

L2 LANGUAGE DEVELOPMENT IN ORAL AND WRITTEN MODALITY

By

Myeongeun Son

A DISSERTATION

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

Second Language Studies – Doctor of Philosophy

2021

ABSTRACT

L2 LANGUAGE DEVELOPMENT IN ORAL AND WRITTEN MODALITY

By

Myeongeun Son

Analysis of second language (L2) learners' language development can allow valid and reliable assessment of the learners' L2 linguistic knowledge, which can enable the teachers and researchers to choose appropriate teaching or experimental materials (Bardovi-Harlig & Bofman, 1989; Wolfe-Quintero et al., 1997). However, most language development research has analyzed spoken production only, so it remains unclear whether learners develop L2 linguistic knowledge in the same way regardless of the modality (Polio, 2017). In addition, L2 writing research has primarily investigated changes in specific elements of L2 writing rather than investigating linguistic development *per se*. This study explores whether L2 learners' language developmental patterns are comparable across oral and written modalities on the basis of processability theory (PT; Pienemann, 1998, 2005).

87 Korean EFL learners with four different proficiency levels from beginner high to advanced (i.e., A2, B1, B2, & C1 based on Common European Framework of Reference) each completed speaking tasks and writing tasks designed to elicit particular morphosyntactic structures predicted by PT in counterbalanced order. The tasks were conducted one-on-one with each participant. The speaking and the writing tasks included open-ended questions and various activities, such as leaving a voice message and writing an email. The speaking and the writing tasks were comparable in order to highlight whether the participants' language development was consistent across the two modalities. In order to encourage the participants to respond to each question simultaneously, designated time constraints were given; the time constraints varied

depending on type of the tasks. After completing either the speaking or the writing tasks, there were intervention tasks to distract participants from the similarity of the task types. The participants' oral responses were transcribed, and the transcribed productions and the written productions were coded. Implicational tables of each modality were created based on the emergence and accuracy of their grammatical structures, and the two implicational tables were statistically compared.

The implicational tables for the speaking and writing tasks were highly correlated with each other, suggesting that the L2 learners' language development is comparable between oral and written modalities. In other words, the theoretical framework of language development can be extended to certain types of L2 writing. In addition, this study found the PT-intrastage development, which some recent studies proposed (Di Biase, Kawaguchi, & Yamaguchi, 2015) not only in earlier stages but also in later stages, and the development within a stage was also comparable between the oral and the written modality. However, accuracy differed between the modalities. The L2 learners reached higher accuracy in the production of morphosyntactic structures earlier in the written than the oral modality. In addition, accuracy in the written modality was more stable. The results provide teachers and researchers with a better understanding of L2 development in oral and written modalities and may help them assess learners' development appropriately.

Copyright by
MYEONGEUN SON
2021

ACKNOWLEDGEMENTS

Without many people's supports throughout my Ph.D. journey, I could not have completed my dissertation. I would like to express my sincere gratitude to them all.

First, I offer my gratitude and appreciation to Dr. Patti Spinner, whose guidance and supervision have helped me enjoy my Ph.D. life from my very first to last steps and become a researcher in the field of SLA. Her academic insights as well as emotional care and encouragement have prevented me from becoming discouraged and kept me concentrated on my dissertation research and writing. I am extremely grateful to have had various opportunities to develop my research ideas under her supervision, which have trained me as a young researcher and prepared me for my future career.

I would also like to thank my dissertation committee members, Drs. Susan Gass, Charlene Polio, and Aline Godfroid for their insights and constructive feedback on my dissertation. Addressing their perceptive comments has strengthened my dissertation. I am particularly thankful to Drs. Aline Godfroid and Peter De Costa for allowing me opportunities to conduct research through which I could harvest two apples. I also appreciate the constant encouragement and support of Dr. Junkyu Lee at Hankuk University of Foreign Studies, which enabled me to realize my reckless dream of completing this Ph.D. journey.

I am deeply grateful to SLS friends: Shinye Lee, Xiaowan Zhang, Melody Ma, Wendy Li, and Michael Wang. They have been great companions, both in joy and in sorrow, throughout my graduate school life. They have always cheered me up and helped me overcome huddles. I am so proud that some have already built their successful careers, and I am sure that the rest will

also do well. I want to say a special thanks to Laura Bowman and Amy Cheadle for their efforts in completing the coding of my data.

Finally, I express my sincere gratitude, respect, and love to my parents, sister, parents-in-law and brother-in-law. With their sacrifice, patience, encouragement, and love, I could start my Ph.D. and finally have the opportunity to write this acknowledgement. Last but not least, I want to deeply thank my family: Jongbong Lee and Dawon Lee. Their existence itself makes me motivated and encouraged. Thank you so much for making me happy all the time.

TABLE OF CONTENTS

LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER 1. INTRODUCTION	1
CHAPTER 2. LITERATURE REVIEW	8
2.1. Issues of oral and written modality	8
2.1.1. Different characteristics of oral and written production.....	8
2.1.2. Cognitive Processes in Speaking and Writing.....	9
2.1.3. Complexity, Accuracy, Lexical Diversity, and Fluency (CALF) in Oral and Written Production.....	11
2.1.3.1. CALF in Task-based language teaching (TBLT)	11
2.1.3.2. Development in CALF.....	13
2.1.4. Developmental Sequence	15
2.2. Processability Theory	18
2.2.1. Intrastage progress.....	24
2.2.1.1. Morphology – syntax interface	24
2.2.2. Accuracy and PT	28
2.2.3. PT in the written modality	31
2.3. Research Questions	33
CHAPTER 3. METHOD	35
3.1. Participants	35
3.2. Materials	36
3.2.1. Speaking tasks	36
3.2.2. Writing tasks.....	37
3.2.3. Questionnaires	38
3.2.4. Brief interview.....	40
3.2.5. Working memory tests	40
3.3. Procedure.....	41
3.4. Measures.....	43
3.4.1. Emergence	44
3.4.2. Accuracy.....	44
3.5. Analysis.....	46
3.5.1. Advance preparation for data analysis	46
3.5.1.1. Data organization.....	46
3.5.1.2. Transcription of spoken responses.....	47
3.5.2. Coding	48
3.5.3. Statistical procedures.....	50
3.5.3.1. Implicational scale	50

CHAPTER 4. RESULTS	53
4.1. Implicational tables	53
4.1.1. Oral responses	53
4.1.2. Written responses	57
4.1.3. Comparison between oral and written responses	61
4.2. Order of emergence within each stage	65
4.2.1. Oral responses	65
4.2.2. Written responses	68
4.2.3. Comparison between oral and written responses	71
4.3. Accuracy	75
4.3.1. Oral responses	75
4.3.2. Written responses	78
4.3.3. Comparison between oral and written responses	81
CHAPTER 5. DISCUSSION	83
5.1. Overview of research questions and results	83
5.2. Research Question 1	84
5.3. Research Question 2	88
5.4. Research Question 3	90
CHAPTER 6. CONCLUSION	92
6.1. Summary and implication	92
6.2. Limitation and future research	93
APPENDICES	96
Appendix A: Speaking and writing tasks designed to elicit S neg V	97
Appendix B: Speaking and writing tasks designed to elicit Plural -s, Plural NP, and Progressive -ing.	102
Appendix C: Speaking and writing tasks designed to elicit Possessive 's	106
Appendix D: Speaking and writing tasks designed to elicit Past -ed	110
Appendix E: Speaking and writing tasks designed to elicit questions and/or Cancel INV	115
Appendix F: Speaking and writing tasks designed to elicit 3 rd person singular -s	123
Appendix G: Background questionnaire	134
Appendix H: Exit questionnaire	135
Appendix I: Question in the brief interview	139
REFERENCES	143

LIST OF TABLES

Table 1. Summary of Previous TBLT Research Results	13
Table 2. Processing Procedure Applied to English (Pienemann, 2005, p. 24)	22
Table 3. ETS Equivalency Table (Papageorgiou, Tannenbaum, Bridgeman, & Cho, 2015)	36
Table 4. Summary of speaking and writing tasks	39
Table 5. Implicational scaling of oral responses.....	54
Table 6. Cs of R for individual structures in the oral responses	57
Table 7. Implicational scaling of written responses.....	58
Table 8. Cs of R for individual structures in written responses.....	61
Table 9. Implicational scaling of oral responses including all target structures	66
Table 10. Implicational scaling of oral responses including all target structures	69
Table 11. Accuracy in implicational scaling for the oral responses	76
Table 12. Summary of accuracy in the oral responses.....	78
Table 13. Accuracy in implicational scaling for the written responses	79
Table 14. Summary of accuracy in the written responses	81

LIST OF FIGURES

Figure 1. Accuracy and Development (Pienemman, 1998, p. 132).....	28
Figure 2. Procedure of the experiment.....	43
Figure 3. Scatterplot of the amount of emergence (based on stage).....	63
Figure 4. The number of emerged stages in the oral and written modalities.....	64
Figure 5. Scatterplot of the amount of emergence (including all target structures)	73
Figure 6. The number of emerged structures in the oral and written modalities	74
Figure 7. Example of the pictures used in Making an announcement A	97
Figure 8. Example of the pictures used in Making an announcement task B.....	98
Figure 9. Example of the pictures used in Creating a flyer task.....	100
Figure 10. Example of the pictures used in Making a guidance task	101
Figure 11. Example of the pictures used in picture description task in the oral modality.....	102
Figure 12. Example of the pictures used in picture description task in the written modality.....	104
Figure 13. Example of the pictures used in picture description task	106
Figure 14. Picture used in Giving directions via phone-call.....	107
Figure 15. Example of the pictures used in picture description task	108
Figure 16. Picture used in Giving direction via online chatting task.....	109
Figure 17. Flyer used in Response to a voice message task	115
Figure 18. Voice message used in Leaving a voice message task.....	116
Figure 19. Flyer used in Leaving a voice message task.....	118
Figure 20. Email used in Respond to an email task.....	119
Figure 21. Flyer used in Writing an email task.....	121
Figure 22. Flyer used in Response to a voice message task	123

Figure 23. Email used in Email response task	129
--	-----

CHAPTER 1. INTRODUCTION

Most language users constantly employ two modes of language production, speaking and writing. The two have different characteristics that affect how second language (L2) learners engage with them. One important difference is that writing gives L2 learners more opportunities to revise and edit their language because writing typically entails limited simultaneous interaction with an audience (Gilabert, Manchón, & Vasylets, 2016), thus allowing more attention to be turned to its production (Byrnes & Manchón, 2014; Williams, 2012). Consequently, writing also provides more opportunities to attend to form than speaking does (Polio, 2016). On the other hand, speaking requires more spontaneous production. The distinct features of each modality have often led L2 researchers to consider them separately.

However, it may not be the case that the acquisition of the two modalities are completely separated. Existing research indicates some overlap between oral and written modalities. For one, there is reason to believe that the cognitive processes involved in speaking and writing are intertwined to some extent. In order to produce oral and written linguistic output, people seem to go through several cognitive processes involving their linguistic knowledge. In this vein, some researchers have posited that language producers employ the same cognitive processes regardless of modality (Brown, McDonald, Brown, & Carr, 1988; Bourdin & Fayol, 1994, 2000; Grabowski, 2010), but see Kellogg, 1996, 2001; Levelt, 1989). Other theoretical models implicitly suggest that some of the cognitive processes involved in speaking and writing may overlap, particularly in the formulation and execution of linguistic output.

Another overlap has been found in measurement of the complexity, accuracy, lexical diversity, and fluency (CALF) of spoken and written performance. CALF measures have been

compared to investigate the extent to which L2 learners' task performance differs depending on the modality of the given task in task-based learning and teaching (TBLT) research. While the comparability of CALF in the two modalities is still debatable, research has provided evidence for some similarities between CALF of speaking and writing (Granfeldt, 2007; Kormos, 2014; Kormos & Trebits, 2009; Kuiken & Vedder, 2011; Tavakoli, 2014; Vasylets, Gilabert, & Manchon, 2017; Zalbidea, 2017, 2020). For instance, using similar tasks in both speaking and writing (e.g., picture description, narration tasks), Granfeldt (2007), Kormos (2014), and Kormos and Trebits (2009) observed no significant difference in the complexity of L2 learners' oral and written production. In sum, linguistic output may be produced through some similar processes and they may share similar constructs regardless of the type of modality.

In addition, a handful of previous studies have attempted to examine whether or not changes in CALF over time (i.e., development of CALF) can be possibly assessed in written contexts. These studies have indicated that speaking and writing performance may show some differences in development in terms of timing, speed, and amount of changes for each construct of CALF; however, they have also confirmed that CALF in L2 learners' performance improves over time in both modalities. For instance, a few studies have compared changes in CALF of spoken production to changes in CALF of written production over time (Bulté & Housen, 2009; Serrano, Tragant, & Llanes, 2012; Weissberg, 2000; Yu, 2000; also see more about accuracy in Polio, 2017).

Some of the theoretical frameworks that address developmental sequence in order of acquisition for linguistic structures are based mainly on L2 speaking but have been applied to L2 writing. These attempts allow fine-grained comparisons that indicate similarities in the order of development between speaking and writing. In particular, Boss (2008), Byrnes and Sinicrope

(2008), and Weissberg (2006) explored developmental progress in L2 writing, although their studies were not based on theoretical frameworks for language development. The results indicate that some parts of the acquisition order of verb morphology (Boss, 2008) and relative clauses (Byrnes and Sinicrope, 2008) in L2 writing are consistent with what previous L2 oral production studies have reported (e.g., Gass, 1980, Pienemann, 1987). In addition, in spite of the consistency in order of acquisition, speed of emergence seems different depending on modality (Weissberg, 2006).

Given such fundamental similarities and/or connections between speaking and writing, it may be that learners develop their L2 and are able to produce it in the same way regardless of modality, although the extent of the development may differ depending on individual differences. However, this issue has remained unclear in the field of second language acquisition (SLA). Most language development studies have analyzed only spoken production. Meanwhile, research on L2 development in writing has been less well theorized, instead of focusing on topics such as learners' changes in knowledge of genres, goal setting, and the writing process (Cumming, 2012; Sasaki, 2004; Tardy, 2012; also see Polio, 2017; Polio & Park, 2016).

Although language development in writing has been underresearched, it is important to understand L2 language development in the written modality as well as in the spoken modality, because written tests are widely used in classroom and research settings for practical reasons such as available technological devices and students' or participants' anxiety about speaking tests (e.g., Aida, 1994; Young, 1986). For instance, a written test of grammatical knowledge rather than a spoken test might be more appropriate as a placement test for writing-oriented courses, and L2 learners' understanding of explicit instruction can be tested through written diagnostic assessment (Cheng, 2019; Green, 2019).

Because concrete and objective indicators of second language (L2) learners' language development allow valid and reliable assessment of L2 learners' grammatical proficiency (Bardovi-Harlig & Bofman, 1989; Gass, 1980; Larsen -Freeman, 1978, 1983; Monroe, 1975; Pienemann, 1998; Spinner, 2011; Wolfe-Quintero, Inagaki, & Kim, 1997), teachers and researchers who may not have access to standardized proficiency tests can use such measurements to easily evaluate students or participants' overall development.. More importantly, teachers and researchers can also understand the extent to which L2 learners' oral and written development are related and whether there are some differences between their oral and written development, enabling further decisions such as the choice of appropriate teaching or experimental materials. In this regard, language development should be examined in L2 learners' writing, with the ultimate aim of contriving a "versatile" barometer of L2 learners' language development that is applicable to both oral and written production. Such a tool would be extremely useful and would enable more practical and accurate measurement of L2 learners' development.

For measuring L2 learners' language development, one theoretical framework is processability theory (PT; Pienemann, 1998, 2005). With this theory, Pienemann proposed a universal implicational order of development that language learners go through while acquiring an L2. Each stage of development contains specific morphosyntactic structures, and the theory predicts that when L2 learners reach a certain developmental stage, they acquire a processing procedure to produce the morphosyntactic structures within that stage. Although the tasks employed in most of the previous studies testing PT have been in the oral modality, PT itself leaves open the possibility of its applicability to L2 writing. In particular, the steadiness hypothesis posits that the developmental order predicted by PT does not change across different

communicative tasks as long as the tasks are designed to tap into the same knowledge in production. Further, research has found that learners' production in diverse contexts, including one-to-one interviews, picture description tasks, and elicited imitation tasks consistently reflects the developmental order predicted by PT (e.g., Ellis, 2008; Pienemann, 1998; Spinner, 2011). The steadiness hypothesis leaves open the possibility of testing PT in written communicative tasks. A few studies have done so. Håkansson and Norrby (2007) were the first to investigate the applicability of PT to written production in L2. Their study's analysis of free writing essays and translation tasks indicates that L2 learners follow the order predicted by PT in their writing development. However, in order to generalize their conclusions, more rigorous research is required. In particular, Håkansson and Norrby's study is limited because they only included advanced learners, and they did not control task difficulty between the oral and the written tasks, which in turn makes it difficult to know if the tasks were comparable.

Moreover, going beyond Pienemann's (1998, 2005) initial proposal that L2 development progresses through ordered stages, recent studies have also provided evidence for development within stages, which might provide a clearer understanding of the comparability of language development in speaking and writing. Specifically, the fact that learners are at the same stage does not necessarily mean they have identical patterns of production. Order of emergence and accuracy may lead to individuals' variability. First, processing procedures for morphology and syntax may not be interchangeable. Syntactic structures usually emerge faster than morphological structures (e.g., Bonilla, 2012, 2015; Dyson, 2009; Itani-Adams, 2009; Qi & Di Biase, 2005). Dyson (2009) posited that syntactic production is dependent on properties of UG and may provide the contexts in which morphological forms emerge. Second, order of emergence may also differ across morphological structures within a certain stage. Learners may

not be able to produce all morphological structures at a certain developmental stage simultaneously (Di Biase et al., 2015). Such different time requirements to produce each structure at a specific stage could be related to soft barriers between each structure that learners may also go through sequentially in a hierarchical order. For instance, Di Biase et al. (2015) showed that after an ESL learner reached PT's "category procedure" stage (i.e., stage 2), she started to produce *-ing* followed by possessive 's, plural s and past *-ed*. Aligned with the idea that progress has intrastage variation is the notion that the amount of time needed to establish high accuracy might also vary across morphological structures (Di Biase et al., 2015; Lee & Spinner, 2017). The forms that emerge later may have relatively lower accuracy, and it may take a longer time for learners to be able to produce them accurately.

These recent findings of learners' variability in language development within a certain stage have been based on L2 oral production; no study has yet investigated intrastage variability in the written modality. Based on the differences in accuracy between oral and written production (e.g., Kormos, 2014; Kormos & Trebits, 2009; Serrano et al., 2012; Tavakoli, 2014; Weissberg, 2000; Zalbidea, 2017) and the greater opportunities to revise and control (or slow down) written production (e.g., Gilabert et al., 2016; Ravid & Tolchinsky, 2002; Williams, 2012), it seems probable that L2 learners' variability within a certain stage would be affected by type of modality. Therefore, comparing variability within a certain stage between oral and written production might allow us to gain better understanding of the similarities and differences of language development in the two modalities.

The primary goal of this dissertation is to investigate L2 language development according to processing requirements in speaking and writing on the basis of PT, focusing on variability in the order of emergence and accuracy of morphosyntactic structures predicted by PT. The study

examines whether PT can be extended to L2 writing and whether L2 learners' developmental patterns are identical across the oral and the written modalities. In particular, this study compares L2 learners' developmental patterns in the written modality to those in the oral modality, based not only on the original developmental order predicted by PT but also on intrastage orders of emergence and accuracy.

CHAPTER 2. LITERATURE REVIEW

2.1. Issues of oral and written modality

2.1.1. Different characteristics of oral and written production

People produce language output in speech and writing. The two modalities have distinctive and contrasting features in many domains. In terms of constraints on online processing, speaking and writing typically differ along three dimensions: (1) the presence or absence of an audience during production; (2) the stability of the language signal; and (3) the degree of control of the language user over the linguistic output (Ravid & Tolchinsky, 2002, p.426). Speaking is usually, but not always, interactive and writing is usually, but not always, not interactive. Listeners provide immediate verbal or nonverbal feedback and responses and speakers can use nonverbal cues that depend on the listeners' presence. In contrast, writers can assume a potential audience for their writing, but they usually do not have immediate interaction with the audience, except in synchronous computer mediated communication (CMC), such as online chatting. The lack of immediate interaction also means that writing is usually completed at the writer's own pace, which is much slower than the pace of spontaneous speech production (Gilabert, Manchón, & Vasylets, 2016). In addition, written production has less variation of accuracy than spoken production, and writers can usually take time to revise and edit their writing before the audience reads it (Granfeldt, 2007). Writers thus generally have more control over their attentional resources and more opportunities to pay attention to their production, compared to speakers (Byrnes & Manchón, 2014; Williams, 2012). In particular, writers can attend to form more than speakers can, which in turn leads to specific structures and vocabulary appearing earlier in L2 learners' writing than in their speaking (Polio, 2016).

2.1.2. Cognitive Processes in Speaking and Writing

There have been different perspectives on the cognitive processes involved in speaking and writing production. On the one hand, some models attempt to explain cognitive processes of either speaking or writing, such as Levelt's (1989) speaking model and Kellogg's (1996, 2001) writing model.

According to Levelt (1989), speech production includes three stages: conceptualization, formulation, and articulation. Speakers first plan their message and formulate a conceptual structure for it; that is the conceptualization stage. Then, they generate a temporary linguistic output in the formulation stage. Next, in the articulation stage, they start to utter the form generated in the formulation stage. This model proposes that language processing for speech production is incremental and speech production is linear. Therefore, language processing for the following component begins while surface morphosyntactic forms are still being constructed, and immediate morphosyntactic forms must be temporarily stored until the current form is completely uttered.

Kellogg's (1996, 2001) model for written production also posits three stages: formulation, execution, and monitoring. The basic processes within each stage are planning and translating in the formulation stage, programming and executing in the execution stage, and reading and editing in the monitoring stage. In the formulation stage, writers plan what they are producing, generate acceptable sentences based on their ideas, and produce inner speech. Then, in the execution stage, writers actually type or write words. Finally, they go through and edit their writing in the monitoring stage. Compared to the stages of the speaking process, the stages of the writing process are more interactive and recursive (Kellogg, 2001; McCutchen, 1996). The later stages of the writing process may be positively or negatively affected by the earlier stages,

and writers can go back and forth among these stages until they complete their production. For instance, if writers find it difficult to construct sentences that explain their ideas, they may try to change their ideas slightly and plan their writing again.

On the other hand, some theoretical models suggest that speaking and writing might use the same cognitive processes, but draw on different subsets of these processes for each modality. For instance, Brown, McDonald, Brown, and Carr (1988) argued that language production includes formulation, execution, and monitoring stages, regardless of modality. They posited that formulation processes might be interchangeable across modalities, but would be combined with different sets of execution processes and different monitoring processes to produce either oral or written language. Bourdin and Fayol (1994, 2000) asserted that speaking and writing require identical cognitive processes but differ in the amount of cognitive resources needed for low-level processes such as transcription until such low-level processes become automatized. Bourdin and Fayol's arguments were based on the results of an experiment in which participants were required to recall some words in the two different modalities (i.e., speaking and writing). They found no difference between the modalities for the adult participants. However, the child participants recalled the words significantly better in the speaking mode than in the writing mode, suggesting that the children's low-level processes in writing had not yet been automatized (see also Grabowski, 2010).

In sum, it has been acknowledged that language is produced after language users go through several stages of cognitive processes with their linguistic knowledge. However, there have been different perspectives on cognitive processes involved in speaking versus writing, although the existing research often tacitly assumes that speaking and writing production may share at least some of the cognitive processes. If speaking and writing do indeed share some

processes, then some characteristics of speaking production may also be found in writing production, and vice versa. Addressing this possibility, the following sections will cover similarities and differences between speaking and writing production by L2 learners in terms of (a) complexity, accuracy, lexical diversity, and fluency and (b) developmental progress.

2.1.3. Complexity, Accuracy, Lexical Diversity, and Fluency (CALF) in Oral and Written Production

2.1.3.1. CALF in Task-based language teaching (TBLT)

TBLT research has investigated whether Robinson's (2001) cognition hypothesis and Skehan's (1998) limited attentional capacity model can be applied to L2 learners' performance (typically measured by complexity, accuracy, lexical complexity and fluency [CALF]). Briefly, these two competing models have different perspectives on the interactions between L2 learners' attentional resources and the cognitive demands of a task. Skehan's model considers L2 learners' attentional resources to be limited, while Robinson's suggests that L2 learners have a variety of attentional resources, which language users draw on differently depending on types of task demands (i.e., resource-dispersing feature vs. resource-directing). In addition, Skehan's model suggests that a complex task leads to increases in fluency and decreases in complexity and accuracy in L2 learners' performance. On the other hand, Robinson's hypothesis indicates that a complex task with a resource-dispersing feature results in decreases in complexity, accuracy, and fluency in L2 performance, whereas a task with a resource-directing dimension leads to increases in complexity and accuracy and decreases in fluency in L2 performance (Johnson, 2017).

The TBLT studies that have touched on the effects of task modality on performance as measured by CALF have had mixed results (Gilabert et al., 2016; Granfeldt, 2007; Kormos &

Trebits, 2009; Kuiken & Vedder, 2011; Tavakoli, 2014; Vasylets, Gilabert, & Manchon, 2017; Zalbidea, 2017, 2020; see Table 1). For instance, Kuiken and Vedder (2011) found that written tasks generated more syntactic complexity than oral tasks, whereas Zalbidea (2017) found the opposite, and other studies observed no significant difference in syntactic complexity between oral and written tasks (Granfeldt, 2007; Kormos, 2014; Kormos & Trebits, 2009; Vasylets et al., 2017).

Kuiken and Vedder's (2011) and Zalbidea's (2017) studies both had a between-subjects design, in which half of the L2 participants completed an oral task and half completed a written task. Kuiken and Vedder required the participants to advise a friend in choosing a holiday destination from five options. In the oral task, the participants left a voice message on an answering machine; in the written task, they wrote a letter. Kuiken and Vedder found that syntactic complexity, measured by clauses per T-unit and dependent clauses per clause was greater in the written task. Zalbidea's study used similar tasks. Participants were instructed to imagine planning a trip with a friend, and they were told the friend's requirements for accommodation. They then either described their choice of accommodation aloud (i.e., oral task) or wrote an email about it (i.e., written task). Zalbidea's results, however, showed that the learners produced more complex output in the oral task, measured by mean length of unit and dependent clauses per unit.

Table 1. Summary of Previous TBLT Research Results

	Complexity	Accuracy	Lexical Complexity	Fluency
Granfeldt (2007)	Oral = Written	Oral > Written	Oral < Written	
Kormos (2014)	Oral = Written	Oral < Written	Oral < Written	
Kormos & Trebits (2009)	Oral = Written	Oral < Written	Oral = Written	
Kuiken & Vedder (2011)	Oral < Written		Oral = Written	
Tavakoli (2014)	Oral < Written	Oral < Written		
Vasylets et al. (2017)	Mixed	Oral = Written	Mixed	
Zalbidea (2017)	Oral > Written	Oral < Written	Oral < Written	

2.1.3.2. Development in CALF

In the field of SLA, there have been two different strands of studies exploring L2 learners' development: developmental sequence studies and developmental index studies (Bardovi-Harlig & Bofman, 1989; Wolfe-Quintero, Ingaki, & Kim, 1998). The developmental index studies (e.g., Larsen-Freeman, 1978, 1983; Monroe, 1975) gauge L2 learners' development in CALF to investigate changes in L2 learners' production rather than acquisition order of particular morphosyntactic forms (e.g., Gass, 1980). By using CALF as a measurement of L2 learners' development, research can provide empirical evidence of how learners' L2 knowledge develops and what they encounter in the developmental process (see Norris & Ortega, 2009). A few developmental index studies have compared L2 development in oral production to that in written production (Bulté & Housen, 2009; Serrano et al., 2012; Weissberg, 2000; Yu, 2000).

For example, Serrano et al.'s longitudinal study examined L2 learners' development in English during a one-year study abroad program. Oral and written data were collected three times (pre-test, while-test, and post-test) by using oral narrative tasks and descriptive essay tasks. The data were analyzed in terms of CALF. The study showed that although both oral and written production developed during the program, the time and speed at which each construct in CALF developed differed across oral and written production. For instance, fluency and lexical diversity in oral production developed earlier, while accuracy in both modalities developed later. Likewise, Weissberg (2000) conducted a case study with five participants for one semester (16 weeks) to examine whether development of L2 syntactic knowledge differs across speaking and writing. The participants were in an intensive ESL program and they were required to complete various oral and written production tasks in the classroom. The study analyzed the emergence of a particular morphosyntactic feature in each modality, but also analyzed longitudinal change in accuracy rates by measuring the percentage of error-free T-units at the beginning and the end of the semester. While the participants' accuracy increased in both modalities, their gains were greater in writing than in speaking.

All in all, modality impacts on CALF are underresearched, and the empirical studies conducted on the topic thus far have had inconsistent results. Based on the existing research, oral and written production may or may not differ in the different aspects of CALF, whether in a single task or in longitudinal development. Additional research is needed before we can draw conclusions regarding how or if modality affects L2 learners' overall performance or certain aspects of their performance.

2.1.4. Developmental Sequence

Many researchers have investigated L2 learners' development by analyzing L2 production to examine the acquisition order of specific morphosyntactic forms (e.g., Gass, 1980; Pienemann, 1998, 2003; Pienemann & Johnston, 1987). As L2 learners gain proficiency, they become more likely to produce a variety of morphosyntactic structures. For instance, Pienemann (1998, 2003, 2005), who proposed processability theory (PT), showed an implicational development in processing procedures, which indicates gradual development in the ability to produce linguistic output (see Section 2.2).

Most developmental sequence studies of L2 learners have investigated and discussed only oral production, and there is currently limited evidence for developmental progress in L2 writing; only few studies on L2 writing take this approach (e.g., Boss, 2008; Byrnes & Sinicrope, 2008; Kyle, Crossley, Verspoor, 2020; Park, 2017; Weissberg, 2006). However, the results of some current L2 writing studies indicate that it may be possible to apply L2 speaking studies' approach to the writing modality.

Boss (2008) investigated the acquisition order of L2 German verb morphology in the written modality. She analyzed writing samples from elicitation tasks conducted with 15 L2 German learners in an elementary German course. Multiple topics (e.g., "My hobbies", and "Letter to a friend") were designed to elicit specific German morphology such as present tense with irregular verbs and perfect tense. The number of verb forms (i.e., inflectional form of the copular verb *sein*, lexical verb *haben*, modal verbs, past participle with irregular verbs, and preterite of verbs) and types of verbs with specific morphological verbs were examined. The results indicated that as L2 learners' knowledge of German developed, the number of present tense conjugations of *sein* and *haben* that emerged increased. The results were aligned with

Pienemann (1987, 1998) which predicted the acquisition order of German morphosyntactic structures based on oral production. The learners' acquisition of past participles of irregular verbs was the only exception, as these forms' emergence did not appear to fit predictions based on Pienemann (1987, 1998).

Likewise, Byrnes and Sinicropo (2008) examined 23 L2 German learners' longitudinal development of relative clauses across four curricular levels (Level 1 to Level 4). The study explored whether the incremental development suggested by the Noun Phrase Accessibility Hierarchy (NPAH; Keenan & Comrie, 1977) applies to written production. Briefly, the NPAH posits a universal learning progression for relative clauses (RCs) from unmarked to marked structures (i.e., with subjective RCs the most acceptable, followed in order of decreasing accessibility by direct object RCs, indirect object RCs, object RCs with prepositions, possessive RCs, and object RCs containing a comparison). The L2 German learners in the study completed a prototypical performance writing task at the end of each curricular level, and Byrnes and Sinicropo assessed their writings to monitor the learners' achievement of learning goals. The researchers extracted 454 RCs from the writings and categorized them by type. The results indicated that most of the learners' use of RCs followed the developmental progression proposed by the NPAH; however, 20 to 30 percent of the RCs in the data contradicted the predictions. In particular, the L2 learners sometimes skipped a certain developmental stage, and their production demonstrated the emergence of only limited numbers of relatively unmarked structures. In addition, although the more advanced learners used more RCs, more were the simplest type (i.e., subject RCs), rather than the more marked relative types.

Both Boss (2008) and Byrnes and Sinicropo (2008) attempted to explore L2 learners' developmental progress in written modalities. These studies found that the learning process in

the written modality may contrast with the learning process in the spoken modality to some extent. However, caution is needed in comparing the results from these studies to the results of studies on learners' development in speaking. As Norris and Ortega (2003) pointed out, researchers attempting to assess the emergence of complex phenomena must be careful not to base their interpretations

on a lack of evidence, as opposed to evidence for the lack of emergence.... [I]t is likely that measurement data are more frequently underinterpreted when researchers do not adequately conceptualize the complexities of measurement behaviors that they intend to elicit (p. 733-734).

In this regard, a lack of evidence may be due not to the fact that L2 learners follow different developmental paths in oral and written modalities but, rather, to experimental task.

In particular, the writing tasks that Boss (2008) used may have led to the results (i.e., on the past participles of weak verbs) that were inconsistent results with predictions based on Pienemann (1987). Although Boss chose specific topics for the writing tasks (e.g., a travel diary) to elicit particular structures (e.g., perfect tense, preterite), only one or two tasks were designed to elicit each particular structure. For each topic, the learners completed a free writing of around one hundred words. The number of tasks and the length of writing for each task may not have been enough to elicit sufficient past participles of weak verbs. In addition, the instructions for the writing tasks left open the possibility that the learners could avoid using certain structures such as those containing irregular verbs. Likewise, Byrnes and Sinicrope (2008) analyzed prototypical performance writings that were not designed to elicit the target structures (i.e., relative clauses) but to evaluate learners' overall achievement at the end of course. Written tasks designed

specifically for the elicitation of target structures might show a developmental progression for writing production that is identical to that for spoken production.

In sum, very few studies have investigated the L2 developmental sequence in the written modality, and those few have not been based on the theoretical frameworks widely used to explore L2 development in the spoken modality. Nevertheless, similar patterns have been found to some extent. External factors such as task type and instructions may have led to findings that show differences between development in oral production and development in written production. Therefore, in order to shed light on L2 learners' development in both oral and written modalities, it may be necessary to investigate whether particular developmental theories can be applied to both modalities with the use of similar tasks (see Polio, 2012, 2017).

2.2. Processability Theory

Pienemann (1998) proposed processability theory (PT) to explain the developmental stages in language learners' interlanguage. PT hypothesizes that, at each developmental stage, language learners acquire a processing procedure that manages information transfer, which is necessary for the production of morphosyntactic structures. Therefore, language learners can only produce the morphosyntactic structures that their processor can deal with at any given stage. The processing procedures described in PT are based on lexical functional grammar's (Kaplan & Bresnan, 1982) feature unification process which posits three processing procedures, in which language users (a) classify the grammatical information in a lexical item, (b) store the information temporarily, and (c) utilize the information at different points in the constituent structure (Pienemann, 1998). In addition, Pienemann (2005) claimed that language learners go through these developmental stages in a universal implicational order. In other words, a learner

cannot skip lower stages and follow a different developmental order. Acquisition of lower level procedures is a prerequisite to function on the next level. This view of development in processing procedures is based on Levelt's (1989) speaking model, which suggests that language processing is incremental and linear: While surface forms are still being constructed, grammatical information is temporarily stored in grammatical memory storage, which ultimately allows automatic processing of the information. According to PT, as language learners reach higher developmental stages, the amount of grammatical information they can process for production increases.

Although PT research has primarily analyzed L2 learners' spoken production to explore their language development, the processing procedure predicted by PT is not limited to a specific modality or specific task type. Pienemann's steadiness hypothesis states that the developmental order predicted by PT does not change across different communicative tasks as long as the tasks are designed to draw on the same production skill. Some research has tested this hypothesis, and reported that learners' production consistently reflects the PT developmental order in one-to-one interviews, picture description tasks, and elicit imitation tasks (e.g., Ellis, 2008; Pienemann, 1998; Spinner, 2011). For instance, Pienemann (1998) compared the results from six different communicative tasks to test the steadiness hypothesis. All the tasks reflected the developmental order of PT in the L2 learners' production, supporting the steadiness hypothesis, although morphology production showed a few mismatches (i.e., in about one percent of the data).

In particular, based on the processing procedures that become available to learners at each stage, PT proposes six specific developmental stages:

1. Lemma access
2. Category procedure

3. Phrasal procedure
4. Verb phrase procedure
5. S-procedure
6. Subordinate clause procedure

In the lemma access stage, learners are limited to producing single words or formulaic expressions. Lacking processing procedures, they are not ready to exchange information, so feature matching or unification does not occur yet. At this stage, nonlinguistic strategies, such as facial expressions play a crucial role in keeping communication going. In the second stage, when category procedures become available, learners are able to assign categorical information to lexical items. At this stage, operations occur only within a single phrasal category. For example, learners are able to put the past tense marker *-ed* on a verb (e.g., *loved*) and the plural marker *-s* next to a noun (e.g., *dogs*). When it comes to syntax, feature matching is not yet available across phrases, but learners start to produce syntactic structures with canonical word order, such as *Mary eats apple*.

Next in the phrasal procedure stage, a single lexical item that involves categorical information can take the role of head of a phrasal category. Within a phrasal category, feature information can be exchanged across elements. For instance, in *these dogs*, the plural feature is consistent within the noun phrase (NP). In terms of syntax, some fronting elements, such as adverbs and prepositional phrases emerge (e.g., *Yesterday, Mary woke up at 7:00 am*). According to the topic hypothesis (Pienemann, Di Biase, Kawaguchi, & Håkasson, 2005), language learners at this stage start to differentiate syntactic topics and functions from the subject. For example, they do not stick to putting a subject in the first position of a sentence as in

a canonical order; adverbs or prepositional phrases that indicate syntactic functions can now occupy that position.

The fourth stage is the verb phrase procedure stage, in which features are matched across phrases. Language learners at this stage can keep the same tense throughout a clause. For example, the past tense is consistently used between two VPs in (1a). In terms of syntax, language learners begin to produce yes-no questions by inverting a subject and a copular verb or other auxiliary verbs, as shown in (1b).

(1) Examples of morphosyntactic structures in the fourth stage

- a. *Mary woke up at 7:00 am and came to the office at 9:00 am.*
- b. *Did you find a book in the library?*

In the fifth stage – the S-procedure stage – grammatical features are matched across phrases. Subject and verb agreement occurs at this stage, as shown in (2a). In (2a), the number and person feature is matched between the subject of the NP, *Mary*, and the verb of the VP, *loves*. Language learners also start to produce *wh*-questions in which *do* or other auxiliary verbs are in second position, as in (2b).

(2) Examples of morphosyntactic structures in the fifth stage

- a. *Mary with brown hair loves these dogs.*
- b. *Where did you find this book?*

In the final stage, the processing procedures for subordinate clauses are developed, such as *wh*-complement clauses in English, as shown in (3a - b).

(3) Examples of *wh*-complement clauses in English

- a. *Mary shows what she made yesterday.*
- b. **Mary shows what did she make yesterday.*

A grammatical *wh*-complement in English does not include auxiliary inversion, which is necessary to make a question. When language learners reach this developmental stage, they can differentiate the syntactic orders of a *wh*-complement and a *wh*-question and produce the *wh*-complement with the correct order.

In short, as processing procedures develop, the morphosyntactic structures that a learner can produce are more complex and various. The processing procedures and morphosyntactic structures that become available to learners at each developmental stage are summarized in Table 2.

Table 2. Processing Procedure Applied to English (Pienemann, 2005, p. 24)

Stage	Processing procedure	L2 process	Morphology	Syntax
6	subordinate clause procedure	main and subordinate clause		cancel INV
5	S-procedure	interphrasal information	SV agreement	<i>do</i> 2nd, AUX2nd
4	VP-procedure	interphrasal information	tense agreement	Y/N inversion, copula inversion
3	phrasal procedure	phrasal information	NP agreement, Neg+V	ADV, <i>do</i> -fronting, topicalization
2	category procedure	lexical morphology	plural, possessive pronoun	canonical order
1	word/lemma	“words”	invariant forms	single constituent

PT has been confirmed by many studies testing several languages, including Arabic (Mansouri, 2005), Chinese (Zhang, 2005), English (Pienemann, 1998; Sakai, 2008; Spinner, 2011), German (Jansen, 2008), Italian (Di Biase & Kawaguchi, 2002), Japanese (Di Biase & Kawaguchi, 2002; Kawaguchi, 2000, 2005), Spanish (Bonilla, 2015), and Swedish (Pienemann & Håkasson, 1999). Most of these studies gathered data through interviews and a variety of communicative tasks, such as picture description and story-telling tasks, in order to elicit specific morphosyntactic structures in a natural way. Although the studies had participants with diverse L1 backgrounds, different age ranges, and different contexts of language acquisition (i.e., foreign language vs. second language; classroom setting vs. naturalistic setting), their results lend support to the implicational stages that PT proposes. In other words, the stages appear to be applicable to typologically different languages.

While PT is concerned with progress across the developmental stages (i.e., interstage progress), recent studies have focused on progress within the developmental stages (i.e., intrastage progress) which provides more fine-grained information on L2 development. Some researchers have reported that structures included in a single stage develop in different orders rather than emerging at the same time; in particular, each stage's syntax is usually produced earlier than its morphology (e.g., Bonila, 2015; Dyson, 2009), and there is also a developmental order of morphological structures in a stage (e.g., Di Biase, Kawaguchi, & Yamaguchi, 2015). Some of the overlaps between speaking and writing may allow for explorations of L2 development in writing on the basis of PT; however, the distinctions between the two modalities also leave open the possibility of differences in their L2 development. In this regard, the qualitative differences within a stage may provide more fine-grained information on L2

development in oral and written modalities. The following sections will provide more detailed information about intrastage progress.

2.2.1. Intrastage progress

2.2.1.1. Morphology – syntax interface

In the original hierarchical order that Pienemann (1998) proposed, each developmental stage is defined by a set of morphological and syntactic structures. Pienemann assumed that identical processing procedures can apply to morphology and syntax. That is, PT posits that when learners reach a particular developmental stage, they develop the designated processing procedure for that stage and become able to apply that procedure to both morphology and syntax. However, more recent studies have investigated the order of development between syntax and morphology. Several studies have demonstrated that progress in syntax is faster than progress in morphology, indicating that the implicational hierarchy of syntax is likely to be separate from that of morphology (e.g., Bonilla, 2012, 2015; Dyson, 2009; Itani-Adams, 2009; Qi, 2011; Qi & Di Biase, 2005).

For instance, Dyson (2009) analyzed two L1 Chinese beginner ESL students' oral production in a set of communicative tasks and interviews, which the participants completed five times. While the results supported PT's developmental stages overall, the study also found that the syntactic structures for each stage emerged earlier than the morphological structures. To give a specific example, one of the participants showed the predicted developmental sequence from the intra-phrasal procedures of Stage 2 to the inter-phrasal procedures of Stage 3 in her production of both morphological and syntactic structures. However, her production also showed the emergence of the morphological structures for each of these stages at a different time than the

emergence of the syntactic structures. For example, in Stage 3, the syntactic structure of negation + verb emerged earlier than the morphological structure of NP + plural -s. In addition, she produced the syntactic structures more frequently and with more diverse lexical items than the morphological structures.

In order to account for this ordered emergence of syntax before morphology, Dyson (2009) proposed the universal properties hypothesis which posits that functional categories in UG as well as the processing procedures predicted by PT develop incrementally. In this view, syntax develops independently with the help of properties of UG; the syntactic structures then provide context for morphology and can be used in the processing procedures for morphology production.

PT states that the emergence of a particular morphosyntactic structure at a certain developmental stage indicates a learner's development of the processing procedure for that stage. For instance, if a learner produces plural -s in an obligatory context, it confirms that he/she has reached the second stage and developed category procedures. However, based on a learner's production, it may still be unclear whether she or he is ready to use category procedures to produce any other second-stage structures or not.

Di Biase et al.(2015) showed that the emergence of morphosyntactic structures at the same stage may not occur simultaneously. Rather, each morphosyntactic structure at a certain developmental stage has a different order of emergence. They analyzed the English oral production of an L1-Japanese child, Kumi, over the first two years after Kumi's arrival in Australia, when she began to learn English as a second language. The data were recorded every two weeks for the first two months, every two months for the rest of the first year, and then every three months for the second year. The results indicated that, of the morphology expected at the

category procedure stage, *ing* emerged first, followed by possessive 's, plural s, and past *-ed* in that order. By the time of the second recording, Kumi produced *-ing* in a variety of contexts, but did not yet distinguish *-ing* forms from verbs with tense and aspect (e.g., *girl looking in the hole*). Based on Kumi's production of *-ing*, Di Biase et al. assumed that she produced *-ing* to fill the grammatical category of a verb, and her production of *-ing* indicated her development of the category procedure. Likewise, possessive 's also emerged quite early, perhaps due to L1 influence, as Japanese has a similar possessive structure. The results are aligned with the developmentally moderated transfer hypothesis (Pienemann, 1998) which posits that relevant L1 features are transferred when L2 learners are ready to acquire a particular morphosyntactic structure in L2. However, compared to *-ing* and possessive 's, the emergence of plural *-s* and past *-ed* was delayed (recording 4 and recording 9, respectively), suggesting that there were "soft barriers" between each structure, which resulted in an intrastage sequence within the category procedure stage. The emergence of past *-ed* occurred even later than the emergence of morphological structures at later stages. Kumi's developmental pattern followed PT if only the morphological structures that emerged earlier are considered, but scalability was only 55% if all the morphological structures are included. Hence, Di Biase et al.'s results indicate that the development of the processing procedure for a developmental stage may not guarantee that all structures at the stage emerge at the same time; rather, there may be an order of emergence within the stage.

In addition to the order of emergence of morphological structures, there is one more difference across the structures at the same stage. Although Di Biase et al. (2015) did not explore this finding in detail, their results show that learners take different amounts of time to establish accurate production of different morphological structures at the same stage. Kumi started to use -

ing accurately in obligatory contexts right after it emerged, whereas she showed inaccurate use of plural -s throughout the data collection period. Likewise, both VP agreement (i.e., agreement between an auxiliary and a verb) and NP agreement (i.e, number agreement within an NP) are Stage 3 phrasal procedures. In Kumi's production, VP agreement (recording 4) emerged slightly earlier than NP agreement (recording 7), and Kumi consistently produced accurate VP agreement after its emergence, while she made some mistakes with NP agreement until recording 14.

The results of this line of research, taken together, suggest that in addition to the interstage developmental progress that PT proposed, there may be intrastage developmental progress in accuracy as well as in emergence. In other words, even when grammatical structures are at the same developmental stage, each structure may emerge at different times and take different amounts of time to become established. With regard to accuracy, Di Biase et al. (2015) indirectly suggested the need to consider accuracy along with emergence to gain better understanding of a learner's development. Considering L2 learners' variation in production as related to accuracy within a stage, this study casts light on the possibility of measuring accuracy together with emergence, unlike previous PT studies, which have only focused on the emergence of grammatical structures. Moreover, the different natures of the two modalities (Section 2.1.1.), such as greater opportunities to revise and attend to form in writing, may lead to differences in accuracy between speaking and writing, which also prompts further investigation of accuracy in PT. The following section will cover PT's perspective on variations in learners' production and describe an empirical PT study that provides evidence for the measurement of accuracy.

2.2.2. Accuracy and PT

PT accounts for language learners' development of processing procedures from lemma access to the subordinate clause procedure. The basic assumption of PT is that language learners proceed step by step and are not able to skip over stages. For instance, if a language learner is at the second stage, he/she is not ready to produce a sentence in which subject and verb agreement takes place. The criterion of development in PT is the emergence of structures, so accuracy is not a primary focus. According to Pienemann (1998, 2005), different morphosyntactic structures have different developmental rates and gradients of accuracy, as shown in Figure 4. The figure illustrates the notion that different grammatical structures, labeled *a*, *b*, and *c*, have different developmental procedures to reach accuracy and different rates to reach a certain point of accuracy. For instance, it takes longer to reach 50% accuracy for structure *a* than structure *b*, but 100% accuracy is reached earlier for structure *a*. Therefore, Pienemann (1998) argued that accuracy, unlike emergence, is not a valid measurement of language learners' development.

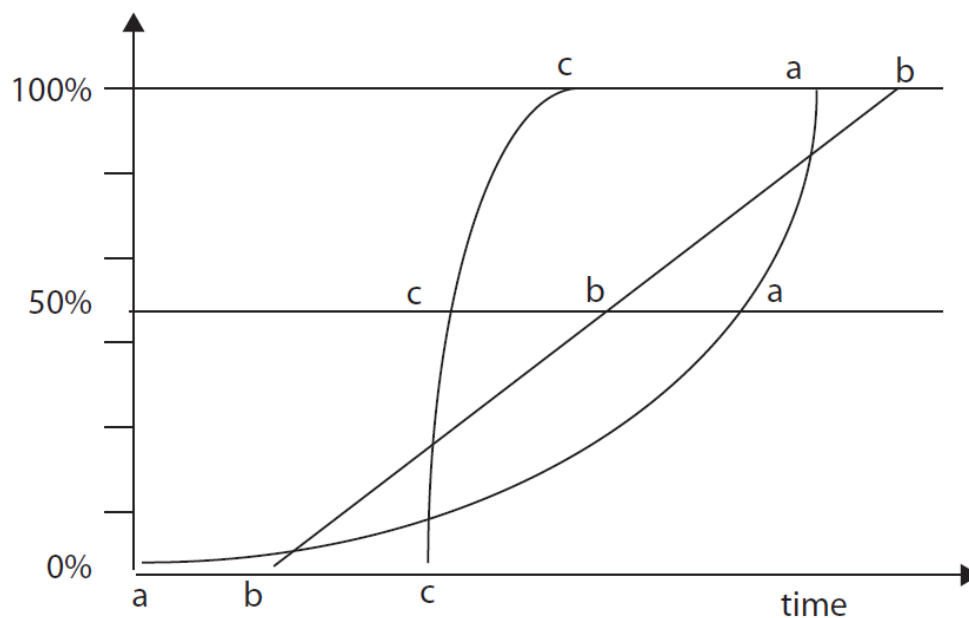


Figure 1. Accuracy and Development (Pienemann, 1998, p. 132)

However, the PT perspective on language learners' variability in production and automatization suggests the possibility of measuring their accuracy in the use of morphosyntactic structures together with emergence. Whenever language learners attempt to produce morphosyntactic structures that are representative of higher stages, they create variations due to their limited processing procedures (Pienemann, 1998). Language learners' variability is constrained by their current developmental stage (Pienemann, 2005). Liebner and Pienemann (2011) pointed out that language learners tend to leave out some required components, include inappropriate components, and simplify structures. For example, when learners at the fourth or lower stages try to produce a question, such as (4a) in conversation, they may try any of the four alternative ways shown in (4b – e), using their available processing capacity.

(4) Variations in learners' production of questions

- a. *When did you arrive in East Lansing yesterday?*
- b. *When Ø (did) you arrive in East Lansing yesterday?*
- c. *When you arrived in East Lansing yesterday?*
- d. *When arrived in East Lansing yesterday?*
- e. *You arrived in East Lansing yesterday?*

A question including an auxiliary in second position is one of the representative morphosyntactic structures of the fifth stage in PT. In order to form this structure, not only must the *wh*-element be fronted but the auxiliary and subject must be inverted. However, learners at the fourth or lower stages have not developed the processing procedures for these operations. Some of them, at the fourth stage, are only able to front the *wh*-element without the inversion of the auxiliary

and subject, whereas others at the lower stages are only able to use the canonical order. Due to their limited capacity, learners who avoid producing (4a) may produce any of the utterances in (4b) to (4e). In other words, based on the learner's developmental stage, his/her interlanguage has structurally limited variations.

In addition, PT indicates that a morphosyntactic structure is acquired, and then the underlying processing routine becomes automatized (Pienemann, 1998). Acquisition of a certain structure includes acquisition of the routine that makes the processing procedure of the structure available, which leads to increased speed and accuracy in production over time. Accurate production of a specific morphosyntactic structure requires such automatization of the underlying routine for it. Learners may have variability in their production of morphosyntactic forms until the routine is automatized for them. Even among morphosyntactic structures situated at the same developmental stage, the accuracy for each structure may differ based on the extent to which the underlying processing is automatized.

Lee and Spinner (2017) conducted elicited oral production interviews with L2 learners. They measured accuracy as well as the emergence of target structures at different stages. The target structures were past *-ed*, possessive *'s*, plural *-s*, and progressive *-ing* in the second stage; objective pronouns and possessive pronouns in the third stage; and the third person singular *-s* in the fifth stage. Accuracy was calculated by means of obligatory occasion analysis: All obligatory occasions for the target structures were identified, and then each occasion was examined to identify whether a correct structure had been supplied (Ellis & Barkhuizen, 2005; Ellis, Sheen, Murakami, & Takashima, 2008). The study showed that accuracy differed across participants and morphological structures. For instance, participants who had reached the fifth stage had different accuracy rates (from 30% to 91%). Participants at the third stage had substantially

lower accuracy (below the criteria of accuracy) for the third person singular -s than those at the fifth stage. In addition, within the second stage, possessive 's had the highest accuracy (77%), whereas past -ed had the lowest accuracy (51%) among the four structures. These results may be due to learners' variability in their production of morphological forms at these developmental stages, and to the different extents to which the underlying routine of the forms had become automatized for different learners. Although the study targeted only several morphological forms, it cast light on the possibility of using the PT framework to investigate accuracy and emergence simultaneously to gain understanding of language learners' development. Further PT research should continue to explore accuracy in production, not only of morphological forms but also of syntactic structures.

2.2.3. PT in the written modality

PT has been primarily supported by L2 oral production studies. Pienemann (1998) proposed the steadiness hypothesis, which posits that L2 learners' ability to process morphosyntactic structures is consistent across different tasks, if these tasks require the learners to produce their L2 using the same skills. The steadiness hypothesis was based on L2 oral production. However, there has been limited evidence for the application of processability theory to written production. Håkasson and Norrby (2007) attempted to investigate this topic with 20 L1 English students of L2 Swedish. The participants completed free writing essays and translation tasks in which they were required to translate an English passage into Swedish, at the beginning of the academic year (Time 1) and again eight months later (Time 2). Oral data were also collected through communicative tasks at Time 2. These tasks were designed to elicit particular morphosyntactic structures. For instance, a participant was given a picture and an interlocutor

asked a particular question (e.g., *Vad finns bredvid päronet?* ‘What is there beside the pear?’) to encourage the participant to describe the picture by using a particular structure.

The results indicated that participants’ development in written production was parallel to their development in their oral production. In other words, participants went through identical processing procedures stages in speaking and writing. It is important to note that there was no time constraint in the writing tasks. Thus, unlike spontaneous speech production, the writing tasks allowed the participants to take additional time to use their declarative knowledge and to plan and monitor their production in writing. In spite of this difference, participants did not produce higher level morphosyntactic structures in writing. However, the participants’ spoken and written production were different in terms of language complexity. For instance, although the participants met the criterion of emergence of subordinate clauses in both the spoken and the written tasks, they tended to avoid subordinate clauses in the spoken tasks. Based on the results, Håkasson and Norrby (2007) concluded that the L2 learners’ processability of structures may be stable, as predicted by the PT, but language complexity may differ across the modalities.

Håkasson and Norrby’s (2007) study was the first to attempt to investigate whether PT can be applied to L2 learners’ written production. However, some elements of this study suggest that the results should be interpreted cautiously, and that further research is needed. First, because the participants all were at fairly high levels of development, it is difficult to conclude that processability is identical across the two modalities based on their production. Future research should include L2 learners with a wider range of proficiency. In addition, one of the written tasks was a translation task. This task may not be as burdensome as other communicative tasks, as translating does not require learners to develop ideas and generate morphosyntactic structures based on their ideas. Rather, the learners needed only to translate “pre-made” English

sentences into Swedish. Given differences in task difficulty, written production in a translation task may not reflect processability well. Employing similar tasks for speaking and writing would provide a better understanding of the application of PT to the written modality.

2.3. Research Questions

The purpose of this dissertation are to assess whether PT predicts the order of emergence for written production as well as oral production and to discover the extent of the similarities and differences in intrastage L2 oral and written production specifically in regard to (a) developmental order between morphology and syntax and (b) developmental order across morphosyntactic structures in a stage; and, third, to assess the accuracy in L2 oral and written production. In order to ensure comparability between the two modalities, this study employs an experiment with similar oral and written tasks including picture-description tasks and narrative tasks.

The research question guiding this study are as follows

1. Do morphosyntactic structures in L2 learners' oral and written production follow the developmental order predicted by PT?
2. Is there an emergence order of morphological structures within each developmental stage in L2 learners' oral and written production? To what extent does task modality affect the emergence order?

3. Does accuracy in morphosyntactic production show parallel patterns with emergence in L2 learners' oral and written production? If so, to what extent does task modality affect the development of accuracy?

CHAPTER 3. METHOD

3.1. Participants

Initially, 93 Korean EFL learners participated in this study. Six learners decided to withdraw their participation in the middle of the experiment; therefore, 87 learners (40 males; 47 female) remained for the final analysis. The participants were recruited from large universities in South Korea. All of them were undergraduate students (age: $M = 22.97$, $SD = 2.53$) who had learned English in instructional settings but rarely used the language outside of the classroom. English was their L2 ($N = 84$) or L3 ($N = 3$). Most of them had not lived for substantial periods of time in English-speaking countries, but 30 had participated in a study-abroad, exchange student, internship, or working holiday program (years: $M = .31$, $SD = .76$).

For the purpose of cross-sectional analysis, the selected participants were categorized as one of four proficiency levels, based on their standardized test scores (e.g., TOEFL, IELTS, TOEIC; see Table #): high beginner (A2 in the Common European Framework of Reference for Languages [CEFR]), low intermediate (B1 in the CEFR), high intermediate (B2 in the CEFR), and advanced (C1 in the CEFR). Only learners who could show their unexpired standardized test scores were selected. The number of participants in each group was almost perfectly balanced; the A2 group had 21 participants, while the other three groups had 22 participants.

Table 3. ETS Equivalency Table (Papageorgiou, Tannenbaum, Bridgeman, & Cho, 2015)

CEFR	TOEFL	IELTS	TOEIC
A2 (n=21)	Below 42	Below 4	Below 550
B1 (n=22)	42 – 75	4 – 5	550 – 785
B2 (n=22)	76 – 94	5.5 – 6.5	786 – 944
C1 (n=22)	Over 95	Over 7	Over 945

3.2. Materials

3.2.1. Speaking tasks

I conducted the speaking tasks one-on-one with each learner. The speaking tasks consisted of a series of short interview-style tasks in which I asked the participants open-ended questions regarding their opinions on familiar topics and about their habitual actions, and a series of speaking tasks with a variety of specific requirements. Each of these 15 tasks took the form of a question designed to elicit a particular morphosyntactic structure that characterizes different stages in PT. Each structure was targeted by two or three questions to give sufficient opportunities to produce the structures.

For each task, I provided an oral prompt by asking the questions; I also provided relevant visual materials such as pictures and flyers, and I told the participant the time limit for that task. The pictures used in the tasks were copyrighted images from Pixabay (<https://pixabay.com>), Open Clip Art (<https://openclipart.org/>), Libreshot (<https://libreshot.com/>), Flickr (<https://www.flickr.com/>), and Pexels (<https://www.pexels.com/>).

Depending on its difficulty, each task had a different time limit, which was selected with the intention of encouraging spontaneous responses. If the allocated time was over in the middle of a participant's sentence, I waited for him/her to finish producing the sentence; the part

produced after the time limit was excluded from the data analysis. I decided the time limits based on time limits for similar tasks in standardized tests (e.g., TOEIC speaking, TOEFL speaking) and the results of a pilot study with five intermediate learners (two at level B1; three at level B2). Table 4 summarizes the task types and their time limits.

As the speaking tasks involved mutual interaction between each participant and me, I sometimes intervened. For instance, on occasion, a participant at the A2 level did not understand the purpose of a task or some vocabulary in the prompt, so I repeated or elaborated it. Or, if the participants' responses were too short, I encouraged them to complete the tasks with the targeted morphosyntactic structures by asking further questions.

3.2.2. Writing tasks

The writing tasks were also conducted one-on-one. I provided the prompt for each task orally, along with any additional guidance needed, and the participant completed the tasks on a computer. The writing tasks were similar to the speaking tasks, with slight adjustments to more closely approximate real-life writing situations. For instance, one oral task asked the participants to leave a voice message to ask some questions to a representative of a hotel in oral task. The corresponding writing task asked the participants to write an email for the same purpose. Because typing tasks longer than speaking, the time limit for each writing task was longer than that of the corresponding oral tasks. Table 4 summarizes the task types and their time limits.

As in the speaking tasks, I sometimes intervened: I provided additional explanation of a prompt if a participant had difficulty understanding it, and I sometimes asked a participant additional questions to encourage them to write more.

The prompts for the writing tasks were oral to minimize the chances of the participants borrowing or copying expressions from the prompts in their writing. However, they were allowed to write down a prompt themselves, as they heard it, if they wished to do so. Along with the prompts, I provided accompanying visual materials and told the participant the time limit for each task.

3.2.3. Questionnaires

Two questionnaires were administered on Qualtrics (<http://www.qualtrics.com>): a background questionnaire and an exit questionnaire. The background questionnaire asked for participants' general demographic information, including their previous English learning experience (see Appendix G). The exit questionnaire was used to gauge participants' perceptions of the experiment (see Appendix H). More specifically, a series of Likert-scale items asked participants (a) to reflect on their performance on the experiment and (b) to report their perceptions of the extent to which their involvement in the first part of the experiment (either the speaking or the writing components) affected their performance on the second part of the experiment (either the writing or the speaking components). In addition, the exit questionnaire included two open-ended questions intended to further elicit participants' perceptions of each mode of task.

Table 4. Summary of speaking and writing tasks

Targeted structures	Speaking task		Writing task	
	Task type	Time limit	Task type	Time limit
S neg V	Making an announcement	Spontaneous	Creating a flyer	Spontaneous
	Making an announcement	Spontaneous	Making a guideline	Spontaneous
Plural -s, Plural NP & Progressive -ing	Picture description task A	1-minute	Picture description task A	2-minute
	Picture description task B	1-minute	Picture description task B	2-minute
Possessive 's	Picture description task	Spontaneous	Picture description task	Spontaneous
	Giving direction via phone-call	Spontaneous	Giving directions via chatting	Spontaneous
Past -ed	Narrative task A	1.5-minute	Narrative task A	3-minute
	Narrative task B	1.5-minute	Narrative task B	3-minute
	Narrative task C	1.5-minute	Narrative task C	3-minute
Questions and/or Cancel INV	Response to a voice message	1-minute	Response to a email	2-minute
	Leaving a voice message A	1-minute	Writing an email A	2-minute
	Leaving a voice message B	1-minute	Writing an email B	2-minute
3 rd person singular -s	Response to a voice message	1.5-minute	Response to an email	3-minute
	Argumentative task A	1.5-minute	Argumentative task A	3-minute
	Argumentative task B	1.5-minute	Argumentative task B	3-minute

3.2.4. Brief interview

After the exit questionnaire, a brief interview was conducted to ask further questions regarding the participant's perceptions of the repetition of similar task types in the oral and written modes (see Appendix I). The interview questions asked the participants to talk about how, and to what extent, doing the first series of tasks (i.e., spoken or written) influenced their performance on the second series of task (i.e., written or spoken). The interview was conducted in Korean.

3.2.5. Working memory tests

After a participant had completed the first task series (either spoken or written), two automated complex working memory span tests – an operation span and a reading span test (Oswald, McAbee, Redick, & Hambrick, 2015) were administered. The tests were created on E-prime and the participants completed the tests individually by clicking a mouse after they were given oral guidance and a written prompt. In addition to assessing working memory, these tests functioned as intervention tasks with the intention of distracting the participants from the similarity of the task types in the spoken and written components of the experiment. In this dissertation, the role of the working memory tests as intervention tasks was primarily focused on; thus, the results of the working memory tests were not analyzed. The participants' working memory will be considered in a future analysis.

Using two working memory tasks had two benefits: (a) increasing the length of the intervention between the two main tasks and (b) providing a better assessment of working memory. According to Conway et al., a single task may be less effective or accurate in assessing

working memory capacity, because different task types reflect individuals' different abilities. For instance, the operation span is linked to mathematical abilities, whereas the reading span is connected to verbal abilities. Therefore, Conway et al. suggested using multiple working memory span tests and averaging their scores.

The Georgia Institute of Technology (<http://englelab.gatech.edu>) gave its permission to translate shortened versions of the operation and reading span tests into the participants' L1, Korean. The reason for using the translated tests was to avoid giving participants extra opportunities to process English, which could improve their performance in the second task and thus interfere with any effects of proficiency or task mode. It was possible that participants were not exposed to opportunities to process English, which may facilitate them to memorize types of tasks that they completed; the differences resulted from participants' L2 proficiency were excluded (Gass & Lee, 2011).

The complex span tasks involve a storage and a processing component (Draheim, Harrison, Embretson, & Engle, 2018). The operation span test required participants to solve arithmetic problems (the processing component) and to recall a letter (the storage component). In the reading span test, participants were asked to read a sentence of 10 – 15 words (the processing component) and to recall a letter (the storage component).

3.3. Procedure

The experiment was conducted in a quiet laboratory room. Individual participants completed the experiment in two sessions with my guidance. I informed the participants of the time limit for each question and provided an automatic stopwatch so the participants could check the remaining time while responding to each question. The order of the speaking task and the

writing task was counterbalanced; thus, half of the individuals completed the speaking tasks on the first day and the other half completed the writing tasks on the first day. The orders of the questions in the tasks were pseudo-randomized; in particular, questions designed to elicit the same target structure were not presented one after the other, but with at least three questions between them. The first day started with a warm-up conversation and the background questionnaire. Then, the participants completed either the speaking tasks or the writing tasks with me, and then took the two working memory tests by themselves. The order of the working memory tests was counterbalanced: Half of the participants started with the operation span test while the other half completed the reading span test first. On the second day, they completed the tasks in the other mode, the exit questionnaire, and the brief interview. They had a short break between each task, and additional short breaks were given if requested. Participants' answers in the speaking tasks were recorded with a Sony ICD-UX560F recorder. The recorder was started before the first question and continued to operate until the participant finished answering the last question in order not to disturb the process or the responses. Microsoft Word was used for the writing tasks, and Inputlog 7.0 (Leiten & Van Waes, 2013), a keystroke logging program recorded the participants' writing. The spelling and grammar check function of Microsoft Word was blocked before the experiment. The entire experiment took approximately two hours. After the completion of all of the components of the experiment, the participants received \$30 as compensation for their participation.

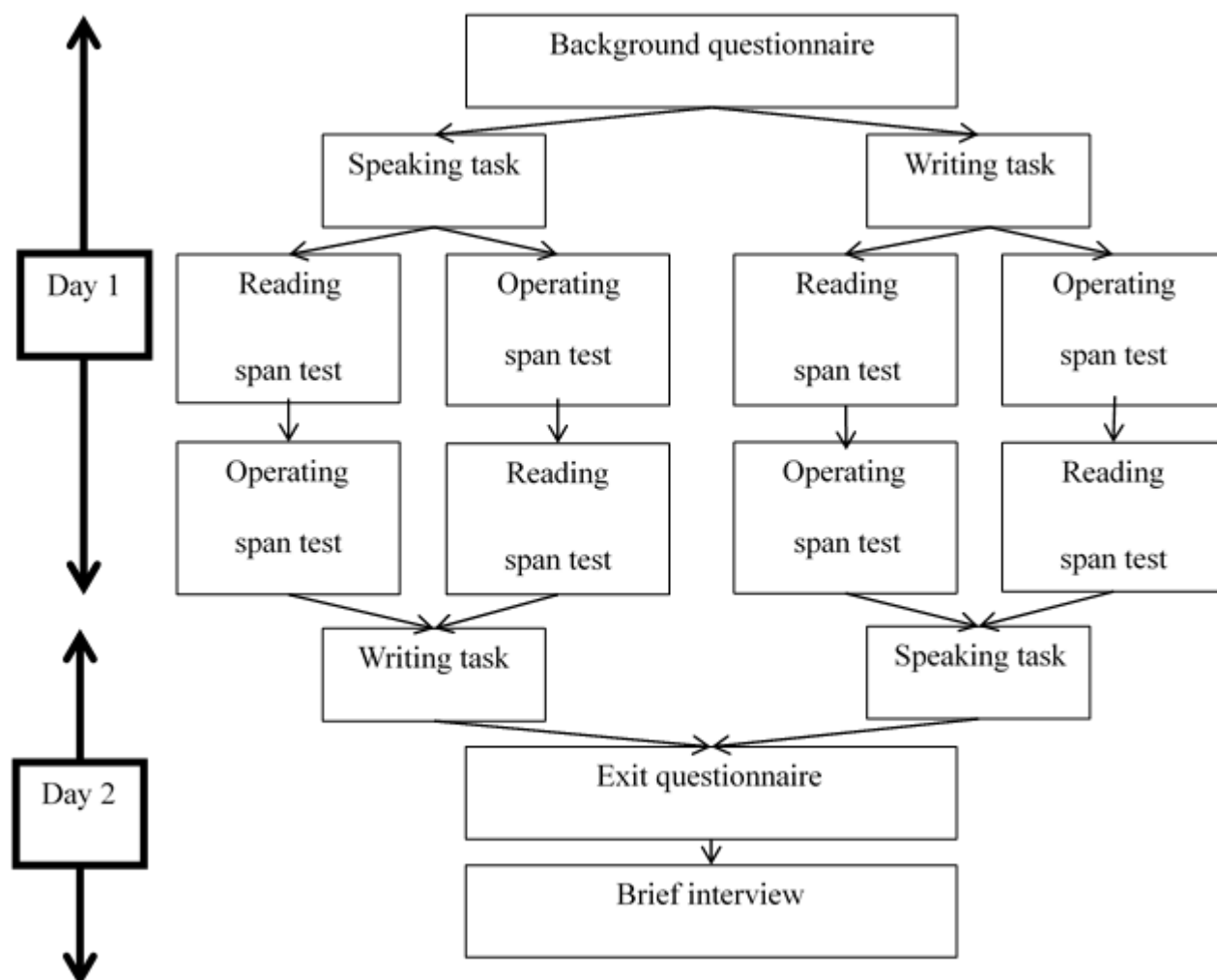


Figure 2. Procedure of the experiment

3.4. Measures

Following the previous research on PT, the emergence and accuracy of each participant's oral and written responses were analyzed. Definitions and criteria are provided in Section 3.4.1 for emergence, and in Section 3.4.2 for accuracy.

3.4.1. Emergence

According to PT, the development of L2 processing should be measured by the emergence of morphosyntactic forms in L2 learners' production rather than their accurate use of structures. As Pienemann (1998) explained, "emergence can be understood as the point in time at which certain skills have, in principle, been attained or at which certain operations can, in principle, be carried out" (p. 138). A single token of a certain morphosyntactic element does not indicate systematic and productive use of the element, which is what determines emergence (Pienemann et al., 1988). Rather, multiple tokens or multiple contexts of use confirm the emergence of morphosyntactic elements in that a sufficient number of tokens or contexts in learners' production shows the regular mapping of form and function in learners' processing (Pallotti, 2007). Therefore, the criteria of emergence for this study are four tokens of each morphosyntactic structure, following Bonilla (2015), Jansen (2008), Pienemann (1998), and Spinner (2011). In addition, if morphosyntactic elements are formulaic or chunked in context, they are excluded (Section 3.2.5.), because formulaic use suggests that the elements may be memorized and may not reflect learners' actual developmental stage (e.g., Myles, 2004).

3.4.2. Accuracy

Following a method used by Lee and Spinner (2017), who were the first to attempt to measure accuracy together with emergence, this study assesses the accuracy of the morphosyntactic structures predicted by PT by means of obligatory occasion analysis (Ellis & Barkhuizen, 2005; Ellis et al., 2008).

Obligatory occasion analysis was conducted for each morphosyntactic structure. Each participant had multiple accuracy scores which were assigned to all the structures. All obligatory occasions for each structure were identified, and each occasion was then examined to investigate whether the correct structure was supplied. Then, the percentage of the correct forms was calculated by using the following formula:

$$\frac{N \text{ correct suppliance in contexts}}{\text{Total obligatory contexts}} \times 100 = \text{percent accuracy}$$

This study defines the “accurate” use of a particular structure in three ways. First, when a structure is related to a verb in a sentence, such as *-ing*, Past *-ed*, and 3rd person singular *-s*, this study considers the correct forms with regard to tense, aspect, and agreement between the subject and verb (Ellis & Yuan, 2004). For instance, if a participant produced *the man riding a horse* to describe a picture, *riding* is considered an inaccurate form of *-ing*. Likewise, if a participant produced *I likes to hang out with my friends*, *likes* in this sentence is be considered an inaccurate form of 3rd person singular *-s* due to incorrect SV agreement. Second, the accuracy of other morphological structures is coded as correct when the participants used the forms accurately and with correct agreement. For example, the number agreement is incorrect in *three bag*, so it is considered an incorrect form of Plural NP. Similarly, because Possessive’s is not used in *Summer (name) shoes*, this case is defined as an inaccurate form of the Possessive ’s. Lastly, the accurate use of syntactic structures is determined by whether the word order of the structures is correct and whether all the required elements are included in the structures. For instance, as there is no subject and dummy *do* in *please not smoke*, this response is categorized as an inaccurate use of S neg V.

Overuse of a particular structure with (in)accuracy was ignored when the structure was used in a single context. For instance, a participant produced the following response to describe one of his/her family members:

My sister's job is teacher. She is good math teaching. She marry. She like XX and sweet food. She like song is classic music a lot. She has one baby. He is very cute.

In this response, the participant did not put 3rd person singular -s to *like* two times. Because these inaccurate forms were produced within the identical context, *like* was counted as one incorrect suppliance. However, if this participant produced *like* instead of *likes* again in different contexts, it was counted as another incorrect suppliance.

In sum, this study's criteria of accuracy go beyond "emergence of accurate form itself" that PT may consider. In particular, in the analysis of *a man riding a horse*, this study assesses that *-ing* emerges based on the PT's proposal, but that it is an inaccurate form of *-ing*.

3.5. Analysis

3.5.1. Advance preparation for data analysis

3.5.1.1. Data organization

Prior to data analysis, the data were organized. As mentioned, the audio-recorder ran for each participant's entire spoken task session; therefore, all of a participant's task responses were in a single file. Each of these files was segmented into sub-files, one for each task. Responses after the time limit were excluded using Audacity; thus, each sub-file contained a response to each question within the time limit. A total of 2,784 sub-files were created. Then, in order to

maintain the greatest possible consistency in transcription and coding, all of the sub-files from all of the learners were grouped by question for coding.

Similarly, the participants used one Microsoft Word document to perform all of the writing task. Each participant's document was divided into sub-files ($N = 2,784$), one for each task response. These sub-files of the written responses were also grouped by question for coding.

3.5.1.2. Transcription of spoken responses

Transcribing the morphosyntactic structures that PT suggests was the primary focus of the transcription. Therefore, it was not necessary to apply detailed transcription conventions, such as those used in conversation analysis (see Lazaraton, 2002).

There were two transcribers. One was a highly advanced L2 speaker with nativelike proficiency, who had earned her master's degree in the U.K. The other was myself. Before starting to transcribe the data, we discussed (a) the target structures, (b) grammatically correct locations of the structures, and (c) possible inaccurate uses of the structures. During this session, we practiced transcription with the use of the data from the pilot study and resolved some discrepancies regarding the target structures. As mentioned, the target structures in the participants' responses were the main focus of the transcription. Therefore, we ignored any differences in our transcriptions of irrelevant elements, such as articles (e.g., *a(n)*, *the*), fillers (e.g., *uh*, *um*), and incorrect pronunciation (e.g., [paiv] instead of [faiv]).

We followed three main principles for the transcription. First, we agreed not to correct any grammatical or semantic errors. Second, if a student produced more than one version of the target structure, for example, in a self-correction, we only transcribed the last structure produced.

For instance, when a student produced the sentence, *the man ride riding is riding a horse*, we transcribed the sentence as *the man is riding a horse*. Third, we transcribed repeated structures only once. For instance, a student produced the sentence, *the man is riding uh riding a horse*. In this case, we transcribed the sentence as *the man is riding a horse*.

Then, considering the even numbers of participants in each proficiency group, we pseudo-randomly selected approximately 30% of the entire data set for norming. Each of us transcribed the selected data individually, and then met again. We discussed some discrepancies regarding the target structures while comparing our transcriptions; we resolved them. After the norming session, we pseudo-randomly divided the rest of the data (i.e., 70% of the entire data set) based on participants' proficiency, after which we transcribed individually.

3.5.2. Coding

Two native speakers of English coded the participants' responses (i.e., the transcriptions of the spoken responses and the written responses) based on the morphosyntactic structures predicted by PT. Both of them were instructors at an English language center at a large university, and had master's degrees in TESOL; in addition, they had multiple previous experiences of coding L2 learners' speaking and writing.

They were asked to count the morphosyntactic structures predicted by PT and judge the accuracy of each usage of the structures in context while reading through the transcriptions of the spoken responses and the written responses. As mentioned, all of the participants' responses to each question were gathered together in order to maintain consistency in coding. Hence, the coders worked through all the responses to one question before moving onto the responses to the

next question. Using Microsoft Excel, they listed all the occurrences of the target structures that they found. One Excel sheet was assigned to one participant's response to one question; thus, the coders filled in a total of 2,784 sheets during the coding.

Each Excel sheet was composed of two tables. In the lefthand table, each row represented one of the morphosyntactic structures predicted by PT; there were two columns, for accurate and inaccurate use, respectively. The two coders listed all of the structures they noted, and they assessed the accuracy of each occurrence of the structures in this table. For instance, if a participant produced the sentence, *the man riding a bike*, the coder classified *riding* as “inaccurate -ing”. The righthand table looked like the lefthand table. Using Excel's COUNTA function, the occurrences of structures recorded in the lefthand table were automatically counted and filled into the assigned cell in the righthand table. Appendix J shows an example of an Excel sheet used in the coding.

The two coders and I had a two-hour practice session. The experimental settings and the target structures were introduced, and then I explained how they would be asked to organize their coding works in the designated Excel sheet. Then, they were trained by using the data from the pilot study and some exemplar inaccurate structures extracted from the main data and discussed their coding. They resolved some discrepancies in their coding through the discussion.

Then, similar to the procedure of the transcription, each of the coders individually counted and assessed the morphosyntactic structures in 30% of the entire set of responses for norming. The intercoder reliability was .87 for the norming session, indicating that the extent to which the two coders agreed on the coding of the part of the responses is acceptable. After calculating the reliability, we had another meeting to discuss and resolve some discrepancies.

After the norming session, half of the rest of the data was assigned to each of the coders, and they worked individually.

After the two coders finished coding, I combined all the Excel sheets for each participant into one sheet and read through the results of the coding for those who produced any of the target structures more than four times to confirm the emergence of the structure and exclude any formulaic uses of the structures (Pienemann et al., 1988; Spinner, 2007, 2011). To be specific, participants sometimes used identical lexical items more than two times to produce a certain morphosyntactic structure throughout the experiment. There were two types of such repetition: non-formulaic and formulaic. In an example of the former, a participant produced *two men are riding a horse* and *the people are riding a bike*. This participant used *riding* two times throughout the experiment; however, in neither case did *riding* seem formulaic, given that the subjects and objects differed. Rather, it seemed to be appropriate vocabulary to describe each picture. Thus, in this case, *riding* was considered as one of the tokens of *-ing*, and the number of emergences was revised accordingly. In the second type, the repeated lexical items were part of apparently idiomatic expressions or formulaic chunks. For instance, a participant repeatedly used *it seems that...* throughout the experiment. In this case, *seems* was not considered as one of the tokens of third person singular *-s*; it was excluded as probable formulaic language.

3.5.3. Statistical procedures

3.5.3.1. Implicational scale

Implicational scaling is a method of visually representing L2 learners' developmental patterns in a table format. It allows for the examination of gradual progress in L2 learning. The first column of an implication table shows all participants arranged in order from less

development in the top rows to greater development in the bottom rows. The rest of the columns indicate each of the target morphosyntactic structures predicted by PT, with structures in the earlier stages appearing in earlier columns (i.e., to the left) and structures in later stages appearing in later columns (i.e., to the right). For instance, plural -s in stage 2 is found in an earlier column than 3rd person singular -s in stage 5. The order of structures that PT predicts to appear in the same stage is determined based on the number of emerged responses for each structure in the data. When a structure fulfills the criteria for emergence in a learner's production, it is marked *e*. When a structure does not fulfill the criteria for emergence in a learner's production, it is marked *N*.

Errors can occur in implicational scaling. In particular, it can show some later stage structures to have emerged while earlier stage structures have not emerged. Based on the number of errors, four calculations are performed to determine the predictability and reliability of an implicational scale (Hatch & Farhady, 1982; Hatch & Lazarton, 1991). First, the coefficient of reproducibility (C of R) indicates the extent to which an implicational scaling predicts each of the participants' development based on his/her rank in the implicational scaling. A C of R over .9 is considered to show that an implicational scaling provides a predictable pattern of participants' development. The same calculation can be performed for each structure in an implicational scaling in order to indicate the predictability of each structure. The C of R is calculated by using the following formula:

$$\text{C of R} = 1 - \frac{\text{Number of errors}}{\text{Total number of responses}}$$

Next, the minimum marginal reproducibility (MMR) and the percent improvement in reproducibility (PIR) are calculated. The following formulae are used:

$$\text{MMR} = \frac{\text{Number of emerged responses}}{\text{Total number of responses}}$$

$$\text{PIR} = \text{C of R} - \text{MMR}$$

Finally, the coefficient of scalability (C of S) is calculated, which indicates the extent to which the results in an implicational scale are reliable. A C of S above .6 is considered to confirm that an implicational scale is scalable. The C of S is calculated by using the following formula:

$$\text{C of S} = \frac{\text{PIR}}{1 - \text{MMR}}$$

Previous research has reported that C of R and C of S are acceptable methods of determining whether an implicational scaling truly represents L2 learners' developmental patterns. In this dissertation, implicational scaling is created for the participants' responses in each modality; in addition, C of R is calculated for each structure in an implicational scaling of each modality.

CHAPTER 4. RESULTS

4.1. Implicational tables

The written responses of one participant (participant ID: par19) were removed, because the data got corrupted and data recovery failed. Therefore, there were 2,816 oral responses from 88 participants and 2,752 written responses from 87 participants. Emergence of the target morphosyntactic structures was analyzed in the participants' responses in each modality, and two implicational tables were created for each modality. The responses to two tasks designed to elicit questions in each modality were excluded due to low numbers of target responses. These four tasks were designed to elicit certain forms of questions, but did not require participants to produce a particular form because doing so would reveal the purpose of the tasks. Because participants did not stick to a certain form, only a small number of responses met the criterion of emergence for the target form.

4.1.1. Oral responses

The implicational scaling for the participants' oral responses is presented in Table 5. PT posits that production of a specific structure in a stage confirms L2 learners' acquisition of the processing procedure in the stage. In this regard, a structure which emerged the most in each stage (e.g., Plural -s in stage 2; Plural NP in stage 3) was selected to create the implicational scaling. There were 435 available instances of emergence in the implicational scaling, and 1 error was found. Based on the results, the C of R was 1 and the C of S was .99, which confirmed that the implicational scaling successfully predicted the participants' L2 development in the oral modality, and that the results reliably reflected L2 developmental patterns in the oral modality.

Table 5. Implicational scaling of oral responses

Group	P.id	Stage				
		S1	S2	S3	S5	S6
B1	par6	e	e	N	N	N
A2	par69	e	e	e	N	N
A2	par89	e	e	e	N	N
A2	par28	e	e	e	N	N
A2	par87	e	e	e	N	N
A2	par71	e	e	e	N	N
A2	par77	e	e	e	N	N
A2	par83	e	e	e	N	N
B1	par50	e	e	e	N	N
B2	par 1	e	e	e	N	N
B2	par24	e	e	e	N	N
A2	par85	e	e	e	N	N
A2	par40	e	e	e	N	N
B1	par36	e	e	e	N	N
A2	par66	e	e	e	N	N
A2	par68	e	e	e	N	N
A2	par75	e	e	e	N	N
A2	par86	e	e	e	N	N
A2	par92	e	e	e	N	N
A2	par93	e	e	e	N	N
B1	par22	e	e	e	N	N
B1	par27	e	e	e	N	N
B1	par29	e	e	e	N	N
B1	par46	e	e	e	N	N
B1	par55	e	e	e	N	N
B1	par60	e	e	e	N	N
B1	par61	e	e	e	N	N
B2	par10	e	e	e	N	N
B2	par12	e	e	e	N	N
B2	par15	e	e	e	N	N
B2	par17	e	e	e	N	N
B2	par19	e	e	e	N	N
B2	par21	e	e	e	N	N
B2	par23	e	e	e	N	N
B2	par4	e	e	e	N	N
B2	par5	e	e	e	N	N
B2	par9	e	e	e	N	N

(cont'd)

C1	par37	e	e	e	N	N
A2	par51	e	e	e	N	N
A2	par90	e	e	e	N	N
A2	par91	e	e	e	N	N
B1	par57	e	e	e	N	N
B1	par58	e	e	e	N	N
B1	par38	e	e	e	N	N
B1	par39	e	e	e	N	N
B1	par42	e	e	e	N	N
B1	par47	e	e	e	N	N
B1	par52	e	e	e	N	N
B1	par64	e	e	e	N	N
B2	par34	e	e	e	N	N
B2	par8	e	e	e	N	N
C1	par18	e	e	e	N	N
A2	par76	e	e	e	e	N
B2	par14	e	e	e	e	N
C1	par41	e	e	e	e	N
A2	par62	e	e	e	e	N
B1	par33	e	e	e	e	N
B2	par3	e	e	e	e	N
C1	par53	e	e	e	e	N
B1	par54	e	e	e	e	N
B1	par35	e	e	e	e	N
B1	par48	e	e	e	e	N
B2	par11	e	e	e	e	N
B2	par13	e	e	e	e	N
B2	par26	e	e	e	e	N
B2	par30	e	e	e	e	N
B2	par43	e	e	e	e	N
B2	par45	e	e	e	e	N
C1	par20	e	e	e	e	N
C1	par32	e	e	e	e	N
C1	par49	e	e	e	e	N
C1	par7	e	e	e	e	N
C1	par72	e	e	e	e	N
C1	par73	e	e	e	e	N
C1	par74	e	e	e	e	N
C1	par78	e	e	e	e	N
C1	par79	e	e	e	e	N

(cont'd)

C1	par80	e	e	e	e	N
C1	par81	e	e	e	e	N
C1	par82	e	e	e	e	N
C1	par84	e	e	e	e	N
C1	par88	e	e	e	e	N
A2	par67	e	e	e	N	e
C1	par65	e	e	e	e	e
C1	par70	e	e	e	e	e
C1	par25	e	e	e	e	e
C1	par31	e	e	e	e	e

Emergence of the target structures followed the PT's prediction: The structures in earlier stages emerged in most of the participants' production, whereas the structures in later stages emerged only in some advanced learners' production. For instance, the emergence of structure in stage 2 can be found in all participants' production, whereas the structures in stage 6 emerged only in five participants' production. the emergence of structures in stage 2.

A set of Cs of R were calculated for individual structures in the implicational scaling (Table 6). The Cs of R for all structures, except Past *-ed*, were over .9, while Past *-ed* showed a lower C of R than the criterion. The C of R for Past *-ed* suggests that reaching a certain developmental stage does not necessarily mean that all possible structures in this stage emerge; this is however consistent with PT theory.

Table 6. Cs of R for individual structures in the oral responses

Stage	Structure	C of R
Stage 1	Words	1
Stage 2	S neg V	0.98
	Plural -s	1
	-ing	1
	Possessive 's	0.91
	Past -ed	0.48
Stage 3	Plural NP	1
	Object pronoun	0.99
Stage 5	3 rd person singular -s	0.99
Stage 6	Cancel INV	0.99

4.1.2. Written responses

The implicational scaling for the participants' written responses is shown in Table 7.

There were 1 error among the 430 available instances of emergence in the implicational scaling.

Based on the number of errors and emerged responses, the C of R and the C of S were calculated. The C of R was 1, and the C of S was 1. These results indicate that the order of the participants' L2 development in their written production could be described accurately and reliably by PT.

Table 7. Implicational scaling of written responses

Group	P.id	Stage				
		S1	S2	S3	S5	S6
A2	par69	e	e	N	N	N
A2	par89	e	e	N	N	N
A2	par77	e	e	N	N	N
B1	par27	e	e	N	N	N
A2	par51	e	e	N	N	N
A2	par90	e	e	N	N	N
A2	par76	e	e	e	N	N
A2	par28	e	e	e	N	N
A2	par71	e	e	e	N	N
A2	par83	e	e	e	N	N
A2	par85	e	e	e	N	N
A2	par87	e	e	e	N	N
A2	par67	e	e	e	N	N
B1	par50	e	e	e	N	N
B1	par6	e	e	e	N	N
A2	par86	e	e	e	N	N
B1	par42	e	e	e	N	N
C1	par73	e	e	e	N	N
A2	par93	e	e	e	N	N
B1	par36	e	e	e	N	N
A2	par40	e	e	e	N	N
A2	par92	e	e	e	N	N
A2	par68	e	e	e	N	N
A2	par66	e	e	e	N	N
A2	par75	e	e	e	N	N
B1	par48	e	e	e	N	N
B1	par38	e	e	e	N	N
B1	par39	e	e	e	N	N
B1	par52	e	e	e	N	N
B1	par60	e	e	e	N	N
B1	par61	e	e	e	N	N
B1	par46	e	e	e	N	N
B1	par22	e	e	e	N	N
B1	par55	e	e	e	N	N
B2	par4	e	e	e	N	N
B2	par9	e	e	e	N	N
B2	par26	e	e	e	N	N

(cont'd)

B2	par10	e	e	e	N	N
B2	par11	e	e	e	N	N
B2	par15	e	e	e	N	N
B2	par17	e	e	e	N	N
C1	par37	e	e	e	N	N
C1	par53	e	e	e	N	N
C1	par20	e	e	e	N	N
B2	par12	e	e	e	N	N
A2	par91	e	e	e	N	N
A2	par62	e	e	e	N	N
B1	par64	e	e	e	N	N
B1	par57	e	e	e	N	N
B1	par47	e	e	e	N	N
B1	par54	e	e	e	N	N
B1	par29	e	e	e	N	N
B2	par1	e	e	e	N	N
B2	par21	e	e	e	N	N
B2	par34	e	e	e	N	N
B2	par30	e	e	e	N	N
B2	par45	e	e	e	N	N
B2	par23	e	e	e	N	N
B2	par24	e	e	e	N	N
C1	par25	e	e	e	N	N
C1	par32	e	e	e	N	N
C1	par41	e	e	e	N	N
C1	par88	e	e	e	N	N
C1	par74	e	e	e	N	N
B2	par3	e	e	e	e	N
B1	par33	e	e	e	e	N
B2	par5	e	e	e	e	N
C1	par65	e	e	e	e	N
C1	par72	e	e	e	e	N
C1	par84	e	e	e	e	N
B1	par58	e	e	e	e	N
B1	par35	e	e	e	e	N
B2	par43	e	e	e	e	N
B2	par13	e	e	e	e	N
B2	par14	e	e	e	e	N
C1	par18	e	e	e	e	N
C1	par80	e	e	e	e	N

(cont'd)

C1	par81	e	e	e	e	N
C1	par82	e	e	e	e	N
C1	par78	e	e	e	e	N
C1	par79	e	e	e	e	N
C1	par7	e	e	e	e	N
C1	par70	e	e	e	e	N
C1	par49	e	e	e	e	N
B2	par8	e	e	e	N	e
C1	par31	e	e	e	e	e

As the table demonstrates, the emergence of the target structures in the participants' written production was similar to that in their oral production. The structures emerged gradually and in the predicted order. The developmental pattern displayed in the table indicates that it may take a longer time to reach a certain developmental stage in the written modality than in the oral modality. For instance, more participants failed to show acquisition of the processing procedure in stage 3 in their written response than in their oral responses. Likewise, fewer participants showed the emergence of 3rd person singular (in stage 5) and the emergence of Cancel INV (in stage 6) in the written modality than in the oral modality.

In addition, the Cs of R were calculated for each morphosyntactic structure in the implicational scaling (Table 8). The results were similar to those from the implication table of the oral responses. Again, the C of R for Past -ed did not reach .9, which suggests that past -ed did not emerge in a predictable way in the written modality.

Table 8. Cs of R for individual structures in written responses

Stage	Structure	C of R
Stage 1	Words	1
Stage 2	S neg V	0.97
	Plural -s	1
	-ing	0.99
	Possessive 's	0.95
	Past -ed	0.48
Stage 3	Plural NP	0.95
	Object pronoun	0.99
Stage 5	3 rd person singular -s	0.99
Stage 6	Cancel INV	1

4.1.3. Comparison between oral and written responses

To explore the extent of similarity and difference between the implicational tables for the two modalities, Pearson correlations were conducted two times with (a) the amount of emergence (out of five stages) in each participant's response and (b) the number of emerged stages in all participants' responses. For instance, emergence occurred at Stage 1, 2, and 3 in a participant's oral response; this participant had three emergences in the oral modality. In addition, among 87 participants' written responses, emergence occurred at Stage 3 in 81 participants' written responses. The number of emerged Stage 3 is 81 in this case. The results of these two correlation analyses indicated that the implicational tables in the two modalities were correlated based on (a) each participant's responses ($r = .50, p < .0001$) and (b) emerged stages in all participants' responses ($r = .99, p = .001$).

Figure 3 indicates the relationship between the amount of emergence (based on stage) in each participant's response in the oral modality and written modalities, respectively. Overall, the amount of emergence in the oral modality seems to align with the amount of emergence in the written modality. However, some participants had a greater amount of emergence in the oral

modality than in the written modality, and vice versa. For instance, a greater amount of emergence in the oral modality ($N = 3$) rather than the written modality ($N = 2$) occurred in some A2-level participants' responses. The number of emerged stages is illustrated in Figure 4. In earlier stages, a higher number of emergence is found, while it decreases in later stages. This trend can be identified in both the oral and written modalities. Although the differences are not significant, the number of emergence is greater in the oral modality than in the written modality in stage 3, 5, and 6. The number of emergence in stage 5 showed greater difference between the oral and the written modality. The differences in stage 1 and 2 were due to the variations in the number of participants in each modality.

In addition, a generalized mixed model was performed with emergence of a structure (coded as emergence and none) as a dependent variable, modality (coded as oral and written) as a fixed effect, and participant as a random effect. Oral modality served as a reference level. Modality was not significant in this model ($\beta = -.21$, $SE = .14$, $z = 1.49$, $p = .14$), indicating that the pattern of the gradual emergence of the target structures was similar in the two modalities.



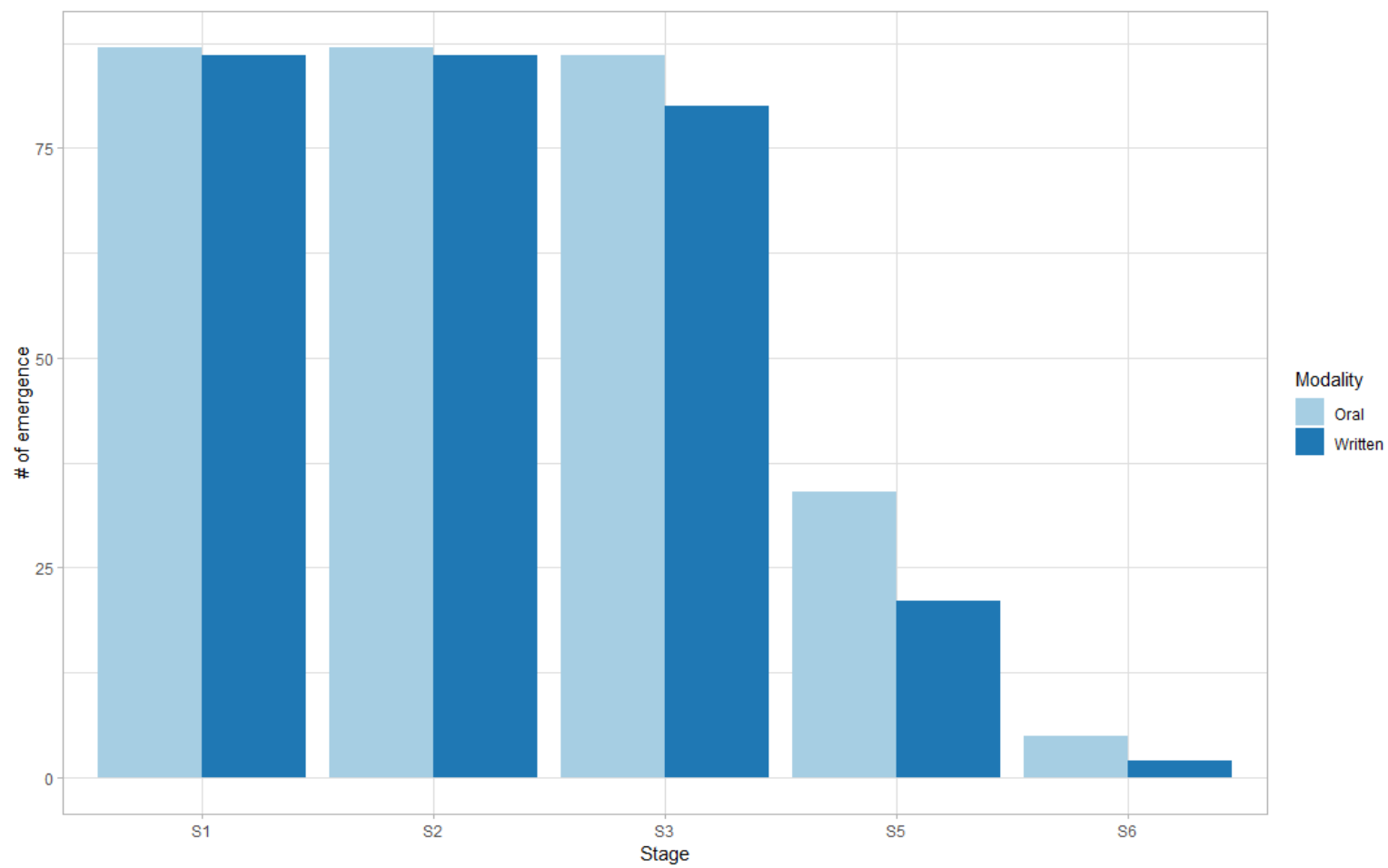


Figure 4. The number of emerged stages in the oral and written modalities

4.2. Order of emergence within each stage

4.2.1. Oral responses

To explore whether structures in each stage emerged at the same time (i.e., intrastage development), a new implicational scaling for the participants' oral responses was created including all target structures (Table 9). There were 870 available instances of emergence in the implicational scaling, and 57 errors were found. Based on the results, the C of R was .93 and the C of S was .73.

Plural NP and Object pronoun were selected as the target structures for stage 3. Although they are in the same stage, these two structures emerged at different times. In particular, all participants but one produced plural NP consistently throughout the oral interview. However, the Object pronoun did not emerge in the production of the 12 participants who appear in the earlier rows. Likewise, S neg V, plural *-s*, *-ing*, and possessive *'s* in stage 2 followed the PT's prediction. Except for a few errors, these structures usually emerged before the structures in the later stages. The implicational scale indicates the following order of emergence in stage 2 from Plural *-s*, *-ing*, S neg V, to Possessive *'s*. However, the emergence of Past *-ed* occurred irregularly, and did not meet the criterion even when structures in the later stages had already emerged. Thus, although it is one of the structures in stage 2, it appears in a later column based on the number of case of its emergence.

Table 9. Implicational scaling of oral responses including all target structures

Group	P.id	Stage / structure									
		S1	S2	S2	S2	S2	S3	S3	S2	S5	S6
		words	-s	-ing	neg	's	Plural NP	Obj. pro	-ed	-s	Cancel INV
B1	par6	e	e	E	e	e	N	N	N	N	N
A2	par69	e	e	E	e	N	e	N	N	N	N
A2	par89	e	e	E	e	N	e	N	N	N	N
A2	par28	e	e	E	N	e	e	N	N	N	N
A2	par87	e	e	E	e	e	e	N	N	N	N
A2	par71	e	e	E	e	e	e	N	N	N	N
A2	par77	e	e	E	e	e	e	N	N	N	N
A2	par83	e	e	E	e	e	e	N	N	N	N
B1	par50	e	e	E	e	e	e	N	N	N	N
B2	par 1	e	e	E	e	e	e	N	N	N	N
B2	par24	e	e	E	e	e	e	N	N	N	N
A2	par85	e	e	e	e	e	e	N	e	N	N
A2	par40	e	e	e	e	N	e	e	N	N	N
B1	par36	e	e	e	e	N	e	e	N	N	N
A2	par66	e	e	e	e	e	e	e	N	N	N
A2	par68	e	e	e	e	e	e	e	N	N	N
A2	par75	e	e	e	e	e	e	e	N	N	N
A2	par86	e	e	e	N	e	e	e	N	N	N
A2	par92	e	e	e	e	e	e	e	N	N	N
A2	par93	e	e	e	e	e	e	e	N	N	N
B1	par22	e	e	e	e	e	e	e	N	N	N
B1	par27	e	e	e	e	e	e	e	N	N	N
B1	par29	e	e	e	e	e	e	e	N	N	N
B1	par46	e	e	e	e	e	e	e	N	N	N
B1	par55	e	e	e	e	e	e	e	N	N	N
B1	par60	e	e	e	e	e	e	e	N	N	N
B1	par61	e	e	e	e	e	e	e	N	N	N
B2	par10	e	e	e	e	e	e	e	N	N	N
B2	par12	e	e	e	e	e	e	e	N	N	N
B2	par15	e	e	e	e	e	e	e	N	N	N
B2	par17	e	e	e	e	e	e	e	N	N	N
B2	par19	e	e	e	e	e	e	e	N	N	N
B2	par21	e	e	e	e	e	e	e	N	N	N
B2	par23	e	e	e	e	e	e	e	N	N	N
B2	par4	e	e	e	e	e	e	e	N	N	N
B2	par5	e	e	e	e	e	e	e	N	N	N
B2	par9	e	e	e	e	e	e	e	N	N	N

(cont'd)

C1	par37	e	e	e	e	e	e	e	N	N	N
A2	par51	e	e	e	e	N	e	e	e	N	N
A2	par90	e	e	e	e	e	e	e	e	N	N
A2	par91	e	e	e	e	e	e	e	e	N	N
B1	par57	e	e	e	e	e	e	e	e	N	N
B1	par58	e	e	e	e	e	e	e	e	N	N
B1	par38	e	e	e	e	e	e	e	e	N	N
B1	par39	e	e	e	e	e	e	e	e	N	N
B1	par42	e	e	e	e	e	e	e	e	N	N
B1	par47	e	e	e	e	e	e	e	e	N	N
B1	par52	e	e	e	e	e	e	e	e	N	N
B1	par64	e	e	e	e	e	e	e	e	N	N
B2	par34	e	e	e	e	e	e	e	e	N	N
B2	par8	e	e	e	e	e	e	e	e	N	N
C1	par18	e	e	e	e	e	e	e	e	N	N
A2	par76	e	e	e	e	N	e	N	N	e	N
B2	par14	e	e	e	e	N	e	e	N	e	N
C1	par41	e	e	e	e	N	e	e	N	e	N
A2	par62	e	e	e	e	e	e	e	N	e	N
B1	par33	e	e	e	e	e	e	e	N	e	N
B2	par3	e	e	e	e	e	e	e	N	e	N
C1	par53	e	e	e	e	e	e	e	N	e	N
B1	par54	e	e	e	e	e	e	e	e	e	N
B1	par35	e	e	e	e	e	e	e	e	e	N
B1	par48	e	e	e	e	e	e	e	e	e	N
B2	par11	e	e	e	e	e	e	e	e	e	N
B2	par13	e	e	e	e	e	e	e	e	e	N
B2	par26	e	e	e	e	e	e	e	e	e	N
B2	par30	e	e	e	e	e	e	e	e	e	N
B2	par43	e	e	e	e	e	e	e	e	e	N
B2	par45	e	e	e	e	e	e	e	e	e	N
C1	par20	e	e	e	e	e	e	e	e	e	N
C1	par32	e	e	e	e	e	e	e	e	e	N
C1	par49	e	e	e	e	e	e	e	e	e	N
C1	par7	e	e	e	e	e	e	e	e	e	N
C1	par72	e	e	e	e	e	e	e	e	e	N
C1	par73	e	e	e	e	e	e	e	e	e	N
C1	par74	e	e	e	e	e	e	e	e	e	N
C1	par78	e	e	e	e	e	e	e	e	e	N
C1	par79	e	e	e	e	e	e	e	e	e	N
C1	par80	e	e	e	e	e	e	e	e	e	N
C1	par81	e	e	e	e	e	e	e	e	e	N

(cont'd)

C1	par82	e	e	e	e	e	e	e	e	e	N
C1	par84	e	e	e	e	e	e	e	e	e	N
C1	par88	e	e	e	e	e	e	e	e	e	N
A2	par67	e	e	e	e	e	e	e	N	N	e
C1	par65	e	e	e	e	e	e	e	e	e	e
C1	par70	e	e	e	e	e	e	e	e	e	e
C1	par25	e	e	e	e	e	e	e	e	e	e
C1	par31	e	e	e	e	e	e	e	e	e	e

4.2.2. Written responses

A new implicational scaling for the participants' written responses which includes all target structures was also created (Table 10). There were 63 errors among the 860 available instances of emergence in the implicational scaling. Based on the number of errors and emerged responses, the C of R and the C of S were calculated. The C of R was .93, and the C of S was .73.

There are some errors in which the Object pronoun emerged earlier than Plural NP in the participants' written responses ($N = 4$). However, in most of the participants' responses, Plural NP emerged earlier than the Object pronoun, as with the oral responses.

When it comes to the structures in stage 2, the order of emergence was more distinct, compared to that shown in the oral response. Plural *-s* emerged first, followed by *-ing*, S neg V, Possessive *'s*, and Past *-ed*. in that order. Past *-ed* thus emerged later than the structures in stage 3.

Table 10. Implicational scaling of oral responses including all target structures

Group	P.id	Stage / structure									
		S1 words	S2 -s	S2 -ing	S2 neg	S2 's	S3 Plural NP	S3 Obj.pro	S2 -ed	S5 -s	S6 Cancel INV
A2	par69	e	e	N	e	N	N	e	N	N	N
A2	par89	e	e	e	N	N	N	N	N	N	N
A2	par77	e	e	e	N	N	N	N	N	N	N
A2	par76	e	e	e	e	N	e	N	N	N	N
B1	par27	e	e	e	e	N	N	e	N	N	N
A2	par28	e	e	e	e	e	e	N	N	N	N
A2	par71	e	e	e	e	e	e	N	N	N	N
A2	par83	e	e	e	e	e	e	N	N	N	N
A2	par85	e	e	e	e	e	e	N	N	N	N
A2	par87	e	e	e	e	e	e	N	N	N	N
A2	par67	e	e	e	e	e	e	N	N	N	N
B1	par50	e	e	e	e	e	e	N	N	N	N
B1	par6	e	e	e	e	e	e	N	N	N	N
A2	par86	e	e	e	N	e	e	N	e	N	N
B1	par42	e	e	e	e	e	e	N	e	N	N
C1	par73	e	e	e	e	e	e	N	e	N	N
A2	par51	e	e	e	e	e	N	e	N	N	N
A2	par90	e	e	e	e	e	N	e	N	N	N
A2	par93	e	e	e	e	N	e	e	N	N	N
B1	par36	e	e	e	e	N	e	e	N	N	N
A2	par40	e	e	e	e	e	e	e	N	N	N
A2	par92	e	e	e	e	e	e	e	N	N	N
A2	par68	e	e	e	e	e	e	e	N	N	N
A2	par66	e	e	e	e	e	e	e	N	N	N
A2	par75	e	e	e	e	e	e	e	N	N	N
B1	par48	e	e	e	e	e	e	e	N	N	N
B1	par38	e	e	e	e	e	e	e	N	N	N
B1	par39	e	e	e	e	e	e	e	N	N	N
B1	par52	e	e	e	e	e	e	e	N	N	N
B1	par60	e	e	e	e	e	e	e	N	N	N
B1	par61	e	e	e	e	e	e	e	N	N	N
B1	par46	e	e	e	e	e	e	e	N	N	N
B1	par22	e	e	e	e	e	e	e	N	N	N
B1	par55	e	e	e	e	e	e	e	N	N	N
B2	par4	e	e	e	e	e	e	e	N	N	N
B2	par9	e	e	e	e	e	e	e	N	N	N
B2	par26	e	e	e	e	e	e	e	N	N	N
B2	par10	e	e	e	e	e	e	e	N	N	N

(cont'd)

B2	par11	e	e	e	e	e	e	N	N	N
B2	par15	e	e	e	e	e	e	N	N	N
B2	par17	e	e	e	e	e	e	N	N	N
C1	par37	e	e	e	e	e	e	N	N	N
C1	par53	e	e	e	e	e	e	N	N	N
C1	par20	e	e	e	e	e	e	N	N	N
B2	par12	e	e	e	e	N	e	e	N	N
A2	par91	e	e	e	e	e	e	e	N	N
A2	par62	e	e	e	e	e	e	e	N	N
B1	par64	e	e	e	e	e	e	e	N	N
B1	par57	e	e	e	e	e	e	e	N	N
B1	par47	e	e	e	e	e	e	e	N	N
B1	par54	e	e	e	e	e	e	e	N	N
B1	par29	e	e	e	e	e	e	e	N	N
B2	par1	e	e	e	e	e	e	e	N	N
B2	par21	e	e	e	e	e	e	e	N	N
B2	par34	e	e	e	e	e	e	e	N	N
B2	par30	e	e	e	e	e	e	e	N	N
B2	par45	e	e	e	e	e	e	e	N	N
B2	par23	e	e	e	e	e	e	e	N	N
B2	par24	e	e	e	e	e	e	e	N	N
C1	par25	e	e	e	e	e	e	e	N	N
C1	par32	e	e	e	e	e	e	e	N	N
C1	par41	e	e	e	e	e	e	e	N	N
C1	par88	e	e	e	e	e	e	e	N	N
C1	par74	e	e	e	e	e	e	e	N	N
B2	par3	e	e	e	e	N	e	e	e	N
B1	par33	e	e	e	e	e	e	N	e	N
B2	par5	e	e	e	e	e	e	N	e	N
C1	par65	e	e	e	e	e	e	N	e	N
C1	par72	e	e	e	e	e	e	N	e	N
C1	par84	e	e	e	e	e	N	e	e	N
B1	par58	e	e	e	e	e	e	e	e	N
B1	par35	e	e	e	e	e	e	e	e	N
B2	par43	e	e	e	e	e	e	e	e	N
B2	par13	e	e	e	e	e	e	e	e	N
B2	par14	e	e	e	e	e	e	e	e	N
C1	par18	e	e	e	e	e	e	e	e	N
C1	par80	e	e	e	e	e	e	e	e	N
C1	par81	e	e	e	e	e	e	e	e	N
C1	par82	e	e	e	e	e	e	e	e	N
C1	par78	e	e	e	e	e	e	e	e	N

(cont'd)

C1	par79	e	e	e	e	e	e	e	e	e	N
C1	par7	e	e	e	e	e	e	e	e	e	N
C1	par70	e	e	e	e	e	e	e	e	e	N
C1	par49	e	e	e	e	e	e	e	e	e	N
B2	par8	e	e	e	e	e	e	N	N		e
C1	par31	e	e	e	e	e	e	N	e		e

4.2.3. Comparison between oral and written responses

Pearson correlations were conducted to compare oral and written modalities' implicational tables that include all target morphosyntactic structures. As earlier, Pearson correlations were conducted based on the number of emergence (out of ten structures) in each participant's response and the number of emerged structures in all participants' responses. The results showed that the implicational tables in the two modalities were highly correlated based on (a) each participant's responses ($r = .69, p < .0001$) and (b) emerged structures in all participants' responses ($r = .99, p < .0001$).

Figure 5 shows the relationship between the amount of emergence (based on each structure) in each participant's response in the oral and written modalities, respectively. For instance, when Plural *-s, -ing*, and Neg emerged in a participant's response, the amount of emergence for this participant is three. In general, when a larger number of structures emerged in a participant's oral response, similar results were found in the participant's written response. Hence, the amount of emergence in each participant's oral and written responses seem consistent with each other. Figure 6 reveals the number of emerged structures in the oral and written modalities. Overall, structures in earlier stages emerged in higher numbers than those in later stages. However, some structures, such as Past *-ed* emerged less than structures in later stage (i.e., Plural NP, Obj. pronoun); structures within the identical stage showed different number of

emergence, which suggests intrastage development. The number of emerged structures is also greater in the oral modality, as is found in the case of the 3rd person singular -s.

In addition, a generalized mixed model was performed with emergence of a structure (coded as emergence and none) as a dependent variable, modality (coded as oral and written) as a fixed effect, and participant as a random effect. Oral modality served as a reference level. Modality was not significant in this model ($\beta = -.19$, $SE = .11$, $z = -1.70$, $p = .09$), indicating that the pattern of the gradual emergence of the target structures was similar in the two modalities.

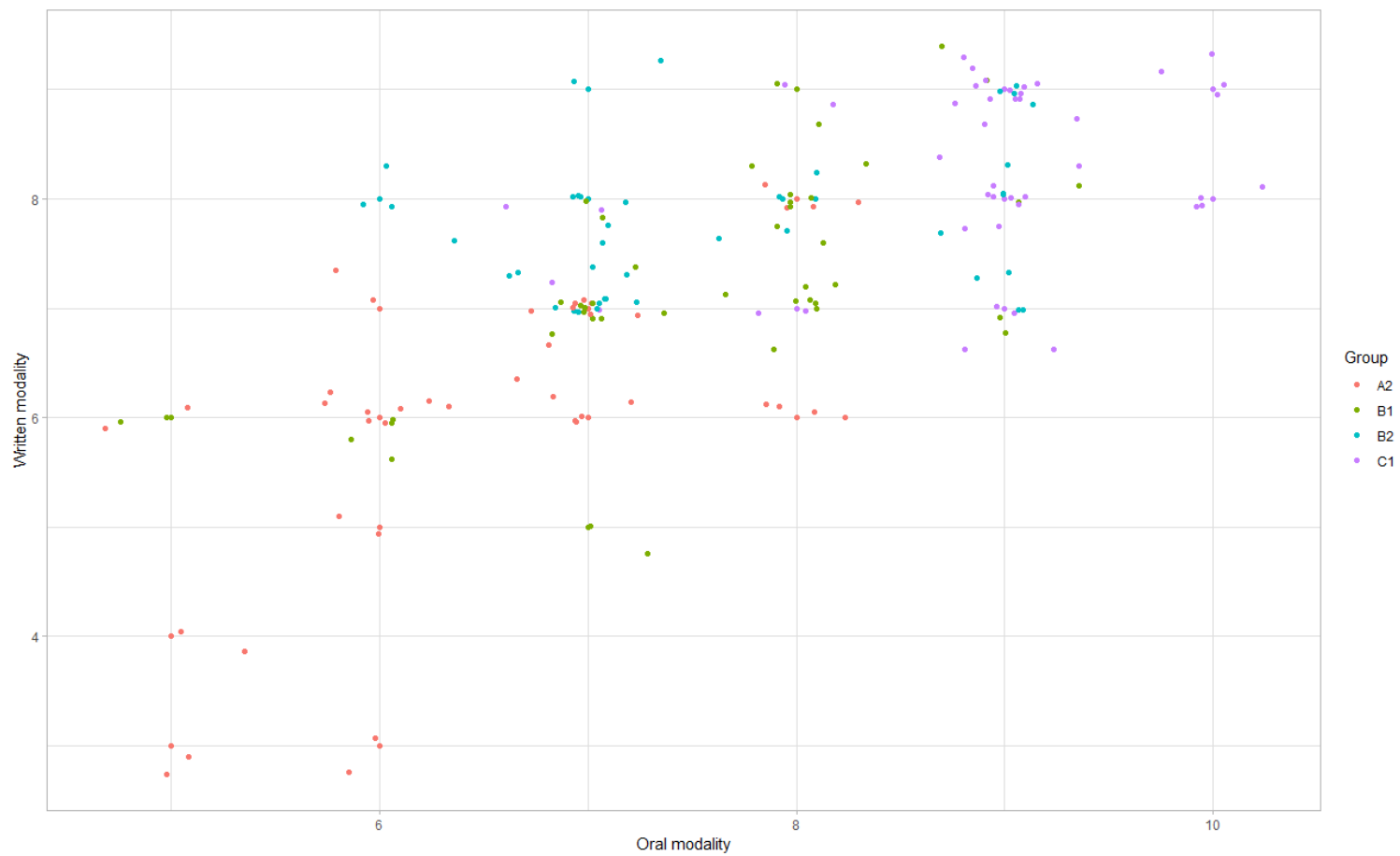


Figure 5. Scatterplot of the amount of emergence (including all target structures)

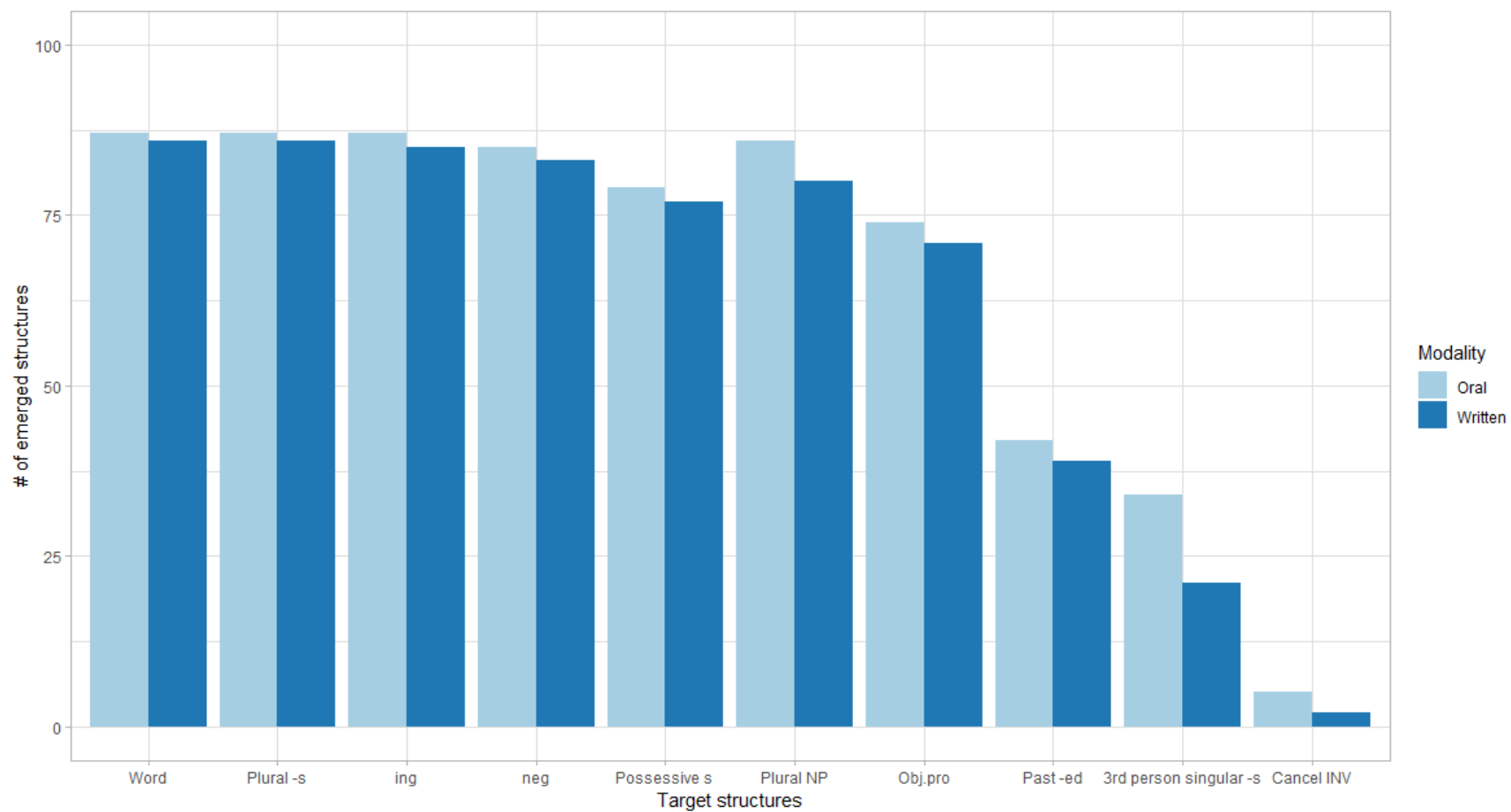


Figure 6. The number of emerged structures in the oral and written modalities

4.3. Accuracy

4.3.1. Oral responses

Following Lee and Spinner (2017), 75% is the criterion of accuracy for this analysis. When the accuracy of a specific structure is below .75, it is considered to indicate that a learner is not yet ready to produce the accurate form of the structure consistently. Table 9 illustrates the accuracy of the participants' oral production together with emergence. The gray cells indicate emerged structures. The number in each cell indicates the extent to which each structure was produced accurately throughout the oral interview, and the number is in bold if it is over .75. Note that empty cells and zeros represent different types of information: empty cells show that the participants did not produce any case of the particular structure (and thus, accuracy could not be calculated); a cell with a zero indicates that all of the participant's productions of the particular structure were inaccurate. In the analysis, only the accuracy of emerged responses was considered, for the reason that high accuracy in a limited number of cases cannot be generalized. For instance, an A2-level student (par89) produced Past *-ed* with a grammatically perfect form, but only once; thus, the participant's perfect accuracy of Past *-ed* may have occurred by chance.

Although the structures in earlier stages, such as plural *-s* and *-ing*, emerged in most of the participants' oral responses, the participants' accuracy with the structures varied. Overall, the participants who showed greater development produced more of the early stage structures more accurately. In other words, it seems that the participants' responses showed a gradually increasing ability to use these structures accurately. In addition, compared to the structures in the later stages, the participants produced the structures in the earlier stages more accurately. For instance, both S neg V (stage 2) and plural NP (stage 3) emerged in most of the participants' responses ($N = 85$ and $N = 86$, respectively). However, 69 participants (78%) produced S neg V

with greater accuracy (i.e., over .75), whereas only 31 participants (34%) reached the same criterion of accuracy in the production of Plural NP (34%). Table 11 presents the number and percentages of the emerged responses that fulfilled the criterion of accuracy.

Table 11. Accuracy in implicational scaling for the oral responses

Group	P.id	Stage / Structure							
		S2 neg	S2 -s	S2 -ing	S2 's	S3 Plural NP	S2 -ed	S5 -s	S6 Cancel INV
B1	par6	0.58	0.65	0.64	0.70	0.35	0.5	0.31	
A2	par69	0.70	0.57	0.09	0.40	0.33	0	0.00	
A2	par89	0.70	0.67	0.53	0.25	0.58	1	0.33	
A2	par28	0.80	0.75	0.15	0.90	0.48	0.67	0.29	
A2	par87	0.83	0.75	0.56	0.82	0.61	0.5	0.38	
A2	par71	0.40	0.70	0.33	0.70	0.75	0.25	0.70	
A2	par77	0.45	0.38	0.41	0.89	0.23	0	0.00	
A2	par83	0.67	0.83	0.38	0.92	0.78			
B1	par50	0.40	0.83	0.44	0.38	0.78		0.17	
B2	par1	0.67	0.74	0.63	0.54	0.67		0.00	
B2	par24	1.00	0.78	0.93	0.92	0.74	1	0.76	
A2	par85	0.80	0.75	0.87	0.91	0.57	0.8	0.55	
A2	par40	0.87	0.86	0.46	0.20	0.50	0	0.55	
B1	par36	0.81	0.56	0.86	0.25	0.43		0.25	1.00
A2	par66	0.83	0.56	0.49	0.68	0.48	0	0.23	
A2	par68	1.00	0.69	0.39	1.00	0.50		0.33	
A2	par75	0.81	0.64	0.27	0.69	0.57	0	0.20	
A2	par86	0.56	0.95	0.35	0.86	0.93	0.5	0.43	
A2	par92	0.80	0.61	0.42	0.67	0.59	0.67	0.00	
A2	par93	0.79	0.89	0.80	0.67	0.71	0	0.60	
B1	par22	1.00	0.80	0.65	0.56	0.62	0	0.38	
B1	par27	0.78	0.66	0.24	0.35	0.50	0.33	0.27	
B1	par29	0.94	0.81	0.53	0.78	0.65	1	0.33	
B1	par46	0.93	0.65	0.83	0.93	0.58	1	0.25	
B1	par55	0.75	0.89	0.85	0.89	0.78	1	0.50	
B1	par60	1.00	0.90	0.86	0.50	0.38	0.4	0.40	
B1	par61	0.60	0.78	0.44	0.81	0.71	0.4	0.43	
B2	par10	1.00	0.70	0.63	0.92	0.70	0	0.67	
B2	par12	0.93	0.59	0.68	0.57	0.73	0.17	0.67	1.00
B2	par15	0.86	0.81	0.70	0.81	0.64	0	0.68	

(cont'd)

B2	par17	1.00	0.71	0.72	0.67	0.79	1	0.60	
B2	par19	1.00	0.88	0.87	0.86	0.81	1	0.80	
B2	par21	0.88	0.97	0.34	0.91	0.87	0.91	0.30	
B2	par23	0.90	0.85	0.78	1.00	0.80	1	1.00	
B2	par4	1.00	0.74	0.56	0.83	0.58		0.36	
B2	par5	0.50	0.71	0.81	0.71	0.53	0	0.29	
B2	par9	1.00	0.56	0.57	0.91	0.52	0.33	0.58	
C1	par37	0.95	0.54	0.53	0.71	0.54	0	0.40	1.00
A2	par51	0.88	0.79	0.52	0.25	0.35	0.71	0.21	
A2	par90	1.00	0.85	0.76	0.90	0.71	0.83	0.50	
A2	par91	0.71	0.82	0.27	0.67	0.60	0.75		
B1	par57	1.00	0.92	0.84	0.33	0.88	0.90	0.67	
B1	par58	1.00	0.64	0.84	1.00	0.81	0.75	0.40	
B1	par38	0.89	0.88	0.62	0.91	0.33	0.25	0.60	
B1	par39	0.62	0.63	0.52	0.89	0.57	0.4	0.22	
B1	par42	0.70	0.80	0.31	0.50	0.58	0.36	0.55	0.50
B1	par47	0.58	0.74	0.51	0.86	0.68	0.60	0.00	
B1	par52	0.92	0.76	0.57	0.76	0.55	0.56	0.08	
B1	par64	1.00	0.79	0.51	0.75	0.68	0.80	0.22	
B2	par34	0.92	0.74	0.75	1.00	0.70	0.77	0.00	
B2	par8	0.93	0.73	0.82	0.77	0.82	0.33	0.33	1.00
C1	par18	0.95	0.81	0.93	0.88	0.67	0.89	0.70	
A2	par76	0.43	0.69	0.21	0.50	0.47		0.67	
B2	par14	1.00	0.90	0.84	0.75	0.79	1	0.69	0.75
C1	par41	0.93	0.73	0.79	1.00	0.74	1	0.71	
A2	par62	0.71	0.62	0.48	0.80	0.26	1	0.33	
B1	par33	0.85	0.75	0.68	0.64	0.33	0.67	0.71	
B2	par3	0.73	0.84	0.63	0.60	0.58	0	0.53	
C1	par53	0.92	0.87	0.91	1.00	0.81	0.75	0.86	1.00
B1	par54	0.88	0.87	0.56	0.91	0.74	0.60	0.64	
B1	par35	0.78	0.86	0.80	0.92	0.59	0.50	0.40	
B1	par48	1.00	0.74	0.81	1.00	0.71	0.71	0.67	
B2	par11	1.00	0.92	0.82	0.79	0.85	0.29	0.63	
B2	par13	0.93	0.70	0.81	0.78	0.75	0.67	0.81	
B2	par26	1.00	0.86	0.77	0.90	0.83	0.88	0.67	
B2	par30	1.00	0.82	0.94	0.88	0.95	1	0.80	
B2	par43	1.00	0.96	0.93	0.89	0.90	1	0.80	
B2	par45	1.00	0.93	0.98	1.00	0.96	1	0.85	
C1	par20	0.92	0.82	0.91	0.91	0.59	0.8	0.64	1.00
C1	par32	0.96	0.79	0.67	0.76	0.66	0.71	0.83	
C1	par49	0.85	0.70	0.86	1.00	0.63	0.33	0.85	1.00
C1	par7	1.00	0.83	0.90	0.91	0.61	0.88	0.85	

(cont'd)

C1	par72	0.85	0.85	0.83	0.91	0.50	0.40	0.64
C1	par73	1.00	0.86	0.84	1.00	0.88	1	0.63
C1	par74	1.00	0.88	0.89	0.88	0.86	1	0.93
C1	par78	0.89	0.86	0.95	0.92	0.90	0.90	1.00
C1	par79	1.00	0.86	0.98	0.89	0.79	1	0.96
C1	par80	0.78	0.96	0.81	0.70	0.84	0.85	0.63
C1	par81	0.86	0.89	0.87	1.00	0.93	0.93	0.89
C1	par82	1.00	0.89	0.92	1.00	0.93	1	0.88
C1	par84	1.00	0.94	0.77	0.86	0.90	1	0.83
C1	par88	1.00	0.93	0.91	1.00	0.79	1	0.50
A2	par67	0.91	0.68	0.25	0.92	0.79	1	0.53
C1	par65	0.93	0.82	0.77	0.86	0.64	0.77	0.50
C1	par70	0.94	0.87	0.77	0.93	0.82	0.46	0.86
C1	par25	1.00	0.80	0.89	0.89	0.67	0.90	0.77
C1	par31	0.95	0.56	0.70	0.93	0.70	0.38	0.57

Table 12. Summary of accuracy in the oral responses

Structure	# of participants who showed the emergence	# of greater accuracy (%)	Structure	# of participants who showed the emergence	# of greater accuracy (%)
S neg V	85	66 (.78)	Past -ed	42	25 (.60)
Plural -s	87	55 (.63)	Plural NP	86	29 (.34)
-ing	87	42 (.48)	3 rd person singular -s	35	16 (.46)
Possessive 's	78	57 (.73)	Cancel	5	2 (.40)
			INV		

4.3.2. Written responses

Similar to the results from the participants' oral responses, accuracy in the production of a particular structure in the written responses increased along with the participants' L2 development. The percentage of the emerged responses that met the criterion of accuracy was

relatively stable, regardless of the difficulty of the structures; in other words, the percentages were similar for structures in earlier and later stages (Table #).

Table 13. Accuracy in implicational scaling for the written responses

Group	P.id	Stage / structure							
		S2 neg	S2 -s	S2 -ing	S2 's	S3 Plural NP	S2 -ed	S5 -s	S6 Cancel INV
A2	par69	0.90	0.67	0.02	0.50	0.20	0.50	0.00	
A2	par89	0.38	0.50	0.63	0.00	0.38	1.00	0.00	
A2	par77	0.60	0.67	0.16	0.45	0.25	0.00	0.00	
A2	par76	0.90	1.00	0.11	0.25	0.73	0.00	0.00	
B1	par27	0.80	0.43	0.68	0.00	0.04	0.33	0.06	
A2	par28	0.78	0.46	0.38	1.00	0.38	0.00	0.19	
A2	par71	0.89	0.83	0.28	0.88	0.86	1.00	0.25	0.00
A2	par83	0.67	0.77	0.35	0.78	0.69	0.50	0.00	
A2	par85	0.69	0.67	0.91	0.83	0.63	0.50	0.67	
A2	par87	0.57	0.89	0.53	1.00	0.71	0.67	1.00	
A2	par67	0.83	0.90	0.69	0.88	0.82	0.33	0.50	
B1	par50	0.70	0.74	0.63	0.42	0.67	1.00	0.14	
B1	par6	0.45	0.67	0.53	0.50	0.45	1.00	0.56	
A2	par86	0.33	0.76	0.40	0.85	0.80	1.00	0.67	
B1	par42	0.94	0.86	0.20	0.63	0.50	0.71	0.67	
C1	par73	1.00	0.97	0.93	0.89	0.94	1.00	0.67	
A2	par51	0.89	0.78	0.83	0.71	0.44	0.67	0.10	
A2	par90	0.50	0.74	0.83	0.58	0.63	0.50	1.00	
A2	par93	0.73	0.82	0.84	1.00	0.92	0.40	1.00	
B1	par36	0.90	0.78	0.92	0.00	0.93	0.67	1.00	
A2	par40	0.67	0.88	0.81	0.93	0.47	0.67	0.63	
A2	par92	0.89	0.71	0.43	0.87	0.46	1.00	0.20	
A2	par68	1.00	0.94	0.26	1.00	0.70	0.00	0.17	
A2	par66	0.93	0.59	0.74	0.85	0.29	0.00	0.56	
A2	par75	0.92	0.93	0.20	1.00	0.63	0.00	0.00	1.00
B1	par48	1.00	0.97	1.00	1.00	0.86	0.50	0.33	
B1	par38	1.00	0.90	0.76	0.91	0.89	1.00	0.71	1.00
B1	par39	0.88	0.93	0.81	0.89	0.55	1.00	1.00	1.00
B1	par52	0.90	0.64	0.67	0.86	0.37	0.67	0.20	
B1	par60	0.71	0.50	0.76	0.74	0.43	0.50	0.33	

(cont'd)

B1	par61	0.73	0.86	0.80	0.67	0.67	1.00	0.60	
B1	par46	1.00	0.85	0.81	0.77	0.96	1.00	0.25	
B1	par22	1.00	0.88	0.83	1.00	0.77	0.50	0.55	1.00
B1	par55	0.78	1.00	0.80	0.72	0.89	1.00	0.33	
B2	par4	0.94	0.72	0.78	0.77	0.44	1.00	0.30	
B2	par9	0.89	0.93	0.74	0.93	0.59		0.71	
B2	par26	0.63	0.91	0.85	0.57	0.88	0.67	0.50	
B2	par10	1.00	0.83	0.45	0.86	0.88	1.00	0.50	1.00
B2	par11	1.00	0.94	0.90	1.00	1.00	1.00	1.00	1.00
B2	par15	1.00	0.78	0.84	0.67	0.61	0.50	0.58	0.33
B2	par17	0.83	0.80	0.80	1.00	0.94	1.00	0.33	
C1	par37	1.00	0.97	0.68	0.78	0.82	0.00	0.40	0.00
C1	par53	1.00	1.00	0.92	1.00	0.97	1.00	0.92	1.00
C1	par20	1.00	0.95	0.98	0.80	0.93	1.00	0.75	0.00
B2	par12	1.00	0.91	0.65	0.25	0.95	0.57	0.75	1.00
A2	par91	0.67	0.77	0.38	0.71	0.67	0.58	0.00	
A2	par62	0.69	0.79	0.73	0.94	0.69	1.00	0.60	
B1	par64	0.93	0.75	0.69	0.60	0.86	1.00	0.33	
B1	par57	1.00	0.91	0.81	1.00	0.97	1.00	0.84	
B1	par47	0.64	0.77	0.73	0.75	0.60	1.00	0.67	0.50
B1	par54	0.88	0.75	0.70	1.00	0.82	1.00	0.83	
B1	par29	0.90	0.92	0.88	0.92	0.73	0.75	0.53	0.33
B2	par1	0.88	0.71	0.69	0.27	0.63	0.57	0.45	
B2	par21	1.00	0.88	0.78	0.60	0.78	0.73	0.14	
B2	par34	1.00	0.96	0.74	0.92	1.00	0.92	0.89	
B2	par30	0.90	0.91	0.96	1.00	1.00	1.00	1.00	
B2	par45	1.00	0.93	0.87	1.00	0.87	1.00	1.00	
B2	par23	0.93	0.94	0.98	1.00	0.92	1.00	1.00	
B2	par24	0.92	0.67	0.95	0.67	0.41	0.83	1.00	
C1	par25	1.00	0.95	0.99	1.00	0.94	1.00	0.70	
C1	par32	1.00	0.95	0.69	1.00	0.77	0.71	1.00	0.00
C1	par41	1.00	0.85	1.00	0.80	0.85	1.00	1.00	1.00
C1	par88	1.00	0.96	0.95	1.00	0.92	0.90	0.78	
C1	par74	1.00	0.98	0.88	0.90	0.87	0.82	0.80	
B2	par3	1.00	0.89	0.83	0.83	0.86	1.00	0.54	1.00
B1	par33	1.00	0.96	0.83	0.89	0.88	0.75	0.63	1.00
B2	par5	0.56	0.72	0.81	0.63	0.80	1.00	0.82	
C1	par65	1.00	0.96	0.95	0.92	0.95	1.00	0.86	
C1	par72	1.00	1.00	1.00	1.00	1.00	1.00	0.89	
C1	par84	0.89	0.75	0.72	1.00	0.87	1.00	0.71	
B1	par58	0.89	0.71	0.94	0.71	0.83	0.67	0.86	
B1	par35	0.89	0.87	0.77	0.71	0.76	0.60	0.71	

(cont'd)

B2	par43	0.73	0.83	0.90	1.00	0.89	1.00	1.00	
B2	par13	1.00	0.95	0.93	0.67	0.88	0.57	0.60	0.00
B2	par14	1.00	0.94	0.88	0.87	0.96	1.00	1.00	1.00
C1	par18	0.88	0.88	0.91	1.00	0.88	0.71	0.75	
C1	par80	1.00	0.83	0.88	1.00	1.00	0.87	0.79	
C1	par81	1.00	1.00	0.94	1.00	0.92	1.00	1.00	
C1	par82	1.00	0.97	0.94	0.93	0.90	1.00	0.82	
C1	par78	1.00	0.97	0.99	0.94	1.00	0.91	0.90	
C1	par79	1.00	1.00	0.97	1.00	0.97	0.94	0.93	
C1	par7	1.00	0.98	0.99	1.00	0.90	1.00	0.94	
C1	par70	1.00	0.95	0.98	0.81	1.00	0.78	0.83	1.00
C1	par49	1.00	0.94	1.00	0.93	0.92	1.00	1.00	
B2	par8	1.00	0.83	0.90	0.92	0.77	1.00	1.00	0.00
C1	par31	1.00	0.81	0.81	1.00	0.85	0.75	1.00	0.67

Table 14. Summary of accuracy in the written responses

Structure	# of participants who showed the emergence	# of greater accuracy (%)	Structure	# of participants who showed the emergence	# of greater accuracy (%)
S neg V	82	67 (.82)	Past -ed	39	30 (.77)
Plural -s	86	68 (.79)	Plural NP	80	54 (.68)
-ing	85	54 (.64)	3 rd person singular -s	21	17 (.81)
Possessive 's	77	58 (.75)	Cancel INV	2	0 (0)

4.3.3. Comparison between oral and written responses

Given that the number of emerged responses is limited in the case of structures that appear on the right side of the table, such as Past -ed, 3rd person singular -s, and Cancel INV, only the accuracy of the emerged responses of the first five structures was considered in the analysis conducted to explore differences in general trends of accuracy between the oral modality and the written modality. Although accuracy of some structures was included in the

analysis, the statistical analysis can reveal the extent to which changes in accuracy and overall accuracy differ between the oral modality and the written modality.

In a linear mixed model, the dependent variable was accuracy of the participants' responses; modality (oral, written) and structure (neg, -s, -ing, 's, Plural NP) were fixed effects; participant was a random effect. The oral modality and Plural NP served as the reference group. Depending on type of structure, accuracy differed. The difference in accuracy of each structure seemed aligned with the order of the designated columns for the structures in the implicational table. In particular, Plural NP which appears in the rightmost of the columns for the five structures in the implicational table among the five structures revealed lower accuracy than S neg V ($\beta = .15$, $SE = .01$, $t = 10.35$, $p < .0001$), -s ($\beta = .09$, $SE = .01$, $t = 5.89$, $p < .0001$), and 's ($\beta = .10$, $SE = .015$, $t = 6.93$, $p < .0001$). However, there was no significant difference in the accuracy of Plural NP and that of -ing ($\beta = -.01$, $SE = .01$, $t = -.83$, $p = .41$). More importantly, the participants produced the target structures more accurately in the written mode ($\beta = .07$, $SE = .009$, $t = 7.08$, $p < .0001$).

CHAPTER 5. DISCUSSION

5.1. Overview of research questions and results

This dissertation primarily aims to investigate whether L2 development in L2 writing can be assessed using Processability Theory; in other words, whether language development in an L2 as predicted by PT is comparable between oral and written modalities. In addition, this study explores whether sequences of development can be observed within a particular stage in each of these two modalities, and whether changes in the accuracy of L2 morphosyntactic structures produced by learners occur predictably in both oral and written modalities.

To address the research questions, this study created two different implicational scales, one for L2 learners' oral production and one for L2 learners' written production. It applied the scales to the participants' task responses and compared the results. The implicational tables that were created illustrated whether each target structure emerged in each participant's responses. The order of the target structures in the tables was determined by considering the ratio of emergence to non-emergence. Hence, structures with lower occurrence of emergence appear to the right while structures with higher occurrence of emergence appear to the left. The resulting tables clearly reveal that structures in the earlier stages according to PT were likely to be produced earlier than structures in the later stages, and that learners become more and more able to produce structures in later stages as their L2 develops. The developmental patterns were comparable between the oral and written modalities, except for a few negligible errors (i.e., non-emerged structure(s) on the left of some emerged structures). In addition, this study found that structures in a particular stage (stage 2, stage 3) did not emerge simultaneously. Rather, each of the structures in a stage emerged at different time; in other words, there is an order of emergence

within stages. This intrastage development was observed in both oral and written modalities, and the developmental pattern was similar between the two modalities.

Next, the study analyzed the accuracy of each structure, exploring the extent to which emerged structures were produced accurately in oral and written modalities and whether accuracy differed depending on the type of structures and the learners' level of development. Accuracy seemed to be aligned with emergence: Structures that emerged earlier were more likely to be produced with higher accuracy than structures that emerged later. In addition, particular structures were not necessarily produced with accuracy when they first emerged; the learners showed a gradual increase in accuracy as their L2 development progressed. This gradual development in accuracy was found in both oral and written modalities. However, the number of emerged structures that met the criterion of accuracy was greater in the written modality; in other words, accuracy was higher and it increased more rapidly in the written modality than in the oral modality.

In the following sections, the results are discussed in terms of how they answer each of the three research questions and in light of previous studies on oral and written modalities and PT.

5.2. Research Question 1

The study's results support the possibility of using PT to investigate L2 language development in writing. They show that PT can be extended to L2 learners' written production and that doing so will lead to better understanding of L2 language development. In particular, morphosyntactic structures emerged in the written modality as well as the oral modality in a predicted way that PT posited. In other words, this study revealed that the developmental

patterns in the written modality were comparable to those in the oral modality. In addition, the results contribute to our understanding of language development itself in writing. Most of the previous studies which investigated L2 development in writing were not based on a theory of language development. Rather, they explored learners' changes in specific elements of L2 writing, such as knowledge of genres, goal setting, and the writing process (Cumming, 2012; Sasaki, 2004; Tardy, 2012).

This study sought to address some limitations in the few previous studies of L2 language development in writing on the basis of one of the theoretical frameworks of language development, and its results confirmed that the theoretical framework, PT, can be applicable to L2 writing, aligned with the results of these studies (Boss, 2008; Byrnes & Sinicrope, 2008). Overall, the previous studies left open the possibility of extending one of the theoretical frameworks of L2 language development, which is based on L2 oral production, to L2 written production, although they also pointed out some inconsistencies with the predictions of the frameworks, which may be due to certain factors: (1) the number of tasks designed to elicit particular structures (i.e., whether the number of a particular task type is enough), (2) task type (i.e., whether a task is communicative), and (3) original purpose of the task (i.e., whether a task is purposely designed to elicit a particular structure). In addition, these previous studies conducted writing tasks only when they sought to apply a certain theoretical framework to written performance; and therefore they did not address whether language development in L2 is comparable between oral and written modalities.

In this regard, to complement the aforementioned studies (e.g., Boss, 2008), and to respond to a call for more diverse research on L2 language development, this study investigated L2 learners' language development on the basis of PT with the use of comparable tasks in oral

and written modalities. Different from previous studies, this study designed a series of communicative tasks that each targeted one of the morphosyntactic structures predicted by PT. The tasks were intended to require the participants to produce the structure in their response. Identical task types were used in both modalities, although some tasks were slightly modified to fit each modality. With this methodological improvement, this study revealed similar developmental patterns between the oral modality and the written modality, and shed new light on L2 language development in the written modality. This study's application of PT's theoretical framework to both oral and written modalities is an important leap in the research on language development as well as PT.

The observed developmental patterns are comparable between oral and written modalities. Although each modality has unique characteristics (e.g., Gilabert, Manchón, & Vasylets, 2016; Williams, 2012), this study demonstrates that a single framework can be used to assess L2 learners' language development in their written production as well as their spoken production. While some differences between oral and written modalities were observed, the overlap between the two modalities showed comparable developmental patterns. Specifically, the participants revised and edited their responses more freely in the writing tasks than in the speaking tasks (Byrnes & Manchón, 2014; Granfeldt, 2007). Although the participants also corrected some of their responses in the speaking tasks, this was limited to when they noticed having made a grammatical error right after uttering it. In this case, the learners usually stuttered, such as *the man ride riding is riding*. Considering this distinction between modalities in regard to revision, it may be often assumed that L2 learners avoid producing complex structures in speaking, although they would use the structures in writing. However, in this study, the participants did not appear to avoid producing more complex structures in the speaking tasks, nor

to attempt to produce more complex structures (over the criterion of emergence) in the writing task. Thus, some differences between oral and written modalities did not affect the overall developmental patterns.

This study's demonstration of the possibility of using PT to investigate L2 learners' development in the written modality also supports the steadiness hypothesis that Pienemann (1998) proposed. He originally demonstrated that L2 learners are able to process morphosyntactic structures consistently regardless of type of communicative task, if the tasks are designed to use the same knowledge in production. The current study suggested that the learners' consistency in oral production that required particular knowledge might extend to their written production. It is assumed in this study that the best way to observe parallels between oral and written production is to use tasks that encourage the use of similar types of knowledge and skills. However, in writing, L2 learners may enable to deploy explicit knowledge and metalinguistic L2 knowledge that they may not be able to deploy in speaking (Ellis, 2003; Schoonen et al., 2009; Wolff, 2000). Specifically, L2 learners may draw on their explicit knowledge to revise their writing (Williams, 2012). Therefore, this study employs time constraints to control the learners' possible access to L2 explicit knowledge in writing and to lead them to rely more on L2 implicit knowledge to complete both tasks. The time constraints facilitate spontaneous written production, which reduces the possibility of using explicit knowledge.

Based on its use of comparable tasks and time constraints, this study encouraged L2 learners to employ in both modalities, in keeping with the steadiness hypothesis, which I argue can now be extended to certain writing tasks. Whether the developmental order of PT is found in other types of writing tasks, such as essays with multiple revisions, has yet to be investigated. The developmental order that PT predicts may be reflected in writing as well as speaking.

Håkansson and Norrby (2007) also examined the applicability of PT to L2 learners' written production. They found that advanced L2 learners followed the developmental order predicted by PT in their writing development. However, the current study's findings provide clearer evidence, for a few reasons. First, Håkansson and Norrby tested only advanced learners, while this study included L2 learners with a wide range of proficiency levels, and found that they followed the developmental order predicted by PT in their writing across proficiency levels. This finding means that the applicability of PT's prediction to L2 written production can be more widely generalized. In addition, following the previous PT studies that conducted various communicative tasks in the oral modality, this study utilized similar communicative oral and writing tasks, rather than using translation tasks to assess writing. To complete the written tasks, the learners had to come up with their own ideas and produce morphosyntactic structures by themselves within the designated time, just as they did in the oral tasks. Thus, in contrast to the previous study, this study used communicative tasks that are comparable between the two modalities. The results demonstrate that L2 learners were able to produce morphosyntactic structures following PT's predicted stages consistently, even in the more complex writing tasks. In sum, based on these findings, Pienemann's steadiness hypothesis can be successfully extended to L2 learners' written production at least with certain tasks.

5.3. Research Question 2

PT posits that reaching a particular stage indicates L2 learners' acquisition of a required processing procedure, and that such acquisition means the learners are able to use the procedure to produce structures at that stage. This study's findings revealed that L2 learners may not be able to produce all the possible structures in a specific developmental stage simultaneously. The

results are aligned with those of Di Biase et al (2015). Their longitudinal study tracked a young L2 learner's progress from right after she arrived in Australia for two years. The morphology in stage 2 emerged as her English developed; however, not all the structures of stage 2 emerged at the same time. Rather, Di Biase et al. reported a particular order of emergence within stage 2.

In line with Di Biase et al. (2015), this study also found an internal hierarchy within some developmental stages, although only the earlier ones. In addition, this study noted the same hierarchies between oral and written modalities. The number of emerged structures differed depending on the type of structure in stage 2 and stage 3. The implication tables showed lower scalability of Past *-ed* than other structures in stage 2, confirming that it emerged much later but also less predictably. Because this study did not analyze the emergence of all the morphological structures in stage 3, such as fronted adverbs (e.g., *Today, he goes to his school*), I cannot fully determine the internal developmental order in stage 3; however, the results show a clear pattern of emergence between Plural NP and Object pronoun. Future research may be able to show intrastage orders in later stages as well.

Specifically, in contrast to the original prediction, Past *-ed* did not emerge at the same as the other morphosyntactic structures in stage 2, but much later. However, this may be due to specific features of the past tense forms in English. Past *-ed* was elicited by two narrative tasks in which L2 learners described a past experience in each modality, and to complete the tasks, the learners relied heavily on just a few irregular verbs, such as *went* (e.g., *I went to X*), *ate* (e.g., *I ate some delicious food with my friends*), and *had* (e.g., *I had a good time*), largely avoiding the regular verbs that would employ Past *-ed*. Unlike other morphological structures in this stage, such as *-ing* and Possessive *'s*, Past *-ed* competes with its opponent, the irregular forms. Thus, it is possible that the high use of common irregular verbs reduced the number of chances to

produce Past *-ed*, which resulted in the unexpected findings. Several of the irregular verbs have high perceptual saliency and occur frequently (Klein et al., 1995; Salaberry, 2000), which may make it more difficult for L2 learners to process and produce Past *-ed*, which, in turn, may encourage them to use irregular verbs more often.

The overuse of irregular verbs may also be due to the time constraints, which were set to encourage the learners to respond to the questions spontaneously. Time constraints can, however, have unintended effects, as when learners fixate on the time limit and try to produce as many sentences as possible as quickly as possible. In consequence, they may forgo complex vocabulary and detailed content, instead choosing simple structures, such as irregular *was* with an adjective (e.g., *It was funny and interesting*) or *did* with a noun instead of a more specific verb (e.g., *I did something fun with my friends last Christmas*).

5.4. Research Question 3

The study's results showed that L2 learners were able to produce morphosyntactic structures more accurately along with their further L2 development, as shown in the emergence of structures, and that their accuracy with structures that appeared earlier in implicational scaling was higher than their accuracy with structures that appeared later. These results were similar between the two modalities. In other words, accuracy increased in a similar way in the learners' oral and written responses. However, the L2 learners reached the criterion of accuracy much earlier in their written production than in their oral production, and their accuracy in the written production was more stable than in the oral production.

These results indicate that while the different characteristics of the oral and written modalities may not have changed the overall pattern of emergence (Research Question 1), they

may lead to differences in accuracy. The written modality both allows and encourages L2 learners to use metalinguistic knowledge to monitor their language and accuracy (Williams, 2012). As Schoonen et al. (2009) argued, spoken language is more tolerant of errors than written language. Due to the different expected tolerance for each modality, L2 learners may monitor their language more thoroughly in written tasks than in oral tasks. In addition, in writing, learners have additional opportunities to plan, check, and edit grammatical errors in their production. Even with time constraints, written tasks may be more plannable than spontaneous speaking tasks (Ochs, 1979, p. 58). The learners could use their limited time to plan, translate, and monitor their writing (Kellogg, 1996). Hence, the unique characteristics of the written modality may have led this study's participants to pay more attention to the structures they produced and to correct more grammatical errors in the written tasks than in the oral tasks.

Although previous TBLT studies have defined and measured accuracy differently than the current study, using measures such as error per T-unit and number of errors (also see, Polio & Shea, 2014), this study's findings are aligned with some TBLT studies that found differences in accuracy between oral and written production (Kormos, 2014; Kormos & Trebits, 2009; Tavakoli, 2014; Zalbidea, 2017) and provide more evidence for the effects of the written modality's characteristics on accuracy. As Zalbidea (2017) pointed out, the writing process is recursive, which allows L2 learners to monitor, evaluate, and edit linguistic forms in their writing, all of which can lead to increased accuracy.

CHAPTER 6. CONCLUSION

6.1. Summary and implication

This study sheds light on the possibility of investigating L2 language development in the oral modality and the written modality on the basis of PT. Because theories of language development have mainly dealt with oral production, it is often unclear whether or to what extent they apply to written production. This study used similar tasks in oral and written modalities and compared the results to provide a better understanding of the applicability of a language development theory, PT, to writing. Extending PT to L2 learners' written production, this study suggests that L2 linguistic knowledge may be developed in similar ways, regardless of modality and that theory-based investigations of language development may be possible in the written modality as well as in the oral modality.

In addition, the study contributes to fine-tuning the PT's original developmental order by showing some evidence of an order of development within a particular stage thus strengthening this theoretical framework (also see Di Biase et al., 2015). In particular, this study revealed that each of possible structures in stage 2 and 3 emerged at different times; that developmental order may be possibly found within a particular stage.

Finally, the study also assessed whether accuracy is aligned with the emergence of morphosyntactic structures to gain a clearer picture of L2 learners' language development. The results demonstrated a gradual progress in accuracy, in which the pattern of the progress seemed similar to that of emergence. However, the written modality exhibited a more consistent progression, which may have been influenced by opportunities to revise and edit in writing.

From a pedagogical standpoint, increased understanding of learners' language development in oral and written modalities can help teachers grasp their learners' developmental patterns and think of appropriate treatments. In addition to knowing the general progress predicted by PT, greater awareness of the internal hierarchies within stages and how accuracy aligns with emergence would give teachers more insight into their learners' development and variability in development. Because teachers need to assess their learners' progress frequently and often quickly, opportunities to appropriately reflect their learners' current status in their teaching and materials are useful. Moreover, it is not always necessary to conduct an oral task to test learners' progress; rather, teachers can decide on task modality depending on the content of their courses. For instance, in a business writing class, written tasks to diagnose students' linguistic knowledge are more in keeping with the course goals and more relevant to the course content.

6.2.Limitation and future research

This study has some limitations, which call for further research on L2 language development in both oral and written modalities from the perspective of PT. First, this study set arbitrary time constraints for both task types. The time constraints encouraged the learners to respond to the questions spontaneously. The time constraints for oral tasks were decided in light of the previous PT studies. However, no previous PT study has provided a principled way of determining appropriate time constraints for written tasks; for each written task, this study doubled the time for the task's oral counterpart, considering the different speeds of typing and speaking. It is possible that the different time constraints between the modalities led to some differences in degrees of spontaneity in the learners' production. Further research should

investigate the issue of time constraints and explore the optimal time constraints for eliciting learners' spontaneous production in both modalities. In addition, it should be explored whether PT can be extended to L2 writing without time constraints.

Next, this study did not analyze L2 learners' online writing behaviors, but doing so may provide concrete evidence for the effects of the different characteristics of the two modalities on accuracy. This study assumed that the unique characteristics of writing affected accuracy. However, measuring online writing behaviors with the use of keystroke logging (e.g., Leijten & Van Waes, 2013), stimulated recall task (e.g., De Silva & Graham, 2015), or a combination of these (e.g., Révész, Kourтали, & Mazgutova, 2017) would provide information on whether and how L2 learners actually do revise and edit their grammatical errors while writing, and whether such behaviors do lead to higher accuracy in their writing than in their speaking.

In addition, L2 learners' keyboarding skills (speed of typing) may have influenced the results. This study asked participants to type their written task responses. In the implicational tables and the accuracy tables, more proficient learners tended to be positioned in later columns (showing more emergence of forms and higher accuracy), whereas most of the lower proficiency learners were positioned in earlier columns (showing less emergence of forms and lower accuracy). These results were taken to indicate that more and more complex structures emerged and accuracy increased as L2 learners' linguistic knowledge was developed. However, differences in participants' keyboarding skills could have a mediation effect on the results (Barkaoui & Knouzi, 2018).

Finally, this study followed Lee and Spinner (2017) in setting 75% as the criterion of accuracy. Because previous PT studies did not explore the accuracy and emergence of morphosyntactic structures together, there is little evidence for whether this criterion is valid.

Thus, the validity of the criterion of accuracy is another issue for future research to address.

Future studies expanding this study are needed to provide additional evidence for validating the criterion of accuracy.

APPENDICES

Appendix A: Speaking and writing tasks designed to elicit S neg V

1. Speaking tasks

a. Making an announcement

Suppose that you are a manager of a movie theater. Before starting a movie, you would like to make an announcement. Based on the following pictures, please inform audiences of what they are allowed or not allowed to do during the movie.

Example of the pictures:



Figure 7. Example of the pictures used in Making an announcement A

Example of the responses:

A2: Please your phone do not use phone in movie theater.

B1: *Don`'t recording video in movie theater.*

B2: *You can`'t take photos of the movies*

C1: *You are not supposed to take picture or video clip of the movie that is being on screen because it`s against the rule of copyright.*

b. Making an announcement

Here`s a list of rules that makes riding an elevator safe for everyone. Suppose that you are a teacher of an elementary school. Based on the following pictures, please explain it in detail to the students.

Example of the pictures:



Figure 8. Example of the pictures used in Making an announcement task B

Example of the responses:

A2: *If elevator is fire, do not ride elevator.*

B1: *When fire, do not ride elevator and use stairs.*

B2: *If fire is emerging elevator, you do not ride in elevator and do not touch any button.*

C1: *You should not take the elevator when it is on fire.*

2. Writing tasks

a. Creating a flyer

Suppose that you are a manager of the dormitory at ABC university. Before starting a new semester, you would like to put flyers on bulletin board on each floor. Based on the following pictures, please complete the flyer with some sentences to inform students of what they are encouraged to do and what they don't/can't do.

Example of the pictures:



Figure 9. Example of the pictures used in Creating a flyer task

Example of the responses:

A2: *Don't somking*

B1: *Please don't smoke.*

B2: *Don't smoke indoor.*

C1: *Please don't smoke in the dormitory building.*

b. Making a guideline

Here's the instructions that come with a box of contact lenses. Based on the following pictures, please complete to write the guidelines to inform customers of dos and don'ts to wear contact lens appropriately.

Example of the pictures:



Figure 10. Example of the pictures used in Making a guidance task

Example of the responses:

A2: Do not wear the contactt lens at the swimming pool.

B1: You never swim when the lense are in your eye.

B2: Please never do swimming wearing contact lens.

C1: Do not swim wearing your contact lenses.

Appendix B: Speaking and writing tasks designed to elicit Plural -s, Plural NP, and Progressive -
ing.

1. Speaking tasks

a. Picture description task A

Please describe the following picture.

- *What are they?*
- *What are they doing?*

b. Picture description task B

Example of the pictures:



Figure 11. Example of the pictures used in picture description task in the oral modality

Example of the responses:

A2: They are playing the basketball. The man holding the ball. He dribbling. Is back, his number is #5. The man is running. His same team holding a ball. His number is 3. Their's opponent.

B1: They are playing basketball. Many crowds watch the playing. There is two team which one is wearing orange uniform and which one is wearing white uniform. Some man has the ball going to the goal. Two men following the man who has the ball.

B2: There are three main athletics who is playing basketball. Two of them are wearing white uniforms. One player is wearing orange uniform. One of the white uniform performers is keeping the ball. He is running very hardly.

C1: There are four basketball players on the picture, #3, 5 and 15 that I can see. One player is holding a ball. He is running through the goal post. There are lots of watchers. The other player is following the player who is holding a ball. Three players are wearing the same uniform, which means they are one team. The other man is wearing red uniform, which means he is the other team. They are all wearing basketball sneakers.

2. Writing tasks

a. Picture description task A

Please describe the following picture.

- *What are they?*
- *What are they doing?*

b. Picture description task B

Example of the pictures:



Figure 12. Example of the pictures used in picture description task in the written modality

Example of the responses:

A2: White team and Yellow team are playing soccer. At the moment, Yellow woman steal white woman's ball.

B1: There are women who are playing soccer game in the ground. Their uniforms have another color. One is yellow, and another is white. White uniform woman is dribbling soccer ball.

B2: There are two women in the picture. They are playing soccer against each other. A woman wearing white uniform is going to shoot the ball to score and the other woman wearing yellow uniform is trying to stop her by tackling. And I can see three more players in the picture. The game is played on the grass.

C1: Some women are playing soccer. Two players are fighting over a ball. All the players are wearing soccer uniforms, each yellow and white. The woman with white uniform, number 12, is stealing a ball from another woman in yellow. They are all blonde and white. They are playing soccer on the grass. In the picture, there are five players in total. Other three are just watching in the back. They are all wearing black sneakers.

Appendix C: Speaking and writing tasks designed to elicit Possessive 's

1. Speaking tasks

a. Picture description task

Please describe the following object.

- *Whose is it?*
- *What color it is?*

Example of the pictures:



Figure 13. Example of the pictures used in picture description task

Example of the responses:

A2: *It's a Summer's shoes.*

B1: *This is Summer's shoes.*

B2: *The toe shoes is Summer's one.*

C1: The toe shoes are Summer's shoes.

b. Giving directions via phone-call

Based on the following map, please describe the direction to X's house.

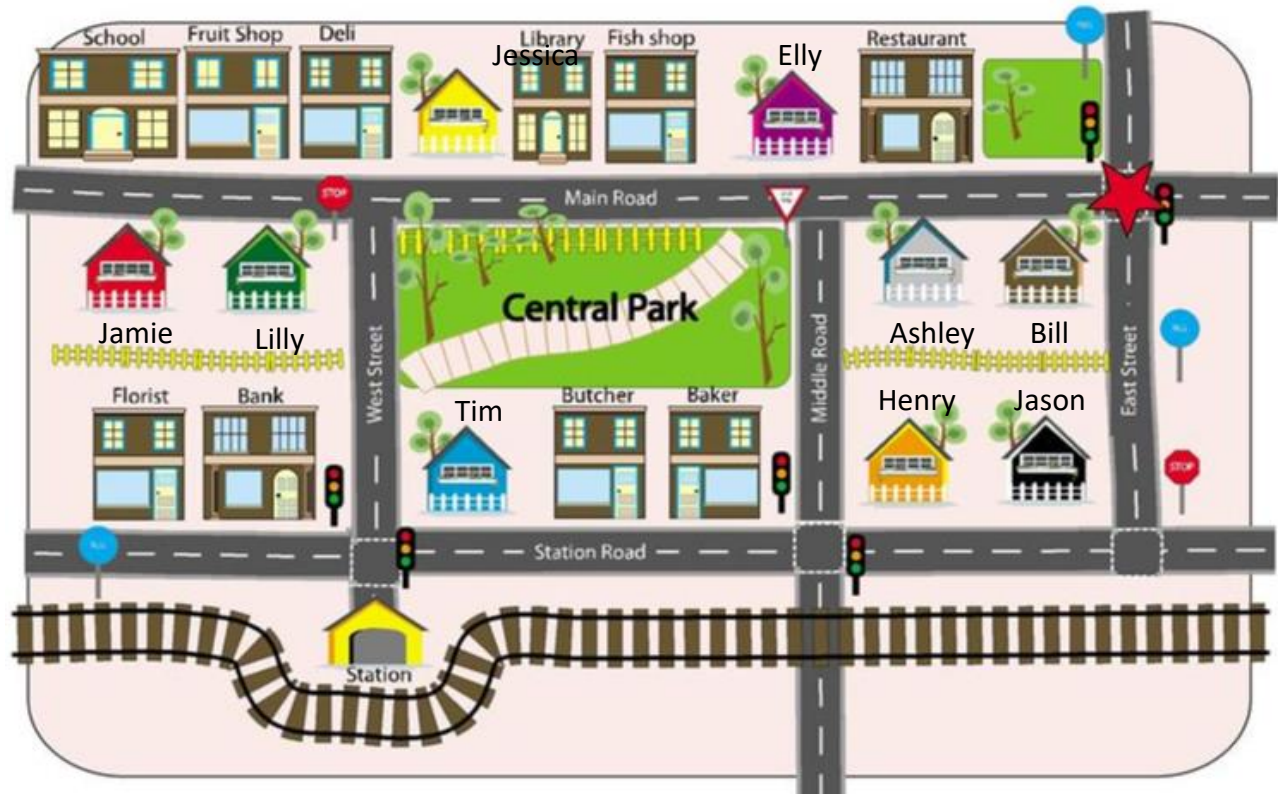


Figure 14. Picture used in Giving directions via phone-call

Example of the responses:

A2: Across the deli shop is Lily's home.

B1: If you watch blue house, that's the Tim's house.

B2: Across from the bank, you can see the Tim's house.

C1: Across the deli shop, there is Lily's house.

2. Writing tasks

a. Picture description task

Please describe the following objects.

- *Whose is it?*
- *What color is it?*

Example of the pictures:

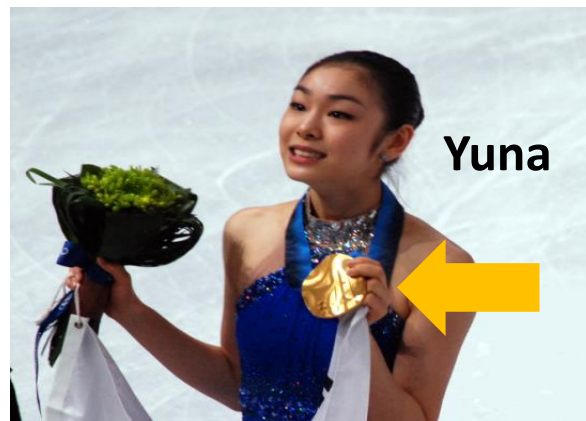


Figure 15. Example of the pictures used in picture description task

Example of the responses:

A2: the object is gold madal and gold madal is yuna kim's

B1: It is gold medal. Yuna Kim is ther owner of the medal.

B2: Yuna Kim won her gold medal. Yuna Kim's gold medal.

C1: This gold medal is Yuna Kim's.

b. Giving directions via online chatting

Based on the following map, please describe the direction to Y's desk.

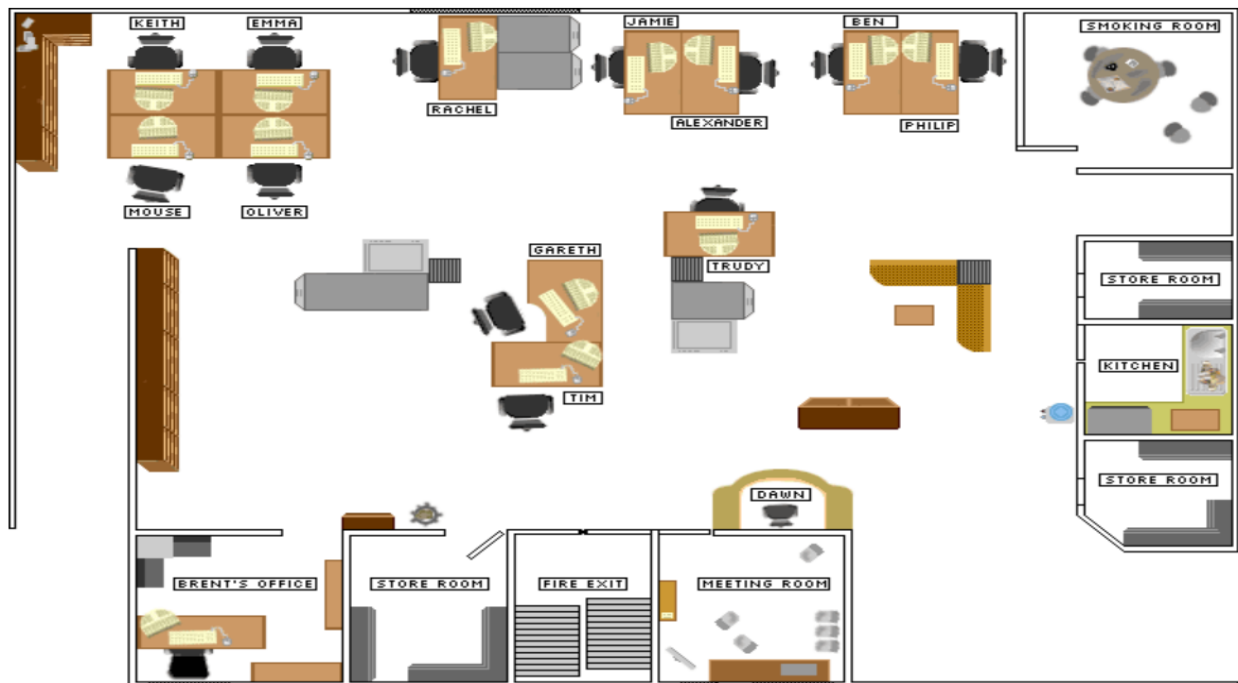


Figure 16. Picture used in Giving direction via online chatting task

Example of the responses:

A2: Go strength and turn right and . OLIVER desk is infront of EMMA desk.

B1: Turn right and go straight at the smoking room. And turn left. Go straight at the Oliver desk.
And turn right. Go and see Emma in left desk.

B2: first of all, turn right and go straight. and then turn left and turn right. You can see EMMA`s desk. In front of her desk , there is a Oliver`s desk.

C1: When you got in the office, just turn right to the edge. Then you turn left and go straight until you get to a table with four chairs. You will find Emma's place on the right side. In front of Emma's desk, there is Oliver's desk.

Appendix D: Speaking and writing tasks designed to elicit Past -ed

1. Speaking tasks

a. Narrative task A

Please describe your best winter holiday in your life.

- *When was it?*
- *What did you do?*
- *Why was it your best winter holiday?*

Example of the responses:

A2: *During last winter vacation, I went to Europe. I went to England first. I go Italy. When I go XX, its sea is very beautiful. Yesterday, I see that too. Walking and walking.*

B1: *Last Christmas, I was dating with my boyfriend. We are go to XX. It was very impressive that we watching firework. There is many lamps. It was very romantic. It was very cold, but I feel very warm heart. We are taking many play and walking and just talking each other.*

B2: *I can't remember fantastic day, but last year's Christmas day, I was in my house and eat chicken with my family. I watched some movie through the television. It was freezing outside. After taking a shower, I went in through the warm room and read some books. It was not very special but it was comfortable experience for me.*

C1: *My best winter vacation was in 2017. I was in Korea and in December 31st, and then we went to Hawaii. We actually had two new year's. When we went to Hawaii, it was really warm.*

That's enough for me to make me a best winter vacation, because it was very warm. It was family trip. We went to two islands, XX and XX. We went to the beach. We also drove some places in XX. We watched a lot of stars. It was my first time to do surfing, although I failed. I really enjoyed it.

b. Narrative task B

Please reflect back on your first semester at XXX university.

- *How did you feel about the new culture, environment, people, etc. back then?*

Example of the responses:

A2: *When I go first my XX university, I know the time going the school. The time is so long, so I very tired. When I come the school, I very tired. I cannot more activity, so I regret that. There is many friends. We just studying and play the game and go to XX.*

B1: *When I was a freshman, I feel a little bit nervous, because I think university and middle school is a different world. When I went to university first, I feel better than my expectation, because I can meet a lot of people and I can do a lot of project. First, I can learn my major what I want. I think it is really helpful for me. I can meet various people. Someone is man and someone is woman and someone is old and someone is young. I can communicate with them. I went to a lot of another place with them, such as downtown and beach.*

B2: *Before I enter to university, I expected that in university, a lot of people are very smart and very adulthood. It's my expectation. Second expectation is I can make some lover. Also, I expected that I can make a lot of friends. I think it was ideal one. In real situation, a lot of people*

in university is not very adulthood people. They are very different from high school students. They are not very good and handsome man I want to love. In university, I can do more free. I can have a lot of private time. I can't get much time with university friends. It was very hard to be familiar with new friends. It was hard for me.

C1: When I entered the university, at first, it was really really good. Everyday was really happy, because I can do whatever I want, like drinking alcohol or having some tea time with my friends in cafe and eating what I want in the good restaurants. There were so many people, so everyday was so happy. There is a bad memory in my first semester of my university. I wanted to enter the club, called XX, so I registered and also I did interview. I entered the club, but that club was so tight. I have to do some homework. There are so many things that I had to do. At that time, I wanted to having some time with my friends. I didn't want to do that homework and I didn't want to go MT with them. Finally, one month later, I have to quit the club. I had bad relationship with them.

2. Writing tasks

a. Narrative task A

Please describe the most memorable event in your life.

- *When was it?*
- *What did you do?*
- *What was it your most memorable event?*

Example of the responses:

A2: At 2015.6.23, I went to military with my mother, older brother, girl friend. But i don't recognized that military well. So, I slept in the

B1: My most memorable experince is going to Jeju island with my best freind. because it was my first to go island. I went sea and the color was very impressive. and I feel ver excited because weather was always good when i was in there.

B2: Last summer, I went to Yeosu alone. It was my first alone-trip. I was so exhausted, so I have to refresh. In Yeosu, I rode cable-car and I can see beach upside. Also, I walked down the empty and calm street, so I felt comfortable.

C1: The most memorable event in my life happened when I was six years old. I was by myself right in front of the door of my apartment. I suddenly realized I did not bring the key. But I wanted to go to the toilet so badly. Since I did not have any place else to go, I just did it on the stairs. It remains as an unforgettable memory because it was so embarrassing.

b. Narrative task B

Please describe your summer vacation.

- What did you do?
- Where did you go?

Example of the responses:

A2: i went to busan with my friends and we went to heawoondae and we swimming in the sea and ate chicken and drunk and we run on the beach.

B1: Last summer vacation I went to seoul for 1month. I stay my friend's room. I was visiting many famous spot in seoul. It was good to stay another city. I went many delicious res

B2: I had summer vacation last summer with my soccer club members. We went to Busan to enjoy summer. we had great food and did some sightseeing. Also we went clubbing at night , having fun. we went to seafood market to have some fresh seafood. we had ocpotus and so on.

C1: I went to Ulsan with my best friend. It was my first to go there even though Ulsan is not so far from Daegu. We visited Ganjolgot and had some nice lunch. We stayed at a private house which is located near the ocean. It was really great but the weather was so bad. Sadly, we couldn't see any sunset or sunries. It was a shame that we couldn't enjoy sunset or sunrise but

Appendix E: Speaking and writing tasks designed to elicit questions and/or Cancel INV

1. Speaking tasks

a. Response to a voice message

Suppose that you found the following flyer of volunteer works and a registration desk of it in the hallway of XXX university's building. You are interested in it so that you would like to ask the representative to provide more detailed information. What questions do you want to ask him/her? Please make at least 3 questions.

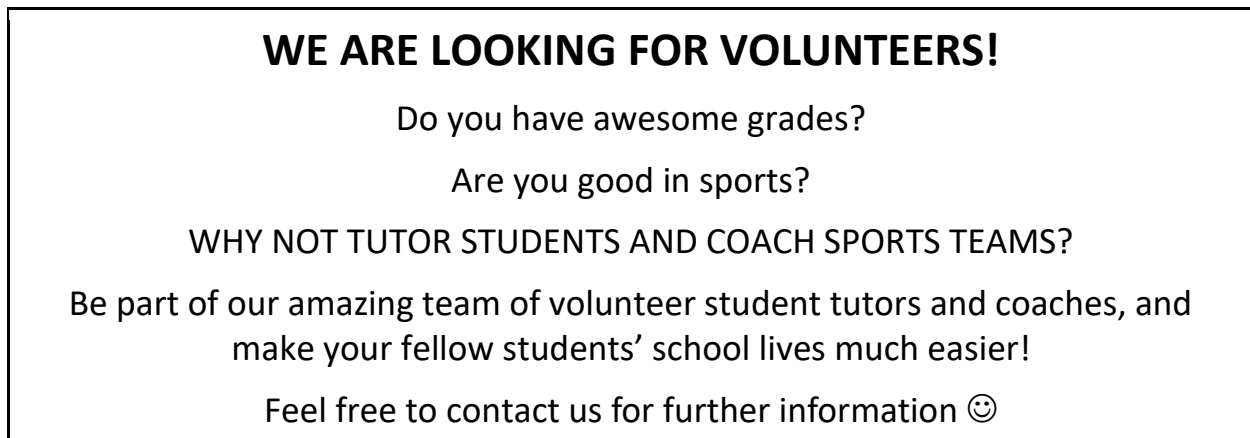


Figure 17. Flyer used in Response to a voice message task

Example of the responses:

A2: What is limit age? What is job with volunteers? Where we go?

B1: When the enrollment day? Any people in grade participate in volunteers? What the tutors do?

B2: *Hi, I want to ask some questions. What kind of volunteering things I will have to do? How many people that already enrolled in this volunteering class? How long I have to be in this volunteering?*

C1: *To be tutor, do I have to have a good grades? What are the limitations for the volunteer works? How many times should I meet my tutees? Are there any places to do tutoring in the university?*

b. Leaving a voice message A

Please respond to the voice message. Respond as if you have currently transferred to ABC school. Please make at least three requests in your response, such as (1) enrollment, (2) information about transportation, and (3) etc.

Hi XXX, this is Laura Ballard on behalf of ABC university. Before everything else, congratulations for your transfer to our school and we are looking forward to meeting you.

I'm calling about your new semester at our university, and I wanted to provide you with some information that may help with your preparation for the new semester. As a facilitator for new students, I would like to help you reach a good transition to this school.

You can reach my by calling 123-456-7890. That number again is 123-456-7890, and ask for reaching me: Laura Ballard. Please feel free to contact me whenever you have any questions.

Figure 18. Voice message used in Leaving a voice message task

Example of the responses:

A2: *I need some guide in my university life. What I learn the course? How can I go university easily? Subway or bus near university? How much the register money?*

B1: *Before transferring ABC university, I took some of subjects in ABC university. My subject which I took previous college granted or not? How to get this university? Is there one subway near ABC university or bus stop? Which semester I have to start? If I transfer ABC university, I start at junior student or second grade student or senior student?*

B2: *I want it to ask you about the transportation. I don't know anything about the transportation. Can you tell me how to get to my campus? I need to register my university. Where should I go to register? I want it to ask you about the money I need to pay for your service fee. How much is it? Let me know if you are free.*

C1: *How can I apply for the classes? Please tell me about date and time for applying for certain classes. I want to know how can I apply for the classes. Please give me some information about transportation. My home is far away from the school. I have to ride a bus or something like that. I don't have a car. What should I ride for getting to the school? I have a question about a period of the semester. When does the semester start? When does end?*

c. Leaving a voice message B

Suppose that you are going to LA with your friends. You have searched for several hotels on the travel agency's website. Before booking one of them, you would like to leave a voice message to the travel agency to ask some questions to get more information on (1) availability of an extra

bed, (2) price of breakfast, and (3) etc. Please make at least 3 questions. Please consider the following information that you found on the website while asking some questions.

ABC Hotel
Available rooms from August 24th - 28th

- King room
- Superior queen room
- Deluxe queen room – City view
- Deluxe queen room – Garden view
- Suite

Breakfast is not included
Facilities

- Parking: \$30 per day
- Pool
- Gym
- Free wifi

For further information, please contact us

Figure 19. Flyer used in Leaving a voice message task

Example of the responses:

A2: How much is breakfast? How about view? Can extra bed?

B1: What I can get the room? How much the breakfast price? Is there pool have sauna?

B2: Hello, I have some questions for booking. Can I get a extra bed? If I can, how much is this?

The page says the breakfast is not included. How much is the breakfast service for the two people? How can I use the gym?

C1: *I'd like to ask some questions. I'm going to hotel with my friend. I'd like to ask about availability of extra bed. Would it be possible to use or rent extra bed if we pay more? According to the flyer, breakfast is not included. I was wondering how much would it be. I'd like to use facilities, such as pool and gym. What is opening time of these facilities?*

2. Writing tasks

a. Response to an email

Please respond to the email. Respond as if you have recently moved to a new city. In your email to the committee, make at least three requests, such as (1) information about supermarkets nearby, (2) setting up the utilities, and (3) etc.

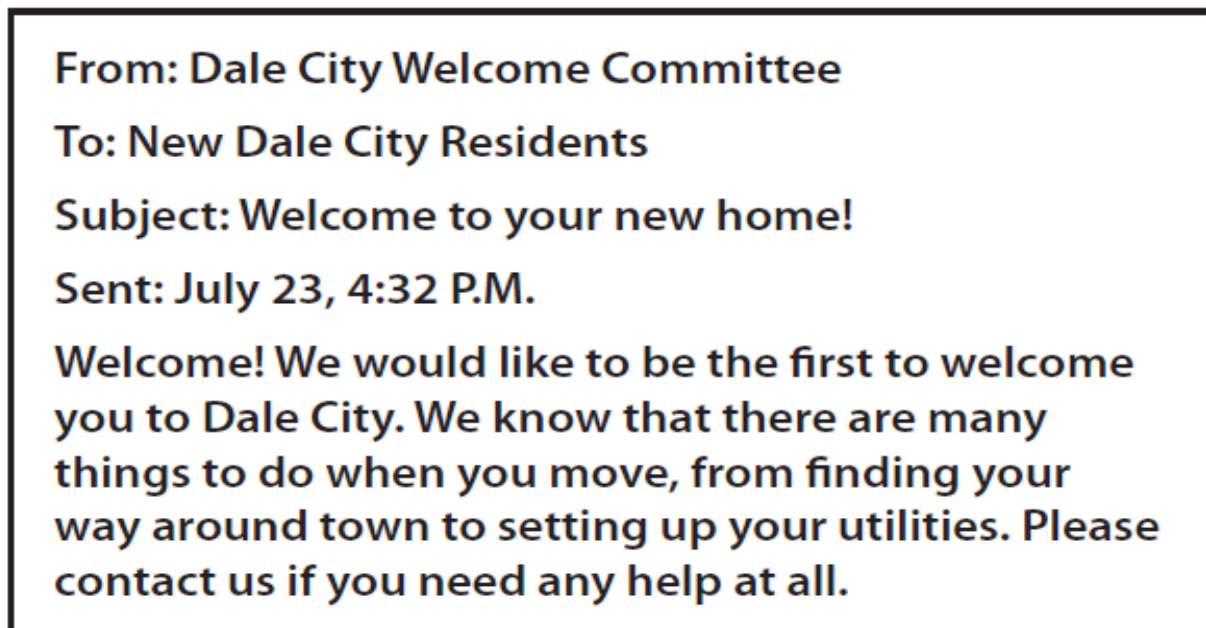


Figure 20. Email used in Respond to an email task

Example of the responses:

A2: *Where is supermarket? What doing myself? How many shops?*

B1: *is there supermarket near the house? is there set up utilities? how much fee per one month?*

B2: *Hello I have some questions for you. first, where is the nearest supermarket??*

next, How can I set up internet and water dispensor and so on? last, when is due date of rent payment??

C1: *Thank you for sending emails to me. I'm really happy to move into the town. I have some questions about settling down in the town:*

1) Is there any supermarket nearby?

2) Where should I contact to get my electricitiy, and water ready?

3) To whom should I call when I have some troubles?

b. Writing an email A

Suppose that you are going to take an express train in Japan with your family. You have found that there is a family ticket that allows you to get an extra discount. Email a representative of the office of railway to ask detailed information on the ticket, such as (1) restrictions of the ticket, (2) price of the ticket, and (3) etc. Please make at least 3 questions.

JAPANESE RAILWAY 
From: Nagoya – Shin-Osaka Date: September 24, 2019 Time: 14:00 – 15: 00
Adults (over 18): \$60 Children (Age: 36 months – 7): \$40 Students (Age: 8 – 17): \$50 Family tickets: DIRECT CONTACT NEEDED
All seats are non-smoking. Wifi-service is available.

Figure 21. Flyer used in Writing an email task

Example of the responses:

A2: *My family from Nagoy to Shin-Osak on September 24,2019 in 14:00 or 15:00 so can i take family tickets? How much the family tickets? Am i take the tickets to railway or send a e-mail?*

B1: *Is there a family tickets that avaible? How much a family tickets? Can I register to non-smoking seats?*

B2: *Hi. I want to buy family tickets. How can I register? And How much that I have to pay for it? At last, where can i get free wifi's password?*

C1: *I would like to know about some restriction when I use family ticket. Is there any limitation of family number? Also, how much is the price of family ticket? Would it be possible to get a discount?*

c. Writing an email B

Suppose that you found a part-time job posting on the website. You would like to get more detailed information on the part-time job so that you email the representative. What questions do you want to ask to him/her? Please make at least 3 questions.

Example of the responses:

A2: Do you give me money that at least pay? If I sick, How can I this situation? My age is lower other people in there?

B1: what time at the working? and what is working first? when working end time?

B2: How long do I have to work each day? How much will you pay for each hour? Where is your company located?

C1: How much do I get paid by time? How long should I work for a week? Do I get paid by week or by month?

Appendix F: Speaking and writing tasks designed to elicit 3rd person singular -s

1. Speaking tasks

a. Response to a voice message

Please respond to the following voice message based on the information. Hello, I'm leaving this voice message about a conference on May 27 that I saw advertised in the newspaper. It's about starting your own business. I was hoping you could give me some information.

- *Could you tell me what time the conference starts and how long it will last?*
- *How much does the conference attendance cost?*
- *I may not be available for the full day. Could you give me information about the activities in the morning, before lunch time?*
- *etc.*



STARTING AND MANAGING YOUR OWN BUSINESS

Date: May 27

Location: Bristol Office Building

Seminars: 9:00 A.M. Financing Your Business, Room 210—*Martha Ross, Certified Public Accountant*
11:00 A.M. How to Promote Your Own Business, Room 312—*Howard Brown, Brown Publishers*
OR
11:00 A.M. Planning for Profit, Room 318—*John Phillips, Phillips Associates*
1:00 P.M. Lunch*
2:00 P.M. Sales Techniques Workshop, Room 246—*Helen King, West Side Consultants*
4:00 P.M. General Discussion

Registration Fee: Individuals, \$95.00
Members of the Business Information Center, \$75.00

*Not included in registration fee.

Figure 22. Flyer used in Response to a voice message task

Example of the responses:

A2: Start is 9 am and last time is 4 pm. Individual cost is \$95. Members of the business information center is \$75. At 9 am, financing your business in room 210, and 11 am how to promote your own business in room 312 and planning for profit in room 318.

B1: The conference start at 9 am to 4 pm. The attendance cost is individuals for \$95 and members of the business information center \$75. Before lunch, first 9 am financing your business room 210 by certified accountant Martha Ross, 11 am how to promote your own business room 312 by Howard Brown Brown publishers.

B2: The conference starts at 9 am. It finishes at 4 pm. Individual is \$95. If you are business information center members, then it's \$75. Before lunch, there is seminar called financing your business at room 201 and how to promote your own business at 312 room. You can choose two seminars. One is how to promote your own business and the other one is planning for profits.

C1: Thank you for expressing your interest in our seminar at first. Seminar starts at 9 am and it ends at 4 pm. That's for 6 hours. If you are trying to attend the seminar alone, you have to pay \$95 for you. If you are member of the business information center, \$20 discount. You have to pay \$75. At 9 am, there is seminar discussing financing your business. The speaker of the seminar is Martha Ross. He is Certified public accountant. At 11 am, there is seminar about the way to promote your own business. The speaker is Howard Brown. You can choose another option. There is a seminar about planning for profit. It's held in room 318.

b. Argumentative task A

Suppose that you are planning a trip to Europe. Before leaving, you would like to create a rough itinerary. There are many ways to get information about countries, accommodation, and etc.:

books, magazine, Internet surfing, and people who have been there. What do you think is the most helpful to prepare for the trip? Give reasons or examples to support your opinion.

Example of the responses:

A2: In my opinion is best choice is Internet searching, because many idea in Internet I choice one plan. Many peoples are really done. Many plan

B1: I think using computer is best way to make itinerary, because in computer, there is many information in blogs or many people's opinion. It is easy to find just put some words in browser. There has many information in browser. There has many people who have been there, so I can connect them easily for SNS. It's easy.

B2: If I have to plan on trip, I would get for help from people who has experience that trip, because they know useful information. They experience them in the first end. I would meet them before my trip and asking about real information, what they experience. But, that's not enough. I would look for some information on the Internet, like basic information, like transportation or culture. To be specific, I'm going to use Youtube who has travel before. Their video show me a lot of real information. I'm going to use, I would say using blog is good way.

C1: There are a plenty of methods to get information when I want to plan for my trip. The best source I think is Internet, because Internet has really best informations. Internet is very time saving method to plan for my trip. It doesn't cost any money or I don't have to spend extra time to get to the library to get books or shops to get magazine. I don't have to spend extra time to meet my friends to get information. Internet is really good source to get information in a very

short time. Also, Internet provides lots of information about some discounts tips and the newest information, such as the price of food for hotels or something.

c. Argumentative task B

Do you think that it's good idea to live with a roommate? Why or why not? Give reasons or examples to support your opinion.

Example of the responses:

A2: *In my opinion, roommate together live is good, because together live not alone. Example, tonight, together sleeping at tonight. In my opinion is good.*

B1: *I think living with roommate is not good, because living with roommate we have to take some rules. It makes me very stress. When roommates make very loud when I was studying, it was very confusion for my study. Roommates sleep very bad that I can't sleep at night. When your roommate is very dirty, you feels very uncomfortable.*

B2: *Living with roommates is not a good idea, because people need privacy. For that reason, living with roommate doesn't guarantee your privacy. Sometimes you need privacy. But I know people need people. Conversation, talk. Sometimes you need to be alone, having your time. That's the reason. If I have to add one more reason, roommate sucks, because sometimes you should bring your friends. If you have roommate, you need to ask your roommate whether you can bring them. You need to think more and care about your roommate. You need to make rules with your roommate. For example, sleeping time, waking time, cleaning days, something like this. It's more complicating.*

C1: I have once live in a dormitory in the first semester of my university. It was not a good experience. Living with someone else requires some rules that I have to abide by, such as “you should not wash on the late time” and “we should talk on the phone very quietly.” and also “you should not watch television or play games in the late nights” Actually, I wake up very early and sleep very early. It’s kind of hard to live with my roommates. The only benefit that I can take when I live with my roommate is that I won’t be lonely.

d. Description task

Please describe one of your family members.

- What does he or she look like?
- What does he or she like to do?
- How often do you meet him/her?
- Why do you like him/her?

Example of the responses:

A2: My mother is in XX. She is working. Her job is the owner of the store. She has short hair. She like to travel.

B1: My mom is professor. She is major in hotel cooking. She likes making food. She also teaching barista. She is very kind for another people. She is very strict for me. She is always good at me. She likes to talk many peoples. She like to earn many moneys. She has big eyes and short hair.

B2: I’m going to describe my mom. She lives in XX, but she likes to come to XX quite often, probably once a week. She is very interested in healthy lifestyle, especially healthy diet. She likes

grocery shopping at organic food store. She doesn't like cooking, so mostly she get some pre-made, microwave food which is relatively healthy. She also likes shopping for clothes a lot.

She is in her late 50s. I thought she has lots of things to talk about.

C1: My older sister is 27 years old. Her major is art. She did art since she was in elementary school. She is studying in Netherlands for her graduate degree. She is currently in Korea, but she is going back next week to finish her study. She is really into fashion. I can just wear what she gives me. She is not but into food, so I could have my favorite food from since we were young. She does yoga. She also drives. Her car is white and her car is very old, so she should be cautious when driving. She likes to drink.

2. Writing tasks

a. Email response

Please respond to the following email.

From: Alexandra Smith

To: ABC museum representative

Subject: Children art exhibition

Sent: July 30, 3:30 pm

Hello, I'm writing to you to request information about a children art exhibition program in your museum. I am planning to attend the program with my students on August 23.

- Could you tell me detailed information about the events on August 23?
- How much does the attendance cost for teenager students and adult?
- Some of my students are disabled children. Do you have any special service for them?

Figure 23. Email used in Email response task

Example of the responses:

A2: *first am 9 to am 10:30 there is two schedule one is graffiti art and another is decorate a bag and graffiti art 's instructor is Amy. children & student for 1day is \$15 at early bird before july 31 and after is \$20 at regular rate and adult our musume giving a bus for them and we are preparing show for fun*

B1: *First your children choose to learn Graffiti art for one and half hour or decorating a bag for it costs 15dollars for one day course if you pay early. it costs 20dollars for one day couser if you don't pay early. first, I will bring wheelchair to someone disabled. and i made fingerwriting fo blind people. I will make elevator or escalator for someone disabled.*

B2: *From 9am to 10:30am, we have graffiti art exhibition and 'decorate a bag'. For teenage students, it costs \$20 for regular rate, \$15 for Early-bird (before July 31). This is one day price. We do have special service for disabled children. We have assistants to help children.*

C1: Thank you for your interest in the event. It has a series of events such as graffiti art, bag decoration, pottery sculpture, and science&art. It starts from 9 a.m. to 3:30 p.m.. I will attach you the detailed time table so you can. The admission fee for teenagers is 15\$, and adults should pay 20\$ for a day. For the students who have physical difficulty, I can give you some helps. There is the letter which can be read by blind students, and I will prepare subtitles for all of the movies in the exhibition for students who can't hear well.

b. Argumentative task A

Suppose that you will graduate in February and you are on the job market. There are many ways to find a job: newspaper, advertisements, internet job search web sites, and personal recommendations. What do you think is the best way to find a job? Give reasons or examples to support your opinion.

Example of the responses:

A2: My opinion is searching internet. because It is very fast and many information in it. So the information that I want.

B1: I prefer ask somebody who get a job. Because I don't know detail information on the newspapers or on the internet. Maybe company posts content their own online page, but I can't get real information or personal experiments. If you ask person who get a job may be they provide suitable information for me.

B2: I think the most useful method is using the internet. Because there are lots of information to find the jobs. And also there are lots of people who gives their personal job information or job finding info. Also It is easy to get that all information.

C1: I think Internet is the best place to find a job that perfectly fits me. First of all, there are lots of information on the Internet. I can find about the salary, work environment, etc. Also, I can learn how to appeal for that company with searching. I mean, I can learn what kind of characteristics that company want, such as socialization and work experience. In addition, Internet is definitely the most convenient way to get information. I don't have to spend time or money.

c. Argumentative task B

Do you think that it's a good idea to have a laptop during a class? Why or why not? Give reasons or examples to support your opinion.

Example of the responses:

A2: In my option is using laptop is bad. Because there can anther action. Example, play games, internet shopping, email.

B1: Using laptop during class is good effection for student. The reason why is, first, it decrease to write what teacher saying and write on the board. Second, students can record the lesson, and it is really helpful to study at home.

B2: I am against using laptop during class. because it is so distracting that student have a tendency that they do something else during class. Furthermore, they make noise which means they are not good for everyone in the class. in conclusion, do not use laptop or any other devices during class.

C1: I believe that students should not use their laptops in class. There are two reasons to support my argument. First, using a laptop during the lecture disturbs the instructor. Instructors

frequently complain that students do not look in their eyes or even the whiteboard during class. Second, students cannot pay full attention in class because they can use the laptop to text friends and Internet shopping. Due to Wifi, students can access the Internet anytime.

d. Description task

Please describe one of your friends.

- *What does he or she look like?*
- *What does he or she like to do?*
- *How often do you meet him/her?*
- *Why do you like him/her?*

Example of the responses:

A2: *She is kind and too much talking. Her job is coffee shop's manager. She likes comfortable clothes.*

B1: *I introduce my friend Yuseon. I met her 3 years ago in the university. She has very white skin. She always hangs out with her friend because she has very outgoing personality. I went to Hong Kong with her 2 years ago, we were so impressed when we saw parade in Disneyland. My friend Yuseon is very*

B2: *I have a friend who is studying at Kyoto. She is majoring a society and she is making an effort to receive her doctor's degree. She is smart, kind and nice person. She has one dog. And she enjoys many society club and doing Yoga to keep her health.*

C1: Her name is Larua and she's French. I met her when I was at university. She was an exchange student from Canada. She's very caring, respectful, friendly, social and open-minded. She loves traveling, Korean food and learning something new. I love her because she's a good listener and understand what I go through during living in other countries as she also lives in Canada. Sometimes we argued because of culr

Appendix G: Background questionnaire

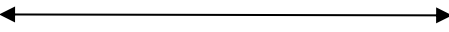
Questions	Response
What is your name?	
How old are you?	
What is your major?	
What is your gender?	
Please provide your scores of TOEFL/IELTS/TOEIC	
Please list all the languages you know in order of dominance:	
How long have you stayed abroad? Why?	1. Duration: 2. Reason(s):
What is your final degree?	
Subjective judgment of English proficiency (0-100 for each)	1. Overall 2. Listening 3. Speaking 4. Reading 5. Writing 6. Grammar 7. Vocabulary

Appendix H: Exit questionnaire

1. How difficult were the tasks to complete?

a. Overall

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very difficult)



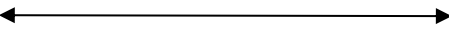
b. Speaking tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very difficult)



c. Writing tasks

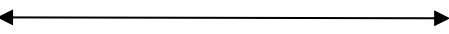
0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very difficult)



2. While participating in the tasks, did you feel frustrated?

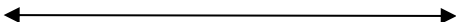
a. Speaking tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very frustrated)



b. Writing tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very frustrated)



3. Do you think that you did well in the tasks?

a. Overall

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very well)



b. Speaking tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very well)



c. Writing tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very well)



4. How interesting was it to participate in the tasks?

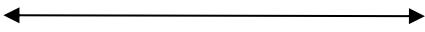
a. Speaking tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very interesting)



b. Writing tasks

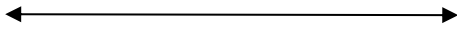
0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very interesting)



5. If you have a chance, do you want to participate in similar experiment once again?

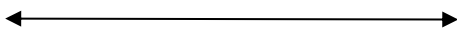
a. Overall

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very certain)



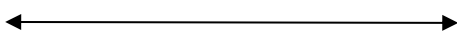
b. Speaking tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very certain)



c. Writing tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very certain)



6. How were you confident while participating in the experiment?

a. Speaking tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very confident)



b. Writing tasks

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (Very confident)



7. Do you think that the first tasks affected the second tasks?

0 (Not at all) 10 20 30 40 50 60 70 80 90 100 (it affected a lot.)



8. Please describe your general thoughts about the speaking tasks.

9. Please describe your general thoughts about the writing tasks.

Appendix I: Question in the brief interview

Please describe your answer of #7 in the exit questionnaire in detail. Do you think that the first tasks affected the second tasks?

Appendix J: Example of the Microsoft Excel sheets used in the coding

1. Example of the Excel sheets that the two coders put the target structures found in each participant's response

Structures	Target		Other	
	Accurate	Inaccurate	Accurate	Inaccurate
S neg V				
Plural (-s)				
Plural NP				
-ing				
Possessive ('s)				
Past (-ed)				
Objective pronoun				
3rd person singular (-s)				
Y/N question				
Wh-copular question				
Aux question				
Cancel INV				

2. Example of the Excel sheets in which the number of target structures is automatically counted.

Structures	Target		Other	
	Accurate	Inaccurate	Accurate	Inaccurate
S neg V				
Plural (-s)				
Plural NP				
-ing				
Possessive ('s)				
Past (-ed)				
Objective pronoun				
3rd person singular (-s)				
Y/N question				
Wh-copular question				
Aux question				
Cancel INV				

3. Screenshot of the modified version of the Excel sheets that the coders used

	A	B	C	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1				Target				Other						Target		Other	
2				Accurate		Inaccurate		Accurate		Inaccurate				Accurate	Inaccurate	Accurate	Inaccurate
3	S neg V												S neg V	0	0	0	0
4													Plural	0	0	0	0
5	Plural												Plural NP	1	0	0	0
6													-ing	2	2	0	0
7	Plural NP		two men										Possessive ('s)	0	0	0	0
8													Past (-ed)	0	0	0	0
9	-ing		referee is trying	two men is trying									Objective pronoun	0	0	0	0
10			they are watching	ball get									3rd person singular (s)	0	0	0	0
11	Possessive ('s)												Do- question	0	0	0	0
12													Be- question	0	0	0	0
13	Past (-ed)												Aux question	0	0	0	0
14													Cancel INV	0	0	0	0
15	Objective pronoun																
16																	
17	3rd person singular (s)																
18																	
19	Y/N question																
20																	
21	Wh-copular question																
22																	
23	Aux question																
24																	
25	Cancel INV																
26																	
27																	

REFERENCES

REFERENCES

- Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The modern language journal*, 78(2), 155-168.
- Bardovi-Harlig, K., & Bofman, T. (1989). Attainment of syntactic and morphological accuracy by advanced language learners. *Studies in second language acquisition*, 11(1), 17-34.
- Barkaoui, K., & Knouzi, I. (2018). The effects of writing mode and computer ability on L2 test-takers' essay characteristics and scores. *Assessing Writing*, 36, 19-31.
- Bonilla, C. (2012). *Testing Processability Theory in L2 Spanish: Can Readiness or Markedness Predict Development?* (Doctoral dissertation, University of Pittsburgh, Pittsburgh).
- Bonilla, C. (2015). From number agreement to the subjunctive: Evidence for Processability Theory in L2 Spanish. *Second Language Research*, 31(1), 53-74.
- Boss, B. (2008). Exploring the acquisition of German verb morphology by instructed learners. *Australian Review of Applied Linguistics*, 31(2), 14-1.
- Bourdin, B., & Fayol, M. (1994). Is written language production more difficult than oral language production? A working memory approach. *International journal of psychology*, 29(5), 591-620.
- Bourdin, B., & Fayol, M. (2000). Is graphic activity cognitively costly? A developmental approach. *Reading and Writing*, 13(3-4), 183-196.
- Brown, J., McDonald, J, Brown, T, & Carr, T. (1988). Adapting to processing demands in discourse production: The case of handwriting. *Journal of Experimental Psychology: Human Perception and Performance*, 14(1), 45.
- Bulté, B., & Housen, A. (2009). *The development of lexical proficiency in L2 speaking and writing tasks by Dutch-speaking learners of French in Brussels*. In Task-based language teaching conference, Lancaster, England.
- Byrnes H. , Manchón R. (2014). Task-based language learning: Insights from and for L2 writing: An introduction . In H. Byrnes & R. Manchón (Eds.), *Task-based language learning-- Insights from and for L2 writing* (pp. 1 -27). Amsterdam, The Netherlands: John Benjamins.
- Byrnes, H., & Sinicrope, C. (2009). Advancedness and the development of relativization in L2 German: A curriculum-based longitudinal study. In L. Ortega & H. Byrnes. (Eds.), *The Longitudinal Study of Advanced L2 Capacities* (pp. 125-154). New York: Routledge.

- Cheng, L. (2019). Diagnostic assessments. In J. Liantas (Editor-in-Chief), *The TESOL Encyclopedia of English Language Teaching*. Hoboken, NJ: Wiley-Blackwell.
- Cumming, A. (2012). Goal theory and second-language writing development, two ways. In R, Manchón (Ed.), *L2 writing development: Multiple perspectives* (pp. 165–190). Boston/Berlin: de Gruyter Mouton.
- De Silva, R., & Graham, S. (2015). The effects of strategy instruction on writing strategy use for students of different proficiency levels. *System*, 53, 47–59.
- Di Biase, B., & Kawaguchi, S. (2002). Exploring the typological plausibility of Processability Theory: Language development in Italian second language and Japanese second language. *Second language research*, 18(3), 274-302.
- Di Biase, B., Kawaguchi, S., & Yamaguchi, Y. (2015). The development of English as a second language. In C. Bettoni & B. Di Biase (Eds.), *Grammatical Development in Second Languages : Exploring the Boundaries of Processability Theory* (pp. 85-115). Retrieved from <http://www.eurosla.org/monographs/EM03/2English.pdf>
- Dyson, B. (2009). Processability theory and the role of morphology in English as a second language development: A longitudinal study. *Second Language Research*, 25(3), 355-376.
- Ellis, R. (2008). Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics*, 18, 4 – 22.
- Gass, S. (1980). An investigation of syntactic transfer in adult second language learners. *Research in second language acquisition*, 132-141.
- Gilabert, R., Manchón, R., & Vasylets, O. (2016). Mode in theoretical and empirical TBLT research: Advancing research agendas. *Annual Review of Applied Linguistics*, 36, 117-135.
- Grabowski, J. (2010). Speaking, writing, and memory span in children: Output modality affects cognitive performance. *International Journal of Psychology*, 45(1), 28-39.
- Granfeldt, J. (2008). Speaking and writing in French L2: Exploring effects on fluency, complexity and accuracy. In S. van Daele, A. Housen, F. Kuiken, M. Pierrard., & I. Vedder (Eds.), *Complexity, accuracy and fluency in second language use, learning & teaching* (pp. 87-98). Belgium: University of Brussels.
- Green, A. (2018). Placement testing. In J. Liantas (Editor-in-Chief), *The TESOL Encyclopedia of English Language Teaching*. Hoboken, NJ: Wiley-Blackwell.

- Håkansson, G., & Norrby, G. (2007). Processability theory applied to written and oral Swedish. In F. Mansouri (Ed.), *Second language acquisition research: Theory-construction and testing* (pp. 81-94). Newcastle: Cambridge Scholars Publishing.
- Itani-Adams, Y. (2009). Development of discourse functions in Japanese and English bilingual first language acquisition. In J. Keßler & D. Keatinge (Eds.), *Research in second language acquisition: empirical evidence across languages* (pp. 41-66). Newcastle: Cambridge Scholars Publishing
- Jansen, L. (2008). Acquisition of German word order in tutored learners: A cross-sectional study in a wider theoretical context. *Language Learning*, 58, 185–231.
- Johnson, M. (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.
- Kawaguchi, S. (2000). Acquisition of Japanese verbal morphology: Applying processability theory to Japanese. *Studia linguistica*, 54(2), 238-248.
- Kawaguchi, S. (2005). Processability Theory and Japanese as a second language. *Acquisition of Japanese as a second language*. 8, 83-115.
- Kellogg, R. (1996). A model of working memory in writing. In C. Levy & S. Ransdell (Eds.), *The science of writing: Theories, methods, individual differences, and applications* (pp. 57-71). Hillsdale, NJ, US: Lawrence Erlbaum Associates, Inc.
- Kellogg, R. (2001). Competition for working memory among writing processes. *The American Journal of Psychology*, 114(2), 175.
- Kormos, J. (2014). Differences across modalities of performance. In H. Byrnes & R. Manchón (Eds.), *Task-based language learning--Insights from and for L2 writing* (pp. 193 -216). Amsterdam, The Netherlands: John Benjamins.
- Kormos, J., & Trebits, A. (2009). *Task-related differences across modes of performance*. In Task Based Language Teaching Conference, Lancaster.
- Kuiken, F., & Vedder, I. (2011). Task complexity and linguistic performance in L2 writing and speaking. In P. Robinson (Ed.), *Second language task complexity: Researching the Cognition Hypothesis of language learning and performance* (pp. 91-104). Amsterdam, The Netherlands: John Benjamins.
- Kyle, K., Crossley, S., & Verspoor, M. (2020). Measuring Longitudinal Writing Development Using Indices of Syntactic Complexity and Sophistication. *Studies in Second Language Acquisition*, 1-32.
- Larsen-Freeman, D. (1978). An ESL index of development. *TESOL quarterly*, 12(4), 439-448.

- Larsen-Freeman, D. (1983). Second language acquisition Getting the whole picture In K. Bailey, M. Long, & S. Peck (Eds), *Second Language Acquisition Studies* (pp. 3-22). Rowley, MA Newbury House.
- Lee, S. & Spinner, P. (2017). *The place of accuracy in Processability Theory*. Paper presentation at 2017 Second Language Research Forum, Columbus, OH.
- Leijten, M., & Van Waes, L. (2013). Keystroke logging in writing research: Using Inputlog to analyze and visualize writing processes. *Written Communication*, 30(3), 358-392.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press
- Mansouri, F. (2005). Agreement morphology in Arabic as a second language: Typological features and their processing implications. In M. Pienemann (Ed.), *Cross-linguistic aspects of Processability Theory* (pp. 117-154). Philadelphia, PA: John Benjamins.
- Monroe, J. (1975). Measuring and enhancing syntactic fluency in French. *The French Review*, 48(6), 1023-1031.
- Norris, J., & Ortega, L. (2009). Towards an organic approach to investigating CAF in instructed SLA: The case of complexity. *Applied linguistics*, 30(4), 555-578.
- Ochs, E. (1979). Planned and unplanned discourse. In *Discourse and syntax* (pp. 51-80). Brill.
- Park, J. (2017). *Syntactic complexity as a predictor of second language writing proficiency and writing quality*. (Doctoral dissertation, Michigan State University, East Lansing)
- Pienemann, M. (2007). *Processability theory: Theory construction and application*. Plenary speech at 17th annual conference of the European Second Language Association, Newcastle, UK .
- Pienemann, M. (1987). Determining the influence of instruction on L2 speech processing. *Australian Review of Applied Linguistics*, 10(2), 83-113.
- Pienemann, M. (1998). *Language processing and second language development: Processability theory* (vol. 15). Amsterdam/Philadelphia: John Benjamins Publishing.
- Pienemann, M. (2003). Language processing capacity. In C. Doughty & M. Long (Eds.), *The handbook of second language acquisition* (pp. 679-714). Malden, MA: Blackwell.
- Pienemann, M. (2005). An introduction to processability theory. In M. Pienemann, *Cross-linguistic aspects of processability theory* (pp. 1-60). Amsterdam/Philadelphia: John Benjamins Publishing.

- Pienemann, M., & Håkansson, G. (1999). A unified approach toward the development of Swedish as L2: A Processability Account. *Studies in second language acquisition*, 21(3), 383-420.
- Pienemann, M., & Johnston, M. (1987). Factors influencing the development of language proficiency. In D. Nunan (Ed.), *Applying second language acquisition research* (pp. 45–141). Adelaide, New Zealand: National Curriculum Resource Center, AMEP.
- Pienemann, M., Di Biase, B., Kawaguchi, S., & Håkansson, G. (2005). Processing constraints on L1 transfer. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 128-153). New York: Oxford University Press.
- Polio, C. (2012). The relevance of second language acquisition theory to the written error correction debate. *Journal of second language writing*, 21(4), 375-389.
- Polio, C. (2017). Second language writing development: A research agenda. *Language Teaching*, 50(2), 261-275.
- Polio, C., & Park, J. (2016). Language development in second language writing. In R. Manchón & P. Matsuda (Eds.), *Handbook of Second and Foreign Language Writing* (pp. 287-306). Boston/Berlin: De Gruyter Mouton.
- Qi, R., & Di Biase, B. (2005). *L1 and L2 patterns in the bilingual language development of a Mandarin-English child*. In 5th International Symposium on Bilingualism. Universitat Autònoma de Barcelona.
- Ravid, D., & Tolchinsky, L. (2002). Developing linguistic literacy: A comprehensive model. *Journal of child language*, 29(2), 417-447.
- Révész, A., Kourтали, N. E., & Mazgutova, D. (2017). Effects of task complexity on L2 writing behaviors and linguistic complexity. *Language Learning*, 67(1), 208-241.
- Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27-57.
- Sakai, H. (2008). An analysis of Japanese university students' oral performance in English using processability theory. *System*, 36(4), 534-549.
- Sasaki, M. (2004). A multiple-data analysis of the 3.5 year development of EFL student writers. *Language Learning* 54(3). 525–582.
- Schoonen, R., Snellings, P., Stevenson, M., & Van Gelderen, A. (2009). Towards a blueprint of the foreign language writer: The linguistic and cognitive demands of foreign language writing. In *Writing in foreign language contexts* (pp. 77-101). Multilingual Matters.

- Serrano, R., Tragant, E., & Llanes, À. (2012). A longitudinal analysis of the effects of one year abroad. *Canadian modern language review*, 68(2), 138-163.
- Skehan, P. (1998a). Task-based instruction. *Annual Review of Applied Linguistics*, 18, 268-286.
- Spinner, P. (2011). Second language assessment and morphosyntactic development. *Studies in Second Language Acquisition*, 33(4), 529-561.
- Tardy, C. (2012). A rhetorical genre theory perspective on L2 writing development. In R. Manchón (Ed.), *L2 writing development: Multiple perspectives*, 165–190. Boston/Berlin: De Gruyter Mouton.
- Tavakoli, P. (2014). Storyline complexity and syntactic complexity in writing and speaking tasks. In H. Byrnes & R. Manchón (Eds.), *Task-based language learning--Insights from and for L2 writing* (pp. 217-236). Amsterdam, The Netherlands: John Benjamins.
- Vasylets, O., Gilabert, R., & Manchon, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67(2), 394-430.
- Weissberg, B. (2000). Developmental relationships in the acquisition of English syntax: Writing vs. speech. *Learning and instruction*, 10(1), 37-53.
- Weissberg, R. (2006). *Connecting speaking & writing in second language writing instruction*. Ann Arbor: University of Michigan Press.
- Williams, J. (2012). The potential role(s) of writing in second language development. *Journal of Second Language Writing*, 21(4), 321-331.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H-Y. (1998). *Second Language Development in Writing: Measures of Fluency, Accuracy, and Complexity*. Honolulu, HI: University of Hawai'i, Second Language Teaching and Curriculum Center.
- Young, D. (1986). The relationship between anxiety and foreign language oral proficiency ratings. *Foreign Language Annals*, 19(5), 439-445.
- Yu, G. (2009). Lexical diversity in writing and speaking task performances. *Applied linguistics*, 31(2), 236-259.
- Zalbidea, J. (2017). 'One Task Fits All'? The Roles of Task Complexity, Modality, and Working Memory Capacity in L2 Performance. *The Modern Language Journal*, 101(2), 335-352.
- Zhang, Y. (2005) Processing and formal instruction in the L2 acquisition of five Chinese grammatical morphemes. In: M. Pienemann (ed.) *Cross-linguistic aspects of Processability Theory* (pp. 155-178). Philadelphia, PA: John Benjamins.