their smartphones, not each others. Like the pictures, some people worry about losing opportunity to communicate each others. [Textbox appears] But what if they are not doing their own work, they are just talking to each other using SNS? What if one of them has a disability to hear or speak, so they use their smartphone to communicate better? The communication using technology is meaningless? Less meaningful than face-to-face communication? Then why? I believe that technology does not make us more alone. Rather, technology enables us to communicate better.

In order to test this observation empirically, I ran regression analyses with the overall quality scores as an outcome variable and the two specific scores to visualization and language, and their interaction, as predictor variables. Three regression models with two predictors and the interaction term were fitted and compared (see Table 13). I included one model with the interaction term of the two predictors given its significant correlation to the overall quality score (r = -.47, BCa 95% CI [-83, .46], p = .01). Model 2 with the two predictors was the most parsimonious model with significant F change ($F(2, 26) = 24.98$, $p < .001$). Model 3 that included the interaction term could increase the adjusted R^2 to 1%, but the F change was not statistically significant ($F(3, 25) = 3.41$, p = .08) and the interaction term was neither significantly contributing to the model (β = -.15, p = .08, BCa 95% CI of the unstandardized B [-.27, .18]). Thus, Model 3 was not the best regression model for the data; Model 2, the best fitting model, showed that 83% of the variance in the overall quality of the multimodal texts was explained by the language use scores (β = .62, $p < .001$) and the visualization score (β = .45, $p <$ .001). The bootstrapped unstandardized betas did not cross zero, indicating that the predictors could accurately explain the variance of the overall quality scores. Therefore, consistent with the observation, the results of the regression analysis pointed to the two independent contributions of the language scores and the visualization scores to the overall performance scores.

Relationships of multimodal task performance to L2 writing proficiency. The bootstrapped Pearson's correlation analyses between multimodal and monomodal writing task scores showed that the overall quality and language scores of the multimodal task performances are correlated with all the subscores and the total score of the writing task performances (see Table 14 and Figures 10, and 11). More specifically, a correlation coefficient between the two language scores respectively for a multimodal task and a monomodal task was the strongest (r =

.59, *p* = .001, BCa 95% CI [.22, .82]). Language scores of the multimodal task were also correlated with vocabulary scores for the monomodal writing performances (*r* = .49, *p* = .008, BCa 95% CI [.17, .75]); they were also correlated with other subscores for mechanics (*r* = 47., *p* = .011, BCa 95% CI [.12, .78]) , content (*r* = .45, *p* = 015., BCa 95% CI [.12, .77]), and organization (*r* = .44, *p* = .018, BCa 95% CI [.07, .75]) with smaller effect sizes. Given that there was only one overarching scoring category for language for the multimodal task performance, these relatively strong correlations between the language scores of the multimodal task performance and the language and vocabulary subscores for the monomodal task performance may indicate that the two tasks measured a uniform language-related competence.

The only area of the multimodal task performance that did not show any correlation to monomodal task scores was visualization. All *p* values for the correlations of visualizations to monomodal task performance scores were over .05, and the biased corrected and accelerated confidence intervals crossed zero. Figure 12 repeats the nonsignificant relationships between the visualization score of the multimodal writing task and all subscores of the monomodal task. Taken together, the nonsignificant relationships between visualization scores and all of the monomodal task scores and the significant relationships between language use and overall quality scores of the multimodal tasks and all monomodal task subscores indicate that the multimodal task performance involves another layer of ability than language ability that is measured in the monomodal writing task. In this study, the ability to use visual elements purposefully was the additional ability of computer-based multimodal writing task required.

Table 13.

Regression Models to Predict Multimodal Text Quality Scores

	B [BCa 95% CI]	*SE*	*β*	*p*	*F*	*df*	*p*	adj. R^2
Model 3					50.29	3	<.001	.84
(Constant)	6.63 [6.45. 6.75]	.09		<.001				
Language use	.52 [.35, .65]	.08	.59	<.001				
Visualization	.38 [.15, .68]	.08	.41	<.001				
Language use*Visualization	-.07 [-.27, .18]	.04	-.15	.08				
Model 2					67.47	2	<.001	.83
(Constant)	6.58 [6.40, 6.73]	.09		<.001				
Language use	.55 [.35, .69]	.08	.62	<.001				
Visualization	.41 [.19, .66]	.08	.45	<.001				
Model 1					58.24	1	<.001	.67
(Constant)	6.58 [6.33, 6.82]	.12		<.001				
Language use	.73 [.48, .95]	.10	.83	<.001				

Note. All predictor variables are centered around the mean; BCa = bias corrected and accelerated; CI = confidence interval.

Table 14.

Bootstrapped Correlations between Multimodal and Monomodal Task Scores.

	Multimodal task								
	Visualization			Language use			Overall		
Monomodal Task	*r*	*p*	BCa 95% CI	*r*	*p*	BCa 95% CI	*r*	*p*	BCa 95% CI
Content	.26	.175	[-.16, .62]	.45*	.015	[.12, .77]	.47*	.011	[.20, .69]
Organization	.21	.278	[-.22, .59]	.44*	.018	[.07, .75]	.47**	.010	[.21, .69]
Vocabulary	.16	.395	[-.23, .54]	.49**	.008	[.17, .75]	.39*	.037	[.07, .65]
Language use	.20	.307	[-.33, .67]	.59**	.001	[.22, .82]	.47*	.011	[.07, .73]
Mechanics	.24	.203	[-.14, .61]	.47*	.011	[.12, .78]	.52**	.004	[.20, .76]
Total	.22	.246	[-.21, .62]	.51**	.005	[.16, .79]	.48**	.009	[.18, .71]

Note. BCa = bias corrected and accelerated; CI = confidence interval.

*$p < .05$

**$p < .01$

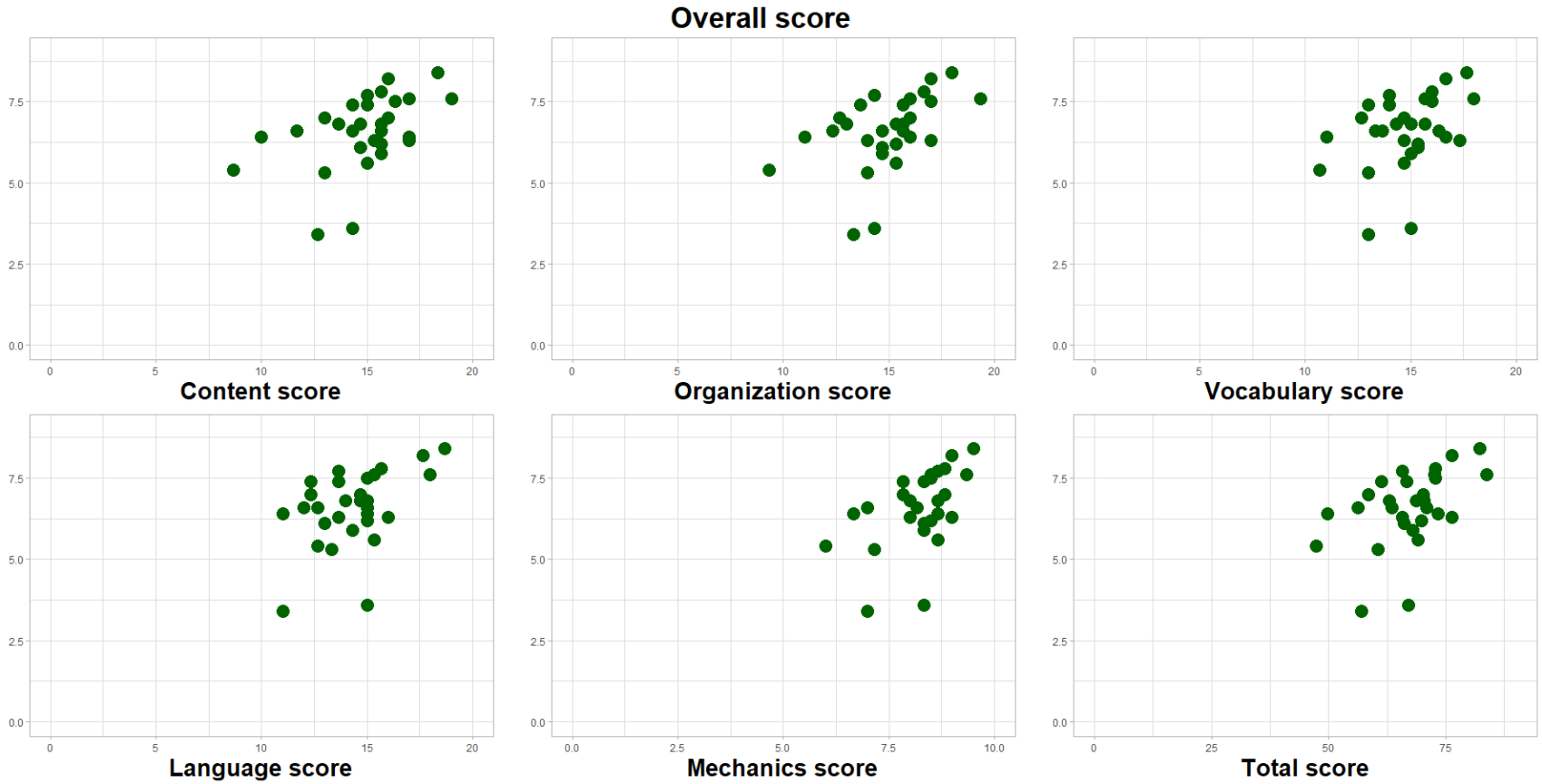

Figure 10. The relationship of the overall quality score of the multimodal texts to the subscores and the total score of the monomodal texts.

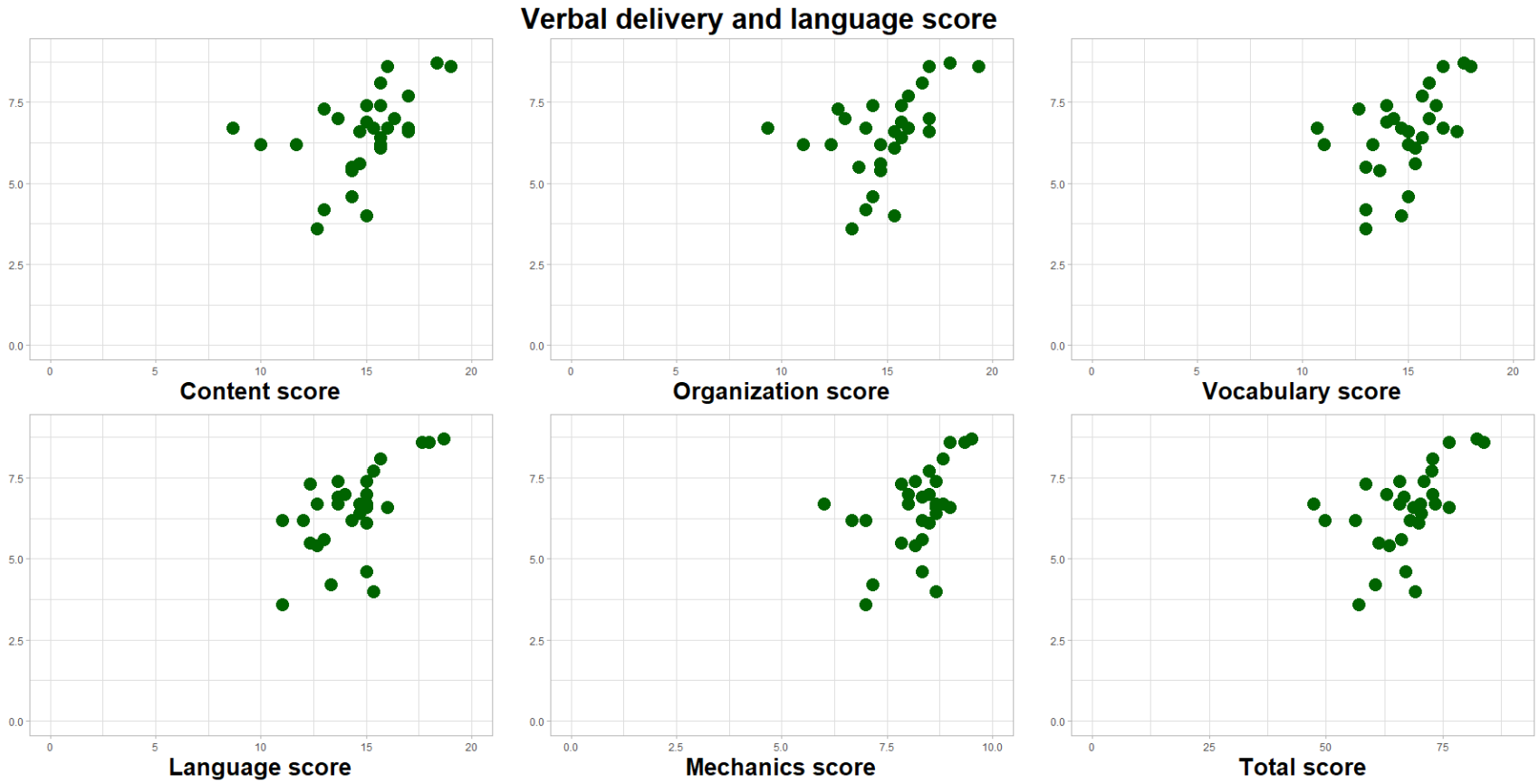

Figure 11. The relationship of the language and verbal delivery score of the multimodal texts to the subscores and the total score of the monomodal texts.

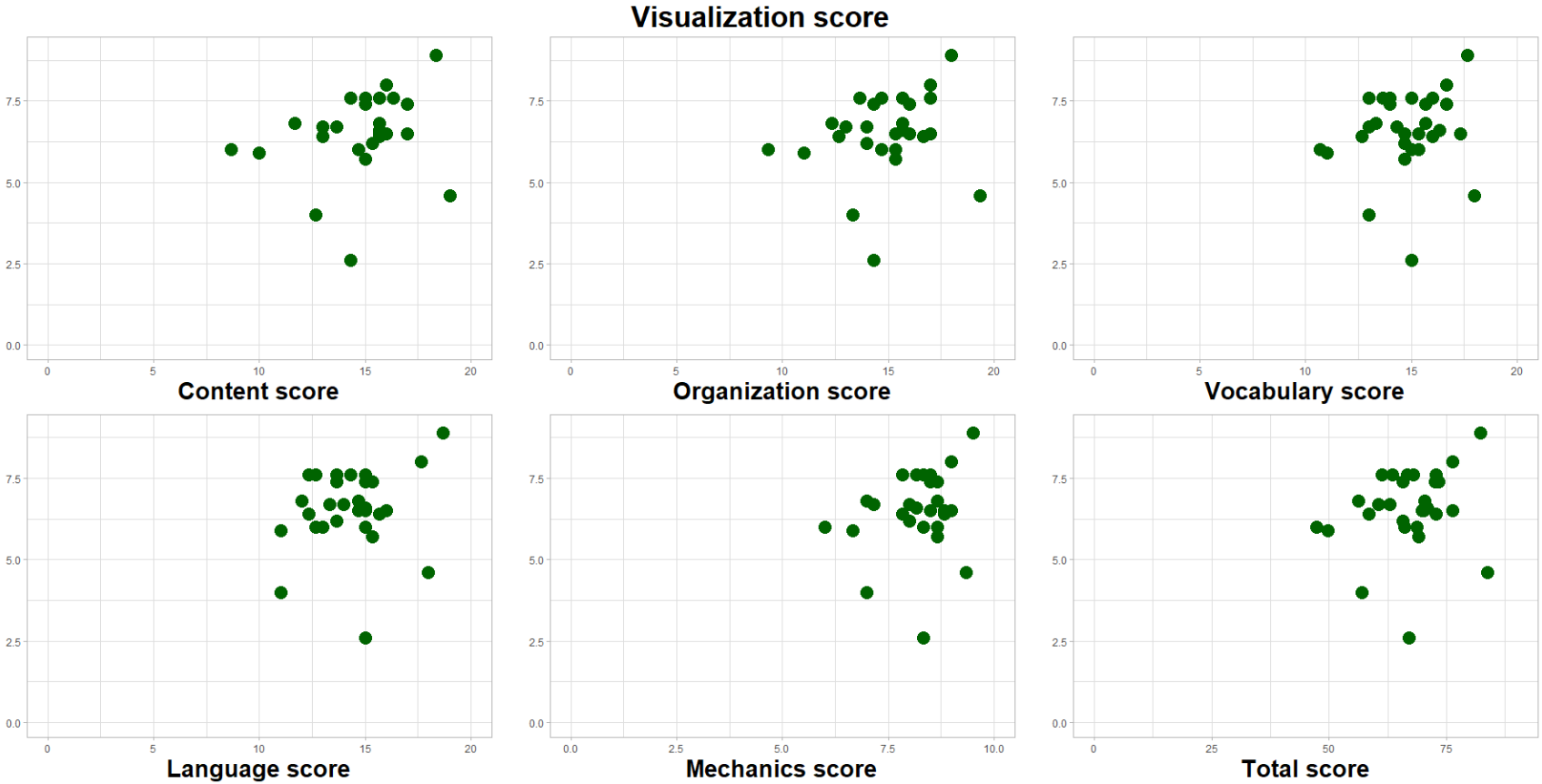

Figure 12. The relationship of the visualization score of the multimodal texts to the subscores and the total score of the monomodal texts.

Multimodal writing processes. In this section, I report twelve focal students' writing processes captured through stimulated recall interviews and screen capture videos of their multimodal writing behaviors. First, I illustrate the patterns in the multimodal writing processes that I found from the two data sets descriptively. Next, I examine their associations with multimodal text quality scores.

Findings from the stimulated recall interviews. From the stimulated recall interviews data, I revealed eight distinct composing processes that can be mapped onto the latest cognitive model of writing by Leijten et al. (2013). Four categories are associated with the writers' own background knowledge on the control level: (1) explaining current plan for processes; (2) explaining current plan for contents; (3) explaining writing schemas; and (4) explaining design schemas.

Most frequently, L2 writers verbalized their current plans for the contents (Total frequency = 85, M = 7.08, SD = 2.35), which changed occasionally throughout the writing processes. Following two excerpts from two writers exemplify their initial plan, or lack thereof. P23 made a concrete plan before really investing efforts in constructing the multimodal texts. On the other hand, P10 verbalized not having developed a specific content but just general ideas.

> *My hands are slow. So I decided to keep one idea throughout the presentation, not giving many supporting reasons. So I thought some narrative format would be good...* [P23]
>
> *I did not have good, clear three reasons. So I just jotted down rough ideas. I did not come up with three reasons, no structures at first. I just kept searching and made three reasons out of reading*. [P10]

Writers frequently explained their plans for the writing processes. It was the second-most frequently verbalized category (Total frequency = 62, *M* = 5.17, *SD* = 3.3). Because the multimodal writing task had a time limit, writers plans were often related to their time management. For example, P01 explained that the time limit determined the way she built the multimodal text. P22 and P26, on the other hand, kept the way they usually do when constructing a text by first outlining their thesis statements and supporting details. Interestingly, P26 deliberately chose to use Korean for the first draft and translate the draft into English because this method makes the L2 text more coherent and logically sound.

> *I didn't know whether I would have enough time to complete the task. I thought, I may not be able to complete it if I make the whole text then record my voice. That is why I decided to record my voice for a slide when I complete each slide.* [P01]
>
> *I wanted to make an outline. I do this step-by-step. So here I thought about my position first.* [P22]
>
> *I could have written things in English first to use time more efficiently, but I wrote in Korean and then translated them into English. Why I did this way, even though it is inconvenient, is because I can see the logical flow more efficiently and fix my logic easily. So, I first wrote everything in Korean roughly and then translated all of them into English.* [P26]

While most of the retroactive verbalization focused on the plans L2 writers proposed at the time of writing, they also substantially explained their background knowledge for the writing and designs. These two categories fall in the writing schemas and design schemas of the Leijten et al.'s (2013) model, which had been discussed as task schemas in previous literature (Hayes, 2012). In total, 68 occurrences were found from the stimulated recall data; on average, each

writer commented on writing schema 35 times, and on design schema for 33 times. For example, the two writers explained how they related the task, which was new to them, with their writing schema:

> *Maybe this is because I am a science person, I am obsessed with finding objective information. Anecdotes can't be on presentation slides.* [P11]
>
> *I thought presentation is objective, but the ideas in my mind are subjective, so it was difficult to organize the ideas.* [P06]

Interestingly, the two writers called the multimodal task as an "objective" genre based on the platform of the writing (i.e., presentation slides). In fact, as found in Study 1, narrated presentation tasks are frequently used to compose a digital story with personal anecdotes. Not because the topic was argumentative, but because they made presentation slides, they related the task to only include scientific and credible information from sources. Given that English presentations are associated with academic and research contexts especially in EFL contexts, L2 writers may have utilized their task schemas to build their own task representation for the current task.

L2 writers' background knowledge on how presentation slides should be composed was also found in the stimulated recall data. For example, P01 substantially edited the appearance of slides, especially for the font colors. She explained that "to give some sort of coherence, I used the blue color for the slide titles". While this writer pinpointed the intention of adding visual enhancement, oftentimes L2 writers did not (or could not) supply as specific information as she provided for why such design elements are preferred. In other words, given the lack of metalanguage for multimodal resources, they could not verbalize why some nonlinguistic resources are conventionally or preferably used. They said, for example:

This layout looks comfortable [P04]

I used dark background colors because the topic was about technology. [Dark colors] look more professional, neat. [P23]

On the process level, stimulated recall data gave evidence for L2 writers' interaction with the task environment by (1) evaluating source multimodal texts; (2) evaluating own text-and-graphic-constructed-so-far; (3) searching for language; and (4) describing technology difficulties. Most frequently among these four categories, writers commented on the source texts including the resources I provided as well as what their web search resulted in. On average, a writer commented 3.92 times on the source texts to evaluate its relevance to their own texts (Total frequency = 47, *SD* = 3.26) with a large variation across writers. Not only they commented on the contents (e.g., appropriateness of the source text as a reference), writers additionally mentioned that they used such source texts to borrow language. For example, P20 discussed:

I first wrote down my ideas and looked for some good expressions here in the reading text. [P20]

Writers also commented on what they had constructed thus far. While this reviewing process provided writers with opportunities to edit their texts, they did not necessarily take time to fix the errors they perceived. For example, P06 and P20 in the follow excerpts decided not to improve their texts. During the stimulated recall, P06 pinpointed the linguistic errors that was noticed at the time of composing and revealed the decision to leave such errors uncorrected. P20 also found that the narration was not satisfying. However, other writers took time to improve the quality of multimodal texts. P11, for example, decided to record the narration once again to remove unnecessary fillers.

I thought, oh there is a grammatical error here, but it is just "s". And, well, it would be really hard to tell. So I just moved on. [P06]

I listened to what I recorded here to see if it sounds okay. Well, actually, it was not okay but I did not have much time and I was a bit tired. [P20]

I re-recorded this slide. [Researcher: Why?] I just had too much uh, uh. It was not smooth. This slide was the least smooth, so I had to redo this slide. [P11]

Because the multimodal task provided writers with three links to external sources including a shot nonverbal video, an article, and a poll website, most of the writers visited those websites. Watching themselves cruising through the provided websites, they recalled what they thought at the time of completing their multimodal writing task. As the following excerpts show, writers evaluated the quality of or the relevance of the source text to their own thoughts toward the topic.

As I read this second text, I found it not very impressive. It just listed pros and cons, but my impression was that it focused more on the negatives. There are so many positive effects to list, but it did not have them. The negative effects mentioned were things that I have heard many times all the time, like people are addicted to these [social networking services] and companies intend to make addicting components using click data. I don't think these negative effects can be resolved by not using technology though. are a matter of use of technology. [P07]

I looked for information that align with my idea. [P10]

While P07 commented extensively on the information that an article provided and evaluated the quality of information, P10 commented that she said she was purposefully looking for content that was coherent with her own ideas. When searching the Internet to develop ideas,

she shifted her purpose of searching from collecting information to choosing phrases that she can use for her multimodal texts:

I cannot make professional language by myself, so I just found some good relevant ones and copied them to organize. [P10]

As such, when a text is provided as a source text, writers used texts to utilize chunks of language in their multimodal texts. This example is coded as Searching for Language which encompasses writers' dictionary searches as well as writers' mental searches to find L1-L2 translate. For examples, P11 commented that for some words he used the online dictionary at the last minute; P26 spent some time to recall the word that she wanted to use after noticing that the word she initially used was not what she felt right.

I eventually used an online dictionary to translate. For words like anonymous and anonymity. [P11]

I couldn't remember how to name this [table of content]. What is it? Index? It sounds awkward. Ah, Contents! [P26]

Lastly, writers commented on the struggles they had with the technology. Three following excerpts display different challenges that three writers faced during multimodal writing through a PowerPoint platform. P23 and P22 recalled that they did not know, or forgot at that time, how to use specific functions thus spent some time to address the technological challenges. P11, on the other hand, expressed an overall frustration of using a computer platform to perform a language task. These challenges shown in the excerpts are somewhat different from what teachers might expect from the current undergraduate students; college students may not be as technology-savvy as teachers imagine.

I tried many times to embed a video clip here. I searched here [link embedding function available in PowerPoint] but there were too many hits. Anyway, I just used the embed function [from Youtube] and inserted the video here. [P23]

I forgot to insert a bullet point. I am trying to figure that out here. [P22]

I am not really familiar with doing something on the Internet [computer] platform. I print materials to read. I am not familiar with doing something on screen. I do handwrite. I am not a slow typewriter, but I feel more comfortable with a paper-and-pencil task. I have a tablet at home but I don't really use it well... My friends only use the Internet [computer] platform, but I still struggle with writing on that platform. [P11]

These eight categories of writing processes observed from the stimulated recall data revealed that writers often explained their current plans for writing processes (Total = 62, *M* = 5.17, *SD* = 3.3) and for content (Total = 85, *M* = 7.08, *SD* = 2.35). As the size of boxes in Figure 13 and the standard deviation in Table 15 show, variation among the writers was larger when they explained the current plans for process. All L2 writers at least once explained their own writing schemas (Total = 35, *M* = 2.92, *SD* = 1.88) and design schemas (Total = 33, *M* = 2.75, *SD* = 1.77). It should be noted that these schemas are relevant to writers' understanding of the task, or task representation, which may limit the number of verbalizations. The low frequency could not conclude that they thought about the task schemas for a couple of times, but the representation may have affected their choices in language and visualization throughout the task.

While these two most frequently verbalized writing processes are on the control level, writers also commented on the process level components such as evaluating source texts (Total = 47, *M* = 3.92, *SD* = 3.26) and evaluating their own multimodal texts they had constructed thus far

(Total = 24, *M* = 2, *SD* = 1.17). Minimum frequencies of the four process-level categories (i.e., evaluating source texts; evaluating own texts-constructed-so-far; searching for language; and describing technological difficulties) were zero, which indicated that not everyone undergoes such multimodal writing processes especially when completing a timed task.

Table 15.

Frequency Statistics of the Writing Processes Reported in the Stimulated Recall Interviews

Process	Total	*M*	*SD*	Min.	Max.	BCa 95%CI of Mean
Explaining a current plan for processes	62	5.17	3.30	1	11	[3.48, 7.17]
Explaining a current plan for contents	85	7.08	2.35	3	12	[5.75, 8.33]
Explaining writing schemas	35	2.92	1.88	1	7	[2.17, 3.92]
Explaining design schemas	33	2.75	1.77	1	6	[1.83, 3.75]
Evaluating source texts	47	3.92	3.26	0	11	[2.42, 5.46]
Evaluating own texts-constructed-so-far	24	2.00	1.71	0	5	[1.08, 2.92]
Searching for language	21	1.75	1.29	0	4	[1.15, 2.42]
Describing technological difficulties	21	1.75	1.87	0	5	[0.92, 2.58]

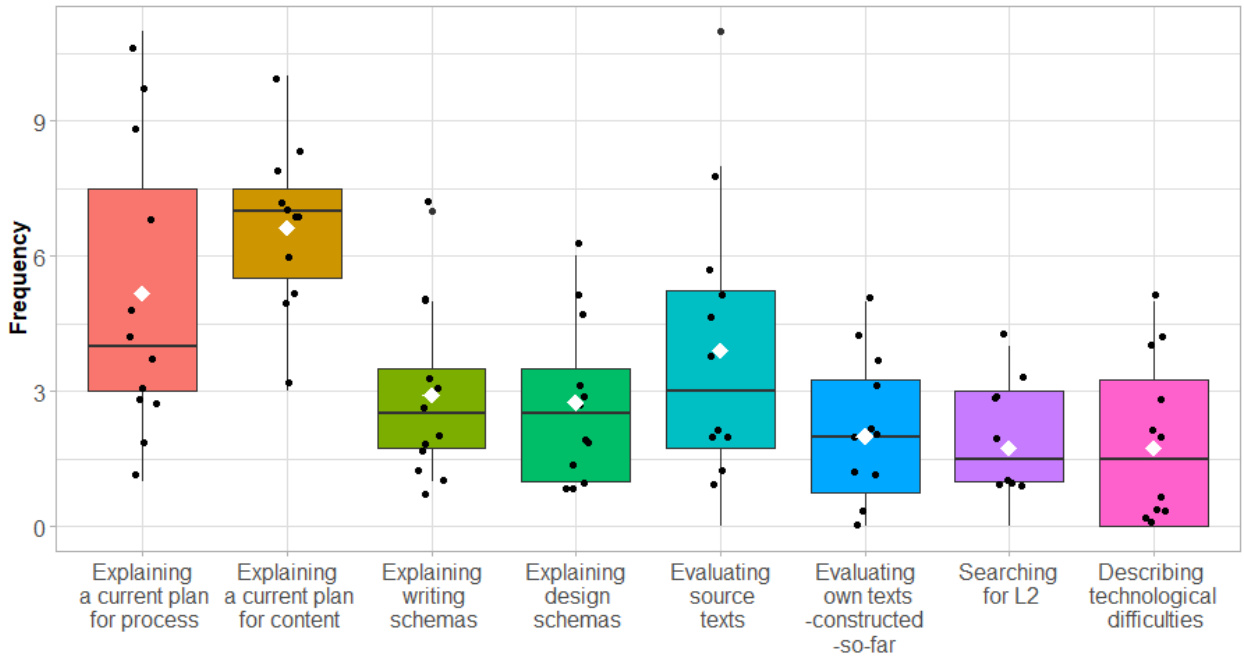

Figure 13. Total frequency of the writing processes reported in the stimulated recall interviews

Findings from the on-screen writing processes. While stimulated recall data revealed what writers reported to have thought during the task, the on-screen writing process data provided what they did and how much time they spent on different multimodal writing processes. As described in the methods section, I used seven categories of multimodal writing processes and reported the descriptive statistics in Table 16 and Figure 14. In Figure 14, y-axis indicates the ratio of each process to total writing time (percentage). In general, L2 writers spent the largest amount of time on editing words on the slide notes (*M* = 23.57, *SD* = 12.19, *Min.* = 2.08, *Max*. = 44.94), followed by accessing provided task resources (*M* = 20.80, *SD* = 13.08, *Min.* = 1.91, *Max*. = 45.59). They also spent about 14.5 % of writing time on searching the internet for the contents and language for their texts (*M* = 14.45, *SD* = 11.85, *Min.* = 0, *Max*. = 31.84). Similar amount time was used for reading their constructed texts (*M* = 13.77, *SD* = 5.46, *Min.* = 6.57, *Max*. = 25.67). They spent about 18.6 % of time to edit words and visuals on the

slides that served as the primary visual information of the final product (For editing words on slides, *M* = 10.97, *SD* = 6.70, *Min.* = 2.50, *Max.* = 25.22; for editing visual elements, *M* = 7.63, *SD* = 5.98, *Min.* = 1.77, *Max.* = 24.21).

Despite the common myths that multimodal writing tasks may limit L2 writers' use of language, findings from the current study tells us a different story. First, L2 writers spent the largest amount of time producing written texts on the slide notes. In addition, they spent much more time on producing written texts either on the slides that appear in the final product or slide notes. As the example in Figure 15 shows, some writers wrote a full script for narration on the slide notes to translate their inner speech to their L2. The script writing process was recursive in that they wrote a few words here and there when ideas and language came up; it was also interactive in that they went on the Internet to search and browse relevant resources. Sometimes, they copied and pasted the language they found helpful for their argument and adjusted the borrowed texts to fit in their own texts. This writing happening in the slide notes is related to the results of the Study 1 that for an academic multimodal task, writing may be hidden from the final product that is evaluated. Speaking in academic multimodal tasks could be related to both "writing" and "speaking". The writing behavioral data provided empirical evidence to supplement the observation from the syllabi data.

Another important finding is L2 writers' use of the Internet to find external resources for content and editing language. While it has been assumed that writers will only search for language by using dictionaries, they rarely used dictionaries. Instead, they used search engines to find relevant articles or graphics that stimulate their ideas and provide chunks of language that they found useful. Interestingly, none of the participants used a search engine in their L1. It is hard to discuss on what grounds they choose one resource to another, but the behavioral data

showed that their Internet searching tends to begin with the contents in the target language and the search results are used for both idea development and language assistance.

On top of the observed patterns in time allocations for the seven writing processes, most noticeable observation was on the considerably large variances across writers. The individual differences spiked for two processes—searching the Internet for resources and editing texts on the slide notes as the lengths of boxes show in Figure 14. Among the categories, the two were somewhat new writing processes to the L2 writers given that often timed writing tasks do not allow writers' use of external resources nor provide alternative spaces that does not show at the final stage. Some writers utilized the newly available features actively, but others might have decided to keep their usual task taking strategies for timed tasks. It is possible that such variances will decrease as writers become familiar with timed multimodal writing tasks, but it is also possible that these processes are more selectively preferred than other processes.

Table 16.

Descriptive Statistics of the Percentage Duration for the Multimodal Writing Processes

Process	*M*	*SD*	Min.	Max.	BCa 95%CI of Mean
Accessing provided resources	20.80	13.08	1.91	45.59	[14.45, 27.76]
Internet searching	14.45	11.85	0	31.84	[9.15, 19.79]
Editing visual elements	7.63	5.98	1.77	24.21	[5.06, 10.70]
Editing words on the slides	10.97	6.70	2.50	25.22	[7.50, 14.95]
Editing words on the slide notes	23.57	12.19	2.08	44.94	[15.95, 31.21]
Reading text-constructed-so-far	13.77	5.46	6.57	25.67	[10.72, 17.32]
Speaking	8.81	4.24	4.73	19.16	[7.11, 10.77]

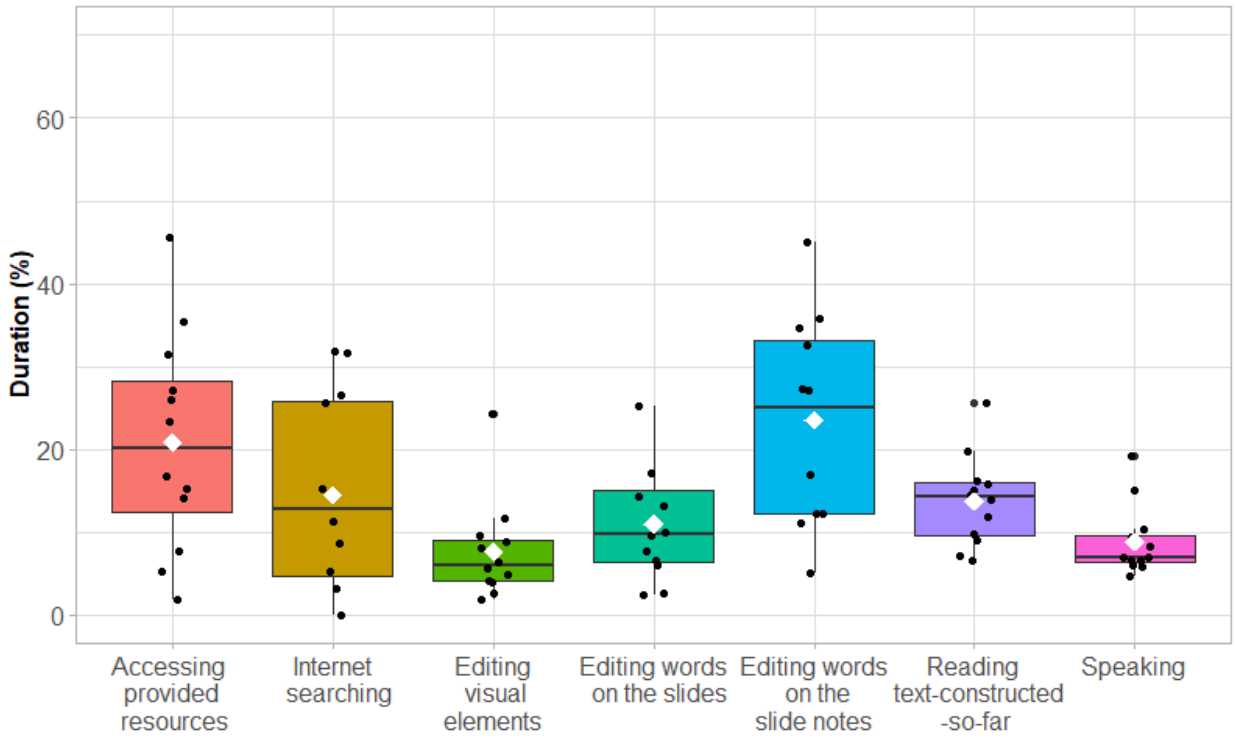

Figure 14. Mean duration of the writing processes from the multimodal writing behavioral data.

.

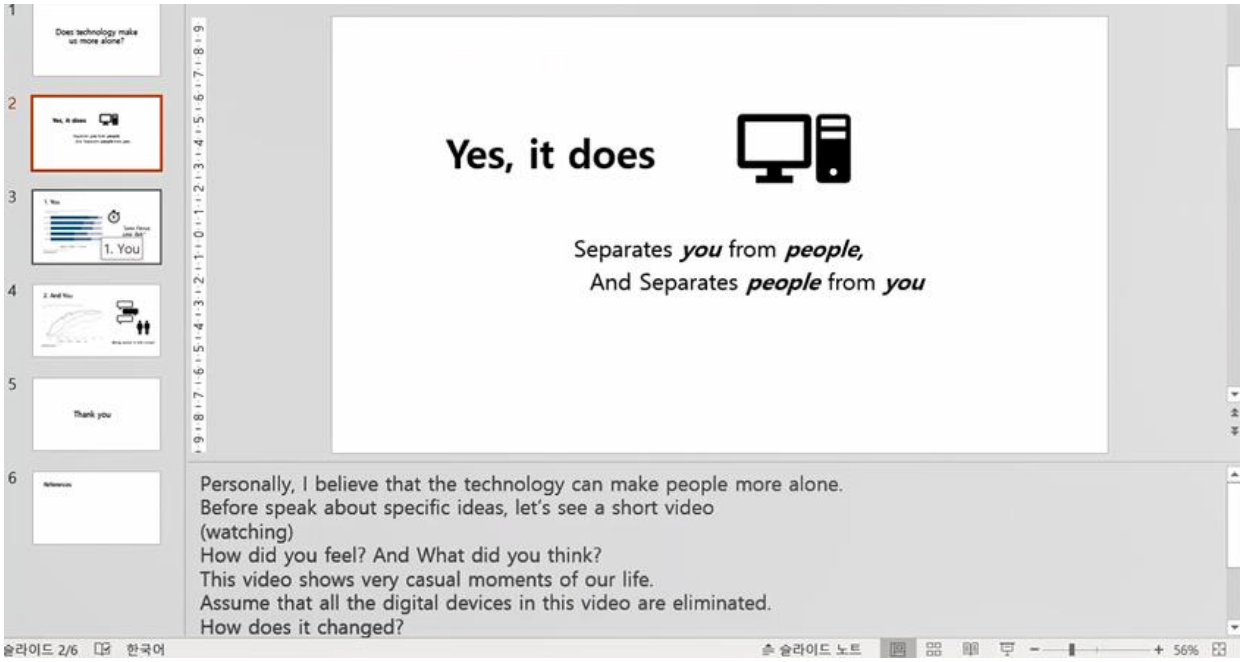

Figure 15. A screenshot of P06's multimodal writing behavioral data

When each writer's processes are divided into five time periods, the seven processes showed different evolving patterns. I describe which multimodal writing processes were observed in each of the time periods. As shown in Figure 16 and Table 18, the first and last time periods were occupied by dominant processes. During the initial stage, L2 writers spent 62.5 % of time to read and explore the provided task materials (period 1, *M* = 62.49, *SD* = 33.95). The last period was mostly invested to record their voices to complete the task (period 5, *M* = 43.02, *SD* = 23.18). What are harder to expect are the writing processes in the middle of task completions (i.e., periods 2 to 4) and how writing processes changed as they constructed multimodal texts.

Writing processes that gradually decreased from time 2 to time 4 are searching the Internet and editing words on the slides. Writers, in general, spent approximately 20 percent of time of period 2 and about 19 percent of time in period 3 on searching the Internet (period 2, *M* = 19.93, *SD* = 19.73; period 3, *M* = 19.06, *SD* = 20.86; period 4, *M* = 10.72, *SD* = 14.02). Internet searching ranked second during periods 2 and 3 but it was dropped to the fourth rank in period 4. The decrease in time from period 4 is a trade off with time needed for the other processes gearing towards voice recording and finishing the task. Another decrease was found on the process for editing words on the slides. In period 2, writers spent 17.7% of time to enter written words on the slides but they spent only 8.26% of time in period 4 (period 2, *M* = 17.70, *SD* = 16.59; period 3, *M* = 12.97, *SD* = 9.84; period 4, *M* = 8.26, *SD* = 5.79). Among the many components to consider for task completion, it was observed that writers first allotted time to develop writing plans by exploring what others shared on the Internet and put their ideas on slides in written words. Thus, the first two periods were mostly driven by online planning and drafting the big pictures of their arguments.

On the other hand, writers strikingly increased the amount of time to edit words on the slide notes from time 2 to 4. In period 4, they spent 42% of time working on their scripts on the slide notes (M = 42.04, SD = 20.50); in the previous periods, they spent 22.8% of time (period 2, SD = 21.13) and 26.54% of time (period 3, SD = 24.90). While they spent less time on editing words on the slide notes during periods 2 and 3, still this category was ranked the first for the two periods. In other words, besides periods 1 and 5 where writers spent time to understand the task and record their voice for narration, writers were occupied with the translating process from ideas to words and the output were placed on the slide notes which were used as scripts. The peak at period 4 reflect that the detailed scripts were developed after they somewhat finished drafting the slides with visual elements and short written words for presentation.

Two processed did not show a linear progression over the three time periods. One is editing visual elements which was most popular in period 3 (period 2, M = 8.70, SD = 9.87; period 3, M = 14.22, SD = 14.71; period 4, M = 12.11, SD = 11.27). Time for visualization was similarly distributed from periods two to five, which is a different from other processes that were distributed differentially across time periods. The other process was reading own texts-constructed-so-far (period 2, M = 16.25, SD = 9.13; period 3, M = 12.79, SD = 7.47; period 4, M = 17.52, SD = 7.96) which had two humps at periods 2 and 4. The second peak at period 4 merits more explanation given its second rank. What could be possible is that as writers produce scripts, they simultaneously evaluated their own texts constructed so far on slides, in terms of logic and their relevance to their original ideas. Also, towards the end of writing time, as texts become longer and closer to the complete composition, writers might have spent more time to go over their texts for final edits.

I thus far illustrated the chronological changes within each multimodal writing process in terms of general patterns focusing on the mean values; however, the variance across writers were was large, which indicate that writers need different amount of time for each process. Among the writing processes captured from the writing behavior data, I found that two of them showed more variance than other processes. Figure 17 shows that two writing processes involved large variance across writers: Internet searching and editing words on the slide notes. This is consistent with the overall pattern reported in Figure 14. The individual differences were also observed from their chronological patterns. Taken together, it could be argued that the new features for the timed writing tasks were heavily or lightly utilized by the writers, but such use is not relevant to the quality of the multimodal texts they compose.

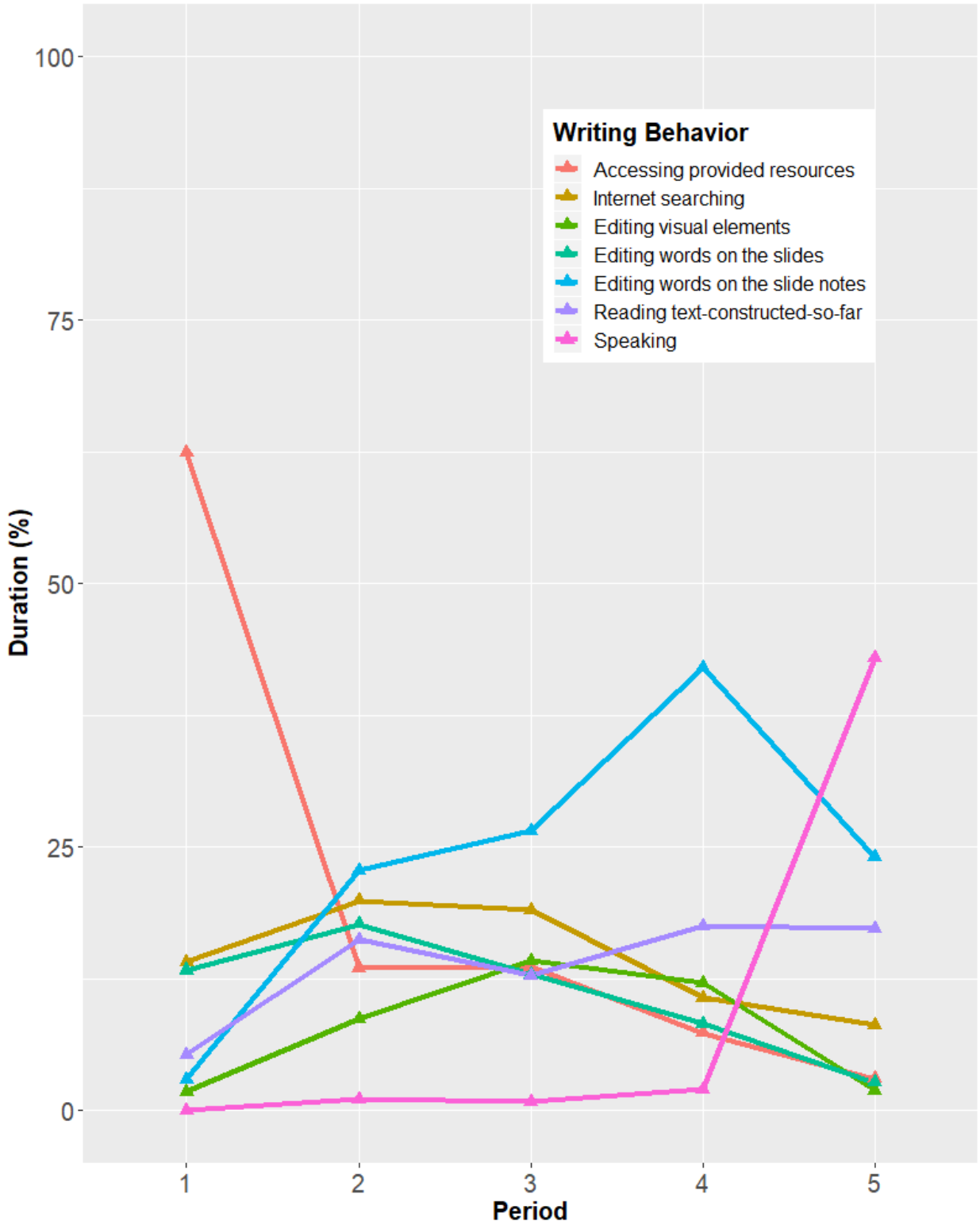

Figure 16. Mean duration of the writing processes from the multimodal writing behavioral data throughout five time periods.

Table 17.

Descriptive Statistics of the Percentage Duration for the Multimodal Writing Processes in Five Time Periods

	Period 1		Period 2		Period 3		Period 4		Period 5	
Process	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Accessing provided resources	62.49	33.95	13.56	18.43	13.60	12.80	7.37	9.71	3.01	5.95
Internet searching	14.11	22.46	19.93	19.73	19.06	20.86	10.72	14.02	8.13	17.59
Editing visual elements	1.80	3.76	8.70	9.87	14.22	14.71	12.11	11.27	1.91	2.43
Editing words on the slides	13.32	17.80	17.70	16.59	12.97	9.84	8.26	5.79	2.65	5.23
Editing words on the slide notes	2.96	4.86	22.80	21.13	26.54	24.90	42.04	20.50	24.01	19.94
Reading text-constructed-so-far	5.31	7.29	16.25	9.13	12.79	7.47	17.52	7.96	17.27	14.02
Speaking	0	0	1.06	3.67	0.82	2.84	1.98	3.90	43.02	23.18

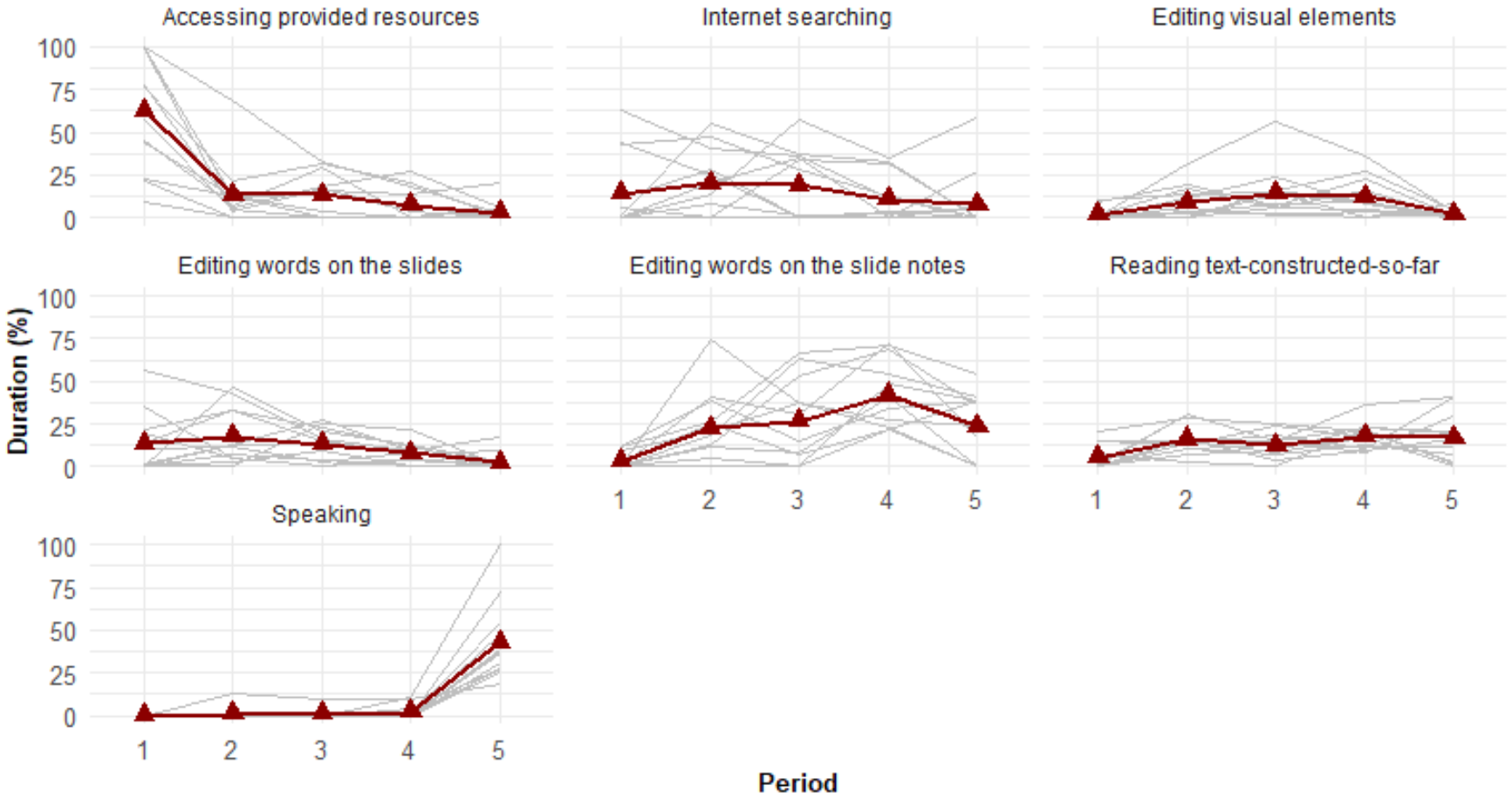

Figure 17. Individual L2 writers' time spent on the seven writing processes throughout five time periods.

Note. Red lines indicate trend lines for each of the multimodal writing processes.

Relationships between multimodal writing processes and performance. I have thus far focused on discussing general patterns in L2 writers' multimodal writing processes. Nonetheless, individual differences were clearly observed. For example, the range of frequencies of processes varied and the confidence intervals are also wide as shown in Table 16. These individual variations could simply reflect writers' individual preferences in multimodal writing. I found that, however, that some of the frequencies have relationships with the text quality at more than a chance level.

The frequencies of three multimodal writing processes were correlated with one of the scores for multimodal text performances, which are explaining current plans for content, evaluating text-and-graphic-created-so-far, and addressing technological difficulties (see Table 18 and Figure 18). Two processes showed positive correlations with the scores. First, writers who received high scores for the overall quality of multimodal texts verbalized their current plans for content more frequently (ρ = .88, p = <.01). Other control-level writing processed did not show systemic relationship with the multimodal text quality.

What is striking in the Table 19 is the negative correlations, significant or marginally significant, between the reviewing process (i.e., evaluating own texts constructed so far) and the scores. Writers who frequently reported to have evaluated their own texts during writing scored lower for the overall quality and language and verbal delivery than others who less frequently commented on their multimodal texts-constructed-so-far (overall quality, ρ = -78, p < .01; language and verbal delivery, ρ = -.61, p = .03); it was also found that the negative correlation between this writing process and visualization quality was approaching the significance level (ρ = -.53, p = .07). L2 writers are often strongly encouraged to review their own texts for accurate and appropriate language, and they did go over their own texts; however, as the examples of P06

and P20 show, they may not make extra efforts to improve the current texts due to time limit or in a hope that readers may not notice the errors. For the timed multimodal writing task, therefore, it was revealed that frequent reviewing processes during writing is not necessarily helpful for improving multimodal text quality.

Lastly, I found another non-significant trending in the predicted direction indicating a positive relationship between the frequency of verbalizing technological difficulties and the visualization score (ρ = .54, p = .07). Thus, the quality of visualization was found to be marginally correlated with two process-level multimodal writing processes with opposite directions; higher scores on the visualization quality were associated with fewer comments on the writers' own multimodal texts constructed so far and more frequent comments on the technological difficulties they had faced during writing. It was observed that when writers explain the technological difficulties, they eventually resolved the challenges (e.g., P20 and P22 in the examples), unlike writers' decisions not to improve their texts despite their awareness of room for improvement.

Taken together, with the stimulated recall data, results revealed that writers who had explicit plans for the contents and actively addressed the technological difficulties utilized visual resources efficiently. Writers evaluated their own texts as they construct multimodal texts, but ones' own critical comments on the texts did not result in improved quality but lowered the overall quality and the subscores. What seems important is whether writers take actions to improve their texts when they know they can make differences.

Table 18.

Spearman's Correlations between the Multimodal Writing Performances and the Frequency of Stimulated Recalls on the Multimodal Writing Processes

	Overall quality		Visualization		Language and verbal delivery	
	ρ	p	ρ	p	ρ	p
Explaining current plan for process	-.05	.88	-.10	.76	-.35	.27
Explaining current plan for content	.44	.16	.88*	< .01	.17	.61
Explaining writing schemas	-.39	.21	-.47	.13	-.32	.31
Explaining design schemas	-.01	.97	.15	.64	-.33	.30
Evaluating source texts	.04	.90	.04	.90	.16	.62
Evaluating own texts-constructed-so-far	-.78*	<.01	-.53	.07	-.61*	.03
Searching for L2	.04	.90	-.31	.32	.30	.34
Describing technological difficulty	.36	.26	.54	.07	.13	.68

* $p < .05$

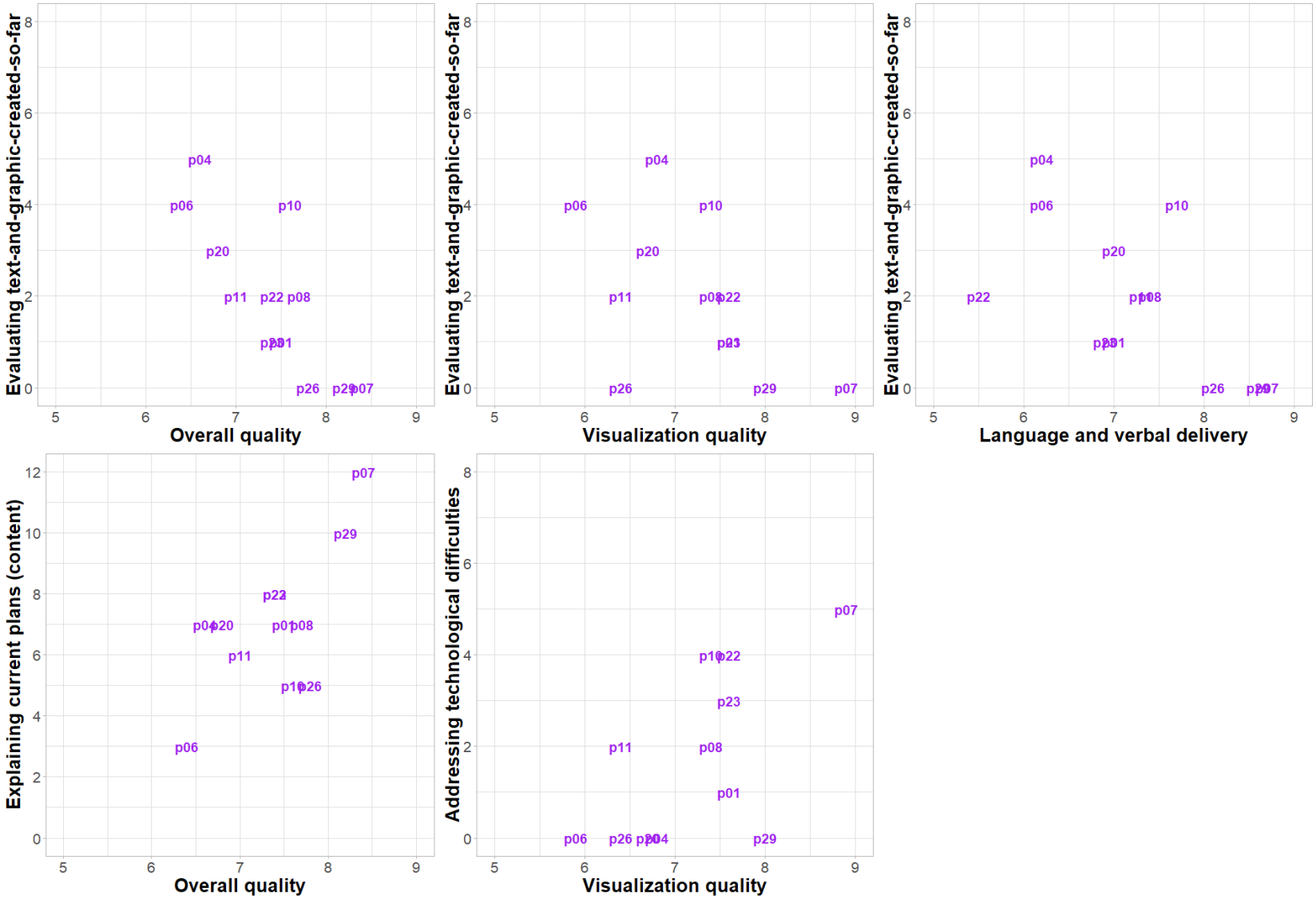

Figure 18. The relationship between multimodal texts quality and writing processes reported in the stimulated recall interviews

Finally, I inspected whether there were any systemic relationships between the time spent on each process and the outcome (i.e., multimodal text quality). As Table 19 shows, none of the Spearman's correlation coefficients were statistically significant. Given I found negative correlations between the process of evaluating own texts-constructed-so-far and the quality scores from the stimulated recall data, I expected a similar pattern for reading text-constructed-so far; however, none of the coefficients were close to significance level (overall quality, ρ =.13, p = .69; visualization, ρ =,.04 p = .91; language and verbal delivery ρ = -.03, p = .93). Thus, it could be concluded that the writing performance is not correlated with the amount of time writers spent on each of the multimodal writing processes and that the two different methodologies paint different pictures.

Table 19.

Spearman's Correlations between the Multimodal Performance Scores and Time Spent on the Multimodal Writing Processes

	Overall quality		Visualization		Language and verbal delivery	
	ρ	p	ρ	p	ρ	p
Accessing provided resources	-.18	.58	-.01	.98	-.27	.39
Internet searching	.23	.48	.26	.41	.35	.26
Editing words on the slides	.21	.51	-.12	.71	.18	.59
Editing words on the slide notes	-.27	.39	-.46	.14	-.04	.91
Reading text-constructed-so-far	.13	.69	.04	.91	-.03	.93
Speaking	-.19	.56	.10	.77	-.28	.38
Editing visual elements	-.28	.38	.25	.44	-.53	.07

L2 writers' perception of the multimodal and monomodal tasks. The descriptive results of the all writers' perceptions on the monomodal and multimodal tasks are presented in Table 20 and Figure 19. Each row of the Table 20 represents writers' response to one item, and scores ranged from 1 to 9. In general, writers' responses for the two writing tasks showed similar variance. When comparing the two tasks with descriptive statistics and data visualizations, means for the multimodal task were higher than those for the traditional monomodal writing task except for two items on confidence (i.e., *I did (or did not do) well on this task.*) that showed lower scores for the multimodal task and motivation (i.e., *I want to (or don't want do) more tasks like this.*) that showed similar means for the two tasks.

Table 20.

Bootstrapped Descriptive Statistics of the Writers' Task Perceptions (n = 31)

	Multimodal			Monomodal		
	M	*SD*	BCa 95% CI	*M*	*SD*	BCa 95% CI
Complexity	6.45	1.98	[5.74, 7.13]	5.35	1.80	[4.77, 5.96]
Difficulty	5.84	1.70	[5.26, 6.39]	4.45	1.67	[3.90, 4.97]
Anxiety	5.35	1.82	[4.74, 5.94]	4.61	1.86	[4.00, 5.23]
Confidence	4.42	1.84	[3.81, 5.03]	4.84	1.46	[4.39, 5.29]
Interest	6.94	1.46	[6.42, 7.42]	6.10	1.76	[5.55, 6.65]
Motivation	6.23	1.54	[5.68, 6.74]	6.16	1.49	[5.68, 6.61]

Note. BCa = bias corrected and accelerated; CI = confidence interval.

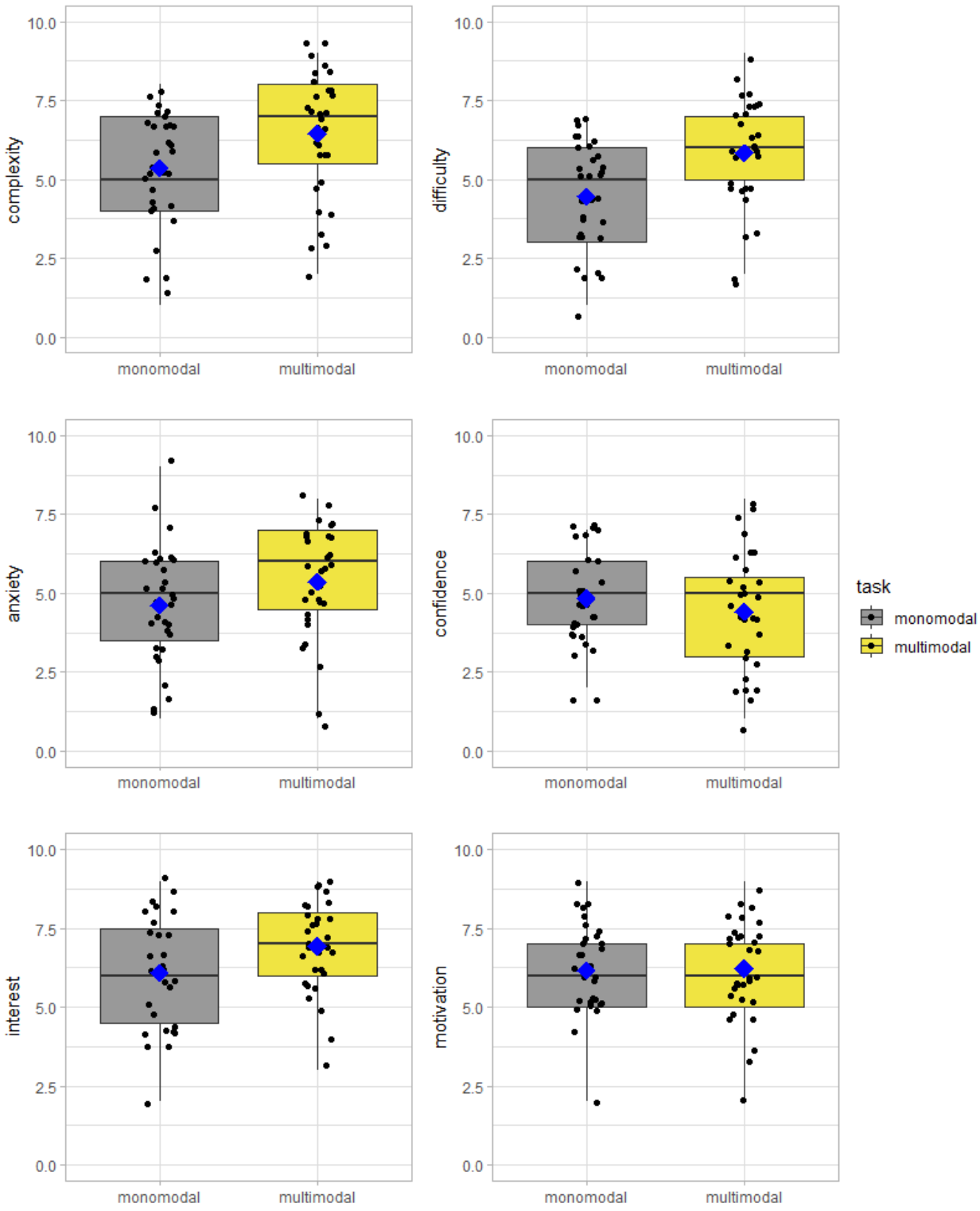

Figure 19. L2 writers' task perceptions across monomodal and multimodal writing tasks. Square points indicate means for each task.

Among the six items, three items showed statistical mean differences with moderate to large effect sizes (see Table 21 for the bootstrapped paired samples t-tests). The three items were complexity (t =3.30, p = 0.002, Cohen's d =0.59), difficulty (t =3.64, p = 0.001, Cohen's d =0.65), and interest (t =3.10, p = 0.004, Cohen's d =0.56). That is, writers perceived the multimodal writing task more complex, difficult, and interesting than the traditional writing task. Even though descriptive statistics showed potential differences that writers find a multimodal made them feel more anxious (for the monomodal task, M = 4.61, SD = 1.86; for the multimodal task, M = 5.35, SD = 1.82) and less confident (for the monomodal task, M = 4.84, SD = 1.46; for the multimodal task, M = 4.42, SD = 1.84) , such differences were not found to be statistically significant (for anxiety, t =1.89, p = 0.069, Cohen's d =0.34; for confidence, t =-1.31, p = 0.201, Cohen's d =0.24). Writers were comparably motivated to do both tasks in the future.

Table 21.

Paired Samples t-tests between ESL Students' Perceptions on the Multimodal and Monomodal Tasks with Bootstrapping

	Mean Difference	*SD*	*t*	*p*	BCa 95% CI	*d*
Complexity	1.10	1.85	3.30*	0.002	[0.45, 1.74]	0.59
Difficulty	1.39	2.12	3.64*	0.001	[0.68, 2.06]	0.65
Anxiety	0.74	2.19	1.89	0.069	[-0.03, 1.55]	0.34
Confidence	-0.42	1.78	-1.31	0.201	[-1.00, 0.19]	0.24
Interest	0.84	1.51	3.10*	0.004	[0.35, .1.35]	0.56
Motivation	0.06	1.69	0.21	0.833	[-0.48, 0.61]	0.04

Note. BCa = bias corrected and accelerated; CI = confidence interval.

*p < .0083 (.05/6).

CHAPTER 5.

DISCUSSION AND CONCLUSIONS

This study aimed to add to the limited amount of research into multimodal writing tasks, especially in terms of the relevance of authentic multimodal writing tasks to pedagogical L2 writing tasks in higher education. Conducting two sequential studies, I explored the valid design of multimodal writing tasks and the role of language in L2 writers' multimodal writing performance and processes. Furthermore, I compared L2 writers' perceptions on the multimodal writing task with those on the traditional argumentative essay writing task.

In this chapter, I first provide a summary of the study results and discuss what these mean to language learners and instructors. Then, I interpret the current findings in terms of research, theory building, and pedagogy. While acknowledging the interdependence of research and teaching implications, I attempted to offer more in-depth, unique suggestions for research by exploring multimodal writing performance and processes, as well as for teaching through the investigation of needs analysis and task perception results.

Integrated Results

From the needs analysis of multimodal writing tasks (Study 1), I could classify multimodal writing tasks being used for undergraduate studies, clearly indicating the flexibility of their formats and expectations. Multimodal tasks in undergraduate courses ranged from those targeting students' creative production without a predetermined format of the outcome to those fostering genre knowledge. Also, some tasks elicited written words combined with nonlinguistic, visual elements (e.g., figures and tables), while others required both written and verbal words as well as other nonlinguistic resources for task completion. For the latter cases, I found that written words were sometimes not explicitly displayed on the multimodal texts but served as written

plans or scripts. Last, individual tasks were more common than collaborative tasks regardless of the goal of multimodal tasks.

In terms of the factors contributive to L2 multimodal performance, L2 writers' ability to use language and visual elements explained the 83% of the variance in the overall quality of the multimodal texts, with language scores being a stronger predictor ($\beta = .62$, $p < .001$) than visualization scores ($\beta = .45$, $p < .001$). Additionally, a series of correlation analyses between the scores of monomodal and multimodal writing tasks revealed that the overall multimodal text quality and language and verbal delivery scores of the multimodal texts were significantly correlated with all sub-scores of monomodal writing performance, which may indicates that the two tasks in fact have large, overlapped constructs, mainly related to language use competence. Thus, these findings point to the importance of writers' good command of language for highly rated multimodal writing performance.

For multimodal task performance, it may not be easy to argue that one language modality is more important than the other (i.e., written or verbal words) because the two were found to be highly intertwined throughout the multimodal writing processes. Specifically, from the multimodal writing process data sets (i.e., stimulated recalls and on-screen writing behaviors), I revealed that L2 writers invested the largest amount of time typing in written words describing visual aids or serving as presentation scripts. Given the role of written words as a detailed plan and rehearsed speech in multimodal writing, I would argue that the main purpose of written words, particularly those on the slide notes, was to translate inner speech into a target language narration.

While course instructors gave multimodal assignments to their students with a clear set of expectations for effective use of visual cues, such expectations appeared not to be always clear to

undergraduate students due to the absence of explicit, predetermined rules. Nonetheless, the result of Study 2 showed student writers' experience with multimodal texts in academic settings might allow them to build their own schemas regarding how to choose and integrate visual elements into multimodal texts. Specifically, the stimulated recall result demonstrated what L2 writers' original plan of using visual cues was and how their multimodal products reflected their endeavor to implement it. For example, P04 found certain layouts more comfortable to see, and P23 noted that dark colors look more professional, which they had not learned through explicit instruction. Therefore, both studies highlight that student writers may be able to acquire the task schemas needed to design multimodal texts through an inductive analysis of recurring patterns across multimodal texts, which is pedagogically in line with genre-oriented instruction.

Next, my analysis focused on the potential link between multimodal writing processes and performance. The stimulated recall result showed that as writers evaluated their own multimodal texts-constructed-so-far more frequently, they tended to obtain lower scores on the overall quality (ρ = -.78, p = < .01) and language and verbal delivery scores (ρ = -.61, p = .03). On the other hand, writers who focused more on their current plans for content tended to perform better on the visualization category (ρ = .88, p = <.01). Other notable trends, although not statistically significant, include the association between technological difficulties and visualization scores (ρ = .54, p = .07), and that between evaluating own text-constructed-so-far and visualization (ρ = -.53, p = .07). Taken together, it appears that writers' excessive evaluation of their own, evolving texts during multimodal writing performance may do more harm than good.

The L2 participants, after task performance, expressed that the multimodal task was more cognitively demanding and difficult, but more interesting than the monomodal essay task.

However, they did not perceive the two tasks differently in terms of anxiety, confidence, and motivation. The result that L2 writers found the multimodal task more interesting than the traditional monomodal task is consistent with previous study findings (e.g., Dzekoe, 2017; Jiang, 2018). It might be the students' experience with various types of academic texts that made them perceive both tasks to be similar with regard to task-inducing anxiety, confidence, and motivation. Higher cognitive demands and difficulty of the multimodal task could be partially explained by evidence from Study 1. According to an Education instructor, undergraduate students' digital literacy may not be as good as we generally assume them to be (e.g., Education Instructor 3's comment: ... *we think the students, they are part of a particular generation and we think that they come in already knowing how to use technology. And actually a lot of our students don't know a lot. Some of our students don't know how to use Google Docs*). Furthermore, despite the students' hands-on practice prior to task performance and experience using the PowerPoint platform, some L2 writers who participated in Study 2 expressed difficulties using this tool and writing on the computer.

The two studies jointly confirmed that successful multimodal writing performance would depend on L2 writers' effective use of both linguistic and nonlinguistic resources. Multimodal tasks in authentic settings generally required students to produce written and spoken words in an integrated way. The timed multimodal task encouraged L2 writers to use written language not only for their final product but for translating their ideas into coherent language for presentation (i.e., scripts). These findings from multimodal writing performance and processes, as well as authentic tasks in undergraduate courses, emphasize the pivotal role of language for the communication of multimodal meaning.

Discussion of Research and Theory Building

Writers' interaction with language during multimodal writing. Writing process data were analyzed and interpreted with reference to the cognitive models of writing (Hayes, 1996, 2012; Hayes & Flower, 1980; Leijten et al., 2013); among the models, I adopted the latest model proposed by Leijten et al. because of its inclusion of the writer's cognitive activities related to external source search and use, and multiple modes of communication. Their work provided grounds for describing multimodal writing processes from the cognitive perspective.

Findings from Study 2 specifically highlight some control-level (i.e., *current plan* and *writing and design schemas*) and process-level elements (i.e., *writing processes* and *task environments*) in the Leijten et al.'s cognitive model of writing. *Searcher*, one of the newly added components in the model, is confirmed from evidence that L2 writers extensively used the Internet to find external resources for content and language editing. In the context of L1 writing, searcher is mainly discussed in terms of idea development and source-based writing; for language, writers are likely to use a dictionary for spell-checking (Leijten et al., 2013). However, L2 writers' use of searcher was to gain access to external sources and then borrow the content and language from the sources to construct their texts. Instead of looking up translated lexical items or synonyms in a dictionary, their prevailing strategy was to find language expressions from relevant texts. *Evaluator* was another process-level element discussed frequently in the stimulated recall data. L2 writers reported that they evaluated web search results, texts accessed on the Internet, and their own text-and-graphics-created-so-far. These are the examples of, based on the model, interactions between the task environment and writing processes.

Findings from the stimulated recall data further confirmed L2 writers' use of task schemas and their dynamic writing plans on the control level. They brought in their task schemas

that consisted of linguistic and visual elements. For design schemas, L2 writers seemed to have awareness of how to include and organize visual cues systematically for multimodal task completion, although we may need further discussion on what should be counted as design schemas (Leijten et al., 2013). Writing schemas have been discussed in terms of linguistic features for rhetorical functions (e.g., Yang, Lu, & Weigle, 2015; Yoon, 2017; Yoon & Polio, 2017) and task representations (e.g., Flower, 1990; Plakans, 2010). Performing the multimodal task on an argumentative topic, L2 writers showed variation in their consideration and application of relevant task schemas. Some writers found a narrative with personal anecdotes sufficiently effective, whereas others used only scientific evidence in supporting their arguments. Thus, despite the observation that writers heavily used their schemas, individual writers displayed different representations for the task, as discussed in previous research (e.g., Nicolás-Conesa, Roca de Larios, & Coyle, 2014; Plakans, 2010; Ruiz-Funes, 2001; Wolfersberger, 2013).

Among the findings, counterintuitive were the negative correlations between reviewing processes and text quality. It has long been believed that good writers would review their text constantly and revise it to improve its quality during the act of writing (e.g., Ellis & Yuan, 2004; Rostamian, Fazilatfar, & Jabbari, 2018), but I argue that the reviewing process itself would not result in better writing performance, but the improved quality of writing can be achieved only when writers respond to the gaps that they found during online reviewing processes. In a mixed methods study, Polio, Tigchelaar, and Lim (2018) examined why L2 writers enrolled in an ESL program do not show language development, especially with regard to linguistic accuracy. We used stimulated recall interviews to investigate ESL students' writing processes. While watching a video replay of their keystroke logs, some participants were able to identify their errors and

correct them immediately. Similarly, some L2 writers in the current study also noticed their errors during the stimulated recall session, but they commented that they decided not to correct them because they were tired or running out of time. This may be a reflection of students' expectations of how their multimodal work will be graded; in performing a multimodal task, students would attend to linguistic accuracy to a lesser extent because there are other, more notable areas (e.g., presentation clarity) to be evaluated. It is also plausible that the timed condition might have made the multimodal task extremely challenging to some students, pushing them to neglect linguistic accuracy. Then, in the case of using multimodal tasks for L2 learners who need language development, we may need to consider three points: (1) sharing a scoring rubric that includes language use explicitly, (2) implementing the task with no time constraints, and/or (3) giving opportunities for language-focused revisions.

Following the procedure from Gánem-Gutiérrez and Gilmore (2018), I examined the on-screen behavioral data, with each participant's writing activity divided into five periods. As multimodal writing proceeded, writers spent more time editing words on slide notes but less time for Internet searching and editig words on slides. On the other hand, Gánem-Gutiérrez and Gilmore found that writers invested more time to access online dictionaries and thesaurus and to revise their texts, but less time for their essay construction. These differences may be rooted in the task differences (i.e., a timed monomodal test task versus a timed multimodal writing task). For example, given the need to look for visual elements and information to strengthen their argument in the current study, writers' use of the Internet for external sources took place in earlier stages, with a focus on gaining access to online texts and graphics. Another reason for the differences may be rooted in the coding schemes. I did not separate revising and transcribing

because it was not possible to discern the two from on-screen writing behavior data, but Gánem-Gutiérrez and Gilmore had separate schemes for those two.

While the participants showed rather consistent patterns of multimodal writing processes, there were two areas that showed notable individual differences: searching the Internet for resources and editing texts on the slide notes. These two processes that are not necessary in conventional writing tasks may not be equally accessible to all student writers. Traditional timed writing tasks do not involve searching for external sources nor using PowerPoint slide notes. Some students whose experience has been limited to traditional monomodal task routines may have great difficulty deploying these new skills, while other students with, for example, greater tolerance of ambiguity would be able to use them with less difficulty, thus building learner-specific task representations.

The role of language in multimodal task performance. The quantitative analysis of the multimodal writing performance has shown that multimodal and monomodal tasks may have overlapped constructs that are related to language use competence; furthermore, it has discovered the additional dimension of multimodal writing performance that is independent from the language use competence. Based on the findings, it could be concluded that a multimodal writing task can be used as a language task with additional support of nonlinguistic resources, which have not yet been discussed to be relevant to language. The nonlinguistic dimension may be a construct-irrelevant variance if the target of measurement is language, but it could become a sub-construct that the traditional monomodal tasks failed to consider if we consider writing proficiency in the current time. Two types of research can address this issue of visualization in contemporary academic writing. The first step is investigating to what extent language learners and instructors find it important to use visualization in completing multimodal writing tasks. The

visual mode of communication is integral to academic writing ability in professional setting (e.g., Lemke, 1998; Tardy, 2005), but it cannot be concluded that it is equally important as language for L2 learners. More focused survey studies on the stakeholders' views on visualization can inform its magnitude in the construct of academic writing. Furthermore, empirical research on the what means good visualization in academic writing can directly inform the construct of visualization. For example, an eye-tracking study of raters' rating behavior can quantify what raters attend to during their evaluation process. Retrospective and introspective verbal recalls of evaluating processes can also shed lights on raters' internal criteria for determining good versus poor visualization performance. Therefore, studies validating the role of visualization in L2 academic texts can advance the knowledge of contemporary L2 academic writing.

Despite overlapping constructs that are related to language use competence, it should be noted that the language modality of the two tasks are fundamentally different. The monomodal writing task only elicited written output, but the multimodal writing task performance required verbal language, mostly speaking based on their written scripts. Language use dimension of the multimodal performance might be related to writers' speaking ability as well. Speaking in academic multimodal tasks requires writers to integrate "writing" and "speaking" as shown in the multimodal writing process result. Much research has focused on receptive and productive skills integration (e.g., Barkaoui, Brooks, Swain, & Lapkin, 2013; Chan, Inoue, & Taylor, 2015; Plakans, 2010; Plakans, Liao, & Wang, 2019), but little has been discussed for productive skills integration such as writing-to-speak tasks and speaking-to-write (Rubin & Kang, 2008).

Another possible interpretation of the overlapping language construct is that language use competence is a universal construct for language performance across of modalities. There have been, of course, a wealth of research that showed within-participant discrepancy between

speaking and writing skills, but it could be due to other mediating factors that come into play at the production stage (e.g., genre knowledge for writing and interpersonal skills for speaking). This is a highly speculative argument, but it might be worth mentioning that in psycholinguistic research measures of L2 linguistic knowledge also adopt different modalities for tests (e.g., speaking in an elicited imitation task and an oral production task, and writing for a metalinguistic knowledge task) and have not yet found conclusive evidence that these tests tap into different constructs; indeed, Vafaee, Suzuki, & Kachisnke (2017) found that all measures may be measuring a universal construct of language. Studies of multimodal writing are copious but most of them focus on how nonlinguistic modes can contribute to meaning making, with little attention to the language modes which should be of the most relevance to learners.

Findings that a good command of language and visualization quality imply that subcategories of an analytic rubric for multimodal writing. Researchers have used some analytic rubrics for multimodal performance evaluation (e.g., Hung et al., 2013; Kang & Kim, 2019), but none of them empirically investigated (1) which categories to include; and (2) to what extent the proposed categories contribute to the multimodal text quality. For example, Hung et al.'s (2013) rubric has five categories (linguistic, visual, gestural, auditory, and spatial designs) with equal weights, but a validation study may further investigate the relative importance of each category to the multimodal writing performance. Kang and Kim (2019) adapted the Burnett et al.'s rubric with five categories (i.e., task fulfillment, content, organization, language and mechanics, and effectiveness of using multi-modes) with equal weight (total possible 5 point for each category). Kang and Kim used this analytic rubric for different multimodal tasks, including argumentative videos, video book reviews, and promotional videos. It is questionable, however, if this analytic rubric can serve to validly interpret writers' multimodal writing ability because it simply added

one subcategory for multimodal communication to a somewhat traditional analytic writing rubric. Additionally, there are two levels in the subcategories. Task fulfillment, content, and organization can be achieved thorough either language or other nonlinguistic modes. For example, transitions could be marked in both linguistic (lexical items) and nonlinguistic modes (e.g., slide transition and the addition of a textbox with a word "but" in Table 13 in Chapter 4).

While the previous studies devised a rubric based on a theory (i.e. Multiliteracies by The New London Group) and a program's goals (Burnett, Frazee, Hanggi, & Madden, 2014), a bottom-up approach can build more valid rubrics for evaluation. One possible approach is to explore the categories for an analytic rubric is by conducting a genre analysis of the academic written texts as D'Angelo, 2010, 2016) and Tardy (2005, 2008) showed. D'Angelo's work, especially, provided more concrete examples of how metadiscourse in academic posters are achieved through linguistic and nonlinguistic modes and how such practice is specific to disciplines. Genre analyses will be able to discover how genre knowledge on the discourse- and metadiscourse-levels, in addition to language knowledge, contributes to a multimodal text quality. It should be noted that skill of composing visual elements is, however, may not be necessarily the matter of language learning. Therefore, future research needs to address (1) what makes good visualization for multimodal texts; and (2) whether this visualization quality is relevant to the construct of a language task or irrelevant to the construct. An empirically developed analytic rubric could be of help to teachers who may not be confident in evaluating multimodal texts (Yi & Angay-Crowder, 2016) and researchers who untilize multimodal writing tasks as language tasks.

Discussion of Teaching and Assessment

Multimodal writing task development. In Study 1, I investigated what types of multimodal writing tasks are needed in undergraduate courses and revealed three themes explaining the types of authentic academic multimodal tasks. First, I found that multimodal tasks in academic contexts serve a wide range of roles depending on the instructional goal. Some tasks are designed to reflect students' needs for effective disciplinary practice (e.g., academic posters and lab reports), while other tasks intend to make students better aware of the wide availability of semiotic ensembles to communicate their meaning effectively (e.g., digital storytelling). Disciplinary multimodal tasks have been found to involve linguistic and nonlinguistic modes of communication that follow their disciplinary conventions (e.g., D'Angelo, 2010, 2016; Rowley-Jolivet, 2002, 2012). Professionals using such disciplinary multimodal tasks expect their students to be accustomed to academic conventions of multimodal texts. Given the value of identifying genre-specific linguistic patterns by text-oriented ESP research (i.e., Swalesian genre research), disciplinary multimodal texts would also merit various lenses of genre analysis to offer suggestions for material developers. As Tardy (2005) noted earlier in her study of academic presentation slides, multimodal genres have been common in professional settings, and these genres merit more attention than they had received.

Together with their imporant role for promoting L2 learners' clear authorial voice and identity (Cimasko & Shin, 2017; Jiang, 2018), multimodal tasks have been found to facilitate language development (Dzekoe, 2017; Vandommele et al., 2017). This finding potentially indicates that, given their multifaceted pedagogical values, expressionistic multimodal tasks for learner agency can also be used in L2 writing instruction with the goal of language development. It should be noted, however, that the issue of the role of multimodal writing tasks for language

development needs more empirical evidence so that multimodal tasks can be designed and implemented with a clear understanding of how they allow students to notice linguistic forms and structures, eventually leading to language development.

Tasks such as an academic presentation or some manifestations of the remix task do not lead students to produce much written alphabetic text, while such alphabetic texts are expected to scaffold the development of multimodal texts (e.g., script writing for presentation and digital storytelling). This observation may indicate that multimodal task performance involves more planning for language formulation and production than we have expected. Furthermore, academic multimodal texts include spoken language that is extensively planned and rehearsed in written language, which may not be captured through conventional speaking tests measuring spontaneous and impromptu speech. This finding offers a pedagogical suggestion of the possibility of viewing writing as a means supporting oral task performance (Rubin & Kang, 2008).

Given the interest in L2 acquisition that views writing as a way to facilitate acquisition (Manchón & Vasylets, 2019; Williams, 2012), the use of monomodal writing as a pre-multimodal task production step, as Jiang (2018) and Dzekoe (2017) did for their participants, might address Manchón's (2017) concern that multimodal tasks may not facilitate acquisition. In fact, multimodal tasks were oftentimes accompanied by pre-tasks and follow-up tasks. The most frequent tasks were reflection essays after completing a multimodal task. For the integration of multimodal tasks in the existing EAP curriculum and sequencing, it would thus be appropriate to include a narrative writing task of the multimodal writing process. By doing so, EAP students would have chances to reflect on their multimodal repertoire. Furthermore, it might help promote

metalinguistic awareness if the instructors encouraged students to consider their language learning process throughout the multimodal composing process.

While there has been little discussion on the sequencing of multimodal tasks for language development, the Multiple Representation Hypothesis (Flower & Hayes, 1984) may provide some ideas about how to sequence multimodal tasks with different amounts of focus on language. For example, if a writing topic heavily requires nonverbal modes of representation in one's mind (e.g., shapes, colors, motion), it would be easier to begin with a multimodal task that involves more nonlinguistic modes of communication. Based on the findings that multimodal tasks are often coupled with monomodal writing tasks, such as reflection essays or a pre-writing essay for a resemiotization multimodal task, the multimodal task can be followed up with multimodal writing tasks to describe the nonlinguistic multimodal ensembles and reflect on the composing processes. Thus, based on the representation modes, a multimodal project can be segmented into smaller units whose amount of language focus increases gradually.

Lastly, individual tasks were far more common for multimodal writing tasks. Interestingly, previous studies of multimodal writing also employed individual tasks (except Vandommele et al., 2017) through which participants expressed clear authorial voice and identity. Given collaborative writing tasks have become increasingly common in ESL classes for peer interaction and scaffolding (e.g., Storch, 2005; Wigglesworth & Storch, 2009; Zhang, 2019), more research on the collaborative multimodal tasks needs to shed light on how learners interact each other for challenges while exploiting various semiotic modes and jointly advance their multimodal writing ability.

Perceptions of the multimodal writing task. This study found that participants perceived the multimodal task more interesting than the monomodal task. This finding is in line what previous studies that somewhat coherently emphasized students' positive evaluation on multimodal tasks (Dzekoe, 2017; Jiang, 2018). What this study adds to research is that the multimodal task was perceived as a more complex and difficult task than the monomodal task. Given that both monomodal and multimodal tasks students completed were on a similar topic with time limit, arguably students' diverging perceptions on the two tasks are rooted in the differences in the available modes of writing such as the number of possible modes to express and technological difficulties in using nonlinguistic modes. On the other hand, it is possible that the increased difficulty and interest is due to the multimodal task representation, or the lack thereof, especially in timed setting. Longitudinal investigation on students' perception could wash out the potential contribution of the "new", which is a compounding effect, to students' perception. In addition, following Jung et al.'s (2019) study, future research can focus on the changes of students' perceptions on the usefulness of multimodal task on language learning.

A caveat is that what students think, however, may not always best represent the needs or goals of instruction. Oftentimes researchers have found gaps between instructors' goals and students' goals (Polio et al., 2018; Yoon, 2019; Zhou et al., 2014). Thus, instructors' perceptions on the multimodal writing tasks need to be further investigated to evaluate and design multimodal writing tasks for learners. Together, the findings of the two studies point to what considerations multimodal writing tasks should be developed for language learning and contextualized into academic language courses.

Implications

Research implications. Methodologically, this dissertation project introduced notable methodological advances. In terms of research design, I integrated two mixed methods designs—an exploratory sequential design (i.e., Study 1 → Study 2) and a concurrent parallel design (i.e., multimodal writing performance, processes, and writers' perceptions on the multimodal task)—to address a larger question: What do instructors and material developers need to consider when using multimodal writing tasks for their language learners? To explore and document the characteristics of multimodal writing tasks for L2 writers, I conducted two studies that focused on the authentic tasks (Study 1) and a timed task developed for an experimental study. Multimodal writing can be examined and explained from different theoretical perspectives including social semiotics, systemic functional linguistics, genre studies, and the cognitive model of writing. Instead of choosing one theoretical framework for research, I chose a pragmatic approach to best answer my research questions. For Study 1, I relied on the concepts from social semiotics that have explained multimodal writing in most of the previous literature; however, I found that that genre studies can explain on what grounds undergraduate course instructors design their multimodal tasks. Study 2, however, is mostly based on the cognitive model of writing. For example, the coding schemes of the writing process data were developed in reference to the latest model by Leijten et al. (2013); task perception data are related to the task schemas within the cognitive model of writing.

Despite the eclectic approach of the current project, this project, specifically Study 2, is the first study to view multimodal writing process from a cognitive perspective. Multimodal writing has been discussed in various research orientations—including social semiotics and systemic functional linguistics that describe and analyze writers' situated orchestration of

resources to best communicate their ideas and a metafunction system of available resources (Bezemer & Kress, 2008; Daly & Unsworth, 2011; Lemke, 1998; O'Halloran, 2004). These studies have added valuable insights as to how L2 writers use the newly available nonlinguistic resources along with language, which has been the dominant mode of communication, and shed light on how they develop their identities with the composing experiences (e.g., Cimasko & Shin, 2017; Jiang, 2018; Pyo, 2016; Tardy, 2005). However, underexplored was a link between the cognitive model of writing, which has influenced much of the current knowledge in L2 writing research. Addressing this gap with data triangulation for a valid interpretation, I revealed and quantified control-level and process-level multimodal writing processes. Using the cognitive model of writing to investigate L2 writers' writing processes, this study further validated the current model by Leijten et al. (2013) that added and changed components for multimodal writing (e.g., transcription technology, design schema, texts-and-graphics-created-so-far).

Methodologically, when looking into writing processes, I utilized both online and offline records of multimodal writing processes. With the online method, I was able to see how much time each writer spent on different writing processes, and the evolutionary trajectories of the seven distinct writing processes as Gánem-Gutiérrez and Gilmore (2018) did. However, given that the on-screen writing behaviors cannot give a clue as to what writers meant to do, the results only provide a picture of observable behaviors. Another layer of results was derived from the stimulated recall data. While stimulated recall interview can only capture what participants remember and verbalize after the writing event is over, it provides valuable information as to their intentions and ideas that may or may not be translated into multimodal texts. By triangulating two data sources, I was able to describe multimodal writing processes with regards to writers' intentions and observable behaviors. Clearly, not all plans were translated into texts or

behaviors due to various reasons as provided in examples. For example, time constraint affected writers' changes in writing plans to downsize; the fact that the writing task does not have a real impact on their academic standing also played a role in not making every effort to improve their texts; hence, they anticipated that their audience would not recognize small mistakes.

Additionally, this study is one of the few studies that analyze quantitative data to investigate multimodal writing (e.g., Dzekoe, 2017; Kim et al., 2019; Vandommele et al., 2017), which aims to reveal systematic patterns across participants. First, I revealed that 83% of the overall quality of multimodal writing performance is predicted by language and verbal delivery quality and visualization quality of the multimodal performance, with the former being a stronger predictor than the latter. Furthermore, the overall quality and the language and verbal delivery scores were significantly and strongly correlated with the L2 writers' essay writing total and subscores. Second, multimodal writing process was also examined qualitatively with excerpts and examples and quantitatively with frequency and time duration data. Therefore, this project adds a new perspective to the inquiry of multimodal writing and provides base for further quantitative studies on the topic.

Pedagogical implications. Considering the findings that revealed the importance of language in both multimodal writing processes and performance, I suggest that multimodal writing tasks can be implemented in language curricula. In what follows, I discuss how instructors can design and implement multimodal writing tasks for L2 writers, and what those tasks contribute to language instruction, which have been the unanswered questions from L2 writing researchers and teachers (Manchón, 2017; Polio, 2019; Qu, 2017).

When it comes to task development, instructors need to first assess their student's learning goals and needs and explore what types of multimodal writing tasks they will encounter

as they use L2 multimodal writing beyond language classrooms. With an assumption that the immediate goal of the language learners in a University's English language courses is to acquire language competence for academic success, I explored the types of multimodal writing tasks that non-language courses across disciplines require. The three themes from Study 1 provide a general idea of task features to task developers. In general, multimodal writing tasks are divided into two categories, those with an emphasis on academic socialization and those focusing on individual ways of using multiple modes for best meaning making. For international graduate students, it may be the case that they immediately need to learn and practice academic multimodal genre conventions. But undergraduate students with limited language and undecided study may benefit from a digital storytelling project that allows them to use different modes, and potentially increases awareness on each of the modes of communication. On a similar note, it would be problematic to use multimodal task if the instructional goal is exclusively on the grammar development. Language was important for multimodal performance, but there was certainly a large amount of contribution by visualization.

Task developers can also consider potential language activities that are likely to occur procedurally as L2 writers complete a multimodal task. Among the diverse underlying cognitive multimodal writing processes, translation processes from ideas to written scripts and rehearsed speech can be a venue for language learning. Here, writing becomes a tool for writing and gives much time for careful language choices. To promote such writing-to-learn activities, tasks can include explicit guideline for writing scripts or written planning. In addition, multimodal writing tasks can be geared toward language by integrating languaging (i.e., self-explaining language problems) or reviewing activities. When L2 writers participated in the stimulated recall interviews, they were able to identify global and local language problems. This indicates that

additional time for reviewing their texts can help writers to use their linguistic knowledge to notice and solve language problems. Suzuki's (2012) study of the instructional effects of written languaging on L2 writing revisions found its facilitative effects even though his participants used their L1 the treatment. But given that L2 writers used English for Internet searching and were able to write L2 for planning purposes, L2 languaging would be a potential option for L2 writers as well. Lastly, instructors should take into account additional communication practice through collaboration. Based on the results of Study 1, it might be more authentic to use individual multimodal writing tasks, but collaborative tasks provide further venues to practice L2 for communication. Practically in the U.S. higher education setting, it is likely that L2 writers don't share their first language for collaboration, and thus use their L2 for communication.

Nevertheless, the emphasis on language does not mean that instructors should avoid discussing nonlinguistic modes of communication in class. As an extract from Participant 27's multimodal performance shows (in Table 13, Chapter 4), a multimodal task introduces another way expressing ideas than language. While visual information can construct messages, it is also possible that L2 learners may use the additional semiotic tools as a complementary strategy. If a course goal is on content, for example Grapin's (2018) case of k-12 science learning, this complementary strategy could be helpful. However, I argue that in the case of EAP courses, instructional focus still needs to be on language development with secondary focus on the nonlinguistic resources. When introducing nonlinguistic modes of communication, genre-oriented instruction might be the most appropriate for language classroom because it aims to raise writers' awareness on multiple modes of communication, such as language and visual resources (e.g., Coccetta, 2018; Molle & Prior, 2008). Furthermore, to explicitly discuss patterns in effective multimodal modes, instructors may need to familiarize themselves with

metalanguage (Shin, Cimasko, & Yi, 2020; Shipka, 2005) and research on how to do genre analysis on multimodal texts (e.g., D'Angelo, 2016).

Additionally, this explicit discussion on the rules underlying multimodal texts between instructors and students will provide a co-constructed grading criteria including language and designing aspects (Yi et al., 2017; Yi, Shin, & Cimasko, 2019). Assessing multimodal task performances has been considered a subjective decision, but interestingly, raters achieved high reliability with their impressionistic scorings for two specific scores (i.e., language and verbal delivery, and visualization) and overall quality scores. While accurately describing raters' internal criteria, rubrics of multimodal writing tasks may need to emphasize language category for increased pedagogical values. Because multimodal writing tasks could mislead writers to focus on nonlinguistic modes at the expense of improving language, a rubric with explicit and detailed expectations for language will help them orient toward language learning.

Together with the findings of the current project on the important role of language in multimodal writing performance and processes, L2 writers' positive perceptions on the multimodal tasks point to the potential values of using multimodal writing tasks in language classes. L2 writers found that the multimodal writing task is more complex, difficult, and interesting. Considering a plethora of research showing learners' positive experiences in using multimodal tasks (e.g., Cimasko & Shin, 2017; Dzekoe, 2017; Jiang, 2018; Smith & Dalton, 2016), using multimodal writing tasks can promote L2 learners' engagements with L2 writing. While developing multimodal writing tasks for language learners may cause burden to instructors and material developers (Yi et al., 2017), learners will be able to learn language through more engaging tasks that are relevant real-world multimodal tasks.

Limitations and Future Research

In this dissertation project, I conducted two studies that respectively explored (1) what multimodal writing tasks L2 writers complete in the target language use domain and (2) how L2 writers perform on a timed multimodal writing task. Based on the findings, I suggested possible ways to use multimodal writing tasks as language tasks and research implications. Nevertheless, there are several limitations that need to be addressed in future research.

This study did not examine how specific linguistic and nonlinguistic features predict multimodal task performance (except for the systemic contribution of visualization to the quality of multimodal texts). Textual and visual features of presentation genres have been discussed in previous multimodal genre studies (e.g., D'Angelo, 2016; Rowley-Jolivet, 2002, 2012; Tardy, 2005), but their focus has mostly been on discipline-specific academic genres that are clearly distinct from general multimodal texts composed under time constraints. Given the need to offer L2 writers some explicit guidelines and rubric descriptors for multimodal tasks, future studies may need to examine what linguistic and nonlinguistic features contribute to better multimodal task performance. Furthermore, drawing on insights from social semiotics and systemic functional grammar studies, further research can examine what components should be included in the analytic rubric of multimodal tasks.

Also, while both studies revealed that nonlinguistic modes play significant roles in constructing multimodal messages, the qualitative analysis of how individual semiotic resources contribute to meaning construction was beyond the scope of this project. Exploring native and non-native speakers' multimodal texts in terms of nonlinguistic elements can offer evidence of what resources L2 writers would need additionally for successful multimodal writing performance. It would also be useful to explore how L2 speakers with different academic

experiences vary in their use of nonlinguistic elements. In addition, investigating how EAP instructors use and evaluate multimodal writing tasks can shed light on the important multimodal elements for language learners. Thus, future research can expand the scope of the study by fully analyzing linguistic and nonlinguistic features in a given target language use domain.

The small sample sizes of the studies may limit the generalizability of the findings. In an attempt to address this sample size issue, in Study 2, I used the bootstrapping method in analyzing multimodal writing performance (n = 29) and task perceptions (n = 31). Nevertheless, it should be acknowledged that, as a statistical procedure of selecting smaller samples out of the sampled participants, bootstrapping cannot fundamentally address the limitations of the small sample sizes. For the writing process data, I did not use the bootstrapping methods because its sample size might be too small for any parametric test (n = 12). Replicating the methods of the current project, future research with a larger sample will be able to provide more generalizable findings and then to advance our understanding of multimodal tasks.

While this study used the impressionistic scoring of multimodal writing, it can be predicted that using an analytic rubric with multiple components would have offered more meaningful results that help clarify student's multimodal writing proficiency. However, because there is little empirical evidence of the subskills of the construct of multimodal writing proficiency, I determined to have expert raters evaluate multimodal texts in terms of their overall perceived quality. As a result, this study could inform us of the relationship between multimodal performance and traditional written language performance. With this finding as a starting point, future studies can be conducted to design a scoring rubric for multimodal tasks. For example, L2 instructors and researchers can use their expertise to identify some sub-constructs of multimodal task performance, and then explore statistically to what extent each of the sub-constructs

explains overall multimodal text quality. Ultimately, we may need to build a model that can predict multimodal text quality using quantifiable linguistic, nonlinguistic, and other multimodal features (e.g., number of graphs, number of slides, number of running words on slides). Finally, the multimodal task designed for this project could be questioned in terms of its authenticity because, in previous research, most multimodal writing tasks have been implemented as long-term projects for which writers are expected to explore and experiment with various semiotic resources (e.g., Cimasko & Shin, 2017; Dzeokoe, 2017; Vandommele et al., 2017). Unlike the majority of previous multimodal tasks, I designed a timed multimodal task that requires writers to use their writing and design schemas, and to use linguistic and nonlinguistic modes for meaning expression. The findings of this study showed that this multimodal task, despite its lower flexibility, still encouraged the L2 writers to make varied use of multiple modes of communication successfully, and also that they found the current multimodal task more interesting than the traditional essay task. It was also revealed that the timed multimodal task allowed the writers to use both spoken and written language for multiple purposes. Given these benefits, I conclude this project by calling for more research that aims to offer valid multimodal writing tasks for language learners. The exploration of multimodal writing from various perspectives and methodologies will enable us to achieve a fuller picture of the relationship of multimodal writing with L2 writing proficiency, task-based language teaching, and social semiotics.

APPENDICES

APPENDIX A.

Interview Questions

1. What is your area of expertise and which undergraduate courses have you taught/ are you teaching?
2. In your classes, what kinds of writing assignments do students need to complete?
3. Do you have multimodal writing assignments in your syllabus?

   Multimodal writing assignments are those ask students to use multiple modes such as audio, video, still images, tables, figures, and so forth, along with some alphabetical text.

   If Yes: *Proceed to Question 4.*

   If No: *Proceed to Question 8.*
4. Describe your multimodal assignments.
5. Do you have explicit guidelines for multimodal assignments?
   a. **If Yes:** What components do you consider in the guidelines?

      If No: Please describe reasons why.

   b. Are you aware of any resources (websites, people) that your students often refer to?
6. Do you have explicit evaluation criteria for those assignments?

      a. **If Yes:** What are your evaluation criteria for the multimodal assignments?

   If No: If you don't have explicit criteria, please describe reasons why.

      b. What makes some multimodal texts score higher than others?
7. Describe how you developed multimodal writing tasks for your students.
8. Why did you decide (not) to include those assignments?
9. Do you think multimodal tasks are different from essay-type tasks? Why?
10. Do you think students who write good essays also compose good multimodal texts? Why?

APPENDIX B.

Table A.1. Study 2 Participants' Background Information

ID	Test	Score	Age	Major	Study Abroad Experience	Counter balancing*
p001	iBT TOEFL	82	23	NA	5 months	Group A
p002	iBT TOEFL	84	22	English language and literature	-	Group A
p003	iBT TOEFL	100	24	Education Technology	-	Group A
p004	IELTS	7	27	Sculpture	-	Group A
p005	iBT TOEFL	85	21	Business	-	Group A
p006	iBT TOEFL	89	23	Chemistry and Nano Science; Business	6 months	Group A
p007	iBT TOEFL	115	19	Physics	4 years	Group A
p008	iBT TOEFL	86	23	Art history	-	Group A
p009	iBT TOEFL	98	23	Chemistry	-	Group A
p010*	iBT TOEFL	110	21	International Studies	9 months	Group B
p011	iBT TOEFL	110	19	Chemistry and Nano Science	-	Group A
p012	iBT TOEFL	103	26	Math Education	2 years	Group A
p013	iBT TOEFL	112	21	International Studies	6 years	Group A
p014	iBT TOEFL	103	24	Science	2 years	Group A
p015	iBT TOEFL	100	24	Economics	-	Group A
p016	iBT TOEFL	97	22	Food Engineering	-	Group A
p017*+	iBT TOEFL	81	23	Food and Nutrition	-	Group B
p018	IELTS	5.5	30	Law	6 months	Group B
p019	iBT TOEFL	107	28	English Education	1 year	Group B
p020	iBT TOEFL	108	23	Politics and International Relationships	-	Group B
p021	iBT TOEFL	92	21	Science Education	-	Group B
p022	iBT TOEFL	92	21	Communication; History	5 months	Group B
p023	iBT TOEFL	82	23	International Office Administration	5 months	Group B
p024	iBT TOEFL	109	25	Economics; Applied Mathematics	9 years	Group B
p025	iBT TOEFL	90	21	Electrical Engineering	6 months	Group B
p026	iBT TOEFL	94	24	Business Administration	4 years	Group B
p027	iBT TOEFL	91	21	English Education	-	Group B
p028	iBT TOEFL	82	21	Environmental Engineering	4 months	Group B
p029	iBT TOEFL	112	19	English Education	5 years	Group B
p030	iBT TOEFL	104	24	Philosophy	9 months	Group B
p031+	iBT TOEFL	105	22	Psychology	-	Group B

* Two participants excluded from writing process data

+ Two participants excluded from writing performance data

APPENDIX C.

Background Questionnaires

1. ID: _____
2. Age, Gender: _____, _____
3. iBT TOEFL score: _____
4. What is your major? _____
5. How many semesters have you studied in the college? _____ semesters
6. Have you studied abroad in English-speaking countries?

   ☐ Yes: when and how long did you stay? _____

   ☐ No
7. How many courses you took were taught in English? _____
8. What kinds of multimodal writing have you done in Korean and English?

	I completed this in Korean	**I completed this in English**
Academic writing without visual component		
Academic writing with visual component		
Written analyses of media		
Digital storytelling		
Video presentation		
Presentations with slides in class		
Poster presentation		
Websites		
Résumés		
Observation note		
Others:		

9. Rank the five most frequently used multimodal assignments from the list above:

1.	2.	3.	4.	5.

10. Mark your level of ability to do the following tasks on your computer.

	extremely challenging								extremely easily
1. insert pictures and graphs in my documents	1	2	3	4	5	6	7	8	9
2. find good images and videos on internet	1	2	3	4	5	6	7	8	9
3. record and insert audio files to a presentation file	1	2	3	4	5	6	7	8	9
4. draw graphic elements on a slide	1	2	3	4	5	6	7	8	9
5. place objects on slides as desired	1	2	3	4	5	6	7	8	9

11. Please mark level of **your ability** to do the followings **in English**:

	extremely challenging								extremely easily
Interpretive Communication (Reading)									
1. read literary texts with ease	1	2	3	4	5	6	7	8	9
2. interpret historical and political documents in the tart language (i.e., archived letters, pamphlets, or speeches).	1	2	3	4	5	6	7	8	9
3. learn how to read and translate texts as they relate to my professional field of interest	1	2	3	4	5	6	7	8	9
4. analyze literary, sociological, and philosophical texts in the target language (i.e., primary sources studied by literary scholars, sociologists, and philosophers)	1	2	3	4	5	6	7	8	9
5. read texts in the target language that encourage abstract modes of thinking	1	2	3	4	5	6	7	8	9
6. understand non-fiction tests such as essays, documentaries, and technical documentation	1	2	3	4	5	6	7	8	9
7. read professional articles in the target language in my field of study	1	2	3	4	5	6	7	8	9
8. read and interpret literary reviews	1	2	3	4	5	6	7	8	9
9. understand the subtleties of political satire in cartoons, essays, or blogs	1	2	3	4	5	6	7	8	9
Interpretive Communication (Listening)									
10. understand and enjoy fiction in various media (i.e., books, films, or TV) with ease	1	2	3	4	5	6	7	8	9

	extremely challenging								extremely easily
11. engage in activities in class that focus on the interpretation of film, commercials, and video.	1	2	3	4	5	6	7	8	9
12. interpret plays and theatrical performances	1	2	3	4	5	6	7	8	9
13. follow the reporting of national or international televised news in the target language	1	2	3	4	5	6	7	8	9
14. follow a lecture on a subject within my field of study	1	2	3	4	5	6	7	8	9
Presentational Communication (Writing)									
15. learn how to write well-constructed compositions and essays in various genres (i.e., narrative, descriptive, or persuasive essays).	1	2	3	4	5	6	7	8	9
16. learn about how to write an analytical essay in the target language	1	2	3	4	5	6	7	8	9
17. learn how to write creatively in the target language	1	2	3	4	5	6	7	8	9
18. learn how to write an in-depth research or position paper in the target language	1	2	3	4	5	6	7	8	9
19. be able to write the content for a multimedia presentation	1	2	3	4	5	6	7	8	9
Presentational Communication (Speaking)									
20. learn about better ways to present information and concepts orally in the target language	1	2	3	4	5	6	7	8	9
21. give presentations on literary or philosophical texts	1	2	3	4	5	6	7	8	9
22. be able to give a clearly articulated and well-structured presentation	1	2	3	4	5	6	7	8	9
Grammar									
23. review and refine my knowledge of various grammatical structures	1	2	3	4	5	6	7	8	9
24. learn to speak and write in the target language without making grammatical mistakes	1	2	3	4	5	6	7	8	9

	extremely challenging								extremely easily
25. analyze and understand the use of grammatical structures in a text	1	2	3	4	5	6	7	8	9
Multimodal Communication									
26. collaboratively use visual and linguistic resources for writing	1	2	3	4	5	6	7	8	9
27. design visual elements to present ideas	1	2	3	4	5	6	7	8	9
28. make graphs when I need to present quantitative findings	1	2	3	4	5	6	7	8	9
29. relate visual information and linguistic information when I read multimodal texts	1	2	3	4	5	6	7	8	9
30. learn about better ways to collaborate different modes to communicate message clearer.	1	2	3	4	5	6	7	8	9

APPENDIX D.

Rubric for the Monomodal Writing Task

	Content		**Organization**		**Vocabulary**		**Language Use**	Score/2	**Mechanics**
20 16	Thorough and logical development of thesis Substantive and detailed No irrelevant information Interesting A substantial number of words for amount of time given	20 16	Excellent overall organization Clear thesis statement Substantive introduction and conclusion Excellent use of transition word Excellent connections between paragraphs Unity within every paragraph	20 16	Very sophisticated vocabulary Excellent choice of words with no errors Excellent range of vocabulary Idiomatic and near native-like vocabulary	20 16	No major errors in word order or complex structures No errors that interfere with comprehension Only occasional errors in morphology Frequent use of complex sentences Excellent sentence variety	20 16	Appropriate layout with indented paragraphs No spelling errors No punctuation errors
15 11	Good and logical development of thesis Fairly substantive and detailed Almost no irrelevant information Somewhat interesting An adequate number of words for the amount of time given	15 11	Good overall organization Clear thesis statement Good introduction and conclusion Good use of transition words Good connections between paragraphs Unity within most paragraphs	15 11	Somewhat sophisticated vocabulary Attempts, even if not completely successful, at sophisticated vocabulary Good choice of words with some errors that don't obscure meaning Adequate range of vocabulary but some repetition Approaching academic register	15 11	Occasional errors in awkward order or complex structures Almost no errors that interfere with comprehension Attempts, even if not completely successful, at a variety of complex structures Some errors in morphology Frequent use of complex sentences Good sentence variety	15 11	Appropriate layout with indented paragraphs No more than a few spelling errors in less frequent vocabulary No more than a few punctuation errors
10 6	Some development of thesis Not much substance or detail Some irrelevant information Somewhat uninteresting Limited number of words for the amount of time given	10 6	Some general coherent organization Minimal thesis statement or main idea Minimal introduction and conclusion Occasional use of transitions words Some disjointed connections between paragraphs Some paragraphs may lack unity	10 6	Unsophisticated vocabulary Limited word choice with some errors obscuring meaning Repetitive choice of words No resemblance to academic register	10 6	Errors in word order or complex structures Some errors that interfere with comprehension Frequent errors in morphology Minimal use of complex sentences Little sentence variety	10 6	Appropriate layout with most paragraphs indented Some spelling errors in less frequent and more frequent vocabulary Several punctuation errors
5 0	No development of thesis No substance or details Substantial amount of irrelevant information Completely uninteresting Very few words for the amount of time given	5 0	No coherent organization No thesis statement or main idea No introduction and conclusion No use of transition words Disjointed connections be-tween paragraphs Paragraphs lack unity	5 0	Very simple vocabulary Severe errors in word choice that often obscure meaning No variety in word choice No resemblance to academic register	5 0	Serious errors in word order or complex structures Frequent errors that interfere with comprehension Many error in morphology Almost no attempt at complex sentences No sentence variety	5 0	No attempt to arrange essay into paragraphs Several spelling errors even in frequent vocabulary Many punctuation errors

REFERENCES

REFERENCES

Adelson, J. L., & McCoach, D. B. (2010). Measuring the mathematical attitudes of elementary students: The effects of a 4-point or 5-point Likert-type scale. *Educational and Psychological Measurement*, *70*(5), 796–807. https://doi.org/10.1177/0013164410366694

Allison, P. D. (1999). *Multiple regression*. Thousand Oaks, CA: Pine Forge Press.

Alyousef, H. S. (2016). A multimodal discourse analysis of the textual and logical relations in marketing texts written by international undergraduate students. *Functional Linguistics, 3*(3), 1–29. https://doi.org/10.1186/s40554-016-0025-1

Anderson, D., Atkins, A., Ball, C., Millar, K. H., Selfe, C., & Selfe, R. (2006). Integrating multimodality into composition curricula: Survey methodolgy and results from a CCCC research grant. *Composition Studies*, *34*(2), 59–83.

Anderson, K. T. (2008). Contrasting systemic functional linguistic and situated lietracies approaches to multimodality in literacy and writing studies. *Written Communication*, *30*(3), 276–299. https://doi.org/10.1177/0741088313488073

Anderson, K. T., Stewart, O. G., & Kachorsky, D. (2017). Seeing academically marginalized students' multimodal designs from a position of strength. *Written Communication*, *34*(2), 104–134. https://doi.org/10.1177/0741088317699897

Archer, A. (2006). A Multimodal Approach to Academic 'Literacies': Problematising the Visual/Verbal Divide. *Language and Education*, *20*(6), 449–462. https://doi.org/10.2167/le677.0

Archer, A. (2010). Multimodal texts in higher education and the implications for writing pedagogy. *English in Education, 44*(3), 201–213. https://doi.org/10.1111/j.1754-8845.2010.01073.xdoi.org/10.1111/j.1754-8845.2010.01073.x

Arnheims, R. (2004). *Art and visual perception: A psychology of the creative eye* (2nd ed.). Oakland, CA: University of California Press.

Barkaoui, K., Brooks, L., Swain, M., & Lapkin, S. (2013). Test-takers' strategic behaviors in independent and integrated speaking tasks. *Applied Linguistics, 34*(3), 304–324. https://doi.org/10.1093/applin/ams046

Bateman, J. (2008). *Multimodality and genre: A foundation for the systematic analysis of multimodal documents*. London, UK: Palgrave Macmillan.

Belcher, D. (2017). On becoming facilitators of multimodal composing and digital design. *Journal of Second Language Writing, 38*, 80–85. https://doi.org/10.1016/j.jslw.2017.10.004

Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication*, *25*(2), 166–195. https://doi.org/10.1177/0741088307313177

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Burnett, R. E., Frazee, A., Hanggi, K., & Madden, A. (2014). A Programmatic Ecology of Assessment: Using a Common Rubric to Evaluate Multimodal Processes and Artifacts. *Computers and Composition*, *31*, 53–66. https://doi.org/10.1016/j.compcom.2013.12.005

Chan, S., Inoue, C., & Taylor, L. (2015). Developing rubrics to assess the reading-into-writing skills: A case study. *Assessing Writing*, *26*, 20–37. https://doi.org/10.1016/j.asw.2015.07.004

Chandler, P. D., Unsworth, L., & O'Brien, A. (2012). Evaluation of students' digital animated multimodal narratives and the identification of high-performing classrooms. *Journal of Literacy and Technology*, *13*(3), 80–127.

Chaudron, C., Doughty, C., Kim, Y., Kong, D.-K., Lee, J., Lee, Y.-G., … Urano, K. (2005). A task-based needs analysis of a tertiary Korean as a foreign language program. In M. H. Long (Ed.), *Second Language Needs Analysis* (pp. 225–262). Cambridge, UK: Cambridge University Press. https://doi.org/https://doi.org/10.1017/CBO9780511667299.009

Chun, D., Smith, B., & Kern, R. (2016). Technology in language use, language teaching, and language learning. *The Modern Language Journal*, *100*, 64–80. https://doi.org/10.1111/modl.12302

Cimasko, T., & Shin, D.-S. (2017). Multimodal resemiotization and authorial agency in an L2 writing classroom. *Written Communication*, *34*(4), 387–413. https://doi.org/10.1177/0741088317727246

Coccetta, F. (2018). Developing university students' multimodal communicative competence: Field research into multimodal text studies in English. *System*, *77*, 19–27. https://doi.org/10.1016/j.system.2018.01.004

Connor-Linton, J., & Polio, C. (2014). Comparing perspectives on L2 writing: Multiple analyses of a common corpus. *Journal of Second Language Writing*, *26*, 1–9. https://doi.org/10.1016/j.jslw.2014.09.002

Creswell, J., & Creswell, J. D. (2018). *Research design: Quantitative, qualitative, and mixed methods approaches* (5th Editio). Los Angeles: SAGE Publication.

Cumming, A., Kantor, R., & Powers, D. E. (2002). Decision making while rating ESL/EFL writing tasks: A descriptive framework. *The Modern Language Journal*, *86*(1), 67–96. https://doi.org/10.1111/1540-4781.00137

D'Angelo, L. (2010). Creating a framework for the analysis of academic posters. *Language Studies Working Papers*, *2*, 38–50.

D'Angelo, L. (2016). *Academic posters: A textual and visual metadiscourse analysis*. Bern, Switzerland: Peter Lang.

Dalton, B. (2012). Multimodal composition and the common core state standards. *Reading Teacher*, *66*(4), 333–339. https://doi.org/10.1002/TRTR.01129

Daly, A., & Unsworth, L. (2011). Analysis and comprehension of multimodal texts. *Australian Journal of Language and Literacy*, *34*(1), 61–80.

Depalma, M.-J., & Alexander, K. P. (2015). A bag full of snakes: Negotiating the challenges of multimodal composition. *Computers and Composition*, *37*, 182–200. https://doi.org/10.1016/j.compcom.2015.06.008

Dzekoe, R. (2017). Computer-based multimodal composing activities, self-revision, and L2 acquisition through writing. *Language Learning & Technology*, *21*(2), 73–95. Retrieved from http://llt.msu.edu/issues/june2017/dzekoe.pdf

Early, M., Kendrick, M., & Potts, D. (2015). Multimodality: Out from the margins of English language teaching. *TESOL Quarterly*, *49*(3), 447–460. https://doi.org/10.1002/tesq.246

Edwards-Groves, C. J. (2011). The multimodal writing process: Changing practices in contemporary classrooms. *Language and Education*, *25*(1), 49–64. https://doi.org/10.1080/09500782.2010.523468

Ellis, R. (2017). Position paper: Moving task-based language teaching forward. *Language Teaching*, *50*(04), 507–526. https://doi.org/10.1017/S0261444817000179

Ellis, R., & Yuan, F. (2004). The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, *26*(01), 59–84. https://doi.org/10.1017/S0272263104261034

Elola, I., & Oskoz, A. (2017). Writing with 21st century social tools in the L2 classroom: New literacies, genres, and writing practices. *Journal of Second Language Writing*, *36*, 52–60. https://doi.org/10.1016/j.jslw.2017.04.002

Field, A. (2013). *Discovering statistics using IBM SPSS* (4th ed.). Thousand Oaks, CA: SAGE Publication.

Flower, L. (1990). The role of task representation in reading-to-write. In Linda Flower, J. Ackerman, M. J. Kantz, K. Mccormick, W. C. Peck, & V. Stein (Eds.), *Reading-to-write: Exploring a cognitive and social process* (pp. 35–75). Oxford, UK: Oxford University Press.

Flower, Linda, & Hayes, J. R. (1980). The Cognition of Discovery: Defining a Rhetorical Problem. *College Composition and Communication*, *31*(1), 21–32.

Flower, Linda, & Hayes, J. R. (1981). A cognitive process theory of writing. *College Composition and Communication*, *32*(4), 365–387.

Flower, Linda, & Hayes, J. R. (1984). Images, plans, and prose: The representation of meaning in writing. *Written Communication*, *1*(1), 120–160. Retrieved from http://journals.sagepub.com/doi/pdf/10.1177/0741088384001001006

Fraiberg, S. (2010). Composition 2.0: Toward a Multilingual and Multimodal Framework. *College Composition*, *62*(1), 100–126.

Gánem-Gutiérrez, G. A., & Gilmore, A. (2018). Tracking the real-time evolution of a writing event: second language writers at different proficiency levels. *Language Learning*, *68*(2), 469–506. https://doi.org/10.1111/lang.12280

González-Lloret, M. (2014). The need for needs analysis. In L. Ortega & M. González-Lloret (Eds.), *Technology-mediated TBLT : Researching technology and tasks* (pp. 23–50). Amsterdam, The Netherlands: John Benjamins.

González-Lloret, M., & Ortega, L. (2014). *Technology-mediated TBLT: Researching technology and tasks*. Amsterdam, The Netherlands: John Benjamins Publishing Company.

Grapin, S. (2019). Multimodality in the new content standards era: Implications for English learners. *TESOL Quarterly*, *53*(1), 30–55. https://doi.org/10.1002/tesq.443

Grapin, S., & Llosa, L. (2020). Toward an integrative framework for understanding multimodal L2 writing in the content areas. *Journal of Second Language Writing*, 1–8. https://doi.org/10.1016/j.jslw.2020.100711

Hagan, S. M. (2007). Visual/verbal collaboration in print: Complementary differences, necessary ties, and an untapped rhetorical opportunity. *Written Communication*, *24*(1), 49–83. https://doi.org/10.1177/0741088306296901

Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of langauge and meaning*. London: Edward Arnold.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Edward Arnold.

Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. New York: Routledge.

Hayes, J. R. (1996). A new framework for understanding cognition and affect in writing. In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 1–27). Mahwah, New Jersey: Lawrence Erlbaum.

Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, *29*(3), 369–388. https://doi.org/10.1177/0741088312451260

Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), *Cognitive processes in writing* (pp. 3–30). NJ: Erlbaum: Hillsdale.

Howell, E., Butler, T., & Reinking, D. (2017). Integrating multimodal arguments into high school writing instruction. *Journal of Literacy Research*, *49*(2), 181–209. https://doi.org/10.1177/1086296X17700456

Hundley, M., & Holbrook, T. (2013). Set in stone or set in motion? Multimodal and digital writing with preservice English. *Journal of Adolescent & Adult Literacy*, *56*(6), 500–509.

Hung, H. T., Chiu, Y. C. J., & Yeh, H. C. (2013). Multimodal assessment of and for learning: A theory-driven design rubric. *British Journal of Educational Technology*, *44*(3), 400–409. https://doi.org/10.1111/j.1467-8535.2012.01337.x

Hyland, K. (2005). *Metadiscourse: Exploring interaction in writing*. London, UK: Continuum.

Jacobs, H., Zinkgraf, S., Wormuth, D., Hartfiel, V., & Hughey, J. (1981). *Testing ESL composition: A practical approach*. Rowley, MA: Newbury House.

Jeon, E. H. (2015). Multiple regression. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 131–158). New York: Routledge.

Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, *32*(1), 241–267. https://doi.org/10.3102/0091732X07310586

Jewitt, C. (2014a). An introduction to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (2nd Ed., pp. 15–30). New York: Routledge.

Jewitt, C. (2014b). Different approaches to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (2nd Ed., pp. 31–43). New York: Routledge.

Jiang, L. (2018). Digital multimodal composing and investment change in learners' writing in English as a foreign language. *Journal of Second Language Writing*, *40*(July 2017), 60–72. https://doi.org/10.1016/j.jslw.2018.03.002

Johnson, M. D. (2017). Cognitive task complexity and L2 written syntactic complexity, accuracy, lexical complexity, and fluency: A research synthesis and meta-analysis. *Journal of Second Language Writing*, *37*, 13–38. https://doi.org/10.1016/j.jslw.2017.06.001

Johnson, M. D., Mercado, L., & Acevedo, A. (2012). The effect of planning sub-processes on L2 writing fluency, grammatical complexity, and lexical complexity. *Journal of Second Language Writing*, *21*(3), 264–282. https://doi.org/10.1016/j.jslw.2012.05.011

Kang, S., & Kim, Y. (2019). A Programmatic Ecology of Assessment: Using a Common Rubric to Evaluate Multimodal Processes and Artifacts. In *Second Language Research Forum*. East Lansing, Michgan.

Kellogg, R. T. (1996). A model of working memory in writing. In (Eds.), (pp. 57- 71). Mahwah, NJ: In C. M. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57–71). Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.

Kellogg, R. T., Whiteford, A. P., Turner, C. E., Cahill, M., & Mertens, A. (2013). Worrking memory in written composition: An evaluation of the 1996 Model. *Journal of Writing Research*, *5*(2), 159–190.

Kim, Y., Belcher, D., & Peyton, C. (2019). Writing to make meaning through multimodal composing: Does it facilitate L2 writing development? In *Symposium on Second Language Writing*. Tempe, Arizona.

Kress, G. (2000). Multimodality: Challenges to thinking about language. *TESOL Quarterly*, *34*(2), 337–340.

Kress, G., & Van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. New York: Routledge.

LaFlair, G., Egbert, J., & Plonsky, L. (2015). A practical guide to bootstrapping descriptive statistics, correlations, t test, and ANOVAs. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 46–77). New York: Routledge.

Larson-Hall, J., & Herrington, R. (2009). Improving data analysis in second language acquisition by utilizing modern developments in applied statistics. *Applied Linguistics*, *31*(3), 368–390. https://doi.org/10.1093/applin/amp038

Lee, O., Llosa, L., Grapin, S., Haas, A., & Goggins, M. (2019). Science and language integration with English learners: A conceptual framework guiding instructional materials development. *Science Education*, *103*(2), 317–337. https://doi.org/10.1002/sce.21498

Leijten, M., Van Waes, L., Schriver, K., & Hayes, J. R. (2013). Writing in the workplace: Constructing documents using multiple digital sources. *Journal of Writing Research*, *5*(3), 285–337. https://doi.org/10.17239/jowr-2014.05.03.3

Lemke, J. (1990). *Talking science: Language, learning, and values*. Norwood, N.J: Ablex Publishing Corporation.

Lemke, J. (1998). Multiplying meaning: Visual and verbal semiotic in scientific texts. In J. R. MArtin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses* (pp. 87–113). New York: Routledge.

Levelt, W. J. (1989). *Speaking: From intention to articulation*. Cambridge, UK: Cambridge University Press.

Li, Z., & Lodge, S. (2017). Disciplinary differences in university lecture slides: A corpus-based and multimodal analysis. In *Canadian Association of Applied Linguistics Conference*. Toronto, Canada.

Lim, J., & Polio, C. (2020). Multimodal assignments in higher education: Implications for multimodal writing tasks for L2 writers. *Journal of Second Language Writing*, *47*(April 2019), 100713. https://doi.org/10.1016/j.jslw.2020.100713

Liu, Y., & O'Halloran, K. L. (2009). Intersemiotic texture: analyzing cohesive devices between language and images. *Social Semiotics*, *19*(4), 367–388. https://doi.org/10.1080/10350330903361059

Long, M. H. (2005). *Second language needs analysis*. Cambridge, UK: Cambridge University Press.

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

López-Serrano, S., Roca de Larios, J., & Manchón, R. M. (2019). Language reflection fostered by individual L2 writing tasks: Developing a theoretically motivated and empirically based coding system. *Studies in Second Language Acquisition*, *41*(3), 503–527. https://doi.org/10.1017/s0272263119000275

Lutkewitte, C. (2010). *Multimodality is…: A survey investigating how graduate teaching assistants and instructors teach multimodal assignments in first-year composition courses*. Ball State University.

Malicka, A., Guerrero, R., & Norris, J. M. (2017). From needs analysis to task design: Insights from an English for specific purposes context. *Language Teaching Research*, 136216881771427. https://doi.org/10.1177/1362168817714278

Manchón, R. M. (2017). The potential impact of multimodal composition on language learning. *Journal of Second Language Writing*, *38*, 94–95. https://doi.org/10.1016/j.jslw.2017.10.008

Manchón, R. M., & Vasylets, O. (2019). Language learning through writing: Theoretical perspectives and empirical evidence. In J. Schweiter & A. Benati (Eds.), *The Cambridge Handbook of Language Learning*. Cambridge, UK: Cambridge University Press. https://doi.org/https://doi.org/10.1017/9781108333603

Martinec, R. (2013). Nascent and mature uses of a semiotic system: the case of image–text relations. *Visual Communication*, *12*(2), 147–172. https://doi.org/10.1177/1470357212471603

Martinec, R., & Salway, A. (2005). A system for image–text relations in new (and old) media. *Visual Communication*, *4*(3), 337–371. https://doi.org/10.1177/1470357205055928

Mills, N., & Moulton, S. T. (2017). Students' and instructors' perceived value of language and content curricular goals. *Foreign Language Annals*, *50*(4), 717–733. https://doi.org/10.1111/flan.12303

Mogull, S. A., & Stanfield, C. T. (2015). Current use of visuals in scientific communication. In *IEEE International Professional Communication Conference*. https://doi.org/10.1109/IPCC.2015.7235818

Molle, D., & Prior, P. (2008). Multimodal genre systems in EAP writing pedagogy: Reflecting on a needs analysis. *TESOL Quarterly*, *42*(4), 541–566.

Morell, T. (2015). International conference paper presentations: A multimodal analysis to determine effectiveness. *English for Specific Purposes*, *37*, 137–150. https://doi.org/10.1016/j.esp.2014.10.002

Nelson, M. E. (2006). Mode, meaning, and synaesthesia in multimedia L2 writing. *Language Learning and Technology*, *10*(2), 56–76.

Nicolás-Conesa, F., Roca de Larios, J., & Coyle, Y. (2014). Development of EFL students' mental models of writing and their effects on performance. *Journal of Second Language Writing*, *24*(1), 1–19. https://doi.org/10.1016/j.jslw.2014.02.004

O'Halloran, K. L. (2004). *Multimodal discourse analysis: Systemic-functional perspectives*. London, UK: Continuum.

Pacheco, M. B., & Smith, B. E. (2015). Across languages, modes, and identities: Bilingual adolescents' multimodal codemeshing in the literacy classroom. *Bilingual Research Journal*, *38*(3), 292–312. https://doi.org/10.1080/15235882.2015.1091051

Palmeri, J. (2012). Creative Translations: Reimagining the Process Movement (1971-84). In J. Palmeri (Ed.), *Remixing Composition: A History of Mulitmodal Writing Pedagogy* (pp. 23–50). Carbondale, Illinois: Southern Illinois University Press.

Plakans, L. (2010). Independent vs. integrated writing tasks: A comparison of task representation. *TESOL Quarterly*, *44*(1), 185–194. https://doi.org/10.5054/tq.2010.215251

Plakans, L., Liao, J. T., & Wang, F. (2019). "I should summarize this whole paragraph": Shared processes of reading and writing in iterative integrated assessment tasks. *Assessing Writing*. https://doi.org/10.1016/j.asw.2019.03.003

Polio, C. (2019). Keeping the language in second language writing classes. *Journal of Second Language Writing*. https://doi.org/10.1016/S1060-3743(08)00005-2

Polio, C., & Friedman, D. (2017). Mixed-methods research. In C. Polio & D. Friedman (Eds.), *Understanding, evaluating, and conducting second language writing research*. New York: Routledge.

Polio, C., Tigchelaar, M., & Lim, J. (2018). Examining linguistic development in ESL writing: A mixed methods approach. In *TESOL Convention 2018*. Chicago, IL.

Pyo, J. (2016). Bridging In-School and Out-of-School Literacies: An Adolescent EL's Composition of a Multimodal Project. *Journal of Adolescent and Adult Literacy*, *59*(4), 421–430. https://doi.org/10.1002/jaal.467

Qu, W. (2017). For L2 writers, it is always the problem of the language. *Journal of Second Language Writing*, *38*, 92–93. https://doi.org/10.1016/j.jslw.2017.10.007

Reid, G., Snead, R., Pettiway, K., & Simoneaux, B. (2016). Multimodal communication in the university: Surveying faculty across disciplines. *Across the Disciplines*, *13*(1). Retrieved from https://wac.colostate.edu/docs/atd/articles/reidetal2016.cfm

Révész, A., Michel, M., & Lee, M. (2019). Exploring second language writers' pausing and revision behaviors. *Studies in Second Language Acquisition*, *41*(3), 605–631. https://doi.org/10.1017/s027226311900024x

Robinson, P. (2005). Cognitive complexity and task sequencing: Studies in a componential framework for second language task design. *IRAL - International Review of Applied Linguistics in Language Teachings*, *43*(1), 1–32. https://doi.org/10.1515/iral.2005.43.1.1

Robinson, P., & Gilabert, R. (2007). Task complexity, the cognition hypothesis and second language learning and performance. *IRAL - International Review of Applied Linguistics in Language Teaching*, *45*(3), 161–176. https://doi.org/10.1515/iral.2007.007

Rostamian, M., Fazilatfar, A. M., & Jabbari, A. A. (2018). The effect of planning time on cognitive processes, monitoring behavior, and quality of L2 writing. *Language Teaching Research*, *22*(4), 418–438. https://doi.org/10.1177/1362168817699239

Rowley-Jolivet, E. (2002). Visual discourse in scientific conference papers: A genre-based study. *English for Specific Purposes*, *21*, 19–40.

Rowley-Jolivet, E. (2012). Oralising text slides in scientific conference presentations: A multimodal corpus analysis. In A. Boulton, S. Carter-Thomas, & E. Rowley-jolivet (Eds.), *Corpus-informed research and learning in ESP: Issues and applications* (pp. 137–165). Amsterdam, The Netherlands: John Benjamins Publishing Company.

Rubin, D. L., & Kang, O. (2008). Writing to speak: What goes on across the two-way street. In D. Belcher & A. Hirvela (Eds.), *The oral/literate connection: Perspectives on L2 speaking, writing, and other media interactions* (pp. 210–225). Ann Arbor, Michigan: University of Michigan Press.

Ruiz-Funes, M. (2001). Task representation in foreign language reading-to-write. *Foreign Language Annals*, *34*(3), 226–234. https://doi.org/10.1111/j.1944-9720.2001.tb02404.x

Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, *9*(3), 259–291. https://doi.org/10.1016/S1060-3743(00)00028-X

Sasayama, S. (2016). Is a 'complex' task really complex? Validating the assumption of cognitive task complexity. *The Modern Language Journal*, *100*(1), 231–254. https://doi.org/10.1111/modl.12313

Serafini, E. J., Lake, J. B., & Long, M. H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for Specific Purposes*, *40*, 11–26. https://doi.org/10.1016/j.esp.2015.05.002

Shin, D. shin, Cimasko, T., & Yi, Y. (2020). Development of metalanguage for multimodal composing: A case study of an L2 writer's design of multimedia texts. *Journal of Second Language Writing*, *47*. https://doi.org/10.1016/j.jslw.2020.100714

Shipka, J. (2005). A multimodal task-based framework for composing. *College Composition and Communication*, *57*(2), 277–306.

Skehan, P. (2016). Tasks versus conditions: Two perspectives on task research and their implications for pedagogy. *Annual Review of Applied Linguistics*, *36*, 34–49. https://doi.org/10.1017/S0267190515000100

Smith, B. E. (2014). Beyond words: A review of research on adolescents and multimodal composition. In R. E. Ferdig & K. E. Pytash (Eds.), *Exploring multimodal composition and digital writing* (pp. 57–71). Hershey, PA: Information Science Reference. https://doi.org/10.4018/978-1-4666-4345-1

Smith, B. E., & Dalton, B. (2016). Seeing it from a different light: Adolescents' video reflections about their multimodal compositions. *Journal of Adolescent and Adult Literacy*, *59*(6), 719–729. https://doi.org/10.1002/jaal.503

Smith, B. E., Pacheco, M. B., & de Almeida, C. R. (2017). Multimodal codemeshing: Bilingual adolescents' processes composing across modes and languages. *Journal of Second Language Writing*, *36*, 6–22. https://doi.org/10.1016/j.jslw.2017.04.001

Storch, N. (2005). Collaborative writing: Product, process, and students' reflections. *Journal of Second Language Writing*, *14*(3), 153–173. https://doi.org/10.1016/j.jslw.2005.05.002

Suzuki, W. (2012). Written languaging, direct correction, and second language writing revision. *Language Learning*, *62*(4), 1110–1133. https://doi.org/10.1111/j.1467-9922.2012.00720.x

Swales, J. (1990). *Genre analysis: English in academic and research settings*. Cambridge, UK: Cambridge University Press.

Tan, L., & Guo, L. (2009). From print to critical multimedia literacy: One teacher's foray into new literacies practices. *Journal of Adolescent & Adult Literacy*, *53*(4), 315–324. https://doi.org/10.1598/JA

Tardy, C. M. (2005). Expressions of disciplinarity and individuality in a multimodal genre. *Computers and Composition*, *22*, 319–336. https://doi.org/10.1016/j.compcom.2005.05.004

Tardy, C. M. (2008). Multimodality and the teaching of advanced academic writing: A genre systems perspective on speaking-writing connections. In D. Belcher & A. Hirvela (Eds.), *The oral/literate connection: Perspectives on L2 speaking, writing, and other media interactions* (pp. 191–208). Ann Arbor, Michigan: University of Michigan Press.

The New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, *66*(1), 60–93. https://doi.org/10.17763/haer.66.1.17370n67v22j160u

Unsworth, L. (2006). Towards a metalanguage for multiliteracies education: Describing the meaning-making resources of language-image interaction. *English Teaching: Practice and Critique*, *5*(1), 55–76.

Unsworth, L. (2007). Image/text relations and intersemiosis: Towards multimodal text description for multiliteracies education. In *The 33rd International Systemic Functional Congress* (pp. 1165–1205).

Vafaee, P., Suzuki, Y., & Kachisnke, I. (2017). Validating grammaticality judgment tests: Evidence from two new psycholinguistic measures. *Studies in Second Language Acquisition*, *39*(59–95), 485. https://doi.org/10.1017/S0272263116000097

Van Avermaet, P., & Gysen, S. (2006). From needs to tasks: Language learning needs in a task-based approach. In K. Van Den Branden (Ed.), *Task-based language education: From theory to practice* (pp. 17–46). Cambridge, UK: Cambridge University Press.

Van Leeuwen, T. (2005). *Introducing Social Semiotics*. New York: Routledge.

Van Leeuwen, T. (2015). Multimodality in education: Some directions and some questions. *TESOL Quarterly*, *49*(3), 582–589. https://doi.org/10.1002/tesq.242 I

Vandommele, G., Van den Branden, K., Van Gorp, K., & De Maeyer, S. (2017). In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch. *Journal of Second Language Writing*, *36*, 23–36. https://doi.org/10.1016/j.jslw.2017.05.010

Vankooten, C., & Berkley, A. (2016). Messy Problem-Exploring through Video in First-Year Writing: Assessing What Counts. *Computers and Composition*, *40*, 151–163. https://doi.org/10.1016/j.compcom.2016.04.001

Wakita, T., Ueshima, N., & Noguchi, H. (2012). Psychological Distance Between Categories in the Likert Scale : Comparing Different Numbers of Options. *Educational and Psychological Measurement*, *72*(4), 533–546. https://doi.org/10.1177/0013164411431162

Walsh, M. (2010). Multimodal literacy: What does it mean for classroom practice? *Australian Journal of Language and Literacy*, *33*(3), 211–239.

Warschauer, M. (2017). The pitfalls and potential of multimodal composing. *Journal of Second Language Writing*, *38*, 86–87. https://doi.org/10.1016/j.jslw.2017.10.005

Wigglesworth, G., & Storch, N. (2009). Pair versus individual writing: Effects on fluency, complexity and accuracy. *Language Testing*, *26*(3), 445–466. https://doi.org/10.1177/0265532209104670

Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, *21*(4), 321–331. https://doi.org/10.1016/j.jslw.2012.09.007

Wolfersberger, M. (2013). Refining the construct of classroom-based writing-from-readings assessment: The role of task representation. *Language Assessment Quarterly*, *10*(1), 49–72. https://doi.org/10.1080/15434303.2012.750661

Yang, W., Lu, X., & Weigle, S. C. (2015). Different topics, different discourse: Relationships among writing topic, measures of syntactic complexity, and judgments of writing quality. *Journal of Second Language Writing*, *28*, 53–67. https://doi.org/10.1016/j.jslw.2015.02.002

Yi, Y. (2017). Establishing multimodal literacy research in the field of L2 writing: Let's move the field forward. *Journal of Second Language Writing*, *38*, 90–91. https://doi.org/10.1016/j.jslw.2017.10.010

Yi, Y., & Angay-Crowder, T. (2016). Multimodal Pedagogies for Teacher Education in TESOL. *TESOL Quarterly*, *50*(4), 988–998. https://doi.org/10.1002/tesq.326

Yi, Y., & Choi, J. (2015). Teachers' views of multimodal practices in K-12 classrooms: Voices from teachers in the United States. *TESOL Quarterly*, *49*(4), 838–847. https://doi.org/10.1002/tesq.219

Yi, Y., King, N., & Safriani, A. (2017). Reconceptualizing assessment for digital multimodal literacy. *TESOL Journal*, *8*(4), 878–885. https://doi.org/10.1002/tesj.354

Yi, Y., Shin, D., & Cimasko, T. (2019). Multimodal literacies in teaching and learning English in and outside of school. *The Handbook of TESOL in K-12*, 163–177. https://doi.org/10.1002/9781119421702.ch11

Yoon, H. (2017). Linguistic complexity in L2 writing revisited: Issues of topic, proficiency, and construct multidimensionality. *System*, *66*, 130–141. https://doi.org/10.1016/j.system.2017.03.007

Yoon, H. (2019). The effects of writing task manipulations on ESL students' performance: Genre and idea support as task variables. In S. Papageorgious & K. M. Bailey (Eds.), *Global perspective on language assessment* (pp. 139–151). New York: Routledge.

Yoon, H., & Polio, C. (2017). The linguistic development of students of English as a second language in two written genres. *TESOL Quarterly*, *51*(2), 275–301. https://doi.org/10.1002/tesq.296

Zhang, M. (2019). Towards a quantitative model of understanding the dynamics of collaboration in collaborative writing. *Journal of Second Language Writing*, *45*, 16–30. https://doi.org/10.1016/j.jslw.2019.04.001

Zhao, C. G. (2012). Measuring authorial voice strength in L2 argumentative writing: The development and validation of an analytic rubric. *Language Testing*, *30*(2), 201–230. https://doi.org/10.1177/0265532212456965

Zhou, A., Busch, M., & Cumming, A. (2014). Do adult ESL learners' and their teachers' goals for improving grammar in writing correspond? *Language Awareness*, *23*(3), 234–254. https://doi.org/10.1080/09658416.2012.758127