DOCTORAL DISSERTATION

Written Performances and Flow Experiences of Chinese Students Studying in Hungary and China across Different Tasks

Qiaoxia Wu

2020
Eötvös Loránd University
Faculty of Humanities

DOCTORAL DISSERTATION

Qiaoxia Wu

Written Performances and Flow Experiences of Chinese Students Studying in Hungary and China across Different Tasks

Doctoral School of Linguistics
Head of the Doctoral School of Linguistics: Dr. Gábor Tolcsvai Nagy
Doctoral Programme of Intercultural Linguistics
Head of the Doctoral Programme: Dr. Rita Brdar-Szabó

Supervisor: Dr. Albert Ágnes

Members of the Assessment Committee:
Dr. Medgyes Péter DSc, professor emeritus, chair
Dr. Katona László, secretary
Dr. Brózik-Pinie Katalin, opponent
Dr. Furcsa Laura, opponent
Dr. Kou Danyang, member
Dr. Brdar-Szabó Rita, member
Dr. Adorján Mária, member
Abstract

Writing plays a major role in Second Language Acquisition (SLA). In foreign language learning, the flow experience has recently attracted increasing interest among researchers, but studies are mainly on its connection with oral tasks (e.g., Kormos & Prefontaine, 2017), creating a research niche with its relevance to writing.

The aims of this dissertation were, on the one hand, to investigate the relationship between task conditions and Chinese students’ English written performances and their flow experiences during the writing process, as well as the correlations between them; on the other, to investigate the possible role that the learning contexts might play in the two aspects. The three task conditions were the teacher-generated content task (TGT), the learner-generated content task (LGT), and the creative task (CT), which were differentiated by manipulating the levels of learner agency/control and the potential for creative language use that the task has, with the TGT having the lowest level of learner agency/control and the LGT having the highest, and the CT having the greatest potential to elicit creativity while the TGT and LGT having no such potential. Participants of the study were two groups of Chinese intermediate English learners studying in Hungary (the CSH) and China (the CSC).

The findings of this study showed that, regarding the written performances, for the CSH and the CSC respectively, certain aspect of the tasks (learner agency, creativity) seems to be linked to some aspects of task performances like the amount of words produced, lexical complexity and syntactic complexity, and accuracy. Besides, for the CSH and the CSC respectively, certain aspect of the tasks (learner agency, creativity) also seems to be associated with some aspects of flow experiences (interest, attention, control, and challenge-skill balance) and anti-flow experiences (boredom, apathy, and anxiety). Regarding the correlations between the two aspects researched, the two aspects seem to correlate differently in each task condition for the CSH and the CSC respectively, indicating the role of English study contexts in these relationships. In addition, in each task condition, there seem to be differences in some aspects of the task performances and flow and anti-flow experiences between the CSH and the CSC at the group level, indicating the role of study contexts in the researched two aspects in each task condition.

Keywords: Chinese students, writing, task conditions, learner agency, creativity, fluency, complexity, accuracy, flow, study contexts
Acknowledgements

First and foremost, I would like to express my heartfelt gratitude to my supervisor, Dr. Albert Ágnes, for her constant guidance and encouragement throughout the writing of this dissertation. I particularly appreciate her meticulous reading of the drafts of this thesis and her invaluable advice and insightful comments concerning the revisions of this thesis. I am also very grateful to her for teaching me how to use SPSS software to analyze data. Her rigorous scholarship and agreeable personalities will be the most valuable treasures and cherished memories in my whole life. I am also indebted to all the teachers at Eötvös Loránd University who guided me in my studies. My sincere thanks go to Dr. Kövecses Zoltán, Dr. Szabó Rita, Dr. Benczes Réka, Dr. Gecső Tamás, and Dr. Csizér Kata for teaching me interesting courses, from which I benefited a lot. Special thanks also go to my colleagues Guo Xinhua and Fen Jing for their useful advice as well as their spiritual support during the process of thesis writing. My immense gratitude is also extended to my friends Chen Mingjie, Hu Miaomiao, Wu Qiaoli, and Zhang Yanlin for their help in recruiting participants, without whom the completion of this dissertation would have been impossible. Last but not least, my inexpressible appreciation is to my dearest family for their loving care for and great confidence in me during the years of my study and thesis writing. Without them, I would never have succeeded in my study and research. My heartfelt thanks to them are beyond words.
# Table of Contents

Abstract .......................................................................................................................... 4

Acknowledgements ........................................................................................................ 5

Table of Contents .......................................................................................................... 6

List of Tables ................................................................................................................. 11

List of Figures ............................................................................................................... 13

Chapter 1: Introduction ................................................................................................. 14
  1.1 Rationale and aim of the dissertation ................................................................. 14
  1.2 Research questions ............................................................................................ 17
  1.3 Overview of the dissertation ............................................................................. 17

Chapter 2: Literature Review and Theoretical Background of the Research ........... 19
  2.1 Introduction ......................................................................................................... 19
  2.2 The current situation and problems of the teaching of English writing in high schools in China ............................................................................................................. 19
  2.3 Student engagement and language learning ..................................................... 22
    2.3.1 Definition of student engagement and measuring student engagement in general education realm .................................................................................................................. 22
    2.3.2 Definition of student engagement and measuring student engagement in language learning ................................................................................................................. 25
    2.3.3 Main studies on student engagement in foreign language learning .......... 30
    2.3.4 Summary ....................................................................................................... 32
  2.4 Maehr’s (1984) personal investment theory ...................................................... 33
    2.4.1 Description of the personal investment theory ............................................ 33
    2.4.2 The relevance of Maehr’s (1984) personal investment theory for student engagement and foreign language learning ................................................................. 35
  2.5 Flow theory .......................................................................................................... 38
    2.5.1 Description of the flow theory ................................................................... 38
    2.5.2 Flow theory and foreign language learning .............................................. 39
2.5.3 Summary: relationships among flow theory, personal investment theory and student engagement in classroom foreign language learning ........................................42

2.6 Task design ........................................................................................................43
  2.6.1 Task complexity ..............................................................................................44
  2.6.2 Learner agency and potential for creativity in language learning tasks ..........46
  2.6.3 Summary ........................................................................................................49

2.7 Empirical research on written narrative tasks ....................................................49
  2.7.1 Empirical research on effects of task complexity on foreign language writing ........................................................................................................50
  2.7.2 Empirical research on student engagement in foreign language or EFL writing ........................................................................................................52
  2.7.3 Empirical research on student engagement in foreign language or EFL writing with affect/flow being taken into consideration .........................................57
  2.7.4 Summary ........................................................................................................61

2.8 Conclusion ..........................................................................................................61

Chapter 3: Research Methods .................................................................................65

3.1 Participants .........................................................................................................65

3.2 Instruments .........................................................................................................69
  3.2.1 Written narrative tasks ....................................................................................70
  3.2.2 Flow questionnaire ..........................................................................................71
  3.2.3 Bio-data questionnaire ....................................................................................72

3.3 Piloting ................................................................................................................72
  3.3.1 Piloting the written narrative tasks ................................................................72
  3.3.2 Piloting the flow questionnaire and the bio-data questionnaire ......................73

3.4 Procedures ..........................................................................................................74
  3.4.1 Data collection in Hungary ............................................................................74
  3.4.2 Data collection in China ..................................................................................75

3.5 Analysis ..............................................................................................................76
  3.5.1 Measures of task performance ......................................................................76
  3.5.2 Measures of flow ............................................................................................83

3.6 Statistical analysis ..............................................................................................85
Chapter 4: Findings concerning the Written Performances and the Flow and Anti-flow Experiences of the CSH

4.1 Results on the relationship between task conditions and the written performances of the CSH ................................................................. 86
4.2 Discussion on the relationship between task conditions and the written performances of the CSH ................................................................. 89
4.3 Results on the relationship between task conditions and the flow and anti-flow experiences of the CSH ................................................................. 91
4.4 Discussion on the relationship between task conditions and the flow and anti-flow experiences of the CSH ................................................................. 93
4.5 Results on the correlations between the flow and anti-flow experiences and the written performances of the CSH in different task conditions while controlling for language proficiency ................................................................. 95
  4.5.1 In the TGT ................................................................. 95
  4.5.2 In the LGT ................................................................. 96
  4.5.3 In the CT ................................................................. 97
4.6 Discussion on the correlations between the flow and anti-flow experiences and the written performances of the CSH in different task conditions while controlling for language proficiency ................................................................. 98
  4.6.1 In the TGT ................................................................. 98
  4.6.2 In the LGT ................................................................. 99
  4.6.3 In the CT ................................................................. 100

Chapter 5: Findings concerning the Written Performances and the Flow and Anti-flow Experiences of the CSC

5.1 Results on the relationship between task conditions and the written performances of the CSC ................................................................. 102
5.2 Discussion on the relationship between task conditions and the written performances of the CSC ................................................................. 104
5.3 Results on the relationship between task conditions and the flow and anti-flow experiences of the CSC ................................................................. 107
5.4 Discussion on the relationship between task conditions and the flow and anti-flow experiences of the CSC ................................................................. 108
5.5 Results on the correlations between the flow and anti-flow experiences and the written performances of the CSC in different task conditions while controlling for language proficiency ................................................................. 109
  5.5.1 In the TGT ..................................................................................... 109
  5.5.2 In the LGT ..................................................................................... 110
  5.5.3 In the CT ..................................................................................... 111

5.6 Discussion on the correlations between the flow and anti-flow experiences and the written performances of the CSC in different task conditions while controlling for language proficiency ................................................................. 112
  5.6.1 In the TGT ..................................................................................... 112
  5.6.2 In the LGT ..................................................................................... 113
  5.6.3 In the CT ..................................................................................... 113

5.7 Summary: comparison between the CSH and the CSC ................................................................. 114
  5.7.1 Results on the differences between the CSH and the CSC concerning the relationship between task conditions and their written performances ................................................................. 115
  5.7.2 Discussion on the differences between the CSH and the CSC concerning the relationship between task conditions and their written performances ................................................................. 115
  5.7.3 Results on the differences between the CSH and the CSC concerning the relationship between task conditions and their flow and anti-flow experiences ................................................................. 118
  5.7.4 Discussion on the differences between the CSH and the CSC concerning the relationship between task conditions and their flow and anti-flow experiences ................................................................. 119
  5.7.5 Results on the differences between the CSH and the CSC concerning the correlations between their flow and anti-flow experiences and their written performances in different task conditions ................................................................. 119
  5.7.6 Discussion on differences between the CSH and the CSC concerning the correlations between their flow and anti-flow experiences and their written performances in different task conditions ................................................................. 122

Chapter 6: Comparison between the CSH and the CSC concerning their Written Performances and their Flow and Anti-flow Experiences in Different Task Conditions ................................................................. 124

6.1 Introduction ..................................................................................... 124

6.2 Differences between the CSH and the CSC concerning their written performances and their flow and anti-flow experiences in the TGT ................................................................. 125
  6.2.1 Results ..................................................................................... 125
6.2.2 Discussion ........................................................................................................................................ 127

6.3 Differences between the CSH and the CSC concerning their written performances and their flow and anti-flow experiences in the LGT ................................................................. 127

6.3.1 Results .............................................................................................................................................. 128

6.3.2 Discussion ........................................................................................................................................ 129

6.4 Differences between the CSH and the CSC concerning their written performances and their flow and anti-flow experiences in the CT ................................................................. 131

6.4.1 Results .............................................................................................................................................. 131

6.4.2 Discussion ........................................................................................................................................ 133

Chapter 7: Conclusions and Pedagogical Implications ................................................................. 134

7.1 Introduction .......................................................................................................................................... 134

7.2 Summary of findings ............................................................................................................................. 135

7.3 Limitations ........................................................................................................................................... 138

7.4 Pedagogical implications and directions for future research ......................................................... 139

References .................................................................................................................................................. 142


Appendix B-Clause guidelines (Based on Polio, 1997, and Lee, 2009) .................................................... 161

Appendix C-Error guidelines (Based on Polio & Shea, 2014, and Arsanyilmaz & Pedersen, 2010) ..................................................................................................................................................... 162

Appendix D-Weighted Clause Ratio (WCR) (Based on Wigglesworth & Foster, 2008) ...................... 163

Appendix E-Flow questionnaire (From Czimmermann & Piniel, 2016) .................................................... 167

Appendix F-Three written narrative tasks ................................................................................................. 171

Appendix G-Bio-data questionnaire ......................................................................................................... 175
List of Tables

Table 1 Key characteristics of engagement with language ........................................26
Table 2 Criteria for identifying engagement with language ........................................27
Table 3 Summary of the research design ....................................................................63
Table 4 Biographical information of the participants ................................................66
Table 5 Characteristics of participants .....................................................................68
Table 6 The number of participants who are engaged in each of the out-of-school English learning activities (OSELAs) .................................................................69
Table 7 Descriptive statistics of each scale of the flow questionnaire and its corresponding reliability coefficient .................................................................74
Table 8 Clause categorization for weighted clause ratio (WCR) ...............................81
Table 9 Scoring the clause categories ......................................................................81
Table 10 Example of categorization and scoring of clauses ......................................83
Table 11 Descriptive statistics of each flow and anti-flow variable in each task ..........84
Table 12 Reliability analysis, PCA of each flow and anti-flow variable ....................85
Table 13 Descriptive statistics of the written performance measures of the CSH in different tasks .....................................................................................87
Table 14 Descriptive statistics of the flow and anti-flow measures of the CSH in different tasks .....................................................................................87
Table 15 Correlations between the written performance measures and flow and anti-flow measures in the TGT controlling for language proficiency ....................96
Table 16 Correlations between the written performance measures and flow and anti-flow measures in the LGT controlling for language proficiency ....................97
Table 17 Correlations between the written performance measures and flow and anti-flow measures in the CT controlling for language proficiency ....................98
Table 18 Descriptive statistics of the written performance measures of the CSC in different tasks .....................................................................................102
Table 19 Descriptive statistics of the flow and anti-flow measures of the CSC in different tasks .....................................................................................102
Table 20 Correlations between the written performance measures and flow and anti-flow measures in the TGT controlling for language proficiency .................108
Table 21 Correlations between the written performance measures and flow and anti-flow measures in the LGT controlling for language proficiency ....................109
Table 22 Correlations between the written performance measures and flow and anti-flow measures in the CT controlling for language proficiency ....................112
Table 23 Differences between the CSH and the CSC regarding correlations between the flow and anti-flow measures and the written performance measures in the TGT.................................120
Table 24 Differences between the CSH and the CSC regarding correlations between the flow and anti-flow measures and the written performance measures in the LGT.........................................................121
Table 25 Differences between the CSH and the CSC regarding correlations between the flow and anti-flow measures and the written performance measures in the CT ..................................................................................122
Table 26 Descriptive statistics of students’ written performance measures and flow and anti-flow measures in the TGT.................................................................126
Table 27 Descriptive statistics of students’ written performance measures and flow and anti-flow measures in the LGT.................................................................128
Table 28 Descriptive statistics of students’ written performance measures and flow and anti-flow measures in the CT.................................................................132
List of Figures

Figure 1 Maehr’s (1984) Personal Investment Theory .........................................................34
Figure 2 Csikszentmihalyi’s (1975, p. 49) Model of the Flow State.................................39
Chapter 1: Introduction

1.1 Rationale and aim of the dissertation

Writing is a major and necessary means for people to communicate. It also plays a major role in Second Language Acquisition (SLA) and has been suggested to be effective in promoting SLA (Brooks & Swain, 2009; Cumming, 1990; Gutiérrez, 2008; Polio, 2012; Williams, 2012). For Chinese students, especially for high school students studying in China, English writing is of more importance in the sense of the following two aspects: (1) It takes up about from 17% to 30% of the total score in their college entrance exam, varying in different provinces; and (2) In China, due to geographical reasons, English learners have few opportunities to use English by speaking it, especially in the western part of Chinese mainland, which leaves them the only way of using English by writing, since speaking and writing are generally considered the two productive skills. Besides, Chinese students, as English as a Foreign language (EFL) learners do not seem to be interested in writing (Yuan, 2012). There are numerous reasons why they do not like English writing, one of which might be that the traditional method of teaching writing is boring and ineffective (Chen, 2018; Liu, 2007). Studies on writing within Task-based Language Teaching (TBLT) might shed light on possible reasons behind students’ negative feelings towards English writing, and would be beneficial and might help improve writing instruction in China.

In this study, narrative writing instead of other genres of writings was chosen because it is important, in the sense that, first of all, it is a frequently used genre by people to narrate their life experiences in their daily communications with others. Additionally, narrative writing tasks are often used in EFL learning classrooms to develop and assess EFL written products. Moreover, in China, narrative writing is also part of the curriculum of writing classes no matter in which grade learners are, from elementary to college level, both in L1 and L2 (Kern, 2000). Therefore, narrative writing was chosen as the research topic of this study.

Since the 1980s, TBLT has gradually become more widespread in English language instruction. Students are asked to perform the tasks that require certain language knowledge and acquire the knowledge and develop their language skills during the process of performing the tasks (Brown, 1994). However, many studies within TBLT (e.g., Ellis, 2009; Kuiken & Vedder, 2007; Préfontaine & Kormos, 2015; Tavakoli & Foster, 2008) have been primarily focusing on the effects caused by task design and task implementation during task performance, but relatively few, if any, investigated the relationship between tasks and student engagement in task performance, which has been gradually recognized as important for language learning achievement.
(Lambert, Philp, & Nakamura, 2017; Lambert & Zhang, 2019). Nevertheless, the most recent studies have furthered this line of research on engagement in task performance by taking into consideration learners’ affective responses to the task and its effects on task performance (e.g., Lambert et al., 2017; Lambert & Zhang, 2019).

Affective characteristics, such as attitudes, motivation, and anxiety are crucial for language learning processes (Dörnyei, 2005). Furthermore, learners’ subjective feelings and responses (i.e., interest, attention) to tasks are believed as a manifestation of learner engagement (Lambert et al., 2017) and tasks of different characteristics might bring about different affective responses (i.e., interest, task-related anxiety, task motivation, and perceived success in task completion) to the task from the language learners, which might facilitate or impede learners’ engagement and their subsequent language development (Kormos & Préfontaine, 2017; Lambert & Zhang, 2019). Therefore, it seems meaningful and significant to explore the relationship between task characteristics, learners’ affective responses, student engagement, and task performance.

The present study will apply the concept of “flow” (Csikszentmihalyi, 1984) to refer to learners’ affective characteristics during task performance. Flow refers to a “peculiar dynamic state—the holistic sensation that people feel when they act with total involvement” (Csikszentmihalyi, 1975, p. 36). It is mostly experienced in autotelic activities like playing chess, playing tennis, singing, and rock-climbing which “give participants a sense of discovery, exploration, problem solution—in other words, a feeling of creativity and challenge” (Csikszentmihalyi, 1975, p. 30). However, flow occurs not only in daily life activities, but has also been found to appear in language learning classrooms (Czimmermann & Piniel, 2016; Egbert, 2003), especially when learners are doing language learning activities or tasks that give them control over what they are doing (e.g., Lambert et al., 2017; Lamber & Zhang, 2018; Phung, 2016; Qiu & Lo, 2016) or constraints to facilitate creative language use (e.g., Tin, 2011, 2012). The current study follows this line of research on flow in foreign language learning and investigates their role in EFL writing.

Taking the above three reasons into consideration, namely the importance of writing, especially its importance for Chinese high school students; students’ negative disposition toward English writing and the teaching method of writing; and the latest research trend of involving affective factors or “flow” in research language learning within TBLT, the current research aimed to enhance studies in the EFL writing area by investigating the relationships among task conditions (teacher-generated content condition versus learner-generated content condition), students’ “flow” experiences during the writing process, and students’ written performances.
Furthermore, many comparative studies have shown that the learning context of studying abroad, as opposed to staying at home has a significant impact on English learning. For instance, studying abroad has been proven to yield much more linguistic gains in learners’ oral production than studying at home (e.g., Dekeyser, 2007; Segalowitz, Freed, Collentine, Lafford, & Diaz-Campos, 2004). However, in these studies, the study-abroad condition is generally viewed as learning a language in a country where the target language is spoken by the majority. Nevertheless, few studies have focused on the effects of the non-target language-speaking foreign language context on L2 learning. Inspired by these studies, the current study also attempted to examine the role of the learning contexts, specifically studying English in a non-English speaking country, Hungary, and studying English in China, on Chinese students’ English writing.

In summary, the present research had four aims: (1) to investigate the relationship between task conditions and Chinese students’ English written performances. The tasks were the teacher-generated content task (TGT), which gave the students no control over the content to write, the learner-generated content task (LGT), which offered the students complete control over what to write, and the creative task (CT), which not only gave the students some control over the content to write, but also consisted of some constraints to encourage creativity and complexity in language use. In a word, they were different in the levels of learner agency/control on learners’ part and the potential to elicit creativity, with the TGT having the lowest level of control and the LGT having the highest, and the CT having the greatest potential to elicit creativity while the TGT and LGT having no such potential, (2) to investigate the relationship between task conditions and Chinese students’ flow experiences during the writing process, which also reflects students’ writing engagement, (3) to research the possible correlations between flow and Chinese students’ English written performances, and (4) to investigate the possible role that the learning contexts might play in Chinese students’ written performances.

The present study is significant in, first, trying to provide a systematic way to evaluate students’ engagement in writing and written performances, and giving references to teachers’ instruction of English writing so as to improve the effectiveness of both learning and instruction of English writing; second, shedding light on the relationships among task conditions, flow experiences and written performances, with the hope to provide teachers and course designers with insights for selecting and designing writing tasks for English writing instruction and English writing assessment; third, attempting to examine the role that the learning contexts might play in Chinese students’ English writing.
1.2 Research questions

The present study is a quantitative research, in which hypotheses are often formulated instead of using research questions. The reasons for not formulating hypotheses but using research questions in the present research are that, although there are some studies on the relationships of task conditions (i.e., the teacher-generated content condition and the learner-generated content condition realized by manipulating levels of learner agency/control on the task) and oral task performance, there are few studies on task conditions and writing, especially flow in writing and the relevance of task conditions to flow experiences in writing. This is a relatively new area. Although I do have some expectations, like learners might produce higher fluency in terms of the number of words if they are in flow while doing the task, it is impossible to expect very specific relationships between them, as well as between the other investigated flow and task performance measures (i.e., complexity and accuracy). Based on the research background and research goals stated above, the following four specific research questions (RQ) were proposed:

1. For Chinese students studying in Hungary (CSH) and China (CSC), how do task conditions relate to their written performances in terms of fluency, accuracy, and complexity?
2. For Chinese students studying in Hungary (CSH) and China (CSC), how do task conditions relate to their flow and anti-flow experience?
3. For Chinese students studying in Hungary (CSH) and China (CSC), what is the relationship between task performance measures and flow and anti-flow measures in different task conditions if students’ language proficiency is taken into consideration?
4. What role do the English learning contexts play in students’ written performances and their flow and anti-flow experiences in each task condition, that is, what kind of differences can be detected between Chinese students studying in Hungary (CSH) and China (CSC) in their written performances and their flow and anti-flow experiences in each task condition?

1.3 Overview of the dissertation

This section of the dissertation presents the overall structure of the study. Chapter 1 serves as an introduction to the description of the study that presents the rationale and aims, as well as the RQs of the dissertation.

Chapter 2 presents the theoretical background of the study. It addresses the fundamental terms of the study, namely engagement, personal investment, flow and anti-flow, TBLT, task, the measures of written products (including fluency, complexity,
and accuracy), and the related empirical studies. The chapter is closed by the discussion of the relation between the key terms and the re-emphasis on the significance of the present study.

Chapter 3 describes the research method, including the participants, the design and the piloting of the instruments, the data collection and the different measures applied. Finally, a brief description of the analysis of the writing samples and flow questionnaire data brings the chapter to an end.

Chapter 4 answers RQ 1, RQ 2 and RQ 3 in connection with the CSH. It presents findings with regard to their written performances, their flow and anti-flow experiences across different tasks, as well as the correlations between the two aspects in each task condition.

Chapter 5 answers RQ 1, RQ 2 and RQ 3 in connection with the CSC. It presents findings with regard to their written performances, their flow and anti-flow across different tasks, as well as the correlations between the two aspects in each task condition.

Chapter 6 answers RQ 4. It provides a general picture of the written performances of Chinese students with intermediate English proficiency as well as their flow and anti-flow experiences in different task conditions. In addition, this chapter also compares the differences between the CSH and the CSC in terms of their written performances and their flow and anti-flow experiences in each task condition to exam the role of study contexts in their task performances and their flow and anti-flow experiences.

Chapter 7 concludes the dissertation. It starts with a summary of the most important findings of the study followed by a discussion of the limitations. Finally, the pedagogical implications of the study and the suggestions for the future research are provided and discussed.
Chapter 2: Literature Review and Theoretical Background of the Research

2.1 Introduction

My research topic is Chinese teenage students’ written performances and their flow experiences in performing writing tasks. Given that the participants in the current research are first year high school students, the literature review is mainly on related studies conducted with high school students. This chapter aims at establishing the theoretical framework of the present research study. It first presents the current situation and problems of the teaching of English writing in high schools in China to show the necessity of researching writing. Then, the main constructs and theories that constitute the main focus of the present research, namely student engagement, the personal investment theory as well as flow theory are presented in detail, with an explanation of the interplay of them in classroom foreign language learning. This chapter is concluded by reviewing the empirical research on written narrative tasks to demonstrate what have been researched and found regarding written narrative tasks and empathize the significance of the present study.

2.2 The current situation and problems of the teaching of English writing in high schools in China

To study the teaching of English writing in high schools, it is important to understand what writing is. It seems tempting to think of writing as a process of making choices of syntactic structures and lexical items once a person wants to express some ideas. However, writing is not purely a simple activity of expressing ideas in words, but also involves “a set of distinctive thinking processes which writers orchestrate or organize during the act of composing” (Flower & Hayes, 1981, p. 366). As a productive skill, writing is different from speaking, which is the other productive skill, in the sense that speaking is a skill inherent in human beings and develops as they grow up through age, while writing must be learned. Just as Raimes (1983) and Myles (2002) argue that writing is not a natural extension of language learning, but a skill acquired through a series of systematic learning and practice, which suggests the importance of the teaching of writing. However, in high schools in China, many studies (e.g., Liu, 2007; Yu, 2017; Zhang, 2011) have shown that the current situation in the teaching of English writing is not beneficial for students and many problems may exist.

Liu (2007) used questionnaires and interviews to study the situation of the teaching of English writing in high schools in China. She discovered that the problems of the teaching of English writing lied in the following four aspects, namely the teachers, the
students, the textbooks, and the teaching methods. Specifically, in the case of the teachers, she found that they pay less attention to writing. Instead, they emphasize the teaching of the language itself, including enlarging vocabulary, memorizing and doing exercises of the sentence structures and the grammar, and rarely practicing writing in class. As regards the students, she pointed out that they lack interest in writing because they get bored of the traditional “students writing, teachers reviewing” pattern, and they are frustrated about not getting systematic training in writing. In terms of the textbooks, her study indicates that the contents of the textbooks are expected to be more varied and authentic, with the cultures of the English-speaking countries being included so as to improve students’ abilities of language use. Finally, concerning the teaching methods, she claimed that different writing methods (i.e., product-oriented method, process-oriented, and the genre-based method) should be combined in the practice of teaching writing in English in China.

The most recent studies on writing instruction in high schools in China also covered the four problematic areas proposed by Liu (2007). Yuan (2012) added that, in terms of the student aspect, students rarely practice writing outside school, indicating that they do not pay much attention to writing nor are interested in it. Cao (2011) pointed out that, concerning the teacher aspect, teachers have been ignoring the training of the English thinking pattern of the students and students rely on their mother tongue to translate and write English. Some other studies also mentioned this issue (e.g., Duan, 2015; He, 2013; Yu, 2017). Cui (2016) added to the teachers aspect that, due to the large size of the class and the busy teaching schedule, the teachers are not able to give detailed feedback on the students’ writings except for a general score, in which case the students would not realize their mistakes and shortcomings in writing so that they may not be likely to improve their English writing.

Chen (2018) applied questionnaire, interview and classroom observations to study the current situation of English writing instruction in high schools in China. She found that there were five problems with writing instruction in high schools: first, the teachers pay little attention to writing instruction; second, the teaching methods for writing instruction are too simple. For example, writing is often used as a way to practice and enforce the learning of certain language expressions and grammatical structures, as a result of which, the teaching of writing strategies and writing skills have often been ignored; third, there is rarely systematic writing instruction in regular classroom English teaching, which leads to little improvement of students’ writing abilities; forth, the teachers’ low professionalism might inhibit the practice of writing instruction and result in incompetence in setting appropriate writing goals and designing meaningful writing
activities; fifth, the students lack interest in writing. Chen (2018)’s findings are mostly compatible with Liu (2007) in the sense that they both reveal the problems with writing instruction in high schools from the aspects of the teachers’ attitude towards writing instruction, their teaching methods as well as the students’ attitude towards writing in English.

There are also several short articles written by high school English teachers (e.g., Chen, 2016; Zhang, 2011) based on their personal feelings, teaching experiences and self-reflection during the process of their teaching practice of English writing, and they hold similar ideas regarding the problems of the teaching of English writing in high schools in China.

From the above overview, we may infer that in the teaching practice of English writing in high schools in China, teachers are the controllers, in the sense that they decide the topics that the students can write about, and when, how and how much they can practice writing and how their writings can be evaluated and given feedback. Students seem to play a passive role, and their interest in writing, their needs, and difficulties in the teaching process of English writing seem to be not noticed or paid enough attention to.

What is more, some studies have been conducted on the differences in the college entrance examination composition from the composition proposition angle in the United States and China, especially the differences between the two countries in the forms and contents of the composition proposition in the college entrance examination, showing that the forms and contents of the composition in the college entrance examination in the United States provide more freedom for learners in terms of what to write and how to write it and encourage creative thinking, while those in China are mainly on the genres of argumentation and letter writing, with few on narration, and mostly with given titles and the required points to elaborate on (e.g., Chen, 2015; Zhang, 2012; Zhang, 2016). Creative writing is also an important part of high school English curricula in the United States (Bennet, 2019). There is no doubt that training experiences with certain genres or forms of writing such as creative writing leads to sophistication in coping with them and higher writing quality in the long run (Ren, 2005).

Given the problems with the current teaching of English writing, urgent changes are needed to be made involving all parties concerned. What also needs mentioning is that, in 2017, the Ministry of Education of the People’s Republic of China released the latest English Curriculum Standards of Senior High school and described that the key issue of education innovation is to change the traditional English course teaching method, which overemphasizes grammar and vocabulary, and fails to cultivate the
students’ practical language competence, to a new teaching method that takes into consideration the students’ interests, life experiences, and cognitive levels. It advocates student participation through various activities. This calling applies equally to Chinese students’ English writing instruction. The present study attempted to offer some insights into solving the problem concerning English writing and the teaching of it by discovering some ways to improve students’ interest and engagement in English writing and enhance their writing development.

2.3 Student engagement and language learning

In the last decades, a lot of research has been focusing on student engagement in different educational research realms, such as education theory, education policy, psychology, and most of these studies have revealed that student engagement has a significant effect on academic performance and achievement. This part of the dissertation presents the definition of student engagement and the ways to measure it, both in general education realm and language learning, as well as the main empirical studies on student engagement in EFL learning.

2.3.1 Definition of student engagement and measuring student engagement in general education realm

The term engagement was first used by Natriello (1984). He said, “Engagement exists when students are participating in the activities offered as part of the school program” (p. 14) while disengagement, the opposite end of the continuum, is defined as “the extent to which students refrain from participating in the activities offered as part of the school program” (p. 14). He argues that disengagement manifests itself in the American high schools as absenteeism, apathy, low-level participation in school, and delinquency (in the form of violence or vandalism). Natriello’s (1984) study is very instructive in unveiling the phenomenon of engagement and disengagement, suggesting that they depend on each other. However, it was disengagement that he explored the most in his studies, rather than engagement.

The construct of student engagement, which, in some studies, is called student involvement, was first put forward by Mosher and McGowan (1985). They defined student engagement as “the attitude leading to, and the behavior of participation in the secondary school’s program” (p. 14). They claim that student engagement is defined by multiple interactive factors that can be categorized into three dimensions: (1) societal, economic, community, and legal factors, (2) family and student characteristics (including psychological characteristics), and (3) school characteristics. Pace (1982) put
forward the theory of *Quality of Effort*. He believes that “Effort is a quality dimension in the sense that some kinds of effort are potentially more educative than others” (p. 2). According to him, in a college environment, the quality of effort is indicated by using college facilities (i.e., libraries, laboratories, student unions) and seizing important opportunities for personal and social development (i.e., contacts with faculty members, experiences in clubs, and organizations). Pace (1982) argues that it is not merely the time that students spend on academic work (indicated by the number of hours a week on study) that is related to progress, grades, and gains from learning, but the time during which they are highly engaged and making great effort that makes a big difference. Astin (1984) proposed the concept of student involvement. He defined student involvement as “the amount of physical and psychological energy that is devoted by the students to the academic experiences” (p. 1). In this view, indicators of a highly involved student are, for example, willingness to spend much time and energy in studying, active participation in student organizations, frequent interactions with faculty members and peers. Compared to Pace (1982), Astin (1984) noted the psychological side of student involvement. However, he seems to restrict this psychological side of student involvement as a kind of willingness to make an effort in learning. Nevertheless, this theory of student involvement distinguished itself from the traditional teaching theories in the sense that it highlights the initiative and active participation of students in the learning process, both physically and psychologically, while the traditional teaching theories consider and treat students as passive knowledge receivers. This theory indicates a positive relationship between the time and the effort that students put into learning and the learning outcomes.

Based on the theory of quality of effort and the theory of the student involvement, Kuh (1991) focused on the learning of undergraduates and undergraduate education and officially posited the concept of *student engagement*. They believe that student engagement has two characteristics: on the one hand, it emphasizes the amount of time and the degree of effort that students devote to learning and other kinds of education activities; on the other hand, universities should distribute learning resources, organize curriculum and provide learning opportunities and services so as to deepen student engagement and finally lead to quality education. Later, Marks (2000) defined student engagement as the psychological investment that refers to the students’ attention, interest, participation, and levels of effort in the learning process. This definition is compatible with Newman’s (1992) and considers student engagement as a two-dimensional construct: behavioral and emotional.
The most recent definition given by Kuh (2003) refers to student engagement as the time and energy that students put into educationally effective activities, both inside and outside the classroom, and the policies and practices that the institutions apply to elicit students’ participation in these activities. This suggests a shift in understanding engagement, from the previous social-psychological approach to engagement research to a perspective that includes socio-cultural and sociological features and processes. Fredricks, Blumenfeld, and Paris (2004) presented a comprehensive review of school engagement regarding its definitions, measures, precursors, and outcomes, and discussed the limitations of the existing research. They divided student engagement into three dimensions, namely the emotional/affective domain (i.e., interest, sense of belonging, the attitude toward learning), the behavioral domain (i.e., participation, academic learning time), and the cognitive domain (i.e., learning objective, self-adjustment).

In general education and school contexts, the most popular instrument concerning the measurement of student engagement is the National Survey of Student Engagement (NSSE) designed by Kuh (2001). However, this measure is too broad in scope and mainly addresses students’ general educational experiences. Later, researchers developed and applied the Student Engagement Instrument (SEI) to measure engagement (Appleton, Christenson, & Raschly, 2006; Betts, Appleton, Reschly, Christenson, & Huebner, 2010). SEI has been widely used by high school teachers to measure the academic or behavioral engagement of students in schooling in a general sense, which specifically manifest themselves by, for example, credits earned, extracurricular participation, homework completion, and attendance.

During the last two decades, inspired by findings on engagement in the psychological area, many researchers have started to borrow the method called the Experience Sampling Method (ESM) from psychology to measure student engagement at school. ESM was first developed by Larson and Csikszentmihalyi (1983) to measure flow. Individuals are given a signaling device, a piece of paper or a Palm Pilot (a personal digital assistant used to keep track of contact information). When they are asked open-and closed-ended questions on their cognitive and affective experiences during the day, they could answer through the device. They are asked to do this at several random times a day, and their scores are tracked over a period of time to measure their engagement in their everyday life (see Hektner, Schmidt, & Csikszentmihalyi, 2007 for a detailed description of ESM). This method has been implemented by researchers to study student engagement in educational areas (e.g., Csikszentmihalyi & Schneider, 2001; Shernoff et al., 2016).
When it comes to classroom student engagement, early studies mainly focus on student behaviors and classroom observations. However, the construct of engagement is multidimensional (Fredricks et al., 2004; Shernoff, 2013). Given that observations can only tell the observer information about the activities that people perform and the contexts in which they perform these activities, it can reveal no information about the cognitive and affective experiences that people are going through the moment in a certain context. In addition to measuring the behavioral dimension of engagement that is indicated by academic performance, its emotional/affective dimension should by no means be ignored, since they interact with each other during the learning process. To illustrate, Epstein and Mecartland (1976) argue that the evaluation of schooling should include quality aspects such as satisfaction and enjoyment with school experiences. This idea is also confirmed by empirical psychological research. For example, Philippe, Vallerand, and Geneviève (2009) found that people engage in activities that produce flow because these activities can elicit enjoyable emotions in them, as a result of which they are willing to invest time and energy, and keep passionate about them.

In summary, the above-reviewed measures of student engagement seem to focus on student engagement in the general educational realm; thus, they seem not very applicable for studies on students’ engagement in language learning context, which brings about the need to find feasible ways to measure engagement in language learning and specifically, in classroom language learning.

2.3.2 Definition of student engagement and measuring student engagement in language learning

With regard to language learning in classroom settings, it is Storch (2008) who first proposed the term of engagement with language to describe “the quality of learners’ metatalk” (p. 99). She claims that the engagement with language links to learners’ attention. Storch’s definition of engagement seems limited in the sense that it includes only the cognitive dimension of engagement in language learning. Drawing on Storch (2008), Svalberg (2009) defined the engagement with language or Engagement as “a cognitive, and/or affective, and/or social state and a process in which the learner is the agent and the language is the object and maybe the vehicle (means of communication)” (p. 244). According to Svalberg (2009), cognitive engagement refers to the individual’s being alert, paying attention and constructing their own knowledge; affective engagement means that the individual confronts with the language-related activities positively, purposefully, willingly, and autonomously; and social engagement refers to the individual’s interaction and initiative during the language-related activities. Svalberg
(2009) offered a principled and manageable framework to identify the characteristics and grade the level of engagement (see Table 1 and Table 2). Svalberg’s (2009) framework of “engagement with language” or “Engagement” (p. 244) is very detailed and useful for measuring engagement in connection with general language exercises and activities. However, Svalberg (2009) seems not to suggest ways of how to put the framework of engagement with language into practice, although it seems to me that the best way is still the class observation, which, nevertheless, is criticized for failing to tap into the emotional states at the moment the learning happens.

Table 1

*Key characteristics of engagement with language*

<table>
<thead>
<tr>
<th>Key characteristics</th>
<th>Cognitive</th>
<th>Affective</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td>State: heightened</td>
<td>State: positive orientation towards the language, the interlocutor, and/or what they represent</td>
<td>Process: maintaining willingness to interact with the interaction</td>
<td>State: behavioral readiness to interact</td>
</tr>
<tr>
<td>alertness, focused attention</td>
<td>Process: focusing reflection and problem solving</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note.* From Svalberg (2009)
Table 2
Criteria for identifying engagement with language

<table>
<thead>
<tr>
<th>Cognitive</th>
<th>Affective</th>
<th>Social</th>
</tr>
</thead>
<tbody>
<tr>
<td>How alert? (Does L seem energetic or lethargic?)</td>
<td>How willing to engage? (Is L withdrawn or eager to participate?)</td>
<td>How interactive (Verbal and other behaviors? Does L use interaction for learning?)</td>
</tr>
<tr>
<td>Does L seem to notice language/interaction features?)</td>
<td>How focused? Is L’s attention on the language (as object or medium) or not? (Does L’s mind seem to wander?)</td>
<td>How supportive? (Verbal and other behaviors? Does L engage in negotiation and scaffolding?)</td>
</tr>
<tr>
<td>How reflective? How critical/analytical? Is L’s reasoning inductive or memory/imitation based?</td>
<td>How autonomous? (Is L’s behavior dependent or independent?)</td>
<td>Leader or follower? (Are L’s interactions reactive or initiating?)</td>
</tr>
<tr>
<td>Does L notice and reflect or simply react? (Does L compare? Does L ask questions? Does L infer/draw conclusions?)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note. From Svalberg (2009)*

Recently, in EFL learning context, student engagement has been studied within the task-based approach. For example, Philp and Duchesne (2016) defined engagement in task performance (or task engagement) as “a state of high attention and involvement, in which participation is reflected not only in the cognitive dimension, but in social, behavioral, and affective dimensions as well” (p. 3). They believe that the four dimensions overlap, interact, and manifest themselves differently in different contexts.

Here, it is necessary to present the definitions of and the measures for different dimensions of student engagement in language learning in general and EFL learning in particular. (1) Behavioral engagement reflects learners’ participation, indicated by how much effort learners put forth, how persistent and actively involved they are (Philp & Duchesne, 2016). Previous related research assessed behavioral engagement by time spent on the task (Gettinger & Walter, 2012), and the number of words or turn-takings in
communication (Dörnyei & Kormos, 2000; Lambert et al., 2017; Mozgalina, 2015; Phung, 2016).

(2) Cognitive engagement is defined as attention and mental effort that learners maintain consistently (Helme & Clarke, 2001), manifested by sharing and evaluating ideas, asking questions, explaining and offering information, arguing, suggesting, reasoning, modifying, and confirming (Lambert et al., 2017; Lambert & Zhang, 2019; Philp & Duchesne, 2016), by the number of language-related episodes (LREs) (Baralt, Gurzynski-Weiss, & Kim, 2016; Svalberg, 2009), which, according to Swain (1996), refers to “any part of the dialogue in which students talk about the language they are producing, question their language use, or other, or self-correct” (p. 70). Cognitive engagement can also be measured by clause counts that express suggestions, propositions, and reasons, and move counts that are connected to the negotiation of meaning demonstrated by confirmation checks, clarification requests, and metalinguistic exchanges (Lambert et al., 2017; Phung, 2016).

(3) Social engagement refers to learners’ affiliation and willingness to participate in the conversation and highlights reciprocity and mutuality (Damon & Phelps, 1989; Storch, 2008). It is measured by the number of backchannelling produced (Lambert et al., 2017; Lambert & Zhang, 2019; Philp & Duchesne 2016; Phung, 2016).

(4) Affective engagement is difficult to define. It is often used interchangeably with emotional engagement. The word affect is “used to label all factors that determine a learner’s disposition toward what goes on in the second language classroom” (Lambert, 1998, p. 85). Pekrun and Linnenbrink-Garcia (2012) used the term, affective responsiveness, rather than affective engagement. They argue that affective response is not a separate dimension of engagement; rather, it is intrinsic to other dimensions of engagement. However, Lambert et al. (2017) equaled the emotion and affect as the same concept in his study. Phung (2016) used the term emotional engagement and said that it generally referred to learners’ positive emotions such as interest, excitement and enjoyment and negative emotions such as boredom, anxiety, and frustration. Emotional engagement is measured in previous related studies by learning learners’ post-task reports about the task and analyzing learners’ interactions (Baralt et al., 2016; Imai, 2010; Mercer, 2015; Phung, 2016). Moreover, as suggested by Lambert et al. (2017), affective engagement can be measured by flow (see section 2.5 for a description of flow theory and the justification of measuring engagement by flow).

In China, there has been some research on student engagement on the basis of the special Chinese context. Kong (2003) published his famous work Student Engagement in the Mathematics Teaching Process, in which he classified student engagement into
three dimensions, namely behavioral engagement, emotional engagement, and cognitive engagement. He referred to behavioral engagement as concentration in class, devotion, and extracurricular time spent on learning. He defined emotional engagement as emotional experiences during learning activities. As for cognitive engagement, according to Kong (2003), it refers to the learning strategies applied by students for learning. Kong (2003) also designed a questionnaire to assess student engagement. Although the questionnaire was based on classroom mathematics teaching, it has been frequently applied and adapted by Chinese researchers to study student engagement in Chinese students’ EFL learning.

Recent research on foreign language learning considers student engagement in classroom learning as a multidimensional construct that includes behavioral, cognitive, social, and emotional/affective dimensions. In the present study, Philp and Duchesne’s (2016) definition of engagement was adopted, but with a modification based on the needs of the present research. In the present study on individual writing, student engagement referred to as a state of high attention and involvement. A highly involved writer showed interest in the task and paid high attention to his or her writing, and experienced flow. The reason for modifying the definition of student engagement given by Philp and Duchesne (2016) is that their definition is more applicable to oral interactive communication, not for individual writing, which is very different from oral production in several perspectives. First of all, no social engagement in the sense of immediate direct verbal communications is involved in individual writing; besides, behaviorally, it is hard to decide whether students are engaged or not when they are not actually writing. For example, seemingly wandering-off might still mean that they are being engaged in a way of thinking. In this sense, writing is an interwoven process composed of writing behavior, cognitive thinking, and emotional experience, and it seems impossible to approach and analyze the three engagement dimensions separately considering that the three measures overlap and correlate (Philp & Duchesne; Stroud, 2017; Swain, 2013). This is also supported by the argument that student engagement is dynamic, fluctuating, context-dependent and interactive (Shornoff et al., 2016). Nevertheless, it seems that student engagement in writing can get inspiration from flow theory, where flow is defined as a state when the individual is intensely concentrated and highly interested in intrinsically enjoyable activities (Seligman & Csikszentmihalyi, 2000). Meanwhile, student engagement is conceptualized as an experience characterized by a simultaneous occurrence of intense concentration, interest, and enjoyment (Marks, 2000; Shernoff, 2013). Therefore, student engagement or student engagement in writing
can be conceptualized from the perspective of flow and can be investigated with the help of flow theory.

In EFL learning, however, there are only a few studies combining student engagement in writing with flow. The present study attempted to extend research in this aspect by applying flow theory to investigate student engagement in EFL writing, which was also one of the original features of this study. Since student engagement, especially the cognitive and affective dimensions of student engagement are intertwined with the construct of flow, the measurement of student engagement in writing in the present study measured by flow will be illustrated in detail in section 2.5.3 together with the description of flow theory.

2.3.3 Main studies on student engagement in foreign language learning

There is a lot of research on student engagement in the general education realm and foreign language learning area. They are mainly concerned with the factors affecting student engagement and its contribution to learning achievement and the correlations between different dimensions of student engagement.

In the general educational context, student engagement has been found to be affected by teacher-related aspects (i.e., teachers’ teaching style, teacher behaviors) (e.g., Robert et al., 2012; Voelkl, 1995), gender (e.g., Harper, Carini, Bridges, & Hayek, 2004; Pavlidou, 2003), and peer relationships and classroom learning conditions (e.g., Cook, Deng, & Morgano, 2007; Juvonen, Espinoza, & Knifsend, 2012; Nelson & DeBacker, 2008). In addition, there are some studies on the influence of emotions (i.e., interest, enjoyment, anxiety) on student engagement. For example, it has been found and argued that positive emotions are likely to enhance student engagement at school and are closely related to academic achievement (e.g., Ainley, 2012; Ainley, Corrigan, & Richardson, 2005; Reschly, Huebner, Appleton, & Antaramian, 2008; Valiente, Swanson, & Eisenberg, 2012) because positive emotions help simulate, build and broaden the resources of the brain to solve the problems confronted (Bolito, Carter, Hughes, Ivanic, Masuhara, & Tomlinson, 2003; Fredrickson, 2004; MacIntyre & Gregersen, 2012), while negative emotions work the other way round (e.g., Linnenbrink, 2007; Rice, Levine, & Pizarro, 2007). However, some researchers have suggested that negative emotions may facilitate learning performance as well (e.g., Gable & Harmon-Jones, 2010).

In a foreign language learning context, gender and language proficiency have also been found to affect student engagement (Schunk & Mulle, 2012). Many recent empirical task-based studies focus on the effects of task factors on student engagement
Since most of these task-based studies specifically focus on student engagement in oral task performance as opposed to it in the writing tasks that the present study is concerned about, in order to show a full picture of how researchers conducted their studies on learners’ engagement in classroom foreign language learning, and what they have found, these studies will be reviewed in detail together with a review of engagement in writing in section 2.8.2.

Some studies in China have also focused on student engagement in English learning in a general sense. Guo (2006) published a descriptive article in which he found that three reasons contributed to the low engagement of college students in EFL learning, and they were the following: (1) anxieties like being afraid of speaking English, making mistakes and losing face in front of the whole class; (2) a lack of confidence caused by the vicious circle of less practice and less improvement; (3) the cultural aspect, specifically the golden mean of the Confucian school that emphasizes behavioral conformity with others. Many studies have been carried out on student engagement in English learning with high school students. Dong (2018) found that language proficiency affected engagement in junior students’ English learning. High achievers were more behaviorally and cognitively engaged than medium achievers. Affectively, medium achievers went through more anxiety and boredom than high achievers. Medium achievers were more behaviorally and cognitively engaged than low achievers and they were more anxious while low achievers felt more bored while learning English. Zhao (2015) also studies student engagement in high school students’ English classroom teaching by applying questionnaires, interviews with the teacher, and classroom observations. It was found that language proficiency affected behavioral, cognitive and affective engagement, among which cognitive engagement was affected by language proficiency the most, indicating that students of higher scores were more engaged in English classroom learning than those with lower scores. Fan (2014), by using questionnaires, interviews with the students and the teachers, and classroom observations, conducted a study on Senior High school students’ English classroom engagement under the teaching model of TBLT and found that teacher-related factors (i.e., teaching style), student-student relationships (i.e., peer pressure), classroom environment, and task factors (i.e., a good match with students’ English proficiency level) affected students’ participation in classroom English learning activities.

As for the contribution that student engagement possibly makes to EFL learning achievement and the correlations between different dimensions of student engagement, Zheng (2010) applied questionnaires, interviews, and class observations to study high
school students’ participation in the English class. It was found that cognitive learning strategies and individual positive affective experiences were significantly positively correlated with learning outcomes. Zhang (2017) used interviews to study the influence of students’ classroom engagement on the learning achievement of English-major undergraduates. The interviewees reported that their classroom engagement had little relation to their learning achievement, but they believed their cognitive engagement and affective engagement affected their behavioral engagement. Yang (2018) conducted a study on junior students’ English learning engagement and the correlation between English learning engagement and English achievement through questionnaires and interviews. He compared students’ grades on English from the mid-term and final-term exams (altogether 2-month study) and found that, generally, there was a significant correlation between students’ engagement and their learning achievement. Specifically, the correlation between emotional engagement and English achievement was the strongest while the correlation between cognitive engagement and English achievement was the poorest. Pu (2017) conducted a study on high school students’ classroom engagement and their EFL learning achievement with the help of engagement questionnaires, class observations, and interviews. He found that there was a positive correlation between behavioral, cognitive and affective engagement and English learning achievement.

2.3.4 Summary

From the above literature review, we can see that studies on student engagement in the foreign language learning or EFL learning context mainly focused on classroom English learning engagement or English learning in the larger school community. However, English learning is a multi-faceted activity that includes the acquisition of different language skills, such as listening, speaking, reading, writing, and translating, indicating that the definitions and the measures of student engagement in these different language aspects might differ from those defined and measured in general classroom or school learning contexts. In addition, the factors that affect student engagement in specific language learning activity (i.e., speaking, writing) and the correlations between student engagement and the learning outcomes of these activities might differ as well. Moreover, recent studies on student engagement in foreign language learning or EFL learning context are under the framework of TBLT and offer some pedagogical implications for foreign language or EFL teaching and learning (e.g., Philp & Duchesne, 2016; Phung, 2016). However, these studies have mainly been conducted on oral tasks. Therefore, the present research adopted Philp and Duchesne’s (2016) definition of
engagement and referred to student engagement as a state of high attention and involvement, and followed the trend and took a task-based approach to student engagement in EFL writing, with the attempt to find the task factors or task conditions that might relate to student engagement in writing, on the one hand, and investigate the relationships between student engagement in writing and their written performances, on the other.

2.4 Maehr’s (1984) personal investment theory

This section presents Maehr’s (1984) personal investment theory, including its contents, development, and application in the general education field, and most importantly the relevance of the Maehr’s (1984) personal investment theory for student engagement and foreign language learning.

2.4.1 Description of the personal investment theory

Maehr (1984) put forward the personal investment theory as illustrated in Figure 1. Personal investment refers to “the extent to which learners are willing to invest their personal resources (i.e., time, talent, energy) into a given task or situation” (as cited in Lambert, 1998). It seems natural that those who put this type of investment into learning or tasks are expected to be more motivated and more likely to succeed. According to the theory, it is the meaningfulness of the task or the situation that results in learners’ investment in the task or the situation. Maehr (1984) argues the meaningfulness that educational tasks and situations have for the learner should have the following four characteristics, namely sense of competence or qualification, perceived behavioral opinions, standards of success, and goal-orientation like task goals, communicative goals, social solidarity goals, and extrinsic goals. He argues that the degree of meaningfulness that a task or a situation has for a learner will, in turn, determine learners’ willingness to invest his or her personal resources (i.e., time, talent, energy) into completing it.
Maehr’s (1984) Personal Investment Theory

**Factors Determining Meaningfulness**

1. Reactive:  
   a. Prior experience  
   b. Socio-cultural context  
2. Potentially proactive:  
   a. Task design  
   b. Roles & expectations  
   c. Information provided

**Meaningfulness**

1. Sense of competence or qualification  
2. Perceived behavioral options  
3. Standards of success  
4. Goal-orientation  
   a. Task goals  
   b. Communicative goals  
   c. Social solidarity goals  
   d. Extrinsic goals

**Investment of Personal Resources**

1. Time  
2. Talent  
3. Energy

**Performance Effects**

1. Direct:  
   a. Direction  
   b. Persistence  
   c. Continued motivation  
2. Combined:  
   a. General activity level  
   b. Level of performance or achievement
According to Maehr’s (1984) personal investment theory, five factors determine the meaningfulness of educational tasks and situations for a learner, namely (1) prior experience; (2) socio-cultural context; (3) task design; (4) roles and expectations; and (5) information provided. The first two factors are out of the teacher’s or the syllabus designer’s control since they are already fixed before the students enter the classroom; thus, they cannot be modified within the temporal and institutional constraints of most L2 programs. Nevertheless, the latter three factors are open to modification. They are important since they are “related to fluctuations in performance via materials design and classroom management” (Lambert et al., 2017, p. 2). In other words, the teacher or the task designer can alter and improve the meaningfulness of the task or the class situation by altering and making good use of these three factors.

According to Maehr (1984), personal investment in the task brings about five motivational behaviors: (1) direction, the extent to which a learner does the assigned tasks rather than something else, (2) persistence, the duration that a learner is willing to work on these tasks, (3) continued motivation, the extent to which a learner returns to previously encountered tasks of his or her own accord, (4) general activity level, the amount of mental and physical activity that learners engage in during the task performance, and (5) level of performance or achievement, the amount of learning that a learner invests into tasks. These five motivational behaviors are also considered by Lambert (1998) as manifestations of a learner’s being personally invested in the learning process. To conclude, on the basis of Maehr’s (1984) personal investment theory, it seems that, to motivate learners to invest more in the task or the situation, it is key to design tasks that are meaningful for learners.

2.4.2 The relevance of Maehr’s (1984) personal investment theory for student engagement and foreign language learning

Maehr’s (1984) personal investment theory is based on human motivation research, intending to account for general educational performance. However, even though it is not specific to foreign language learning, it “provides a theoretically-motivated foundation for claims about the relationship between task factors and affective performance variables” (Lambert, 2002, p. 14). Therefore, the theory seems applicable to L2 or foreign language task design to facilitate student engagement in task performance.

Maehr’s (1984) personal investment theory is very much in line with SLA theories in their views on student engagement in the L2 learning or engaged L2 learners. For example, Crookes and Schmidt (1991) defined classroom motivation in language
learning as learners’ willingness to become “productively engaged in learning tasks, and sustain that engagement, without the need for continual encouragement and direction” (p. 490). Crookes and Schmidt (1991) suggested that motivation controls engagement in and persistence with the learning task. This notion of engagement suggested by Crookes and Schmidt (1991) seems to be consistent with Maehr’s (1984) notions of direction and persistence, which are considered as two manifestations of learners’ personal investment in tasks. These effects (i.e., learners’ direction and persistence) brought about by personal investment in the learning process from learners are consistent with the characteristics of engaged language learners. For example, Reeve (2012) described engaged learners as “actively enriching” during the task interaction process, not just “passively receiving” (p. 151) and Svalberg (2009) described engaged learners as having a “willing, and autonomous disposition” towards the language or the language-related activities and being “interactive and initiating socially” (p. 247). In other words, learners’ engagement in learning languages can be manifested by the degree to which they invest their personal resources (time, talent and energy) in learning the language, indicating that the concept of engagement in language learning and the concept of investment in the personal investment theory (Maehr, 1984) are consistent.

In one of the latest studies on student engagement in language learning tasks, Philp and Duchesne (2016) classified student engagement into four dimensions, namely behavioral, emotional, cognitive and social engagement. Maehr’s (1984) personal investment theory is in accordance with Philp and Duchesne’s (2016) three dimensions of engagement (social engagement excluded), with his notion of “direction” and “persistence” corresponding to the behavioral engagement and the notion of “general activity level” partly corresponding to cognitive engagement. Therefore, it is safe to say that Maehr’s (1984) personal investment theory is largely consistent with the various definitions of engagement in EFL in the sense that they both highlight learner agency (Lambert et al., 2017).

Regarding the relevance of Maehr’s (1984) model for L2 or foreign language education, according to Lambert (1998), two issues should be taken into consideration: on the one hand, identifying and operating specific task design factors that are hypothesized to promote personal investment; on the other, claiming the measurements for the effects of personal investment on learners’ L2 production. Regarding the first issue, three factors of Maehr’s (1984) personal investment theory, namely task design, roles and expectations and information provided are relevant to L2 education since they are under the direct control of teachers and syllabus planners, and are the starting point for designing foreign language learning tasks. For the second issue of measuring a
Lambert et al. (2017) offered an example in their study on the effects of a specific task condition (the learner-generated content condition) on Japanese English learners’ engagement in language production. They measured three types of engagement in terms of learners’ language performance or performance behaviors. Specifically, behavioral engagement was measured in terms of the number of words produced, the amount of time invested in task performance; cognitive engagement was measured in terms of the number of clauses used to expand on the semantic content, such as suggestions, propositions, elaborations, reasons, and opinions, as well as the number of moves related to the negotiation of meaning; and social engagement was measured in terms of the number of backchannelling produced. Although Lambert et al.’s (2017) method of measuring task engagement is comprehensive and manageable, it focused on oral communicative tasks, the performance process of which is available to observe. Along with the language product, the researchers can get a clear picture of learners’ engagement during task completion. However, in individual writing, it seems impossible to use this method to measure learners’ engagement, especially their cognitive and emotional engagement. Therefore, it seems necessary to find a different method to measure learners’ engagement in writing, which also constitutes one of the research aims of the current study.

As for the role played by personal investment and engagement in L2 learning, Robinson’s (2001a, 2005, 2007) model suggests that task performance and L2 learning are affected by cognitive task demands through the retrieval of interlanguage resources, such as their existing knowledge and experiences about the task activities as well as their existing language knowledge that might be needed to complete the task. This retrieval process requires effort and engagement from the learner and can be manipulated by task designing. Therefore, we assume if we can manipulate the retrieval process of these interlanguage resources by manipulating the meaningfulness of the task, then task performance of learners are expected to be affected accordingly through their engagement and investment during task performance since tasks with different meaningfulness to the learner require different degrees of effort during the retrieval (Maehr, 1984). In other words, tasks that differ in meaningfulness to the learner are assumed to require different cognitive effort for interlanguage resource retrieval and lead to different personal investment on the learner’s part, thus leading to different task performance. Moreover, on the basis of Maehr’s (1984) personal investment theory, the meaningfulness of the task is the key that task designers or teachers should consider when designing tasks for classroom L2 or foreign language learning. The word meaningful suggests that teachers and task designers should design tasks based on
information that learners will find meaningful, rather than what they feel will be meaningful or important to learners (Lambert, 1998). In other words, to facilitate good task performance, what matters most for L2 pedagogical task design is to provide “personally-invested” (Lambert, 1998, p. 95) task conditions in which learners are active in performing the task. The effects of these task conditions can be measured by how much the learner is engaged in performing the task, as how Lambert et al. (2017) did in their study (see section 2.3.2 for the description of the ways of measuring student engagement in language learning).

On the basis of the above literature review on personal investment theory as well as its relevance to student engagement and foreign language learning, it can be argued that tasks characterized by different features might lead to differences in students’ engagement in performing the task. Therefore, it would be interesting to investigate if and how the task features in terms of the levels of control on learners’ part and the potential to elicit creative language use included in the task would relate to the students’ written performances and their writing engagement.

2.5 Flow theory

This part of the dissertation introduces flow theory and its application in foreign language learning as well as presents its relevance to Maehr’s (1984) personal investment theory and student engagement in the present study.

2.5.1 Description of the flow theory

Flow is a term used in positive psychology, referring to “a peculiar dynamic state-the holistic sensation that people feel when they act with total involvement” (Csikszentmihalyi, 1975, p. 36) and “what people feel when they enjoy what they are doing when they would not want to do anything else” (Csikszentmihalyi, 1989, p. 183). According to Csikszentmihalyi (1975), flow is characterized by: (1) the tasks are feasible for a person; (2) there is a limited stimulus field for a person to concentrate, leading to a “narrowing of consciousness” or a “giving up the past and the future” (Maslow, 1971, p. 63-65); (3) a person goes through “loss of ego”, “self-forgetfulness”, “loss of self-consciousness”, and even “transcendence of individuality” and “fusion with the world” (Maslow, 1971, p. 65, 70), becoming more intensively aware of internal processes; and (4) a person is in control of his actions and the environment, sensing that he is skillful enough to meet the environmental demands. Csikszentmihalyi (1975) posited the Model of the Flow State (see Figure 2). The model shows that if challenges for action are greater than the actor’s skills, worry occurs; if skills increase but
challenges do not, a state of boredom arises; and anxiety appears when the degree between the challenges for action and the actor’s skills greatly differs. In other words, flow occurs “when challenges for action are in balance with the actor’s skills” (Csikszentmihalyi, 1989, p. 49), in which context, people “report feeling more active, alert, concentrated, happy, satisfied, and creative although not necessarily more cheerful or sociable” (Csikszentmihalyi & LeFevre, 1989, p. 816).

Figure 2
Csikszentmihalyi’s (1975, p. 49) Model of the Flow State

Flow is important and beneficial for people’s life and growth. Csikszentmihalyi (2003) proposed the model of the Growth of Complexity through flow within the framework of flow theory. This model suggests that learning is a spiral process in which flow experiences stimulate learning and new challenges appear during the learning activities. To sustain flow, learners have to develop greater levels of skill to master the challenge. Once they master the challenge, they continually identify a more complex challenge and create an ideal match for their skill in order to experience flow. Flow, thereby, enhances and evokes development and growth. Flow, not only benefit people’s lives in a general sense, but also contribute to their learning.

2.5.2 Flow theory and foreign language learning

Although the concept of “flow” first appeared in positive psychology, it has always been applied within the educational context. Egbert (2003) claims that flow also exists in the foreign language classroom. She, on the basis of flow theory by Csikszentmihalyi (1975, 1989), described the hypothetical relationship between flow and language learning. Egbert (2003) divided task-specific flow into four dimensions:

(1) The balance between challenge and skill that makes the participants willing to explore and complete the task;
Attention from the participants on the task. Egbert (2003) suggested that unintentional focus or attention is closely related to flow and may contribute to language learning by supporting higher fluency;

Interest from the participants derived by finding the task intrinsically interesting or authentic. This is in line with Maeher’s (1984) claim that learners are more likely to invest time and effort if they find the task meaningful and Hidi and Renninger’ (2006) definition of individual interest which can predict whether someone would be willing to re-engage with a certain task. Besides, concerning task interest, Hidi (1990) claimed that interest determines how persistent and how attentive learners can be with the task. Egbert (2003) also argues that creative and open-ended tasks may get the students interested in the task and inspire them to actively confront the task;

A perceived sense of control. Control, as defined by Egbert (2003), mainly refers to the control over the learning situation in an interactive learning situation. While, in the present study, control not only meant a perceived sense of the overall control over the completion of the task, but most importantly, it included the agency and freedom given to the students in terms of what to write and how to write it during the writing process. Control in this sense has been argued to be beneficial in creating a positive emotional environment and encouraging students to write about what of personal interest, which helps foster student engagement (Bruning & Horn, 2000).

Flow, as a state of feeling characterized by interest, attention, and enjoyment, has usually been measured by ESM in psychology, as mentioned in section 2.3.1. However, as for measuring of flow in foreign language learning, Egbert (2003) designed a seven-point Likert scale questionnaire to assess learners’ flow in classroom foreign language learning on the basis of the Four-Channel Flow Model (Csikszentmihalyi, 1975). The questionnaire includes 14 items that are related to the four dimensions of flow (challenge and skill balance, attention, interest, and control). On the basis of Csikszentmihalyi’s (1975) theory and Egbert’ (2003) flow questionnaire, Czimmermann and Piniel (2016) proposed a more comprehensive five-point Likert scale to measure flow which, besides the already mentioned four flow components, also includes three anti-flow components, namely apathy, anxiety, and boredom. Their flow questionnaire includes 24 items, with interest, challenge-skill balance, control, boredom, apathy each having 4 items, attention having 3 items, and anxiety having 1 item. The questionnaire provides a comprehensive framework for research on task-specific flow in the EFL context. Since Egbert (2003) and Czimmermann and Piniel (2016) have not described the three anti-flow dimensions, it is a necessity to briefly introduce and explain the three anti-flow aspects in the EFL context here.
(1) Anxiety in psychology refers to an unpleasant emotional state when one feels tense and nervous (Spielberger, 1972). MacIntyre and Gardner (1994) defined language anxiety as the feeling of apprehension and fear that are specifically associated with L2 or foreign language contexts when a learner is expected to perform language output, like in speaking or writing. Studies have shown that foreign language anxiety exerts a negative effect on language learning and communication (Dewaele, 2002; Robinson & Gilabert, 2007). However, there has been suggestion that anxiety might also have facilitating effects on foreign language learning (Dörnyei, 2005; Hewitt & Stephenson, 2012).

(2) Boredom is associated with decreased attentiveness and slower thinking processes, leading to being less able to control attention, vigilance, thus making performance decline. Generally, boredom has been considered to arise by a mismatch between the students’ high ability and their low expectations towards the task or the activity. Students are also likely to become bored when they are in a situation where there is a lot of information to deal with or they feel a lack of ability to solve the task or to follow the on-going discussion (Lee Do & Schallert, 2004; Pekrun, Goetz, Daniels, Stupnisky, & Perry, 2010).

(3) Apathy is generally indicated by reduced motivation and passivity about the task a person is doing. Apathy is considered as the antithesis of flow caused by failing to sense challenges and abilities (Piniel & Albert, 2017). Apathy has been found to be negatively related to language learning motivation and self-efficacy (Piniel & Albert, 2017). Nevertheless, this aspect has been under-researched in the foreign language learning context.

In conclusion, foreign language learning generates both positive and negative emotions as the process unfolds. In L2 or EFL learning, many researchers have been conducting studies on emotions (positive emotions and negative emotions) or affect (Dewaele, Chen, Padilla, & Lake, 2019; Dewaele & MacIntyre, 2014; Horwitz, 2010; MacIntyre & Gregersen, 2012). However, studies on flow and language learning are indeed scarce (see Piniel & Albert, 2019 for the theoretical overview of flow in language learning). Besides, investigating task-specific affective factors (i.e., situational interest, anxiety) might help understand students’ engagement in the tasks. Although emotions, more specifically task-specific emotion of flow have already been studied in connection with oral tasks (Kormos & Prefontaine, 2017), few, if there is any, have been done in EFL writing. Therefore, the present study aimed to take a comprehensive look at task-specific flow and anti-flow in EFL writing and reveal its relation to students’ written task performance.
2.5.3 Summary: relationships among flow theory, personal investment theory and student engagement in classroom foreign language learning

First of all, with regard to the relationships between flow and engagement, based on the above literature, we know that in foreign language learning classrooms, learners’ engagement in the task is characterized cognitively by attention and concentration, behaviorally by the effort and persistence in task completion; socially by learners’ affiliation and willingness to participate in the conversation, and emotionally by learners’ affective and emotional responses to the tasks. It seems that flow and engagement conceptually overlap. Steele and Fullagar (2009) categorized the three similarities that flow and engagement share: (1) the two concepts both have cognitive, affective, and physical components, and are both enjoyable experiences affectively and entail physical innervation; (2) both concepts are linked with states of extreme concentration and involvement in the activity; and (3) they are states that are intrinsically motivating. Steele and Fullagar (2009) claim that these similarities between the two constructs make it legitimate to apply flow theory to understand engagement. However, Steele and Fullagar’s (2009) claim is based on flow and engagement in general academic activities.

When it comes to specific tasks in classroom language learning, optimal experience or flow termed by Csikszentmihalyi (1984) has been described as the “ultimate task engagement” (Philp & Duchesne, 2016, p. 59). This indicates that when ultimately engaged in the task, learners tend to experience flow. Based on flow theory, a learner is experiencing flow when they pay attention to, concentrate on, feel interested in, and enjoy an activity (Csikszentmihalyi, 1997). Therefore, it is key to know what task characteristics can generate flow. Maehr’s (1984) personal investment theory offers the answer to this question. According to Maehr (1984), learners are more likely to be engaged when they find the task interesting and meaningful; there, they are more likely to invest time, energy, and effort, which are likely to bring about optimal experience or flow. In this sense, flow can be used to understand learners’ task-specific engagement in foreign language learning. The stronger the flow that the students experience, the more engaged in the task they are.

In the present study on individual writing, it seems that students’ behavioral engagement, cognitive engagement, and affective engagement are involved. Since flow is a state of dynamic mental states, it may influence and also be manifested by students’ behavior, cognition, and emotional experiences at the same time during their task performance. In other words, the present study considers the variables contributing to flow as a form of students’ writing engagement, with the attempt to provide a new way to get access to student engagement in writing and measure it.
Concerning the relationship between flow and personal investment in the task, it is acknowledged that task or activity features, such as challenges and skill balance, affect learners’ flow or emotions (i.e., interest, attention, anxiety). These psychological factors, in turn, affect learners’ linguistic performance on the task (Aubrey, 2017; Egbert, 2003). This is consistent with Maehr’s (1984) personal investment theory which also highlights the crucial role of task features in motivating learners’ investment of time and effort and their engagement in the task. That is to say, task features are likely to affect the flow experiences as well as student engagement in the task. According to Maehr (1984), a task that is meaningful for learners enhance their engagement in the task. Therefore, it is important for task designers to explore what task features or task characteristics elicit flow or emotional experiences that can facilitate learners’ engagement in the task.

The present study hypothesized that writing tasks features (learner agency, potential for creative language use) would possibly influence students’ writing engagement measured by their flow experiences, and aimed to find their roles in students’ written performances as well as their flow experiences.

2.6 Task design

Task refers to “an activity in which meaning is primary; there is some relationship to the real world; task completion has some priority, and the assessment of task performance is in terms of task outcome” (Skehan, 1998, p. 38). Tasks serve as the basic unit of instruction in TBLT, which is a pedagogical approach that started to gain ground in L2 teaching and learning since the 1980s. TBLT attaches great importance to the task-based syllabus to achieve the balanced development of English learners’ linguistic forms and meanings. This attracts many researchers’ interest in order to discover the role of tasks in foreign language instruction. Among the various TBLT areas of study, the current dissertation discusses task research from a pedagogic perspective, that is, materials development and task design. Generally, most task-based foreign language studies have focused on the cognitive dimension of task design and foreign language performance (e.g., Kormos & Trebits, 2012; Ojima, 2006; Peng, 2018; Préfontaine & Kormos, 2015; Ruiz-Funes, 2015). Recent task-based foreign language studies have developed an increasing interest in the effects of learner agency and the potential to elicit creativity that the tasks have on learners’ task performance, on the one hand, and on their engagement in doing the tasks, on the other (e.g., Lambert et al, 2017; Lambert & Zhang, 2019; Tin, 2011, 2012, 2013). Control in these studies usually refers to the control over the content to express, and potential for creativity refers to “the need to say something new” (Tin, 2013, p. 385), both of which lead to student engagement in
performing the tasks as well as affect task performance. This section of the dissertation presents how task complexity, learner agency, and the potential for creativity in task design are realized and experimented with in recent research in the EFL learning context.

2.6.1 Task complexity

The most recent and popular theories on task complexity are Robinson’s (2001a, 2005, 2007) Triadic Componential Framework and Skehan’s (1998) limited-attentional resources model. Robinson’s Triadic Componential Framework argues that there are three factors that should be taken into consideration during task design. The first is task complexity, which refers to the processing burden the task imposes on the learner and which should serve as the primary basis of sequencing decisions. The second factor, task difficulty is about differences brought about by individual differences between learners; these learner factors include affective variables such as motivation, anxiety and ability variables such as aptitude, proficiency, and intelligence. The third factor, task condition is related to interactive factors, on the one hand, and participation and participant variables, on the other. This framework places the largest emphasis on components of task complexity, as these describe features that make any given task more or less difficult for any learner. Moreover, Robinson claims that task complexity can be manifested along two dimensions: a resource-directing dimension (e.g., the number of task elements, the complexity in reasoning) and a resource-dispersing dimension (e.g., planning time). The two aspects are supposed to affect attentional resource allocation in significant ways during L2 task completion; therefore, a cognitively more complex task will always be more demanding for any learner than a cognitively simple one.

Robinson (2001b, 2003, 2005) also put forward the Cognition Hypothesis, an elaboration of the Triadic Componential Framework (2001a, 2005, 2007). According to Robinson’s Cognition Hypothesis, an increase in task complexity along the resource-directing aspect leads learners’ attention to more difficult concepts, therefore directing their attention to their linguistic system and certain linguistic codes to meet the cognitively more complex conceptual demands, thus leading to greater accuracy and complexity and lower fluency. In contrast, increasing task complexity along the resource-dispersing dimension by, for example, offering no planning time or background information, does not direct learners’ attention to any particular aspect of the linguistic system or the linguistic code but simply increases the overall processing burden, therefore resulting in a decrease in fluency, accuracy, and complexity.
Robinson’s (2001b, 2003, 2005) model is in some ways in contrast with Skehan’s (1998) limited-attentional resources model, according to which increasing a task’s cognitive complexity always results in decreases in accuracy, complexity, and fluency. According to him, there are two language systems, namely the exemplar-based system and rule-based system, and the three aspects of language production, namely fluency, accuracy, and complexity, draw on them. It is assumed that different aspects of linguistic performance rely on different systems (Skehan, 1998). To be more specific, fluency requires learners to access their exemplar-based system and use ready-made chunks, and when a problem arises, they apply communication strategies to manage it. However, accuracy and complexity rely on learners’ rule-based system, thus demanding syntactic processing. Moreover, accuracy requires learners to control the existing resources to avoid errors, while complexity involves the re-constructing of the existing language resources, which promotes risk-taking. Consequently, the three aspects cannot be achieved simultaneously due to limited attentional capacity (Skehan, 1998; Skehan & Foster, 1999). Skehan holds that this trade-off effect will occur between the complexity and fluency of speech in performance and might be task-induced, which means certain task factors possibly elicit more complexity at the cost of fluency and vice versa.

Since Robinson’s (2001b, 2003, 2005) and Skehan’s (1998) models are grounded on and mostly applied to studies of oral production, these frameworks might be limited in their implications for modeling task complexity in writing as the same model of task complexity might affect task performance differently in different modes, such as oral versus written (Tavakoli, 2014). When it comes to writing, writing task complexity is supposed to be connected to the writing process. Kellogg (1996) put forward his model of the writing process which includes three phases, namely formulation, execution, and monitoring. Formulation, on the one hand, entails planning when the writer establishes goals for the writing and generates ideas to reach these goals; on the other hand, it also includes organizing the structure and selecting lexical and syntactic frames that are needed to encode those generated ideas to make the linguistic units ready for execution. Execution requires programming, where the output from translation is converted into production schema for handwriting or typing, a process of executing or the actual production of sentences. Monitoring consists of reading and editing. Therefore, for EFL writing, the cognitive task complexity of the task can be understood and analyzed regarding the different stages of the composing process. To be more specific, since writing tasks set in motion a series of complex mental processes that a writer goes through during all the steps of composing, the task is believed to affect the allocation of students’ limited attentional and memory resources, which are likely to influence the
different stages of the writing process (Kormos, 2012), thus leading to written texts of different qualities. However, research about writing along this line, no matter whether in L1 or L2, is very scarce.

2.6.2 Learner agency and potential for creativity in language learning tasks

In TBLT, tasks of different characteristics are considered to affect learners’ interest, effort, and L2 production; therefore, it is essential to design tasks that evoke learners’ interest, facilitate them to make effort and produce better L2 writing in task-based writing instruction. It has been suggested by many studies that learners are expected to be more engaged and invested in tasks operated on personally-invested task condition (e.g., Lambert et al., 2017; Maehr, 1984), which brings out the call for personally-invested L2 task design and learning (Lambert & Minn, 2007). Students are not likely to fully engage themselves if they do not find the task interesting or meaningful. In other words, the meaningfulness that the task has for learners is crucial as it will determine the degree to which they are willing to become personally invested in performing it (Lambert et al., 2017).

Inspired by Maehr’s (1984) personal investment theory, some studies on task engagement apply tasks that are operated on two conditions, the teacher investment (TI) condition and the personal investment (PI) condition (Lambert & Minn, 2007; Lambert et al., 2017). The TI and the PI condition are compatible with agency/autonomy/control, which have been found and argued to contribute to learners’ learning since a very early time in the framework of positive psychology (Noels, 2001; Oldfather & West, 1999; Williams & Burden, 1997), showing that they can influence learners’ motivation to learn. The three concepts share the meaning that the learner makes his own decision about the pace and process of the learning activities; therefore, the three concepts are used interchangeably in the current study (see Huang & Benson, 2013 for a detailed discussion of them). In addition, many studies have showed that agency is associated with students’ school-related engagement in academic activities (Wong, 2000), students’ engagement in classroom learning (Reeve, Jang, Carrell, Jeon, & Barch, 2004; Tanaka, 2007), as well as students’ experience of flow in academic activities (Steele & Fullagar, 2009).

In addition to learner agency, the potential for creativity that a task has is another task feature that has been found to influence students’ engagement in tasks and task performance. In applied linguistics, creativity is defined “as the playful use of language to construct new/unknown meanings, transforming one’s current linguistic and conceptual world and involving several types of creative thinking” (Tin, 2011, p. 216).
Tin (2011) argues that creative tasks that include high formal constraints can result in the emergence of complex language. Here, the concept of constraint is different from the concept of control. Constraint is a key feature of a creative task, referring to “any restriction on freedom that limits the number of possible solutions available for solving the problem at hand, including rules, goals, and limitations on choice, boundaries, and scarcity” (Joyce, 2009, p. 5). Control, however, means that learners can make their own decision concerning what language to use (Lo & Hyland, 2007), and it is assumed to encourage learners to activate their real-world language use (Ellis, 2003). Tin (2012) suggests that many communicative tasks, such as those that encourage reliance on personal experiences, thus giving control to students without constraints, can only elicit language that expresses “known meaning”, “meaning known to self although unknown to the interlocutors” (Tin, 2012, p. 178). She claims that extensive freedom/control over the task on the learner’s part without constraints often results in “less creative combinations of available resources and lower use of unusual resources” (Tin, 2012, p. 178), leading to a lack of desire to use complex language. In line with the above, Tin (2012) argues that creative tasks with constraints that require learners to construct new meanings or “unknown meaning” (p. 178) can facilitate creative language use, resulting in the production of more complex language.

However, not all constraints result in creative language use and the constraint that facilitates creativity is called “constraint desirable for creativity” (Tin, 2012, p.179), referring to task features that can promote learners to construct new meaning by searching in their linguistic resources for riskier, unknown language expressions and combinations of already mastered utterances and structures (Tin, 2011, 2012). Tin (2012) came up with three dimensions of manipulating constraints in creative tasks, namely goal constraints, input constraints, and outcome constraints. Specifically, goal constraints mean that in a creative task, the goal is presented abstractly and partially, rather than concretely and fully as is done in communicative tasks or controlled tasks, so as to encourage learners to generate unknown meaning as the task continues. Input constraints refer to limitations on the amount and the forms of input information, which, in a creative task, can be the pre-inventive information by learners, such as some random words the learners might be required to write down at the beginning as opposed to, for example, the completely presented picture prompts in a communicative task or a teacher-controlled task. Finally, output constraints mean that form and meaning are equally emphasized in a creative task, for example, the analogy used by learners in writing poems is considered a manifest of creativity (Tin, 2011), unlike in communicative tasks where meaning is prioritized. As for the timing of giving
constraints, Tin (2013) pointed out that constraints can be given to learners either in “the idea-generation phase” or the “idea-exploration phase” (Finke, Ward, & Smith, 1992). For example, the teacher could ask the students to write down names of objects in the idea-generating phase, and then, in the idea-exploration phase, the teacher could give the constraint by asking them to write sentences with the sentence structure “If I were a/an (insert the word generated above), I would…”. Tin (2011, 2012, 2013) suggests that creative tasks with such formal constraints given at the proper time are expected to elicit creative language use.

Recently, some researchers have begun to develop an interest in creative writing in foreign language learning. “Learning to express oneself creatively is a high individual process deeply rooted in one’s inner imaginative and emotional life. At the same time, it also calls for a heightened perception and awareness of the world outside of oneself” (Lutzker, 2015, p. 134), and it is argued that the creation of the opportunities for the pupils to take their own initiatives together with their broader search for meaning and orientation makes a meaningful education (Greene, 2001). In the context of classroom English writing and writing task design, Lutzker (2015) and Greene’s (2001) ideas about creative writing are in consistent with Tin (2012)’s idea of constraint for eliciting “unknown meaning” (p. 178) and “constraint desirable for creativity” (Tin, 2012, p.179) in the sense that they all seem to highlight that creative writing tasks should be designed in the way that provokes interest and drive learners beyond themselves. Writing tasks of this kind is most likely to enhance language task performance as well as language development, as is argued by Maley (2006), the challenges and constraints integrated into the creative task require the disciplined engagement, “the fascination of what’s difficult” (p. 35) to develop one’s creative writing skills. According to Maley (2006), this disciplined engagement refers to the affective and cognitive engagement in creative writing. He also argues that higher degree of affective and cognitive engagement in aesthetic creation, such as creative writing, as opposed to expository writing, foster learners’ language capabilities at all levels, such as grammar, vocabulary, phonology, and discourse since they attempt to express their unique personal meanings in interesting and demanding ways; thus, it is possible that “they engage with the language at a deeper level of processing than with expository texts” (p. 36). Maley (2006)’s idea of challenges and constraints of creative writing is also consistent with Tin’s (2012) idea of constraints and control on learners’ part to elicit creative language use. In a word, in the task-based classroom English writing context, it seems interesting to investigate how the manipulation of control and constraint of creative writing tasks or potential for creative language use relates to writing engagement as well as writing task performance.
The current study continues with this line of research on foreign language learning. The tasks applied in the current study belonged to two types: the teacher-generated content task and the learner-generated content task. Specifically, the first picture narration was a teacher-generated content task (TGT), in which the learners were given no control since the content to write was predetermined. The second task, the personal story narration and the third task, the creative task (CT) were learner-generated content tasks (LGT) because the students had higher levels of agency and could generate their contents. What distinguished the LGT and the CT was that the CT included the constraint desirable for creative language use along the input dimension (Tin, 2012) because the students were required to come up with something new by using the given unrelated words. In essence, the three tasks were different in the levels of agency with the TGT having the lowest level of agency and the LGT having the highest. In addition, they were also different in eliciting creativity, with the CT having the greatest potential to elicit creativity while the TGT and LGT having no such potential.

2.6.3 Summary
In summary, these two task dimensions, learner agency, and the potential for creativity in language learning tasks, are likely to lead to learner engagement. One of the aims of the current research was to see if and how learner engagement investigated with the help of flow that the students experienced was affected by the levels of learner agency and the potential for creativity in tasks, and what kind of language that the students produced in tasks characterized by different levels of agency and different potential for creativity. What is important to mention is that the present study applied the framework of task complexity (Robinson; 2001b, 2003, 2005; Skehan, 1998), Maehr’s personal investment theory (1984) and Kellogg’s (1996) writing model to help understand the results with regard to the students’ written performances and their flow and anti-flow experiences in different tasks.

2.7 Empirical research on written narrative tasks
Many recent task-based studies have been carried out on the relations between task types and learners’ task performances. The tasks adopted in these studies, regardless of the writing genre (i.e., argumentation, narrative, description), or task mode (spoken or written), are mostly designed based on Robinson’s (2001a, 2005, 2007) Triadic Componential Framework and distinguished in terms of task complexity. Since the current study focuses on narrative writing, only empirical studies on the influence of
2.7.1 Empirical research on effects of task complexity on foreign language writing

In terms of studies with task complexity manipulated along the resource-directing dimension, Kuiken and Vedder’s (2008) research addressed the question of how the number of content elements that students have to consider in a letter-writing task impacts the linguistic characteristics of writing tasks. 91 Dutch university students of Italian and 76 students of French performed two writing tasks with prompts of differing cognitive complexity. Their results showed that students were more accurate in the cognitively more complex task in which they had to consider a larger number of elements, but in terms of the syntactic complexity and lexical variety, no differences were observed. The study provides some support for Robinson’ Cognition Hypothesis (2001b, 2003, 2005) in the sense that more task complexity leads to greater accuracy in language use.

Kormos (2011) investigated the effects of task complexity manipulated by providing different levels of cognitive complexity in terms of more or less demand for plot conceptualization on students’ narrative writing. Two tasks were applied, the cartoon description task and the unrelated picture description task that were different in whether they have a clear storyline or not, which were expected to require different cognitive demands from the writers. The cartoon description task included six pictures in order and had a clear storyline so that the students did not need to conceptualize the content of the story, which eased the complexity demands in the phase of planning, that is, selecting and ordering the relevant ideas (Kellogg, 1996; Skehan, 1998). Nevertheless, the task also required learners to use certain low-frequency words in L2, which increased the processing load in the linguistic encoding phase (Skehan, 1998) or transcribing phase (Kellogg, 1996). When doing the unrelated picture description task, however, learners had the freedom to design their own stories, but they could tailor them to match their linguistic resources, which might lead to increased conceptualization effort during planning but a potentially reduced load in the transcribing process. The results showed no substantial difference in linguistic features of the written output produced in the two tasks.

In China, some studies have also been carried out on the influence of task complexity on written task performance. Li (2010) conducted a study on the effects of task complexity on the English writing of non-English majors. Three narrative tasks were provided for the participants, namely a cartoon description with some strips,
narration with six related pictures in order, and a narration with six unrelated pictures. The three narrative tasks were considered having different cognitive load in the conceptualization phase and linguistic encoding (Kellogg, 1996; Skehan, 1998), with the first one having the lowest cognitive load on the students and the third one the highest. The results showed that, with the increasing of task complexity, learners generated less accurate, more grammatically and less lexically complex writing, and there was no significant difference regarding fluency.

In an attempt to investigate EFL writing in China, some studies have been conducted on the effects that tasks of different genres have on students’ written performance. Shao (2003) investigated how task types (personal information introduction, picture description, and argument presentation) affected non-English majors’ EFL written performance. The three tasks were different in cognitive load provided on the students, with the personal information instruction providing the least cognitive load on the students and the argument presentation the most. The study found that, in terms of fluency, there was a tendency that the personal information introduction task led to the highest fluency while the argument presentation task led to the lowest, no matter whether there was a time limit or not. Concerning accuracy, the personal information introduction task led to the highest accuracy while the argument presentation task led to the lowest in both conditions with a time limit and without a time limit. Moreover, the personal information introduction task led to the highest complexity. Cai (2009) also investigated the relationships between discourse genre (argumentation and narration) and task performance. The results showed that argumentative writing, which possessed a higher cognitive load, resulted in significantly higher syntactic complexity than narrative writing. This supports the prediction of Robinson’s Cognition Hypothesis (2001b, 2003, 2005) that the increase of complexity along the resources-directing leads to more complex performance. Song (2016) conducted a study with English majors on the influence of task types (argument task, expository task, and narrative task) on their written performance in terms of syntactic fluency (T-unit length and clause length) and accuracy. The three tasks were considered different in cognitive load on the students, with the argument task being the most difficult and the narrative task the least. The study showed that, with regard to syntactic complexity, the students preferred using more complex sentences with many subordinations in the argument task and expository task while using simple short sentences in the narrative task.

In conclusion, these studies on writing, specifically on written narrative tasks analyzed the effects of task complexity on written performance. Their results were often
contradictory. However, they seem to share the notion that “task properties have a significant impact on the nature of performance” (Skehan & Foster, 1997, 2002). However, in task-based language learning and teaching, task complexity is not the only factor that affects task performance. Recent studies have found that task performance (e.g., student engagement) is related to agency that the task provides for learners (Lambert et al., 2017; Phung, 2016; Poupore, 2014) (see section 2.8.2 for a review of the empirical studies on learner agency and student engagement in task). Recent studies have also found that the potential to elicit creativity that the task owns affect students’ engagement and their task performance (Tin, 2011, 2012). However, empirical research in this line is scarce. Most of these studies are on oral production and few, if there is any, have been conducted on writing. Moreover, these studies have been conducted with college or university students, thus offering little information about different populations, like middle or high school students. The present research aimed to fill the gap by investigating how EFL writing tasks characterized by different levels of learner agency and potential for creative language use would affect Chinese teenage students’ engagement in writing as well as their written performance.

2.7.2 Empirical research on student engagement in foreign language or EFL writing

In the foreign language learning area, some empirical studies have been carried out on student engagement in performing the tasks of different complexity within the framework of TBLT. Although this section should focus on student engagement in writing, I would like to start by reviewing empirical studies on student engagement in oral communication for two reasons: on the one hand, the number of relevant studies on student engagement in the writing realm is small; on the other, a comprehensive review of student engagement in both oral and writing contexts can provide a fuller understanding of the role of the construct in foreign language learning areas.

Lambert and Minn (2007) explored how task conditions affected L2 task performance. Two Japanese females of intermediate level English proficiency studying in university were required to do three oral tasks (the instructional task, the narrative task, and the opinion task) in the personal investment (PI) condition and teacher investment (TI) condition. These tasks were designed based on Yule’s (1997) typology and were considered different in discourse complexity, with the instructional task being the least complex and the opinion task the most. In the PI condition, learners needed to come up with their own contents and resources for the task, which, however, were already provided as part of the task in the TI condition. Post-performance protocols were
applied to learn learners’ metacognition, and three questionnaires (motivation questionnaire, anxiety questionnaire, and a feedback questionnaire) were applied to investigate their motivation and anxiety as well as their feedback (i.e., the perceived enjoyment, usefulness, difficulty, and confidence) in different tasks. The results showed that, concerning task performance, on the basis of raw data, in the instructional task and the narrative task, complexity and accuracy produced were somewhat higher in the PI condition than in the TI condition. However, it was the opposite case in the opinion task, with higher complexity and accuracy being produced in the TI condition than in the PI condition. As for fluency, the difference only occurred in the opinion task, with higher fluency in the TI condition than in the PI condition. This is because the participants were more involved in the PI condition by making full use of their existing knowledge. However, as task complexity increased (from the instructional task to the narrative task to the opinion task), complexity and accuracy progressively increased while fluency progressively decreased in both TI and PI condition. Regarding learners’ motivation and anxiety experienced in the tasks, it was found that tasks operated on the PI condition created more positive affect on learners than tasks operated on the TI condition. For instance, learners reported having found the tasks in the PI condition more enjoyable, more useful for their English improvement, less difficult, and slightly increased their confidence in doing them in the future than those in the TI condition, inferring that the tasks operated on the PI condition was more flow-generating.

Poupore (2014) conducted a study with 38 adult Korean English learners of intermediate proficiency concerning the influence of content-related conditions on their engagement in performing an oral task. It was found that tasks with the content of personal life themes were reported being more interesting than those with content irrelevant or of little relevance to their personal life. Consequently, the participants were more intrinsically motivated and engaged when doing the tasks with content covering personal experiences.

Lambert et al. (2017) investigated how narrative oral tasks characterized by different amounts of control for learners, namely learner-generated (LG) content task and teacher-generated (TG) content task, affected their engagement in L2 use. The participants were 32 university English majors who were divided into three groups based on their language proficiency. It was found that all three groups of learners were more engaged in the LG content task behaviorally, cognitively, socially, and emotionally than they were in the TG content task. Specifically, behaviorally, they produced more words and invested more time in the LG content task; cognitively, they made more effort to elaborate and negotiate the semantic content of the task to a greater
extent; socially, they produced a larger number of backchanneling; emotionally, they reported being more positive about the LG content task, indicating that they were more interested in the LG content task and enjoyed doing it more. They also found that language proficiency did not affect their engagement in the task.

Qiu and Lo (2016) investigated engagement in two kinds of oral narrative tasks with 60 Chinese undergraduates. One task required them to tell a story on the basis of the given pictures while the other required them to share their personal experiences without any pictures. Each task type included two tasks that were different in topic familiarity. It was found that learners were more cognitively engaged but less socially engaged in the familiar content task than in the unfamiliar content one, which was indicated by more elaborate clauses and fewer self-repairs made that were produced in the former task than in the latter. As for emotional engagement, the researchers interviewed 21 participants immediately after they finished the task. The results revealed that the participants reported having more positive feelings (i.e., interest, confidence, less nervousness, and less embarrassment) that facilitated task engagement in the familiar content task. However, the study did not show any difference in task performance on the picture narration task and the personal experience narration task.

Phung (2016) carried out a study with 21 adult learners of English from various cultural and language backgrounds at two universities in the United States and investigated their engagement in two opinion-gap oral tasks. In the first task, they were asked to propose nine cultural artifacts that represented American culture and were asked to choose the three most representative ones after discussing them; in the other task, learners were given the nine activities and were asked to discuss them and choose the three most preferred ones. The findings showed that learners preferred the first task; thus, they were more engaged behaviorally, as shown by the larger number of words and turn-takings that appeared in the task and more time spent on the task; cognitively, as indicated by the larger number of overlaps and turn completions and backchannels produced; and socially, as shown by the larger number of sequences of negotiation of meaning and form and the number of self-repairs. The post-task interview revealed that they found the first task more interesting and enjoyable. What is more, the post-task interview showed that the first task was more preferable because: (1) They were more personally relevant and emotionally engaging to learners; (2) They offered learners the possibility to generate the task content themselves which might enhance the affective engagement since they were less likely to encounter unfamiliar content, thus leading to more positive affective responses; (3) They aroused a genuine need to communicate by
speaking with interlocutors of different first language (L1) backgrounds; and (4) They were perceived by learners as being less difficult and more controllable.

Phung’s (2016) findings are consistent with Qiu and Lo’s (2016), Poupore’ (2014) and Lambert et al.’ (2017) findings that learners prefer to do tasks that were operationalized on a learner-generated content condition, in which they could talk about topics of personal relevance and interest. These studies benefit our understanding about what is going on with the learners when they are involved in oral communication activities in the teacher-generated and learner-generated content conditions. However, these studies failed to capture the associations between task-specific engagement with pedagogic task performance (Lambert, 2017).

Lambert and Zhang (2019) conducted a case study on the effect of TG and LG conditions on student engagement and English oral task performance. The participants were four female Japanese second-year undergraduates, two of whom were English majors, and the other two were Chinese majors. Their first language (L1) was Japanese and they had advanced English proficiency. The participants were asked to do three tasks (instruction, narration and opinion) in two task conditions, the teacher-generated content (TGC) condition and the learner-generated content (LGC) condition. The result showed that, regarding student engagement in language use, the participants were more interactive, more elaborated on their ideas and contents expressed, and used more affiliative backchannelling while doing the tasks in the LGC condition, showing that they were more emotionally and socially engaged in the LGC condition; regarding task performance measured by fluency, complexity, and accuracy, on all the three tasks, the two pairs spoke more fluently and accurately but produced less complexity on tasks in the LGC condition. Lambert and Zhang argued that the participants’ engagement was associated with their task performance. The participants reported using know rather than novel linguistic resources in the LGC condition, which might also contribute to their higher fluency and accuracy in the LGC tasks. Moreover, the study showed inconsistent findings between the two pairs of participants concerning anxiety and motivation in performing the tasks. Specifically, the English learners reported feeling less anxious and more motivated when performing the LGC tasks, while for the Chinese learners, it was just the opposite. Lambert and Zhang explained that the Chinese learners felt more anxious in the LGC tasks might be caused by their inexperience with communicative tasks in their formal instruction, and as for the difference in motivation between the two pairs, it was their preference for tasks operated on the TGC condition that might have caused their higher motivation in the TGC tasks. However, the study failed to cover the effects of task types on student engagement and task performance.
Concerning student engagement in foreign language writing or EFL writing, Ren (2005) used questionnaires, classroom observations, and after-class interviews to study the influence of student involvement on learning outcomes. It was found that for the scores of the opening answer questions about writing, behavioral involvement had no significant effect on it while cognitive, especially deep-level cognitive involvement and affective involvement (i.e., a sense of enjoyment) played a vital role in writing scores. Mutwarasibo (2014) conducted a study on engagement in EFL group writing with 34 undergraduate students in Rwanda. From the open-ended interview, the students reported being encouraged and even challenged to talk in English and express their ideas. They also reported benefiting from the group writing through developing interpersonal and social skills as well as building confidence and increasing specific knowledge in EFL. In other words, it was the sense of satisfaction and success in the process of group writing that enhanced their engagement in group writing. Wang (2016) investigated the relationship between university students’ engagement and writing outcomes. She had two groups of students, the experimental and the control group. The experimental group was given 12 writing classes (45 minutes per class) over six weeks and at each class, the teacher did interventions to enhance the students’ learning engagement. The interventions were such as changing the teaching procedures, exchanging the roles of the teacher and the students, the teacher’s intervening on group discussions, peer reviewing, and teacher’s changing teaching style (i.e., being humorous, encouraging students, and creating a relaxing and pleasant atmosphere). However, in the control group, the students were taught in a traditional way that the teacher gave the instructions and explained the key writing points to the students, students wrote, and the teacher chose some of the students’ writings, showed, and made comments on them to the whole class. It was found that the three dimensions of student engagement in the experimental group all had positive effects on their writing proficiency, among which, behavioral engagement had the strongest effect. This study indicates that many factors may affect writing engagement in classroom writing instruction, such as classroom atmosphere, teaching methods, and teaching activities. In addition, learner characteristics such as L1 and L2 learning experiences, have also been found having an influence on engagement (mainly cognitive engagement) in the third language (L3) writing since learners might apply the writing strategies and the writing skills that they have been trained and have practiced in L1 and L2 writing when they are writing in L3 (Liu & Kong, 2014; Vogel, 1992).

To sum up, these studies, in essence, address the role of affect in task design and task performance and devote to finding the independent factors in task design that can
enhance learner engagement in performing the tasks (Lambert, 2017). Fortunately, they seem successful in the sense that they found that learners prefer doing tasks operated on learner-generated condition in which they can come up with task content on their own or they have background knowledge about or are familiar with the tasks. Tasks with such characteristics are more personally relevant and emotionally engaging to them, and thus bring about more affect/positive emotions during task performance. These findings about student engagement in the learner-generated content task condition are consistent with Maehr’s (1984) argument that meaningfulness of a task for learners leads to personal investment in task performance.

2.7.3 Empirical research on student engagement in foreign language or EFL writing with affect/flow being taken into consideration

In recent years, many empirical studies have suggested that the psychological process of flow exists and serves as a basis for good performance (e.g., Macquillan & Conde, 1996) in sports (e.g., Jackson & Roberts, 1992), in work and leisure (e.g., Csikszentmihalyi, 1989), business (e.g., Csikszentmihalyi, 2003), and classroom learning (Shernoff, Csikszentmihalyi, Shneider, & Shernoff, 2003). The common idea these studies seem to present is that the participants’ or the learners’ flow experiences enhance engagement in the task or the situation when the perceived challenge of the task and their skills are balanced.

In the context of the foreign language or EFL learning, flow has been found existing in the foreign language learning activities and foreign language learning classroom both on a general and a task-specific level. For example, Schmidt and Savage (1994) investigated flow experiences with 16 Thai students of English in a voluntary English-language program. They found that flow did occur in many contexts in their English learning experiences both within and outside the classroom. Egbert (2003) carried out a study with Spanish language learners during an 8-week long, 3-hour daily block Spanish course to exam whether flow would appear in classroom foreign language learning. The teacher and researcher designed 7 tasks (reading, discussion, electronic chats, and email exchange) in different modes (individual, small group, and whole group) for each class. Egbert (2003) adopted observation, a perception survey, and interviews to investigate learners’ flow experiences. She reported that flow did exist in the foreign language learning classroom. She suggested that four conditions operate to facilitate flow: (1) a perceived balance of task challenge and participant skills; (2) opportunities for learners to focus their attention; (3) an intrinsic interest in doing the task; and (4) an ability to exercise control over the task process and outcome. She concluded that these
conditions could affect the nature of task performance and ultimately language learning. However, her study seems to fail to point out that task mode, in terms of oral or written, might make a difference to learners’ engagement in doing the task. Czimmemann and Piniel (2016) investigated the flow experiences of advanced language learners in the Hungarian EFL learning classroom. They examined the factors that supported flow and those that suppressed flow during the process of EFL learners completing a narrative task in individual, pair, and group modes. The findings showed that the majority of the participants tended to experience moderate to high level of flow in the classroom EFL learning in a general sense. The results of the study also suggested that task-specific flow was significantly negatively correlated with anti-flow experiences of boredom, apathy, and anxiety. Interestingly, there were no differences in flow appearing in tasks that were in different task modes, indicating that the number of interlocutors involved in a task did not affect students’ flow experiences during task performance.

More recently, affect/flow has mostly been studied for its relevance to student engagement and learning outcomes. MacIntyre and Gregersen (2012) argued that in a language learning environment where learners were given meaningful and interesting tasks especially featured to elicit and maintain flow, they were likely to experience language learning in a positive way, which ultimately enhanced their language learning and language development. Kormos and Préfontaine (2017) examined 40 French EFL learners in terms of how affective factors (i.e., interest, task-related anxiety, task motivation, perceived success in task-completion) influenced their fluency on three oral narrative tasks, namely the unrelated picture narration, the story retell, and the related picture narration, which were considered different in conceptualization and formulation demands for the participant, with the unrelated picture narration being the most difficult task and the related picture narration the least. The participants’ retrospective interview comments were analyzed to identify the main task characteristics that they considered as either enabling or disenabling fluent performance. They found that cognitive demands of the tasks were related to learners’ affective response to the tasks, and, specifically, the perceived task-related anxiety and perceived success in task completion were greatly related to the fluency produced by learners. This study sheds some light on our understanding of how task characteristics are linked to psycholinguistic processing and speech production.

Aubrey (2016) investigated the relationship between task engagement and engagement in language use by applying the construct of flow. He asked EFL learners to perform two tasks with a culturally familiar interlocutor and a culturally unfamiliar interlocutor respectively. He found that inter-cultural contact had a significant positive
effect on flow and one aspect of engagement in L2 use (turn-taking). In a subsequent study, Aubrey (2017) studied flow qualitatively by analyzing learner diaries written after each task (5 tasks for each group in 5 weeks) to investigate learners’ flow. There were two groups of participants, one doing the intra-cultural oral task and the other the inter-cultural oral tasks. It revealed that learners in the inter-cultural task-based interactions reported significantly more flow-enhancing experiences and significantly fewer flow-inhibiting experiences than learners who performed the same tasks in the intra-cultural condition. Learners preferred being paired with partners from different cultural and L1 backgrounds. Such learners might have less shared background knowledge about the assigned topics, and this could generate a more genuine need to communicate to develop a shared understanding. This finding is in line with Aubrey’s (2016) finding that inter-cultural contact in a task-based context seemed to have increased the likelihood of learners achieving a flow state. Aubrey (2016) divided the flow dimensions into 6, namely challenge-skill balance, control, interest, enjoyment, sense of accomplishment, and attention. He reported that enjoyment was connected to a sense of playfulness, doing simpler and more familiar activities and interest was associated with something meaningful and valuable. This finding is in line with Maehr’s (1984) suggestion that the meaningfulness of the task is important in generating learners’ interest in the task and making them put more personal effort into performing the task. Another significant finding of Aubrey (2017) is that each flow dimension contributed differently to shaping the flow-related experiences of learners to varying degrees. Sense of accomplishment was the most frequently occurring flow-enhancing dimension in the inter-cultural group while the challenge-skill balance was the most prominent in the intra-cultural group.

With regard to foreign language or EFL writing, there is some research suggesting that flow exists in writing and acts as an either driving or distraining force for writing performance. For instance, Abbert (2000) did a case study of two fifth-grade students’ intrinsic motivation for writing. The study showed that children who were highly willing to write experienced flow. Abbert argued that it was the sense of flow that might explain why some children persisted in writing when faced with a challenge, spent much time and effort engaging in writing activities that they found interesting, and why some children did not. It was also found that flow occurred when students had control over important aspects of writing, such as ownership, genre, style, and length. In other words, flow occurred when they were engaged with the writing that featured with self-determined interests. Lo and Hyland (2007) found that the writing activities and topics that were interesting and relevant to their real personal life and have genuine
audiences enhance young writers’ engagement and motivation to write and built their confidence in writing. This finding is consistent with findings of studies on engagement in oral task performance that oral tasks that are relevant to the participants’ or learners’ personal life seem to enhance their engagement in performing the tasks (Lambert & Minn, 2007; Poupore, 2014).

In addition, some studies have conducted on the factors that facilitate or inhibit flow in EFL writing engagement and written performance. Liu (2006) conducted a study with 20 English sophomores on their emotional experienced during the writing process and its relevance to writing engagement. The students were asked to report what they were feeling while writing and were also interviewed after writing about their emotional experiences. The students reported that the writing topic was the factor that easily affected their emotions. If the topic was relevant to their real life or about the hot social issues at the time being in which they were interested, they got involved in writing more easily. In contrast, topics that did not allow them to choose the writing materials that they were interested would make them bored, and thus, they were not likely to engage the students to write. This finding is in accordance with similar studies on student engagement in performing oral tasks (Lambert & Minn, 2007; Poupore, 2014) in the sense that learners prefer and are more likely to experience positive emotions or flow in doing tasks that are of personal relevance and personal interest for them. Li (2007) conducted a study on flow in English writing with 95 college non-English major sophomores by applying classroom observations, flow questionnaires adapted from Egbert (2003), interviews, and reviewing the journals written by the participants. It was found that, first of all, most of the participants, with an average mean of 67.9% experienced flow during the writing process, among which the score for the interesting aspect of flow was the highest, indicating that the participants were likely to get into the flow state more easily when they got interested in the writing activity they were doing at the time. In addition, it was also found that the relatively low language proficiency had a negative effect on students’ flow experiences. Wu (2010) conducted an empirical study on the teaching of writing on the basis of the flow theory in a high school English class. She argued that teachers could influence students’ flow by helping the students set clear and appropriate goals, providing necessary input and helpful language resources (i.e., vocabulary and sentence structures), and encouraging output by giving feedback or asking the students to discuss and give feedback to each other. Flow theory seems applicable for writing instruction.

What is more, some recent studies focus on the role of specific flow elements or specific affective factors in EFL writing. For example, Guo (2019) conducted a
quantitative study with high school students on the relevance of their English writing anxiety to their writing strategies and writing achievement. It was found that students feeling high anxiety seldom used English writing strategies (i.e., purposefully using some complex sentence structure, avoiding making mistakes in spelling) while low anxiety students often used writing strategies. The study also showed that writing anxiety was negatively correlated with writing achievement. Ren (2019) focused on motivation in writing (i.e., learner’ attitude towards writing, interest, goal, and devotion to writing) and found that motivation for writing had an effect on university students’ writing proficiency at the beginning and the end of the semester. However, devotion to writing or effort made for writing was significantly correlated with their writing performance positively.

2.7.4 Summary

From the above literature review on foreign language and EFL writing, it can be seen that learners’ engagement in writing and their flow experiences are closely related. However, there are few systematic studies conducted on the relevance of flow to writing engagement and writing task performance in foreign language or EFL writing with learners’ affective responses toward the tasks or flow being taken into consideration. In addition, most studies on task-specific flow in foreign language and EFL learning mainly focus on engagement in oral tasks with college or university students, leaving teenage students as an under-represented age-group. The present study attempted to extend the study scope in this line by focusing on Chinese teenage students’ engagement in EFL writing within the TBLT approach.

2.8 Conclusion

From the above literature review, it can be concluded that in the task-based foreign language or EFL learning context, flow theory has proved to help understand learners’ engagement in and performance on both oral tasks and writing tasks in the classroom language learning context as well as outside the classroom in a general sense. However, studies on flow in foreign language or EFL writing are quite rare, so there is a research niche to be filled. Although TBLT attaches great importance to the teaching practices in the classroom to lead to good language learning results, for example, task sequencing based on cognitive complexity has often received a lot of attention in applied linguistics, currently, task-based studies seem to develop an increasing interest in designing meaningful and creative tasks to facilitate good language performance as well as elicit
flow experiences during task performance (e.g., Lambert et al., 2017; Qiu & Lo, 2016; Tin, 2011, 2012). The present study extended research in this line.

In conclusion, the present study focus on four aspects: first, how task conditions relate to students’ task performances on narrative writing; second, how task conditions relate to students’ writing engagement measured by their flow and anti-flow experiences in performing the task; third, for the CSH and the CSC respectively, what is the relationship between task performance measures and flow and anti-flow measures in different task conditions if students’ language proficiency is taken into consideration; and forth, what is the role of study contexts in the students’ written performances and their flow and anti-flow experiences in each task condition (see Table 3 for the summary of the research design).
### Table 3

**Summary of the research design**

<table>
<thead>
<tr>
<th>Research questions (RQ)</th>
<th>participants</th>
<th>Instruments</th>
<th>Data analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>RQ1. For Chinese students studying in Hungary (CSH) and China (CSC), how do task conditions relate to their written performances in terms of fluency, accuracy, and complexity?</td>
<td>20 Chinese students studying in Hungary and 20 Chinese students studying in China, with language proficiency of intermediate level</td>
<td>Three narrative written tasks, namely the teacher-generated content task (TGT), the learner-generated content task (LGT), and the creative task (CT), a flow questionnaire (a five-point Likert-scale), and a bio-data questionnaire</td>
<td>Statistical analyses: descriptive analyses, repeated measures ANOVA</td>
</tr>
<tr>
<td>RQ2. For Chinese students studying in Hungary (CSH) and China (CSC), how do task conditions relate to their flow and anti-flow experiences?</td>
<td></td>
<td></td>
<td>Statistical analyses: Cronbach’s alpha correlation coefficients, principal component analysis, descriptive analyses, repeated measures ANOVA</td>
</tr>
<tr>
<td>RQ3. For Chinese students studying in Hungary (CSH) and China (CSC), what is the relationship between task performance measures and flow and anti-flow measures in different task conditions if students’ language proficiency is taken into consideration?</td>
<td></td>
<td></td>
<td>Statistical analysis: partial correlational analysis while controlling for language proficiency (reflected by the participants’ placement test scores)</td>
</tr>
<tr>
<td>RQ4. What role do the English learning contexts play in students’ written performances and their flow and anti-flow experiences in each task condition, that is, what kind</td>
<td></td>
<td></td>
<td>Statistical analyses: descriptive analyses, independent samples t-tests</td>
</tr>
</tbody>
</table>
of differences can be detected between Chinese students studying in Hungary (CSH) and China (CSC) in their written performances and their flow and anti-flow experiences in each task condition?
Chapter 3: Research Methods

This chapter of the dissertation presents the methods employed in this empirical study. This study is quantitative in nature because it uses quantitative tools to gather quantitative data on task experiences and aims to reveal relationships between variables on the samples examined. In the following section, first, the participants and the instruments are presented, which are followed by a detailed description of the piloting of the instruments. The subsequent section describes the procedure of data collection. Finally, an introduction to the data analysis and statistical procedures for the study brings an end to the chapter.

3.1 Participants

Since one of the aims of the present study was to explore the role of language learning contexts, namely the Chinese English learning context and the Hungarian English learning context in Chinese students’ English learning, two groups of Chinese students were chosen as participants. One group included Chinese students studying in Hungary, abbreviated as CSH and the other group of participants consisted of Chinese students studying in China, abbreviated as CSC. Although my participants in Hungary differ in grades at school, their English language proficiency is the same as my participants in China, which makes them comparable in the aspects researched in this study. Therefore, in order to avoid ambiguity, they are all called Chinese teenage students in the present study. Since this study focused on Chinese teenage EFL learners, not specifically restricted to either middle school students or high school students, the issue of the exact grades in which the students were was not a concern in this study. The only requirement for choosing participants was that their language proficiency was at the intermediate level (B1 and B2) according to the Common European Framework Reference for Languages (CEFR). The reason behind this was that students with intermediate-level language proficiency were expected to produce readable and good quality writing texts. What is of importance to mention is that, for the CSH, the participants must have been studying in Hungary for at least 3 years by the time of the data collection. The reason for this requirement was that if the students were newly-transferred or they immigrated from China to Hungary a very short period of time ago, it is more likely that the Hungarian English learning context has not started to influence their English learning. It is also possible that they might still be going through the time of fitting in the new language learning environment, in which case, it is impossible to examine the effects of the Hungarian English learning context on their English learning.
Table 4 presents the biographical information of the participants, namely the total number (N) of participants, the gender, the age and the proficiency scores of the two groups of participants. The ages and the proficiency scores were presented with the mean age or the mean score (Mean), the minimum age or the minimum score (Min.), and the maximum age or the maximum score (Max.).

Table 4

**Biographical information of the participants**

<table>
<thead>
<tr>
<th></th>
<th>Gender</th>
<th>Age</th>
<th>Proficiency score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Male</td>
<td>Female</td>
<td>Mean</td>
</tr>
<tr>
<td>CSH</td>
<td>10</td>
<td>10</td>
<td>13.75</td>
</tr>
<tr>
<td>CSC</td>
<td>8</td>
<td>12</td>
<td>15.95</td>
</tr>
</tbody>
</table>

*Note.* N=Number; Min.=Minimum; Max.=Maximum; CSH=Chinese Students Studying in Hungary; CSC=Chinese Students Studying in China.

For the CSH, the participants were recruited in two ways. For students who were going to a private Chinese school at weekends every week in Budapest, those whose English was good were recommended by their English teacher. As a result, 23 students were recommended. Then they were assembled in a classroom and take the Oxford Placement Test 1, New 2004 edition (OPT). The placement test includes two parts, namely listening and grammar, each taking up 100 points out of 200. It took 2 weeks to finish testing them because not all of them went to the school every weekend. As a result, 18 of them were found to have an intermediate proficiency level (120-149 out of 200) according to the OPT Language level, which also corresponds to level B1 (120-134) and B2 (135-149) according to CEFR. In the case of those students who did not go to the school at weekends during the time of data collection, 9 students were reached by adopting convenience and snow-ball sampling (Dörnyei, 2007), and did the placement test either at their homes or in coffee shops. Five of them met the requirement. What needs to be pointed out is that the three narrative written tasks of the study were piloted with 3 of them; therefore, they were removed from the final list of participants. As a result, 20 students whose total scores ranged from 121 to 165 and whose grammar scores ranged from 50-77 were chosen as the participants for the current study. There were 10 males and 10 females, with an age range from 12 to 17 and a grade range from 6 to 12.

For the second group of participants, the CSC, two classes of first-year students in two public high schools in my hometown city, Yinchuan, Ningxia, China were chosen.
The first-year students in high school were chosen for two reasons: on the one hand, their ages were the most similar to the participants in Hungary; on the other, they were expected to be able to produce quality writings because they had finished three years’ official and systematic English study in middle school and almost one year in high school. The placement test was given to the whole class by their English teachers and the test was also invigilated by them. It turned out that a total number of 40 students were found to have intermediate level English proficiency. Then, only 20 out of them were chosen on the basis of their grammar scores to be parallel with those in Hungary. As a result, their total scores ranged from 121 to 147 with their grammar scores ranging from 51 to 76. There were 8 males and 12 females, and their ages were from 15 to 17 (see Table 4 for the biographical information of the participants).

One important point needs to be mentioned here: although I tried to create similar samples in the two English learning contexts by only accepting the students whose English was at intermediate level as participants, I failed, because it turned out that the two groups of Chinese students were significantly different in their placement test scores (p<0.05). Therefore, this is a limitation of the current study. However, the t-test showed that this significant difference was caused by the students’ listening scores, while their grammar scores (mean=61.15 for the CSH and mean=64.50 for the CSC) were found not significantly different (p>0.05). Now that the present study is on writing where grammar probably made a greater difference in their written production while listening skills did less so or not at all, it is safe to say that the two groups of Chinese students were comparable regarding the overall effect of the students’ English learning contexts on their written performance and their flow experiences on the writing tasks. Table 5 presents the characteristics of the participants, specifically in terms of the teaching languages of their English classes (TLECs), the out-of-school English learning activities (OSELAs) (see Table 6 for the number of students who are engaged in every specific OSELA), and the number of languages spoken by the students (LsSS). What’s important to mention is that the numbers in the brackets indicated the number of participants.
Table 5

Characteristics of participants

<table>
<thead>
<tr>
<th></th>
<th>Number of students</th>
<th>TLECs E</th>
<th>E</th>
<th>E</th>
<th>H</th>
<th>E</th>
<th>&amp;</th>
<th>&amp;</th>
<th>&amp;</th>
<th>H</th>
<th>G</th>
<th>C</th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
<th>Number of OSELAs</th>
<th>Mean</th>
<th>Min.</th>
<th>Max.</th>
<th>Number of LsSS E</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSH</td>
<td>20</td>
<td>13</td>
<td>3</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>5.15</td>
<td>1 (1)</td>
<td>10</td>
<td>2.35</td>
<td>2 (3)</td>
<td>6 (1)</td>
<td>(1)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CSC</td>
<td>20</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>20</td>
<td>3.10</td>
<td>1 (2)</td>
<td>5 (5)</td>
<td>1.05</td>
<td>2</td>
<td>3</td>
<td>(19)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. TLECs=Teaching Languages of English Classes; OSELA=Out-of-school English Learning Activities; LsSS= Languages Spoken by Students; E=English; H=Hungarian; C=Chinese; Min.=Minimum; Max.=Maximum; CSH=Chinese Students Studying in Hungary; CSC=Chinese Students Studying in China.

Firstly, concerning TLECs (see Table 5), for the CSH, 13 participants reported that their English subjects were taught in English (E), 3 participants reported that their English subjects were taught in a combination of English and Hungarian (E and H), 2 in a mix of English and German (E and G), and 2 mostly in Hungarian (H); for the CSC, they all reported that their English was taught in a combination of English and Chinese (E and C). This difference seems to infer that the CSH have more chance to expose themselves to English in school than the CSC.

Secondly, concerning the OSELAs (see Table 5), for the CSH, it can be inferred that the majority of the students (18 out of 20) reported doing 2 to 9 kinds of the OSELAs that were listed in the questionnaire (see Appendix G for the students’ bio-data questionnaire). Fourteen of them reported doing more than 4 kinds of the OSELAs (see Table 6), and 4 of them reported doing some other activities that had not been listed in the bio-data questionnaire, such as traveling, going to summer camps in English-speaking countries, and joining in international student organizations. For the CSC, we can see that the majority (13 out of 20) reported doing only 2 to 5 kinds of the OSELAs for their English learning, which was a narrower range compared to the 2 to 9 types of the OSELAs reported by the CSH. In addition, for the CSC, only 5 students reported doing more than 4 types of the listed OSELAs and none of them reported doing any other unlisted OSELAs (see Table 6). In addition, from Table 6, we can see that the
OSELAs indicated by the letter A (Talking with foreign friends or relatives), B (Talking with classmates from other countries), D (Reading English websites), G (Playing English video games), and J (Others) seem to be much more popular with the CSH than the CSC. Besides, the types of OSELAs represented by the letter C (Reading English novels or magazines), E (Learning English from English learning Apps), F (Watching English films, plays or TV series), H (Listening to English songs), and I (Taking English lessons outside school) seem to be just as popular with both the groups of Chinese students. The comparison between the two groups of Chinese students show that they differ a lot in their English learning experiences outside of school, with the CSH taking various kinds of out-of-school English learning activities for their English learning than the CSC.

Table 6
The number of participants who are engaged in each of the out-of-school English learning activities (OSELAs)

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>Number of OSELA &gt;4</th>
</tr>
</thead>
<tbody>
<tr>
<td>CSH</td>
<td>9</td>
<td>10</td>
<td>9</td>
<td>11</td>
<td>8</td>
<td>13</td>
<td>9</td>
<td>15</td>
<td>7</td>
<td>4</td>
<td>14</td>
</tr>
<tr>
<td>CSC</td>
<td>1</td>
<td>1</td>
<td>7</td>
<td>1</td>
<td>6</td>
<td>14</td>
<td>3</td>
<td>19</td>
<td>3</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>N</td>
<td>10</td>
<td>11</td>
<td>16</td>
<td>12</td>
<td>14</td>
<td>27</td>
<td>12</td>
<td>34</td>
<td>10</td>
<td>6</td>
<td></td>
</tr>
</tbody>
</table>

Note. CSH=Chinese Students Studying in Hungary; CSC=Chinese Students Studying in China.

Finally, concerning LsSS (see Table 5), for the CSH, only 3 of them reported speaking only Chinese and English, and 1 student reported speaking as many as six languages, which indicates that the majority of them (16 out of 20) speak 3 to 5 languages. For the CSC, the majority of them reported speaking only Chinese and English (19 out of 20) and only 1 participant reported speaking 3 languages, Chinese, English, and Japanese. In other words, the two groups of students are very different in their foreign language learning experiences.

3.2 Instruments
The instruments in the current study consisted of three written narrative tasks, a flow questionnaire that was used to measure the student engagement in doing the tasks as well as a bio-data questionnaire mainly about the students’ biographical information and their English learning experiences.
3.2.1 Written narrative tasks

The three written narrative tasks adopted in the present study were the picture description task, the personal experience narration task, and the story-creating task (see Appendix F for the three narrative tasks). The three tasks were different in learner agency and potential for creative language use, which was realized by manipulating the conditions on which the tasks were operated.

The picture description task required the students to narrate a story based on 6 related pictures in order. The pictures (see Appendix F) were about a family going to a park at the weekend and the boy falling into the pond when playing with ducks and being rescued by a young man. The students were asked to contain all the information that the pictures presented. Since the content was predetermined by the teacher or the task designer, the task was called the teacher-generated content task (TGT) and the task condition was called the teacher-generated content (TG) condition. The reason why the scenario of a family going to a park and the boy falling into the pond was chosen is that it is familiar to all Chinese teenagers, no matter if they live in Hungary or China. The story was also supposed to be fun and could encourage students to write. The 6 pictures were in color, given the fact that pictures in color are more interesting to teenagers than those in black and white. With a sequence of pictures being provided, the students were very restricted in terms of possible content or messages to write; therefore, this task was assumed to result in the least writing engagement from the students.

The personal experience narration task required the students to recall their past experiences and write about a memorable event that had happened in their life. In this task, the students were in absolute control and they could generate the task content and write about whatever they liked; therefore, it was labelled as the learner-generated content task (LGT) and the task condition was labelled the learner-generated content (LG) condition. This task that empowered the students with complete control over their writing was assumed to enable them to make strong connections between writing and their own experiences, thus leading to the most writing engagement from the students.

The story-creating task introduced some constraints while also providing some control for the students. The students were provided with 10 unrelated words in order to make sure the same genre and the same topic of writing could be produced; at the same time, the students still had freedom to create their own stories as long as they used at least 6 of the 10 given words; therefore, it was called the creative task (CT) and the task condition was called CT condition. For this task, the most challenging part had been to decide what words should be given. In the end, 10 words describing objects, events, and environments were chosen, and they were the following: “desert, island, museum, plane,
blood, ceremony, storm, dragon, bomb, and flying carpet”. These 10 words were assumed to possibly to result in interesting stories and creative language use since these words might enable the students to make associations and use various words so as to produce good written performance (Wang, 2012). Since the participants were teenagers and their level of English was intermediate, in order not to cause them difficulty in understanding the meaning of the prompts, the 10 words were checked against the Corpus of Contemporary American English (COCA) at https://corpus.byu.edu/ to ensure that they were among the first 3000 most frequently-used words, and the Chinese equivalents of the words were also offered. This task was half-constrained and half-controlled for the students and was expected to facilitate the most creative and complex language use.

In a word, among the three tasks, the TGT gave the least control to the students while the LGT gave the most control to the students. Besides, in terms of the potential to elicit creativity, the CT was more likely to elicit creativity than the TGT and the LGT. The title “A memorable event” was given to the three tasks since this topic was very general and the students were supposed to have some knowledge about it and also have something to say about it. Besides, in order to lead them to a story scenario, the starting sentence was also offered by the prompt like “This is something that I will never forget….”. Moreover, the suggested time and word limits were the same for all the three tasks. Students were recommended to finish writing in 30 minutes and were asked to produce a minimum of 100 words. The task instructions were in Chinese to avoid difficulty in understanding them.

3.2.2 Flow questionnaire

Flow was first demonstrated to exist in the foreign language classroom by Egbert (2003). She (2003), based on Csikszentmihalyi’s (1975) flow theory, designed the flow questionnaire to study EFL learners’ flow experiences during EFL task performance. The questionnaire consists of 14 items reflecting four dimensions, interest, control, focus, and attention. However, this questionnaire only concentrates on flow, without mentioning anti-flow experiences. Besides, focus and attention seem to overlap. Czimmermann and Piniel (2016) developed a flow questionnaire based on Egbert’s (2003). Their questionnaire includes two parts, the flow part which includes four categories, namely interest, attention, control, and challenge and skill balance, and the anti-flow part which includes boredom, apathy, and anxiety, with a total of 27 items. However, the last three items attached to the questionnaire address the perceived level of abilities required by the task, the perceived difficulty of the task, and the perceived level
of creativity required by the task respectively. The present study adapted Czimmermann and Piniel’ (2016) flow questionnaire by keeping the seven flow and anti-flow dimensions, but excluded the last three items about the perceived abilities, difficulty, and creativity required by the task for the reason that only one item for each construct make them less reliable in measuring the construct of interest.

The adapted flow questionnaire for the pilot included 35 items, with each construct containing 5 items. Besides, considering that the participants were teenagers and they might not be able to identify the nuance in the seven choices in each item of the scale due to their relatively low language competence, the questionnaire was kept as a 5-point Likert scale. Then, the modified flow questionnaire was translated into Chinese by me, and a friend of mine who is also a Ph.D. student in the English language pedagogy program translated it back to English. Then, we compared the translated version with the English one, discussed the ambiguous items and came to a final agreement on all the items and their Chinese translation. Finally, the flow questionnaire was piloted (see section 3.3.2 for the result of the flow questionnaire pilot).

3.2.3 Bio-data questionnaire

This is a comparative study in terms of the English learning contexts, in the sense that the two groups of Chinese students were living and studying in two very different countries. Therefore, it was hypothesized that the different English study contexts might play a role in their linguistic performance and their flow experiences on the different writing tasks. Therefore, some biographical questions were designed. The questions included four aspects, namely personal data (i.e., age, gender, the school they go to, grade level), information on English instructional situations (i.e., the teaching language of English classes), English learning experiences (i.e., length of learning English), and English activities outside of school (see Appendix G for the bio-data questionnaire).

3.3 Piloting

There were two kinds of instruments applied in this study, the writing tasks and questionnaires. This section describes the validation of the two types of instruments in their Chinese versions.

3.3.1 Piloting the written narrative tasks

The TGT was piloted with a Chinese student studying in Hungary regarding the time needed for task completion, length requirements, and the contents of the pictures, as well. As a result, the contents of the pictures were clear to her and she wrote 123
words in 40 minutes. She pointed out that 30 minutes was not enough for her to complete the task. This might have been caused by her low English proficiency since her placement test score was lower than 120 (the minimum score for intermediate level). Then the task was piloted with another student whose placement test score was above 120, and it turned out that the task worked well, and he produced 132 words in 21 minutes. The LGT and the CT were piloted in the same way with another two students whose placement test scores were above 120. As a result, the student who did the LGT produced 183 words in 24 minutes and the student who did the CT produced 118 in 19 minutes. In a word, the three tasks were successful in terms of writing time, length requirement, and instruction. It needs to be pointed out that these three students were not included in the final sample of the study.

3.3.2 Piloting the flow questionnaire and the bio-data questionnaire

The flow questionnaire was first piloted with one student (not one of the participants of the present study) to make sure that the contents were clear and not confusing. She was asked to write a short piece about anything she liked. After finishing the writing, she was requested to complete the flow questionnaire. She was asked to think aloud while doing the questionnaire. It is important to mention that she was trained to do the think aloud before performing the task. While doing the flow questionnaire, she was required to read out each item and speak out her thoughts and doubts, and the whole process was recorded. The time she finished doing the questionnaire, we listened to the recording and discussed and clarified the words and items that had caused her confusion and misunderstanding. Besides, the bio-data questionnaire was also piloted with this student after we finished everything with the flow questionnaire, and it turned out to work well. Then, I asked one of my colleagues who was teaching high school English in a county of Ningxia, China. She asked all the students in her class to write a short composition on “Friendship” and do the flow questionnaire once they finished writing. The written texts of 25 students who were recommended by their teacher as being good at English were accepted and they were analyzed with SPSS for the reliability of the flow questionnaire. As a result, Cronbach’s alphas for all the scales are above 0.60, as shown in Table 7. Besides, as shown by Table 7, there left only 3 items under the subscale of control. Therefore, 2 new items were added to it before applying it with the actual participants. The 2 items were “I felt that I had no control over what was happening during the task” and “I could not decide how to deal with the task”. Therefore, the final version of the flow questionnaire consisted of 31 items (see Table 11 for the reliability result of the final version of the flow questionnaire and see Appendix E for the
final version of the flow questionnaire). Then, the same questionnaire was attached to each task and the data collection was administrated for my main study.

Table 7
Descriptive statistics of each scale of the flow questionnaire and its corresponding reliability coefficient

<table>
<thead>
<tr>
<th>Scales</th>
<th>Mean</th>
<th>SD</th>
<th>Cronbach’s alpha</th>
<th>N of items</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>3.60</td>
<td>0.80</td>
<td>0.666</td>
<td>5</td>
</tr>
<tr>
<td>Attention</td>
<td>4.02</td>
<td>0.80</td>
<td>0.636</td>
<td>4</td>
</tr>
<tr>
<td>Control</td>
<td>3.67</td>
<td>1.01</td>
<td>0.511</td>
<td>3</td>
</tr>
<tr>
<td>Cs balance</td>
<td>3.75</td>
<td>0.74</td>
<td>0.737</td>
<td>4</td>
</tr>
<tr>
<td>Boredom</td>
<td>1.80</td>
<td>0.76</td>
<td>0.624</td>
<td>4</td>
</tr>
<tr>
<td>Apathy</td>
<td>2.44</td>
<td>0.99</td>
<td>0.718</td>
<td>4</td>
</tr>
<tr>
<td>Anxiety</td>
<td>2.37</td>
<td>0.78</td>
<td>0.780</td>
<td>5</td>
</tr>
</tbody>
</table>

Note: SD=standard Deviation; N=Number; Csblance=Challenge-skill balance.

3.4 Procedures

Data collection was carried out from the end of April till the end of August 2018. The tasks were given to the participants at least one week apart, and the tasks for all the participants were given in different orders. Each group of the participants was divided into three sub-groups when completing the tasks each time, with the first sub-group doing the tasks in the order of the TGT, the LGT, and the CT, the second doing the tasks in the order of the LGT, the CT, and the TGT, and the third sub-group doing the tasks in the order of the CT, the TGT, and the LGT. The reason for this was to minimize the effects of task order on students’ task performance.

What is important to mention is that the flow questionnaire was attached to each writing task and students were asked to do the questionnaire immediately after they finished writing. The bio-data questionnaires, for all the participants, were attached to the final task and were completed after the flow questionnaire.

3.4.1 Data collection in Hungary

The data collection in Hungary started from the end of April till the end of August 2018. The data collection was difficult since, even though many Chinese students studying in Budapest, they were either too young, studying in primary schools, under 12
years old, or they were not qualified in terms of their English language proficiency, or they have been studying in Budapest for fewer than three years.

Data collection was carried out by me in person in the case of 15 participants who went to the Chinese school at weekends. For the other 5 participants who did not go to the Chinese school at weekends, one task was sent to them each time with the clarified ask requirements and a recommended deadline. They printed them out and finished the task on their own by hand-writing. Their writings were collected by one of my students and given to me in person. One week later, the second task was sent to them and they went through the same procedure as they did in the case of the first one. After their written texts were collected, the last task was sent to them one week later.

### 3.4.2 Data collection in China

Data collection in China took place from the beginning of June till the beginning of July 2018, within the academic period of high schools in China. All the data were collected personally. Before going back to Yinchuan, China, I got in touch with two teachers, one was an English teacher and the other was a Chinese teacher, both of whom were classroom teachers of two classes in two different high schools. The Oxford Placement Test 1, New 2004 edition (OPT) was sent to them by email and they printed them out and tested their students in their presence. The grading of the placement papers started once I arrived on 6th, July 2018. It turned out that 21 students met the requirements and the top 20 students were chosen as the participants of my research.

As for the real data collection, since one of the schools is a boarding school, the students have evening exercise classes from 7:30 pm to 9:00 pm. During this time, they are allowed to study anything out of their own will. I was allowed to go into the classroom and ask the students to do the tasks during this period when my colleague was supervising the exercise class. The other school is not a boarding school, which means students go home after school at 5:30 pm. Therefore, the tasks was administered by me either during the last class on Monday when the class meeting was held by the classroom teacher or during the Chinese reading class from 7:30 to 8:00 am for the reason that if 30 minutes were not enough, the students could continue writing till they finished it because the next class was Chinese, and they did not need to hurry thanks to the fact that the Chinese class was taught by my friend.

In both classes, the tasks were given to the whole class in order not to disrupt normal class procedures, but only the written texts of those students who were chosen as participants were analyzed as final data for the present study. Tasks were administered in different orders, the same as how they were conducted in Hungary.
3.5 Analysis

This section of the dissertation describes the steps of data analysis. The analysis of the task performance is discussed by describing the three task performance measures, namely fluency, complexity, and accuracy. Besides, the measures of the flow in the present study are also summarized.

3.5.1 Measures of task performance

This section of the dissertation presents the measures that were used to measure the Chinese students’ task performance. The students’ task performance in the current study was examined by describing the linguistic features of their written products in terms of fluency, complexity, and accuracy for comparability reasons with previous TBLT literature (e.g., Kormos, 2011; Kuiken & Vedder, 2008; Michel, Kuiken, & Vedder, 2012; Sasayama, 2016; Tavaloki & Foster, 2008).

3.5.1.1 Measures of fluency

In task-based studies, utterance fluency is often assessed as breakdown fluency, which is about learners pausing behavior, repair fluency (measured by the frequency of repetitions and self-corrections), and speed fluency (reflecting the speed with which speech is delivered) (for a recent discussion, see Bosker, Pinget, Quene, Sanders, & De Jong, 2013; De Jong, Steinel, Florijn, Schoonen, & Hulstijn, 2012).

Regarding the fluency of writing, it is usually measured by the length of text produced within a time limit (Larsen-Freeman, 1978; Wolfe-Quintero, Inagaki, & Kim, 1998), writing time with the planning time excluded, total time on task including time for planning (Ong & Zhong, 2010), or words per T-unit, words per clause, and the length of text (Ishikawa, 2006). Since this study is on writing, and there was no time limit for writing, fluency was measured by the total number of words produced by the participants. Variability in this regard was expected since only a minimum word limit was set, in which case, students could write as much as they wanted. The rationale behind this decision was to have some insight into learners’ productivity as one regard of their written production (Ma, 2005).

3.5.1.2 Measures of complexity

Complexity in linguistic performance usually consists of lexical complexity and syntactic complexity. Lexical complexity refers to the lexical richness of the language. It includes lexical density, lexical diversity and lexical sophistication (Read, 2000; Skehan, 2003). Lexical diversity is about the vocabulary size, the number of different word types
that a text contains; lexical sophistication, which is also referred to as lexical variety, is about the frequency of the words used by learners; and lexical density refers to the ratio of the number of the content words and the number of function words in a produced text. In the current study, lexical density is not used as a metric for lexical complexity since the written samples are short, ranging from 101 to 505 words, with a mean of 184.30, so the requirement with regard to this measure might be too strict for the participants.

Lexical diversity is usually measured by the type-token ratio (TTR) (Crookes, 1988; Hoover, 2003). Token refers to the total number of words in a text; types refer to the different words in a text. However, TTR is criticized for being sensitive to text size and is not stable and reliable when the text gets longer. For example, Richards (1987) found that the “type-token ratio falls rapidly as the number of tokens increases” (p. 205). As a substitute, D-value has been found to be able to measure lexical richness in L2 texts in a reliable way (Jarvis, 2002). It can plot the curve of TTR against increasing token size on the basis of a mathematical probabilistic model. The present study used D-value to measure the lexical diversity of the linguistic performance since the texts produced by the participants are relatively short and D-value is more reliable in measuring short texts than TTR. Lexical variety in the current study is measured by P_Lex (Meara, & Bell, 2001). It is a software that measures lexical sophistication and is claimed to be able to handle relatively short texts (Kojima & Yamashita, 2014; Meara & Bell, 2001). It operates by, firstly, dividing the text into segments of ten words and then counting the number of difficult words that are not included in the list of the 1,000 most frequent English content words in the segments. Then, it calculates a figure, termed as P_Lex which indicates the likelihood of the occurrence of difficult words. The higher the figure is, the more likely the rare words are used. Both variables can be calculated with online software available at http://www.lognostics.co.uk/tools/D_Tools/D_Tools.htm. Before applying the software, I edited the original writings of the students by correcting the wrong spellings (e.g., thier/their), changing the contracted forms (e.g., I’m/I am) and abbreviations into their full forms (e.g., WWI/World War I) (see Appendix C for the error guidelines), since, if not, P_Lex software considers them as infrequent even though they are not when typed out in their right and full forms in the software.

Syntactic complexity measures have been used to assess task-related variation in L2 writing to investigate differences in L2 written texts across different proficiency levels over periods (Ortega, 2003). Norris and Ortega (2009) summarized syntactic complexity measures by reviewing 16 empirical studies on task-based language learning. They classified these measures into four dimensions: overall complexity (e.g., mean length of T-unit/AS unit), complexity by subordination (e.g., clauses per T-unit/AS unit,
S-nodes per T-unit), phrasal complexity (e.g., mean length of clauses), and variety of forms (e.g., frequency of tensed forms, modals, passive forms, etc.). Among the four dimensions of syntactic complexity, complexity measured by subordination seems to be the most employed. In the L2 writing domain, mean length of T-unit, which belongs to the overall dimension of complexity, seems to be the single most employed complexity measure (Ortega 2003; Wolfe-Quintero et al. 1998); thus it is often prioritized over other dimensions of the construct (Norris & Ortega, 2009). For the syntactic complexity of written data, T-unit has often been used as a basic unit of analysis. According to Norris and Ortega (2009), a T-unit is a single independent clause plus any subordinate clauses attached to it or embedded in it. To understand the concept of T-unit, we need to know what a clause is. A clause refers to a syntactic construction containing a subject and predicate and forming part of a sentence or constituting a whole simple sentence (Norris & Ortega, 2009). Therefore, in the current study, the mean length of T-unit (MLT), mean number of clauses per T-unit (MNCT), and mean length of clause (MLC) were used to acquire a comprehensive understanding of the syntactic complexity of the participants’ written texts. Before analyzing the data, the guidelines were established for marking the boundaries of T-units and clauses. Despite having invested great effort into making the definitions of T-units and clauses as clear as possible, judging a T-unit and a clause was a complicated work, especially when handling the EFL learners’ written texts that seemed to contain many errors. Therefore, before finalizing the guidelines for counting T-units and clauses, I examined the students’ written texts for ambiguous cases that might create difficulty for coding and came up with specific rubrics for the categorization of all cases. All the example sentences in both T-unit and clause guidelines were obtained from authentic pieces of students’ writings for the current study (see Appendices A, B, and C for the guidelines for identifying T-units, clauses, and errors). Besides, what needs to be pointed out is that in order to maintain reliability, another rater (who has been teaching English linguistic courses for many years in the university) and I rated 12.5% of the samples for the number of T-units, the number of clauses, the number of right T-units, and the number of right clauses on the basis of T-unit, clause, and error guidelines, and the result was that for each written text produced by the participants, in more than 95% of the cases we came up with the same numbers of T-units, clauses, correct T-units, correct clauses and the number of each error type. We discussed the confusing points until the final agreement on the rating systems (see Appendix A, B, C and D) was reached, with which I rated the rest of the sample texts on my own.
Both the T-unit guidelines and the clause guidelines were a combined and modified version of Polio’s (1997), Polio and Shea’s (2014), and Lee’s (2009). The rationale behind the decision of combining the guidelines of the three researchers was to make them as comprehensive and inclusive as possible. Besides, in the process of adapting their guidelines, I changed some items and added several new items to their guidelines based on the characteristics of the written data of my own study.

After establishing the T-unit and clause boundaries, T-units and clauses needed to be counted. Although software was experimented for these calculations, the counting results produced by the software and I by hand were not consistent, which might due to the differences in criteria for counting T-units and clauses. Therefore, I counted them by hand on my own. After segmenting the T-units and clauses, the SPSS 22.0 for Windows was used to calculate the MLT, the MLC, and the MNCT to show the syntactic complexity of the written texts.

3.5.1.3 Measures of accuracy

The most commonly used measures of accuracy in L2 performance are local measures and global measures. Local measures aim to track the use of designated grammatical features in the L2 such as verb and noun morphology, while global measures focus on the overall level of accuracy in an oral or written product. According to Foster and Wigglesworth (2016), local measures may be appropriate for measuring the development of specific morphological features over time. However, they are problematic considering that students acquire grammatical features not in a concurrent way, and it is also difficult to make sure that a specific grammatical feature occurs frequently in the data to make its measurement meaningful (Foster & Wigglesworth, 2016).

By contrast, global measures avoid the above problem since they essentially divide the data into segments and calculate the error rate. The specific accuracy measures included in global measures are mean errors per hundred words, the number of errors in syntactic units (such as mean errors per T-units, per clause), and the proportion of the units that are error-free (such as percentage of error-free clause or T-units). Here, a T-unit is a single independent clause plus any subordinate clauses attached to it or embedded in it, and a clause refers to a syntactic construction that contains a subject and a predicate, forming part of a sentence or constituting a whole simple sentence (Norris & Ortega, 2009). Global measures of accuracy are more preferred by researchers (Lee, Joo, Moon, & Hong 2007; Wigglesworth & Storch, 2009; Wolfe-Quintero et al., 1998).
Among the variety of methods for measuring accuracy, Foster and Wigglesworth (2016) claim that, although clause as a unit is more preferable, it has two problems: on the one hand, more complex sentences entail multiple clauses, which makes them unfairly disadvantaging for the students who produce longer complex sentences with more clauses because they are more likely to make more errors in them; on the other, this way of measuring accuracy may also result in a problem that, when clauses are binarily distinguished as error-free and not error-free, not error-free ones bear the same responsibility for the accuracy of the whole clause no matter how minor or serious the error is, which, in another words, means that the not error-free clause might not be blamed enough if it contains serious errors. In order to make a more fine-tuned distinction, some researchers categorized the errors into different levels based on the effects they have on the comprehensibility of the clause (Nas, 1975) or their impact on communicative success or adequacy (Kuiken & Vedder, 2008). However, even though they sorted the errors into levels of error severity, they only provided information about the proportion of the data characterized by these errors, not combining and integrating them into the data set as a whole.

Wigglesworth and Foster (2008) proposed a new method to measure accuracy, which is the weighted clause ratio (WCR). It combines all the errors in a global score by categorizing each clause according to the gravity of error within it. Evans, Hartshorn, Cox, and Jel (2014) examined the validity of the WCR. Their study is the first empirical study applying the WCR to measure accuracy. They tested the validity of the WCR and claimed that WCR is much more precise than other measures, such as error-free T-units, error-free clauses. Table 8 and Table 9 give a detailed explanation of how units of analysis (clauses) are categorized for this measure and the procedure of weighing and scoring errors.
Table 8

Clause categorization for weighted clause ratio (WCR)

<table>
<thead>
<tr>
<th>Clause category</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entirely accurate</td>
<td>The clause is accurately constructed.</td>
</tr>
<tr>
<td>Level 1</td>
<td>The cause has only minor errors (e.g., in morphosyntax) that do not compromise meaning.</td>
</tr>
<tr>
<td>Level 2</td>
<td>The clause contains serious errors (e.g., verb tense, word choice or word order) but the meaning is recoverable, though not always obvious.</td>
</tr>
<tr>
<td>Level 3</td>
<td>The clause has very serious errors that make the intended meaning far from obvious and only partly recoverable.</td>
</tr>
</tbody>
</table>

*Note.* From Wigglesworth and Foster (2008).

Foster and Wigglesworth (2016) used the accuracy measure, WCR on both oral and written data produced by students of different L2 proficiency levels. They reported that WCR potentially reflects the researcher’s information about what kinds of errors that learners make at different levels of proficiency, at different times, and in different conditions that other measures do not, allowing easier cross-study comparisons. Foster and Wigglesworth (2016) believe that this new measure, grounded in a comprehensive review of prior practice in the field, has the advantages of being relatively easy to use, measuring accuracy rather than error, and being very sensitive to small changes in accuracy.

Based on the review about the accuracy measure, the WCR, together with the error-free T-unit ratios (%EFT) and error-free clause ratios (EFC), will be applied to measure accuracy of the participants’ written texts in this research.

Table 9

Scoring the clause categories

<table>
<thead>
<tr>
<th>Clause Category</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entirely accurate</td>
<td>1.00</td>
</tr>
<tr>
<td>Level 1 error</td>
<td>0.80</td>
</tr>
<tr>
<td>Level 2 error</td>
<td>0.50</td>
</tr>
<tr>
<td>Level 3 error</td>
<td>0.10</td>
</tr>
</tbody>
</table>

*Note.* From Wigglesworth and Foster (2008).
As for the WCR, Wigglesworth and Foster’s (2008) provided criteria for categorization and scoring of clauses according to the error types the clauses contain (see Appendix D for the error types that make each clause type). Clauses can be categorized into the following four categories, namely entirely accurate clause, level 1 clause, level 2 clause, and level 3 clause, and different types of clauses are given different scores (see Table 8 for definitions of each clause type and Table 9 for the ways of scoring the clause categories). Since I marked the level of error gravity based on Wigglesworth and Foster’s (2008) criteria for clauses categorization and scoring (see Table 8 and Table 9) when identifying and counting the number of clauses earlier, the last step left was to do the calculations. First, I calculated the score of entirely accurate, level 1, level 2 and level 3 clauses by multiplying them by the specific score attributed to that level (1, 0.8, 0.5, 0.1 respectively) (see Table 9). Then, I used SPSS to calculate WCR by dividing the sum of the clause scores by the total number of the clauses. Table 10 presents an illustration showing how the samples were rated.
Table 10

*Example of categorization and scoring of clauses*

<table>
<thead>
<tr>
<th>Clauses</th>
<th>Error level</th>
<th>Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 On one evening, I had an interesting dream.</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>2 I dreamed</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>3 that I would go to the place by plane.</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>4 I wanted to find a museum</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>5 because my friend invited me to join a ceremony.</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>6 She said</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>7 too many people would come to join.</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>8 So I was very happy.</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>9 On the way to the museum, a terrible thing happened.</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>10 I got lost in the storm.</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>11 Last, I was setting on a island.</td>
<td>1</td>
<td>0.8</td>
</tr>
<tr>
<td>12 I was very afraid</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>13 and I wanted to go home now</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>14 In this time, a dragon turned up in the sky.</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>15 It seems to take me home.</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>16 I believed it.</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>17 And I arrived home in five minutes.</td>
<td>2</td>
<td>0.5</td>
</tr>
<tr>
<td>18. I thought I was very lucky.</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>19 When I woke up,</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>20 I found</td>
<td>Accurate</td>
<td>1.0</td>
</tr>
<tr>
<td>21 it is just a dream.</td>
<td>2</td>
<td>0.5</td>
</tr>
</tbody>
</table>

\[
\text{WCR} = \frac{(13 \times 1.0 + 4 \times 0.8 + 4 \times 0.5)}{21} = 0.87
\]

*Note:* WCR=Weighted Clause Ratio

### 3.5.2 Measures of flow

In the current study, flow was measured by a five-point Likert scale adapted from Czimmermann and Piniel (2016) but was modified for the current study. The final flow questionnaire consisted of the 7 scales proposed in Czimmermann and Piniel (2016) with 31 items, namely interest (5 items) (item 1, 5, 11, 16, 27), attention (4 items) (item 3, 8, 14, 28), control (5 items) (item 2, 6, 12, 17, 30), challenge-skill balance (4 items) (item 4, 10, 15, 22), boredom (4 items) (item 7, 18, 25, 29), apathy (4 items) (item 19,
21, 23, 24) and anxiety (5 items) (item 9, 13, 20, 26, 31) (see Appendix E for the flow questionnaire).

Even though the flow questionnaire was piloted with a group of students who were not the participants for the current study (see Table 7 for the result of the flow questionnaire pilot), a new reliability analysis was carried out on my actual participants. What is important to mention is that the two added items under the subscale of control after piloting worked well; therefore, the final number of items for control was 5. Table 11 presents the descriptive statistics of all the scales in each task. Principal component analysis was also adopted to measure the consistent reliability of the scale. The new reliability figures (Cronbach alphas) are reported in Table 12 for the 40 Chinese students. Cronbach’s alphas for all the scales are above 0.60. Although there were minor inconsistencies for the scale of attention and the scale of anxiety, indicating these two were potentially problematic and maybe more investigation is needed, the majority of the scales worked successfully, showing a high reliability of the scale.

Table 11

Descriptive statistics of each flow and anti-flow variable in each task

<table>
<thead>
<tr>
<th>Measures</th>
<th>N</th>
<th>TG</th>
<th>LG</th>
<th>C</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>Interest</td>
<td>5</td>
<td>2.94</td>
<td>0.72</td>
<td>3.10</td>
</tr>
<tr>
<td>Attention</td>
<td>4</td>
<td>3.23</td>
<td>0.73</td>
<td>3.33</td>
</tr>
<tr>
<td>Control</td>
<td>5</td>
<td>3.53</td>
<td>0.82</td>
<td>3.80</td>
</tr>
<tr>
<td>Csbalance</td>
<td>4</td>
<td>3.40</td>
<td>0.73</td>
<td>3.60</td>
</tr>
<tr>
<td>Boredom</td>
<td>4</td>
<td>2.67</td>
<td>0.82</td>
<td>2.51</td>
</tr>
<tr>
<td>Apathy</td>
<td>4</td>
<td>2.38</td>
<td>0.74</td>
<td>2.38</td>
</tr>
<tr>
<td>Anxiety</td>
<td>5</td>
<td>2.33</td>
<td>0.71</td>
<td>2.30</td>
</tr>
</tbody>
</table>

Note. N=Number; TGT=Teacher-Generated Content Task; LGT=Learner-generated content Task; CT=Creative Task; SD=Standard Deviation; Csbalance=Challenge-skill balance
Table 12

Reliability analysis, PCA of each flow and anti-flow variable

<table>
<thead>
<tr>
<th>Measures</th>
<th>N</th>
<th>TGT</th>
<th>LGT</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cronbach’s alpha</td>
<td>PCA</td>
<td>Cronbach’s alpha</td>
</tr>
<tr>
<td>Interest</td>
<td>5</td>
<td>0.80</td>
<td>1</td>
<td>0.85</td>
</tr>
<tr>
<td>Attention</td>
<td>4</td>
<td>0.72</td>
<td>2</td>
<td>0.75</td>
</tr>
<tr>
<td>Control</td>
<td>5</td>
<td>0.84</td>
<td>1</td>
<td>0.76</td>
</tr>
<tr>
<td>Csbalance</td>
<td>4</td>
<td>0.72</td>
<td>1</td>
<td>0.63</td>
</tr>
<tr>
<td>boredom</td>
<td>4</td>
<td>0.80</td>
<td>1</td>
<td>0.88</td>
</tr>
<tr>
<td>Apathy</td>
<td>4</td>
<td>0.74</td>
<td>1</td>
<td>0.73</td>
</tr>
<tr>
<td>Anxiety</td>
<td>5</td>
<td>0.75</td>
<td>2</td>
<td>0.72</td>
</tr>
</tbody>
</table>

Note. PCA=Principle Component Analysis; N=Number; TGT=Teacher-Generated Content Task; LGT=Learner-generated content Task; CT=Creative Task; Csbalance=Challenge-skill balance

3.6 Statistical analysis

The software SPSS 22.0 for Windows was used to analyze data that were derived from the analytical procedures. Cronbach’s alphas were calculated to establish the reliability of the flow questionnaire. In addition, for each group of Chinese students, ANOVAs with repeated measures were used to analyze the differences in their written performances (as measured by fluency, complexity, and accuracy) and their flow and anti-flow experiences (measured by interest, attention, control, challenge-skill balance, apathy, boredom, and anxiety) across different tasks. In addition, for each task performed by the two groups of Chinese students respectively, correlations were used to explore the paired correlation between written linguistic performances and students’ flow and anti-flow experiences, while controlling for language proficiency (as measured by the placement test scores) to remove its influence on the variables of interest. Finally, independent samples t-tests were conducted to compare the students’ written performances and their flow and anti-flow experiences in each task condition between the two groups of Chinese students, the CSH and the CSC, to see the role played by English study contexts in their English writing and their flow experiences.
Chapter 4: Findings concerning the Written Performances and the Flow and Anti-flow Experiences of the CSH

This part of the dissertation presents the descriptive statistics regarding the written performances of the CSH and their engagement as shown by their flow and anti-flow experiences in different task conditions. Then repeated measures ANOVAs were used to explore the effects of different task conditions on task performance and task-specific flow and anti-flow. This line of study might provide some insights for teachers and writing task designers in designing writing tasks to enhance students’ engagement in doing the task and to elicit good written linguistic performances.

4.1 Results on the relationship between task conditions and the written performances of the CSH

This section of the dissertation presents the results of the current study concerning the written performances of the CSH. The students’ written performances were assessed in terms of fluency, complexity, and accuracy. Table 13 presents the descriptive statistics for the means and standard deviations (SD) of the written performance measures in different tasks.
Table 13

*Descriptive statistics of the written performance measures of the CSH in different tasks*

<table>
<thead>
<tr>
<th>Measures</th>
<th>TGT</th>
<th>LGT</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>TW</td>
<td>20</td>
<td>194.10</td>
<td>57.89</td>
</tr>
<tr>
<td>D-value</td>
<td>20</td>
<td>64.70</td>
<td>12.10</td>
</tr>
<tr>
<td>P_Lex</td>
<td>20</td>
<td>1.08</td>
<td>0.34</td>
</tr>
<tr>
<td>MLT</td>
<td>20</td>
<td>8.94</td>
<td>1.51</td>
</tr>
<tr>
<td>MNCT</td>
<td>20</td>
<td>1.23</td>
<td>0.13</td>
</tr>
<tr>
<td>MLC</td>
<td>20</td>
<td>7.31</td>
<td>1.14</td>
</tr>
<tr>
<td>%EFC</td>
<td>20</td>
<td>68.25</td>
<td>21.85</td>
</tr>
<tr>
<td>%EFT</td>
<td>20</td>
<td>62.48</td>
<td>25.55</td>
</tr>
<tr>
<td>WCR</td>
<td>20</td>
<td>0.88</td>
<td>0.09</td>
</tr>
</tbody>
</table>

*Note. CSH=Chinese Students studying in Hungary; TGT=Teacher-Generated Content Task; LGT=Learner-Generated Content Task; CT=Creative Task; N=Number; SD=Standard Deviation; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clause per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.*

Regarding fluency, using Pillai’s trace, for the CSH, repeated measures ANOVA revealed no relationship between task conditions and fluency (indicated by TW), $V=0.14$, $F(2, 18)=1.42$, $p=0.269$, partial $\eta^2=0.14$. However, the standard deviations of fluency produced in the three different task conditions were 57.89, 58.87, and 111.78 respectively, showing that there was a large variation among the students in this group concerning fluency in the CT condition. It seems that while doing the task in the CT condition, some students became extraordinarily excited and engaged, and wrote much more words than their peers.

Regarding lexical complexity indicated by D-value and P_Lex, using Pillai’s trace, repeated measures ANOVA revealed that the task conditions were not associated with D-value, $V=0.13$, $F(2, 18)=1.35$, $p=0.28$, partial $\eta^2=0.13$. On the contrary, repeated measures ANOVA showed a significant relationship between task condition and P_Lex, $V=0.32$, $F(2, 18)=4.28$, $p=0.03$, partial $\eta^2=0.32$. Cohen (1988) suggested that partial $\eta^2$ at 0.01, 0.09, and 0.25 stand for small, medium, and large effects respectively (Tabackhnick & Fidell, 2007, p. 55). Follow up comparison indicated that the pairwise difference between the TGT and the CT was significant ($p<0.05$). The students produced
a higher P_Lex in the TG condition than in the CT condition (Mean=1.08 for the TGT and Mean=0.77 for the CT respectively), indicating that students used more rare words in the TG condition.

In terms of syntactic complexity indicated by MLT, MNCT, and MLC, using Pillai’s trace, repeated measures ANOVA revealed that the task condition were not associated with: (1) MLT, V=0.03, F (2, 18)=0.25, p=0.79, partial η2=0.03, and (2) MLC, V=0.06, F (2, 18)=0.61, p=0.56, partial η2=0.06. However, repeated measures ANOVA revealed a relationship between task conditions and MNCT, V=0.43, F (2, 18)= 6.78, p=0.01, partial η2=0.43. Follow up comparison indicated that the pairwise difference between the TG condition and the LG condition was strongly significant (p<0.01). The students produced a lower MNCT in the TG condition than in the LG condition (Mean=1.23 for the TGT and Mean=2.43 for the LGT respectively), showing that the students wrote more simple sentences or subordinations in complex sentences in the LG condition than in the TG condition. Besides, the pairwise difference in MNCT between the LG condition and the CT condition was also significant (p<0.05). The students produced a higher MNCT in the LG condition than in the CT condition (Mean=2.43 for the LGT and Mean=1.27 for the CT respectively), which means that the students produced a greater complexity in the LG condition compared with that in the CT condition.

Regarding accuracy, using Pillai’s trace, repeated measures ANOVA showed that the task conditions were not related to: (1) %EFC, V=0.20, F (2, 18)=2.31, p=0.13, partial η2=0.20, (2) %EFT, V=0.12, F (2, 18)=1.21, p=0.32, partial η=0.12, and (3) WCR, V=0.02, F (2, 18)=0.17, p=0.84, partial η2=0.02. This seems to indicate that for the CSH, there was no main association between task conditions and accuracy.

In summary, it seems to show that an increase of learner agency (from the TGT to the LGT) was not associated with fluency and accuracy, but brought about a significant increase in syntactic complexity in the sense of MNCT, but not in the sense of MLT and MLC. Besides, an increase of the potential for creativity (from the LGT to the CT) seemed to result in a significant decrease of syntactic complexity (in the sense of MNCT, but not in the sense of MLT and MLC), but was not related to fluency and accuracy. Moreover, an increase of both learner agency and potential for creativity (from the TGT to the CT) seemed to lead to a decrease of lexical variety, but did not influence fluency, syntactic complexity, and accuracy.
4.2 Discussion on the relationship between task conditions and the written performances of the CSH

This section of the dissertation discusses the research results described in section 4.1. The effects of the task conditions on the written performances of the CSH were analyzed in detail.

In terms of fluency, for the CSH, no statistically significant difference was found in terms of fluency across the three tasks. This finding is inconsistent with some similar studies on the effects of TG condition and LG condition on oral task performances (e.g., Lambert et al., 2017; Poupore, 2014) where learners produced higher fluency in the LG condition, showing that higher learner agency and more potential for creativity in writing tasks do not necessarily lead to higher fluency. Besides, this finding seemed to be contradictory to Maehr’s (1984) personal investment theory in the sense that meaningfulness and the potential to elicit creativity that a task has does not necessarily lead to better performance in terms of fluency. However, the standard variation score for fluency produced in the CT was much larger than that produced in the other two task conditions even though the difference in fluency across the three tasks was not statistically significant (SD=57.89, 58.89 and 111.78 in the TGT, LGT and CT respectively), which indicated that there was larger individual variation regarding the number of words produced by the participants in this group in the CT. Therefore, it can be inferred that fluency in the CT condition was greatly individual-dependent. Since the students in this group were quite homogenous concerning their language proficiency, we might attribute this variation in fluency to individual preferences for this task.

In terms of lexical complexity, for the CSH, task conditions were found to be related to lexical complexity of the students’ written texts. First, between the TGT and the CT with regard to lexical variety (indicated by P_Lex), it was found that words that students used in the TG condition were more difficult and rare than those used in the CT condition. A reason for this might be that the predetermined content illustrated by the pictures forced the students to use certain words to express certain ideas that were not necessarily easy, while in the CT, the students could decide to use the words they were sure about and avoid the difficult ones that they were not sure about. This characteristic of the TGT might benefit their lexical complexity (Kormos, 2011; Lambert & Zhang, 2019). Another possible explanation for this finding might be that, according to Robinson’s (2001b, 2003, 2005) Cognition Hypothesis, the CT task was more difficult and complex than the TGT along the resource-directing dimension since it actually involved two tasks, inventing a story and telling it in English, thus increased the demand on the part of the learner for conceptualization (Skehan, 1998) and planning (Kellogg,
1996), as a result of which, the students had to use more attentional resources on the meaning of their message rather than on the form of the language. This finding partially contradicted Robinson’s Cognition Hypothesis that argues that more task complexity along the resource-directing dimension leads to more complexity in task performance.

In terms of syntactic complexity, the students produced greater complexity in terms of mean number of clauses per T-unit in the LG condition than that in the TG condition, and the difference was statistically significant. It seems that the complete control over what to write and how to write it in the LGT enables the students to focus more on the formulation phase of writing, where they, not only generate content, but also encode the language for later use (e.g., organize the structure and select lexical and syntactic frames that they might need in the actual writing phase) (Kellogg, 1996). Therefore, they might write more smoothly using more complex sentence structures during the writing process. In the TGT, nevertheless, the students seemed to translate the content presented on the pictures in an orderly fashion, so they used shorter sentences with more simple sentence structures. Besides, the students might become more willing to do the LGT and more engaged in the LGT since the students were given more control over the content to write and they could write anything of personal interest and personal relevance, which might arouse their intrinsic interest and helped foster their writing engagement, thus leading to a better performance (Liu, 2006; Maehr, 1984; Reeve et al., 2004; Ren, 2017; Reschly et al., 2008; Tanaka, 2007). This finding is partially consistent with the results of some similar studies on the effects of TG and LG conditions on oral task performance which found that greater complexity was produced in the LG condition (Lambert & Minn, 2007; Lambert et al., 2017; Poupore, 2014; Shao, 2003). Moreover, looking at it from the perspective of Robinson’s Cognition Hypothesis, the LGT was more complex and resulted in greater complexity. Therefore, the finding supports Robinson’s claim that increased task complexity along the resource-directing dimension increases linguistic complexity.

Besides, when comparing the LGT and the CT, the students produced greater complexity in terms of mean number of clauses per T-unit in the LGT. Interpreted from Robinson’s Cognition Hypothesis, the CT was supposed to be more cognitively demanding since the students were required to read the words and choose at least six of the ten and make connections to produce a complete story. However, the LGT was supposed to be comparatively easier since the students could rely on their personal experiences and write about them, therefore, they might have written about something that they had already experienced or talked about before, so the content was probably familiar to them. So, in this sense, the LGT was easier than the CT because it likely
involved topic familiarity, while the CT was new to them, requiring more effort at conceptualization and less attentional resources for formulation (Kellogg, 1996; Skehan, 1998), thus leading to lower syntactic complexity. It seems that, in the CT, a chance for word association and creativity, instead of encouraging variability and diversity in lexical choices as well as creative language use, probably added more burden in content conceptualization on the students’ part.

In terms of accuracy, no statistically significant difference was found among the three different task conditions. This might be due to the participants’ intermediate-level English proficiency, which made their text accuracy stand independent from the effects of task conditions. The students might have interpreted their involvement in the study as a testing situation, and they might have tried to avoid taking risks of using difficult language so as not to make mistakes.

4.3 Results on the relationship between task conditions and the flow and anti-flow experiences of the CSH

Table 1 presents the descriptive statistics of flow and anti-flow measures that showed the flow experiences of the CSH in different tasks. Using Pillai’s trace, repeated measures ANOVA revealed no relationship of the task conditions and (1) interest, $V=0.16$, $F(2, 18)=1.68$, $p=0.22$, partial $\eta^2=0.16$; (2) attention, $V=0.06$, $F(2, 18)=0.60$, $p=0.56$, partial $\eta^2=0.06$; and (3) control, $V=0.28$, $F(2, 18)=3.54$, $p=0.05$, partial $\eta^2=0.28$. On the contrary, repeated measures ANOVA revealed that the task conditions were associated with challenge-skill balance, $V=0.31$, $F(2, 18)=4.13$, $p=0.03$, partial $\eta^2=0.31$. Follow up comparison indicated that the pairwise difference between the TGT and the CT was significant ($p<0.05$), indicating that the students felt a higher balance of challenge and skill in the CT than in the TGT (Mean=3.95 for the CT and Mean=3.58 for the TGT respectively).
Table 14

Descriptive statistics of the flow and anti-flow measures of the CSH in different tasks

<table>
<thead>
<tr>
<th>Measures</th>
<th>TGT</th>
<th>LGT</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>Interest</td>
<td>2.99</td>
<td>0.70</td>
<td>3.12</td>
</tr>
<tr>
<td>Attention</td>
<td>3.36</td>
<td>0.63</td>
<td>3.53</td>
</tr>
<tr>
<td>Control</td>
<td>3.79</td>
<td>0.69</td>
<td>4.13</td>
</tr>
<tr>
<td>Csbalance</td>
<td>3.58</td>
<td>0.61</td>
<td>3.84</td>
</tr>
<tr>
<td>Boredom</td>
<td>2.78</td>
<td>0.67</td>
<td>2.45</td>
</tr>
<tr>
<td>Apathy</td>
<td>2.28</td>
<td>0.56</td>
<td>2.18</td>
</tr>
<tr>
<td>Anxiety</td>
<td>2.08</td>
<td>0.58</td>
<td>2.05</td>
</tr>
</tbody>
</table>

Note. CSH=Chinese Students studying in Hungary; TGT=Teacher-Generated Content Task; LGT=Learner-Generated Content Task; CT=Creative Task; N=Number; SD=Standard Deviation; Csbalance=Challenge-skill balance

Moreover, using Pillai’s trace, repeated measures ANOVA revealed no association between task conditions and boredom, $V=0.15$, $F(2, 18)=1.57$, $p=0.24$, partial $\eta^2=0.15$. On the contrary, repeated measures ANOVA revealed that the task conditions were associated with apathy, $V=0.30$, $F(2, 18)=3.92$, $p=0.04$, partial $\eta^2=0.30$. Follow up comparison indicated that the pairwise difference between the TGT and the CT was significant ($p<0.05$). The students went through more apathy in the TG condition than in the CT condition (Mean=2.28 for the TGT and Mean=1.90 for the CT respectively). Besides, using Pillai’s trace, repeated measures ANOVA revealed that the task conditions were associated with anxiety, $V=0.40$, $F(2, 18)=5.93$, $p=0.01$, partial $\eta^2=0.40$. Follow up comparison indicated that the pairwise difference between the TGT and the CT was significant ($p<0.05$). The students felt less anxiety in the CT condition than in the TG condition (Mean=1.75 for the CT and Mean=2.08 for TGT respectively). Besides, follow-up comparison indicated that the pairwise difference between the LGT and the CT was significant ($p<0.05$). The students felt less anxiety in the CT condition than in the LG condition (Mean=1.75 for CT and Mean=2.05 for LGT respectively).

However, according to the mean values for each variable, a tendency seemed to appear that, across the three tasks, the CT evoked the highest flow experiences while the TGT did the least. In terms of anti-flow, consistently with the students’ flow experiences, it seemed that the CT was the least anti-flow provoking while the TGT was the most anti-flow provoking.
In summary, the results showed that for the CSH, an increase in learner agency in tasks (from the TGT to the LGT) did not impact the students' flow and anti-flow experience. Besides, an increase concerning control and the potential for creativity together in the tasks (from the TGT to the CT) seemed to result in less apathy and less anxiety on the students. Moreover, an increase in the potential for creativity in the tasks (from the LGT to the CT) was likely to bring about less anxiety on the students as well.

4.4 Discussion on the relationship between task conditions and the flow and anti-flow experiences of the CSH

For this group of Chinese students, an interesting finding was that the CT gave the students a greater sense of challenge-skill balance than the TGT. A possible explanation for this might be that the TGT resulted in an unfavorable blend of low skills and low challenges for the students since all that they needed to do was to translate what was presented on the pictures. This was also the cause of the appearance of more apathy in the TGT, too. To be more specific, a significant difference was detected between the two tasks concerning apathy, showing that the students felt more apathy in the TGT than in the CT. As has been shown in many studies (e.g., Csikszentmihalyi, 1989; Czimmermann & Piniel, 2016; Dewaele & MacIntyre, 2014), apathy occurs because the task is less challenging or the students do not own enough skills for the task. Nevertheless, the CT which was half-constrained and half-controlled for the students turned out to be better matched and balanced in task challenge and the students’ writing skills.

Besides, the CT seemed to be less anxiety inducing than the TGT. Comparatively speaking, the CT was more difficult than the TGT, no matter with regard to either task complexity (Robinson, 2001b, 2003, 2005), or the content-planning (Kellogg, 1996). A possible explanation for this might be that the CT which gave them a better match between task challenge and skill might lead to favorable emotions, such as an interest in the task and enjoyment brought out by confidence of being capable to complete the task. This explanation can be supported by the fact that, according to the mean values of raw data, in the CT, the flow variables (interest, attention, and control) were higher than those in the TGT, indicating that the students were more likely to experience flow in the CT. Therefore, they sensed less negative emotions such as anxiety in the CT.

Moreover, between the LGT and the CT, the students experienced higher anxiety in the LGT than in the CT. It seemed that the potential for creativity that a task had for the students was likely to compensate for the potential anxiety which was likely to associate with task constraint. A possible explanation might be that in the CT, the opportunities to
create ideas by imagination was an important factor that determined learners’ preference for the task and made them pleased, thus making them more engaged while doing the task. When they were engaged in what they were doing, they could free themselves from anxiety. Nevertheless, the LGT task that provided the students with more autonomy about what to write seemed more boring and less exciting for the students, as can been seen from the lower mean values for the flow variables (interest, attention, and control) and the higher mean values for anti-flow variables (boredom and apathy), showing that they were not very engaging in performing the task. When they were not engaged and involved in writing, instead, they might have felt anxious because they might worry if they could meet the task requirement and complete the task, or they might be afraid of disappointing the teacher.

In addition, for the CSH, no significant difference was observed concerning control across the three tasks, which was unexpected. However, from the mean values of the raw data, the TGT seemed to contain the least control for the students, while the CT contained the most. That the differences in control across the three tasks were not significant might be due to the small sample size of the present study.

Finally, it seems that the students’ flow and anti-flow experiences were not significantly different between the TGT and the LGT. However, the mean values for flow and anti-flow measures seemed to show a tendency that the LG condition was more flow-generating and less anti-flow generating than the TGT. A possible explanation for this might be that the TGT had a clear and fix content, which prevented the participants from generating their own ideas; therefore, this task characterized by little learner agency was less personally meaningful and less emotionally engaging for them (Maehler, 1984). This finding, although not significantly different possibly due to the small sample size of the present study, seemed consistent with the results of earlier studies on the effects of task conditions on the university students’ or adults’ oral task performance which found that the students reported having or were found experiencing more positive feelings (i.e., interest, confidence, motivation) in the LG condition where the topics were more familiar and personally relevant, and of more personal interest (e.g., Lambert & Minn, 2007; Lambert et al., 2017; Phung, 2016; Poupore, 2014; Qiu & Lo, 2016). This seems to infer that when it comes to the effects of the TG condition versus the LG condition on learners’ task performance, the task mode (oral versus written), as well as the English study contexts of the participants in the current study seem not to link to the relationships between task conditions and task performance since no difference in flow and anti-flow was observed across the two tasks at a group level.
4.5 Results on the correlations between the flow and anti-flow experiences and the written performances of the CSH in different task conditions while controlling for language proficiency

After examining the relationships between task conditions and the written performance measures and flow and anti-flow measures respectively, it might be interesting to see how they correlate with each other. In order to remove the possible effect of proficiency on the relationships between task output measures and flow and anti-flow measures, partial correlations were used to analyze the relationships between students’ performances on the tasks and their flow and anti-flow experiences in each task while controlling for their proficiency score. This section presents the findings.

4.5.1 In the TGT

Table 15 presents the correlations between the written performance measures and flow and anti-flow measures as independent variables in the TGT while controlling for language proficiency. It can be seen that MLT and MNCT (indicating syntactic complexity) and WCR (showing accuracy) seemed not to be correlated with any flow and anti-flow measures. However, fluency indicated by TW was found to moderately correlate with apathy positively (r=0.46, p<0.05). As for the two measures indicating lexical complexity, D-value and P_Lex, they changed with some of the flow and anti-flow measures, but mostly in a moderate way. For instance, the lexical diversity shown by the D-value was slightly correlated with interest negatively (r=-0.47, p<0.05), while weakly correlated with both boredom (r=0.56, p<0.05) and anxiety (r=0.46, p<0.05) positively, indicating that when the students were less interested in and more bored and anxious during the task performance, they produced more diverse vocabulary in their writing, which is unexpected. Meanwhile, the other measure of lexical complexity indicated by P_Lex was also found to mildly correlate with anxiety (r=0.51, p<0.05) positively, suggesting that when the participants felt more anxious when doing the task, they produced more rare words. With regard to syntactic complexity, MLC was found to mildly positively correlate with control (r=0.58, p<0.05), showing that the students produced longer sentences of simple sentence structures and longer subordinations in complex sentences when they sensed more control over the task. In terms of accuracy, %EFT moderately correlated with attention (r=0.47, p<0.05) positively, suggesting that more attention was likely to result in higher ratio of accurate T-units. Besides, the %EFC that indicated accuracy was slightly negatively correlated with anxiety (r=-0.48, p<0.05), indicating that more anxiety seemed to lead to a lower ratio of accurate clauses.
Table 15

*Correlations between the written performance measures and flow and anti-flow measures in the TGT controlling for language proficiency*

<table>
<thead>
<tr>
<th>Measures</th>
<th>Interest</th>
<th>Attention</th>
<th>Control</th>
<th>Csbalance</th>
<th>Boredom</th>
<th>Apathy</th>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td>0.10</td>
<td>0.03</td>
<td>-0.26</td>
<td>-0.10</td>
<td>-0.01</td>
<td>0.46*</td>
<td>0.27</td>
</tr>
<tr>
<td>D-value</td>
<td>-0.47*</td>
<td>-0.42</td>
<td>-0.32</td>
<td>-0.01</td>
<td>0.56*</td>
<td>0.38</td>
<td>0.46*</td>
</tr>
<tr>
<td>P_Lex</td>
<td>-0.06</td>
<td>-0.38</td>
<td>0.12</td>
<td>-0.04</td>
<td>0.43</td>
<td>0.22</td>
<td>0.51*</td>
</tr>
<tr>
<td>MLT</td>
<td>-0.16</td>
<td>-0.07</td>
<td>0.43</td>
<td>-0.28</td>
<td>0.30</td>
<td>0.20</td>
<td>-0.09</td>
</tr>
<tr>
<td>MNCT</td>
<td>-0.06</td>
<td>0.09</td>
<td>-0.09</td>
<td>-0.35</td>
<td>0.03</td>
<td>0.18</td>
<td>-0.01</td>
</tr>
<tr>
<td>MLC</td>
<td>-0.15</td>
<td>-0.12</td>
<td>0.58*</td>
<td>-0.03</td>
<td>0.29</td>
<td>0.08</td>
<td>-0.12</td>
</tr>
<tr>
<td>%EFC</td>
<td>0.22</td>
<td>0.46</td>
<td>0.19</td>
<td>0.30</td>
<td>-0.32</td>
<td>-0.32</td>
<td>-0.48*</td>
</tr>
<tr>
<td>%EFT</td>
<td>0.25</td>
<td>0.47*</td>
<td>0.09</td>
<td>0.40</td>
<td>-0.36</td>
<td>-0.34</td>
<td>-0.42</td>
</tr>
<tr>
<td>WCR</td>
<td>0.33</td>
<td>0.30</td>
<td>0.14</td>
<td>0.27</td>
<td>-0.33</td>
<td>-0.29</td>
<td>-0.31</td>
</tr>
</tbody>
</table>

*Note.* TGT=Teacher-Generated content Task; N=Number; Csbalance=Challenge-skill balance; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.

4.5.2 *In the LGT*

Table 16 presents the correlations between the written performance measures and flow and anti-flow measures of this task while controlling for language proficiency. First of all, fluency was found negatively correlated with boredom ($r=-0.46, p<0.05$). With regard to complexity, MLC was positively correlated with anxiety ($r=0.54, p<0.05$), showing that the students were likely to produce longer sentences of simple sentence structures or longer subordinations in complex sentences when they were feeling more anxious. Besides, %EFC which indicated accuracy was correlated with anxiety negatively ($r=-0.47, p<0.05$), indicating that students produced a higher ratio of error-free clauses when they were less anxious.
Table 16

Correlations between the written performance measures and flow and anti-flow measures in the LGT controlling for language proficiency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Interest</th>
<th>Attention</th>
<th>Control</th>
<th>Csbalance</th>
<th>Boredom</th>
<th>Apathy</th>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td>0.40</td>
<td>0.39</td>
<td>-0.09</td>
<td>0.30</td>
<td>-0.46*</td>
<td>-0.09</td>
<td>-0.04</td>
</tr>
<tr>
<td>D-value</td>
<td>-0.08</td>
<td>-0.06</td>
<td>-0.41</td>
<td>-0.20</td>
<td>0.05</td>
<td>0.12</td>
<td>0.31</td>
</tr>
<tr>
<td>P_Lex</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.18</td>
<td>0.25</td>
<td>-0.12</td>
<td>-0.39</td>
<td>-0.41</td>
</tr>
<tr>
<td>MLT</td>
<td>0.15</td>
<td>0.01</td>
<td>0.26</td>
<td>0.25</td>
<td>-0.18</td>
<td>-0.27</td>
<td>-0.05</td>
</tr>
<tr>
<td>MNCT</td>
<td>-0.11</td>
<td>0.09</td>
<td>0.25</td>
<td>-0.02</td>
<td>-0.07</td>
<td>-0.13</td>
<td>-0.06</td>
</tr>
<tr>
<td>MLC</td>
<td>0.20</td>
<td>-0.16</td>
<td>-0.25</td>
<td>0.03</td>
<td>0.18</td>
<td>0.41</td>
<td>0.54*</td>
</tr>
<tr>
<td>%EFC</td>
<td>0.11</td>
<td>0.077</td>
<td>0.01</td>
<td>0.23</td>
<td>-0.10</td>
<td>-0.149</td>
<td>-0.47*</td>
</tr>
<tr>
<td>%EFT</td>
<td>0.01</td>
<td>0.137</td>
<td>0.02</td>
<td>0.10</td>
<td>-0.05</td>
<td>-0.09</td>
<td>-0.34</td>
</tr>
<tr>
<td>WCR</td>
<td>0.09</td>
<td>-0.3</td>
<td>-0.12</td>
<td>0.19</td>
<td>0.10</td>
<td>-0.16</td>
<td>-0.39</td>
</tr>
</tbody>
</table>

Note. LGT=Learner-Generated content Task; N=Number; Csbalance=Challenge-skill balance; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.

* p<0.05. ** p<0.01.

4.5.3 In the CT

Table 17 presents the correlations between the written performance measures and flow and anti-flow measures of the CT while controlling for language proficiency. It can be seen that syntactic complexity indicated by MLT, MNCT, and MLC, and all the three accuracy measures seemed not to change with the flow and anti-flow measures. Nevertheless, fluency (reflected by TW) had a relatively low positive correlation with attention ($r=0.48$, $p<0.05$), showing that more attention might be associated with higher fluency. Concerning lexical complexity, lexical diversity indicated by D-value negatively correlated with boredom ($r=-0.46$, $p<0.05$). Meanwhile, P_Lex was positively correlated with challenge-skill balance ($r=0.49$, $p<0.05$). This finding indicated that the higher challenge-skill balance and less boredom the students had during task performance, the greater the lexical complexity they were likely to produce.
Table 17

Correlations between the written performance measures and flow and anti-flow measures in the CT controlling for language proficiency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Interest</th>
<th>Attention</th>
<th>Control</th>
<th>Csbalance</th>
<th>Boredom</th>
<th>Apathy</th>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td>0.36</td>
<td>0.48*</td>
<td>-0.29</td>
<td>-0.04</td>
<td>-0.35</td>
<td>-0.11</td>
<td>-0.02</td>
</tr>
<tr>
<td>D-value</td>
<td>0.32</td>
<td>0.23</td>
<td>0.24</td>
<td>0.43</td>
<td>-0.46*</td>
<td>-0.30</td>
<td>-0.26</td>
</tr>
<tr>
<td>P_Lex</td>
<td>0.26</td>
<td>0.40</td>
<td>0.32</td>
<td>0.49*</td>
<td>-0.40</td>
<td>-0.18</td>
<td>-0.14</td>
</tr>
<tr>
<td>MLT</td>
<td>-0.17</td>
<td>0.17</td>
<td>0.07</td>
<td>0.04</td>
<td>0.19</td>
<td>0.28</td>
<td>0.25</td>
</tr>
<tr>
<td>MNCT</td>
<td>0.07</td>
<td>0.20</td>
<td>-0.40</td>
<td>0.07</td>
<td>-0.17</td>
<td>0.13</td>
<td>-0.04</td>
</tr>
<tr>
<td>MLC</td>
<td>-0.27</td>
<td>0.07</td>
<td>0.36</td>
<td>-0.01</td>
<td>0.36</td>
<td>0.25</td>
<td>0.30</td>
</tr>
<tr>
<td>%EFC</td>
<td>0.06</td>
<td>0.13</td>
<td>0.11</td>
<td>0.17</td>
<td>-0.27</td>
<td>-0.14</td>
<td>-0.42</td>
</tr>
<tr>
<td>%EFT</td>
<td>0.06</td>
<td>0.13</td>
<td>0.11</td>
<td>0.17</td>
<td>-0.27</td>
<td>-0.14</td>
<td>-0.42</td>
</tr>
<tr>
<td>WCR</td>
<td>0.10</td>
<td>-0.24</td>
<td>-0.42</td>
<td>0.01</td>
<td>-0.05</td>
<td>-0.02</td>
<td>-0.05</td>
</tr>
</tbody>
</table>

Note. CT=Creative Task; N=Number; Csbalance=Challenge-skill balance; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clause; WCR=Weighted Clause Ratio.

* p<0.05. ** p<0.01.

4.6 Discussion on the correlations between the flow and anti-flow experiences and the written performances of the CSH in different task conditions while controlling for language proficiency

This section of the dissertation discusses the correlations between flow and anti-flow measures and the written performances of the CSH, with proficiency set as a controlling variable.

4.6.1 In the TGT

In the TGT, first of all, fluency seemed to be positively correlated with apathy. This finding is very counterintuitive and needs further research. Regarding lexical complexity, the students produced more diverse words when they were less interested in the task but more bored and anxious while doing the task, showing that less flow experienced was likely to lead to greater lexical diversity. A possible explanation for this finding might be that the task content (going to the park and falling into the pond) made the students uninterested and bored. However, they probably compensated for this by trying to write about this boring event in a more interesting way by using a greater variety of words. Besides, more anxious students might be more likely to do so since they might want to
show their language competence and impress the teacher. This might also be true in the case of lexical variety, which was also found to positively correlate with anxiety. In other words, it might be possible that the anxious students took the task more seriously and wanted to impress the teacher in the way of using more varied vocabulary. Another support for this finding is that negative emotion which is low in intensity might broaden the students’ attentional focus and improve their peripheral memory (Gable & Harmon-Jones, 2010), leading to a high cognitive engagement manifested by a high arousal of the students’ linguistic knowledge, which might bring about the use of diverse and various vocabularies. This infers that low level negative emotions which are actually low level anxiety, low level apathy, and low level boredom in my study seemed to profit the students’ cognitive engagement and were beneficial for their task performance. Regarding the syntactic complexity, the students produced longer subordinations in complex sentences and longer simple sentences in their writing when they had better control. This finding is reasonable since flow occurs when students have control over the important aspects of writing (i.e., ownership, genre, length) (Abbot, 2000). In this task, the statements about control in the questionnaire might be understood by the students as the control over the content, the predetermined content in this TGT, which released their cognitive load in generating ideas to write, in which condition, they could focus more on the language form and produce greater syntactic complexity. In terms of accuracy, it was found that more attention contributed to a higher %EFT and less anxiety led to a higher %EFC. This might be explained in the sense that positive emotions like attention improve the students’ ability to recall relevant recourses (Linnenbrink, 2007; Rice et al., 2007; Ryan & Pintrich, 1997), indicating an interaction between emotion and cognition (Swain, 2013). This finding supports flow theory (Csikszentmihalyi, 1975; Egbert, 2003) in language learning in the sense that students are likely to boost their performance when they are in flow.

4.6.2 In the LGT

Regarding fluency, it was found that less boredom seemed to lead to higher fluency. This might be because, on the one hand, the students might be more willing and behaviorally engaged to do the task when they feel interested instead of feeling bored (Egbert, 2003; Maehr, 1984); on the other, less boredom might benefit them by making them stay attentive for a longer time as shown by several studies on flow and student engagement (e.g., Lambert et al., 2017; Phung, 2016; Shernoff et al., 2003).

In terms of syntactic complexity, a lower level of anxiety was likely to profit syntactic complexity reflected by writing longer sentences of simple sentence structures.
and longer subordinations in complex sentences. A possible explanation for this might be that the anxiety that the students felt made them take the task more seriously and worked harder on the language forms. Besides, their intermediate language proficiency maybe also contributed to their greater syntactic complexity since they might have a good mastery of language knowledge. In a word, a serious attitude toward the task and better language proficiency were likely to contribute to syntactic complexity. This finding again showed that lower anxiety was likely to exert a facilitating effect on task performance by making the students more cognitively engaged in doing the task (Dörnyei, 2005; Gable & Harmon-Jones, 2010; Hewitt & Stephenson, 2012).

With regard to accuracy, the students seemed to produce a higher ratio of Error-Free Clauses when they were less anxious. This might be because when the students were less anxious, they could be more attentive and make the best use of their acquired linguistic knowledge to produce greater accuracy. This seems to be easy to understand since less anxious students were more likely to free up their cognitive resources and, instead of being nervous, they devoted more intensive effort to the task (Ainley & Richardson, 2005), and they might use various writing strategies such as using well-mastered language knowledge rather than the difficult and unfamiliar ones, as suggested by Guo (2019). Moreover, this finding is supported by flow theory (Csikszentmihalyi, 1975; Egbert, 2003) in language learning in the sense that students are likely to boost their performance when they are in flow.

4.6.3 In the CT

Regarding fluency, more attention was likely to lead to higher fluency. This finding is in accordance with flow theory (Csikszentmihalyi, 1975; Egbert, 2003). A possible explanation for this might be that the attentive students tended to be more engaged behaviorally and write longer texts since cognitive engagement was likely to affect behavioral engagement (Zhang, 2017).

In terms of lexical complexity, boredom was found to be negatively correlated with lexical diversity and challenge-skill balance was positively correlated with lexical variety. This finding also supports flow theory (Csikszentmihalyi, 1975; Egbert, 2003) and Maehr’s (1984) personal investment theory in the sense that, while doing a task where they felt a higher balance between challenge and skill, the students were more likely to reach a flow state and become more engaged emotionally, cognitively, and behaviorally. When they felt interested and confident that they could complete the task, they would make the best use of their linguistic knowledge (i.e., various and diversified vocabulary). Syntactic complexity seemed to be not correlated with the flow and
anti-flow measures. A possible explanation might be that the CT task provided not only potential for creativity, but also constraints, that is, the students had to use at least six out of the ten given words. Therefore, they might have devoted most of their attention and effort to generating the content of the story as well as integrating these words in the story, rather than risking themselves by using more complex sentences.

Accuracy also seemed to be independent of the flow and anti-flow measures. This might be due to the participants’ intermediate-level English proficiency. For a student of high language proficiency, the selection of lexical and syntactical structures becomes more automatic and these students are expected to be more accurate (Kellogg, 1996; Kormos, 2012); therefore, their accuracy might not be easily affected by their subjective emotions.
Chapter 5: Findings concerning the Written Performances and the Flow and Anti-flow Experiences of the CSC

This part presents descriptive statistics for each of the written performance measures and flow and anti-flow measures for the CSC. Repeated measures ANOVAs were used to show the differences with respect to these variables across the three different tasks. A study like this could provide a full picture of the effects that the task conditions had on learners’ written performances and their flow and anti-flow experiences.

5.1 Results on the relationship between task conditions and the written performances of the CSC

This part of the dissertation presents the findings concerning the effects of task conditions on the written performances of the CSC. The written performances are assessed in terms of fluency, complexity, and accuracy, with proficiency set as a controlling variable. Table 18 presents the descriptive statistics for the means and standard deviations (SD) of the written performance measures in different tasks.

Table 18
Descriptive statistics of the written performance measures of the CSC in different tasks

<table>
<thead>
<tr>
<th>Measures</th>
<th>N</th>
<th>TGT</th>
<th>LGT</th>
<th>CT</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Mean</td>
<td>SD</td>
<td>Mean</td>
</tr>
<tr>
<td>TW</td>
<td>20</td>
<td>166.00</td>
<td>48.45</td>
<td>145.20</td>
</tr>
<tr>
<td>D-value</td>
<td>20</td>
<td>64.50</td>
<td>14.42</td>
<td>60.90</td>
</tr>
<tr>
<td>P_Lex</td>
<td>20</td>
<td>0.88</td>
<td>0.35</td>
<td>0.88</td>
</tr>
<tr>
<td>MLT</td>
<td>20</td>
<td>8.78</td>
<td>1.68</td>
<td>8.51</td>
</tr>
<tr>
<td>MNCT</td>
<td>20</td>
<td>1.29</td>
<td>0.13</td>
<td>2.38</td>
</tr>
<tr>
<td>MLC</td>
<td>20</td>
<td>6.84</td>
<td>1.18</td>
<td>6.75</td>
</tr>
<tr>
<td>%EFC</td>
<td>20</td>
<td>56.59</td>
<td>20.49</td>
<td>68.23</td>
</tr>
<tr>
<td>%EFT</td>
<td>20</td>
<td>53.23</td>
<td>20.10</td>
<td>68.17</td>
</tr>
<tr>
<td>WCR</td>
<td>20</td>
<td>0.83</td>
<td>0.01</td>
<td>0.87</td>
</tr>
</tbody>
</table>

Note. CSC=Chinese Students studying in China; TGT=Teacher-Generated Content Task; LGT=Learner-Generated Content Task; CT=Creative Task; N=Number; SD=Standard Deviation; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.
In terms of fluency, using Pillai’s trace, repeated measures ANOVA revealed that the task conditions were associated with fluency, $V=0.29$, $F(2, 18)=3.71$, $p=0.05$, partial $\eta^2=0.29$. Follow up comparison indicated that the pairwise difference between the TGT and the LGT was significant ($p<0.05$). The students wrote more words in the TGT than in the LGT (Mean=166.00 for the TGT and Mean=145.20 for the LGT respectively).

Regarding lexical complexity indicated by $D$-value and $P_{Lex}$, using Pillai’s trace, repeated measures ANOVA showed no relationship between task conditions and $D$-value, $V=0.08$, $F(2, 18)=0.78$, $p=0.48$, partial $\eta^2=0.08$, and $P_{Lex}$, $V=0.14$, $F(2, 18)=1.47$, $p=0.26$, partial $\eta^2=0.14$. Regarding syntactic complexity indicated by the MLT, MNCT, and MLC, repeated measures ANOVA revealed no relationship between task conditions and MLT, $V=0.14$, $F(2, 18)=1.50$, $p=0.25$, partial $\eta^2=0.14$; (2) MLC $V=0.09$, $F(2, 18)=0.93$, $p=0.41$, partial $\eta^2=0.09$. On the contrary, repeated measures ANOVA revealed that the task conditions were related to MNCT, $V=0.44$, $F(2, 18)=6.98$, $p=0.01$, partial $\eta^2=0.44$. Follow up comparison indicated that the pairwise difference between the TGT and the LGT was strongly significant ($p<0.01$). The students produced lower MNCT in the TG condition than in the LG condition (Mean=1.29 for the TGT and Mean=2.38 for the LGT respectively). Besides, the pairwise difference between the LGT and the CT was strongly significant ($p<0.01$). The students produced higher MNCT in the LG condition than in the CT (Mean=2.38 for the LGT and Mean=1.29 for the CT respectively).

In terms of accuracy, using Pillai’s trace, repeated measures ANOVA revealed no relationship between task conditions and $\%\text{EFT}$, $V=0.24$, $F(2, 18)=2.84$, $p=0.08$, partial $\eta^2=0.24$. On the contrary, it revealed that the task conditions were associated with $\%\text{EFT}$, $V=0.36$, $F(2, 18)=4.94$, $p=0.02$, partial $\eta^2=0.36$. Follow up comparison indicated that the pairwise difference between the TGT and the LGT was significant ($p<0.05$). The students produced a higher $\%\text{EFT}$ in the LG condition than in the TG condition (Mean=68.17 for the LGT and Mean=53.23 for the TGT respectively). Besides, repeated measures ANOVA revealed that the task conditions were related to WCR, $V=0.32$, $F(2, 18)=4.29$, $p=0.03$, partial $\eta^2=0.32$. Follow up comparison indicated that the pairwise difference between the TGT and the LGT was significant ($p<0.05$). The students produced a higher WCR in the LG condition than in the TG condition (Mean=0.87 for the LGT and Mean=0.83 in the TGT respectively). In a word, the students produced greater accuracy indicated by higher ratio of Error-Free T-units and weighted clause ratio in the LGT than they did in the TGT.
In summary, it seems that an increase of learner agency (from the TGT to the LGT) was associated with a significant decrease in fluency, but a significant increase in syntactic complexity and accuracy. Besides, an increase of the potential for creativity the task (from the LGT to the CT) seemed to be related to a significant decrease of syntactic complexity (in the sense of MNCT, but not in the sense of MLT and MLC), but had no influence on fluency and accuracy. Moreover, an increase of both learner agency and the potential for creativity the task (from the TGT to the CT) seemed to have no significant influence on fluency, complexity, and accuracy.

5.2 Discussion on the relationship between task conditions and the written performances of the CSC

This section of the dissertation analyses in detail why the written performances of the CSC differed across different task conditions.

With regard to fluency produced by the CSC, there was a statistically significant difference between the TGT and the LGT, with higher fluency being produced in the TGT. For the CSC, it seems that having more control over what to write and how to write it did not lead to a higher fluency. From the perspective of Kellogg’s (1996) writing model and Robinson’s (2001a, 2003, 2005) Triadic Componential Framework, this might be caused by the different amounts of the time needed for each writing phase and the different degrees of cognitive effort required from the students by the two tasks. In the LGT, the participants had to spend more time and devote more effort to the planning phase of writing because they needed to recall their past experiences and generate the content, while in the TGT, they did not need to spend time on planning the content and the structure of their writing and could have more time for language transcribing. It seems that a trade-off effect appeared in this case due to limited attentional capacity (Skehan, 1998). That is to say, an increase in task complexity seemed to result in greater linguistic complexity and accuracy possibly at the expense of fluency. This explanation is possible because it was found that the participants exhibited higher complexity and accuracy in the LGT than in the TGT and the differences between the two performance measures were statistically significant (see section 5.1.2 and 5.1.3). However, this finding is in contrast with the results of similar studies on effects of the TGT condition and the LGT condition on oral task performance which found that students produced higher fluency in the LG condition (Lambert & Minn, 2007; Lambert et al., 2017; Poupore, 2014; Shao, 2003), showing that task mode might affect fluency on the tasks that operated on the TG condition and the LG condition due
to the different processes of speaking and written outputs and the cognitive demands required during the processes.

Regarding complexity, for the CSC, task conditions seemed to be associated with the complexity of the students’ written texts. Between the TG condition and the LG condition, the students produced higher syntactic complexity in the LGT than in the TGT. This might be because the LG condition gave the participants more freedom of and control over the contents to write, making the task more meaningful to them, which helped enhance their engagement to write, and drive them to make the best use of their existing linguistic knowledge to produce greater complexity. Interpreted from Maehr’s (1984) personal investment theory, this finding supports the theory in the sense that the task that was meaningful for the learner was likely to drive the students to invest more time and effort to do it. This finding is partially consistent with studies on university students’ or adults’ oral performances in which learners produced higher complexity in the LG condition (e.g., Lambert et al., 2017; Lambert & Zhang, 2019; Poupore, 2014), which infers that task mode (oral versus written) might not interfere the relationships between task performance in terms of complexity in language use and task conditions. Moreover, this finding also supports Robison’s (2001b, 2005, 2007) Cognition Hypothesis which claims that more task complexity along the resource-directing dimension leads to greater complexity in language use.

Between the LGT and the CT, a significant difference was observed concerning syntactic complexity. The sentences that the students wrote in the LG condition were more complex with regard to sentence structures. In a word, the potential for creativity in the CT did not seem to necessarily lead to more creative language use or greater complexity in language use, which is against the Maley (2006)’s idea that creative writing seems to foster learners’ engagement with language so as to produce greater complexity. To explain this, first, it should be kept in mind that, from the perspective of Robinson’s (2001a, 2005, 2007) Triadic Componential Framework, the LGT was less cognitively complex than the CT, because the content to be produced was supposed to be familiar to the students, thus reducing their cognitive load for conceptualization of the content, while it was the other way round in the CT in which the students needed to create a story out of several words by reasoning or imagining. To be more specific, in the LGT, the students were expected to write a story about a memorable event they had experienced; thus, it might require less time and effort to get access to these familiar experiences. However, in the CT, the students were expected to create a story with several words that seemed unrelated, so their priority was to produce a story and complete the task, rather than to focus on language form. Besides, this finding can also
be understood from the perspective of Kellogg’s (1996) writing model and Skehan’s (1998) limited-attentional resources model. The two tasks differed in the formulation stage of the writing process, which consists of planning ideas and organizing the structure and selecting vocabulary as well as grammar knowledge. In the CT, generating ideas might take students more time and cognitive effort than that in the LGT, which might be the reason for a lower complexity in the CT. What is more, besides cognitive factors, affective factors (i.e., anxiety, motivation) might contribute to the greater syntactic complexity in the LGT. Specifically, as shown by Table 19 in the following section 5.3, we can see from descriptive statistics of the two tasks concerning flow and anti-flow that, except for the variable of interest, all the other flow variables were higher in the LGT than in the CT. However, all the three anti-flow variables were lower in the LGT than in the CT. It seems that the LGT had the tendency to lead to higher flow and lower anti-flow for this group of students. The positive emotions that the students had in the LGT might motivate them to seek supplemental resources and relevant language knowledge to perform well on the task (Linnenbrink, 2007; Rice et al., 2007; Ryan & Pintrich, 1997) and get more engaged to do the task well (Egbert, 2003; Liu, 2016; Lo & Hyland, 2007; Maehr, 1984).

In terms of accuracy, for the CSC, task conditions seemed to be associated with the accuracy of their written products. Students produced higher accuracy in the LG condition than in the TG condition. It seems that having more control over the content of the writing benefit the students’ written performance in terms of accuracy. This finding supports Maehr’s (1984) personal investment theory. A possible explanation for this might be that the TGT had a predetermined storyline and left little room for the students to adjust the contents to their linguistic resources (Kormos, 2011). Moreover, they probably did not have the chance to purposefully avoid the vocabulary or syntactic constructions that they had not mastered so well. Therefore, they produced lower accuracy levels in the TGT than in the LGT. However, in the LGT, the learners were given opportunities to tailor their language according to their language resources reservoir and avoid those they were not so sure about, which might contribute to greater accuracy. Besides, since the personal story narration might neither be practiced in their daily English output nor at English writing classes, they might have performed the writing task more carefully and use more attentional resources in expressing their ideas into proper English.

In summary, the influence of the TG condition and LG condition was found on all the three measures of linguistic performance of the CSC, with higher fluency in the TG condition while greater complexity and accuracy in the LG condition. This finding is in
accordance with Robinson’s Cognition Hypothesis in the sense that the students produced lower fluency and greater complexity and accuracy when the task complexity increases along the resource-directing dimension. However, the lower complexity that the students produced in the CT as opposed to that in the LG condition contradicts Robinson’s Cognition Hypothesis since the CT was considered more difficult along the resource-directing dimension with more elements and higher reasoning demands, but still failed to lead to higher complexity.

5.3 Results on the relationship between task conditions and the flow and anti-flow experiences of the CSC

Table 19 presents the descriptive statistics of the flow and anti-flow measures that showed the flow experiences of the CSC in different tasks. Using Pillai’s trace, repeated measures ANOVA revealed no relationship between task conditions and (1) interest, $V=0.24$, $F(2, 18)=2.82$, $p=0.09$, partial $\eta^2=0.24$; (2) attention, $V=0.12$, $F(2, 18)=1.26$, $p=0.31$, partial $\eta^2=0.12$; (3) control, $V=0.10$, $F(2, 18)=0.98$, $p=0.39$, partial $\eta^2=0.10$; and (4) challenge and skill balance, $V=0.09$, $F(2, 18)=0.86$, $p=0.44$, partial $\eta^2=0.09$. However, from the original data, it seems that the LGT facilitated the highest level of flow in terms of interest, control, and challenge and skill balance while the TGT facilitated the least in terms of the same measures. Besides, between the TGT and the LGT, the LGT seemed to be more flow-generating. Moreover, between the LGT and CT, the students seemed to be more attentive while doing the CT than the LGT, while they got more interested, felt more control, and sensed a higher balance of challenge and skills when performing the LGT, showing that the CT that provided more potential to elicit creative language use was more likely to generate flow than the LGT that offered students complete control over the content to write.

Concerning anti-flow, using Pillai’s trace, repeated measures ANOVA revealed no relationship of the task conditions and (1) boredom, $V=0.07$, $F(2, 18)=0.66$, $p=0.53$, partial $\eta^2=0.07$; (2) Apathy, $V=0.07$, $F(2, 18)=0.63$, $p=0.55$, partial $\eta^2=0.07$; or (3) anxiety, $V=0.09$, $F(2, 18)=0.92$, $p=0.42$, partial $\eta^2=0.09$. However, from the mean values of the anti-flow measures, it can be found that, for the CSC, the CT seemed to have generated the most anti-flow while the LGT generated the least.
Table 19

Descriptive statistics of the flow and anti-flow measures of the CSC in different tasks

<table>
<thead>
<tr>
<th>Measures</th>
<th>TGT N</th>
<th>Mean</th>
<th>SD</th>
<th>LGT N</th>
<th>Mean</th>
<th>SD</th>
<th>CT N</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interest</td>
<td>20</td>
<td>2.88</td>
<td>0.76</td>
<td>20</td>
<td>3.08</td>
<td>0.73</td>
<td>20</td>
<td>3.01</td>
<td>0.87</td>
</tr>
<tr>
<td>Attention</td>
<td>20</td>
<td>2.09</td>
<td>0.81</td>
<td>20</td>
<td>3.13</td>
<td>0.75</td>
<td>20</td>
<td>3.29</td>
<td>0.68</td>
</tr>
<tr>
<td>Control</td>
<td>20</td>
<td>3.26</td>
<td>0.87</td>
<td>20</td>
<td>3.45</td>
<td>0.76</td>
<td>20</td>
<td>3.36</td>
<td>0.56</td>
</tr>
<tr>
<td>Csbalance</td>
<td>20</td>
<td>3.21</td>
<td>0.81</td>
<td>20</td>
<td>3.36</td>
<td>0.78</td>
<td>20</td>
<td>3.28</td>
<td>0.60</td>
</tr>
<tr>
<td>Boredom</td>
<td>20</td>
<td>2.56</td>
<td>0.96</td>
<td>20</td>
<td>2.56</td>
<td>0.86</td>
<td>20</td>
<td>2.73</td>
<td>0.83</td>
</tr>
<tr>
<td>Apathy</td>
<td>20</td>
<td>2.49</td>
<td>0.89</td>
<td>20</td>
<td>2.59</td>
<td>0.74</td>
<td>20</td>
<td>2.68</td>
<td>0.69</td>
</tr>
<tr>
<td>Anxiety</td>
<td>20</td>
<td>2.58</td>
<td>0.75</td>
<td>20</td>
<td>2.54</td>
<td>0.63</td>
<td>20</td>
<td>2.72</td>
<td>0.63</td>
</tr>
</tbody>
</table>

Note. CSC=Chinese Students studying in China; TGT=Teacher-Generated Content Task; LGT=Learner-Generated Content Task; CT=Creative Task; N=Number; SD=Standard Deviation; Csbalance=Challenge-skill balance.

5.4 Discussion on the relationship between task conditions and the flow and anti-flow experiences of the CSC

For the CSC, task conditions had no significant effect on the participants’ flow and anti-flow experiences. However, from the descriptive statistics, it appeared as if certain tendencies could be identified concerning the students’ flow and anti-flow experiences while performing different tasks. For instance, between the TG condition and the LG condition, the participants were likely to experience lower flow in the TG condition than they did in the LG condition. This seemed to be the same as the case for the CSH. This finding seems to be consistent with the results of earlier studies on the effects of task conditions on oral task performance which found that the participants experienced more positive emotions in the LG condition (Lambert & Minn, 2007; Lambert et al., 2017; Phung, 2016; Poupore, 2014; Qiu & Lo, 2016). This difference, although did not reach statistical significance, seemed to infer that, for the CSC, tasks with more learner agency was more likely to generate flow.

Besides, between the TGT and the CT, the CT was likely to elicit greater flow, which might be because the CT was more interesting and meaningful to them since they were offered the freedom to generate their own stories (Maeher, 1984). However, the CT also triggered more anti-flow. This finding can be explained in the sense that the students might want to fulfill the task requirements and meet the teacher’s expectations, and also did not know how their performance would be assessed, and thus caused them anti-flow.
What is more, between the LGT and the CT, the mean values of the variables indicated that there was a tendency that the CT triggered higher anti-flow experiences and less flow experiences indicated by less interest, less control, less challenge-skill balance. The CT which offered a better chance to be creative seemed to be more likely to make the students bored and anxious. The reason behind this might be that they had rarely practiced similar tasks in their writing class (Ren, 2005; Zhang, 2012; Zhang, 2016). Therefore, compared with the LGT, they seemed to sense less control over the CT and also felt a lack of the skills to meet the task challenges, which might cause them anxiety (Csikszentmihalyi, 1975) and less engagement to perform the task (Fan, 2004).

5.5 Results on the correlations between the flow and anti-flow experiences and the written performances of the CSC in different task conditions while controlling for language proficiency

For the CSC, the correlations between the flow and anti-flow measures and task performance measures were investigated in a way to disregard the differences in language proficiency among the students, just as it was done in the case of the CSH. Therefore, correlations were calculated to investigate the relationships between students’ written performances on the tasks and their flow and anti-flow experiences in each task. This section presents the findings.

5.5.1 In the TGT

Table 20 presents the correlations between the written performance measures and flow and anti-flow measures in the TGT. It seems that the students’ written performances concerning lexical variety (P_Lex), syntactic complexity (MLT, MNCT, and MLC), and accuracy (shown by %EFT, %EFC, and WCA) did not correlate with their flow and anti-flow experiences while doing the task. However, for this group of students, their fluency in the task was positively correlated with challenge-skill balance \((r=0.54, p<0.05)\), indicating that the students wrote more words when they perceived that their skills and task challenges were better matched. Besides, the lexical diversity was found negatively correlated with apathy \((r=-0.47, p<0.05)\) and anxiety \((r=-0.47, p<0.05)\), indicating that the students employed a larger variety of words when they felt less apathy and less anxiety while doing the task.
Table 20

Correlations between the written performance measures and flow and anti-flow measures in the TGT controlling for language proficiency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Interest</th>
<th>Attention</th>
<th>Control</th>
<th>Csbalance</th>
<th>Boredom</th>
<th>Apathy</th>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td>0.30</td>
<td>0.26</td>
<td>.28</td>
<td>0.54*</td>
<td>-0.33</td>
<td>-0.37</td>
<td>-0.44</td>
</tr>
<tr>
<td>D-value</td>
<td>0.02</td>
<td>-0.08</td>
<td>-.09</td>
<td>-0.04</td>
<td>-0.35</td>
<td>-.47*</td>
<td>-.47*</td>
</tr>
<tr>
<td>P_Lex</td>
<td>0.45</td>
<td>0.29</td>
<td>0.19</td>
<td>0.39</td>
<td>-0.01</td>
<td>0.14</td>
<td>0.31</td>
</tr>
<tr>
<td>MLT</td>
<td>0.13</td>
<td>-0.11</td>
<td>0.01</td>
<td>0.04</td>
<td>0.21</td>
<td>0.17</td>
<td>0.15</td>
</tr>
<tr>
<td>MNCT</td>
<td>0.23</td>
<td>-0.10</td>
<td>0.11</td>
<td>0.11</td>
<td>-0.02</td>
<td>0.25</td>
<td>-0.29</td>
</tr>
<tr>
<td>MLC</td>
<td>0.02</td>
<td>-0.09</td>
<td>-0.03</td>
<td>-0.01</td>
<td>0.23</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>%EFC</td>
<td>-0.14</td>
<td>-0.10</td>
<td>-0.06</td>
<td>-0.02</td>
<td>0.02</td>
<td>0.09</td>
<td>-0.11</td>
</tr>
<tr>
<td>%EFT</td>
<td>-0.14</td>
<td>-0.12</td>
<td>-0.18</td>
<td>-0.03</td>
<td>-0.07</td>
<td>0.02</td>
<td>0.04</td>
</tr>
<tr>
<td>WCR</td>
<td>-0.15</td>
<td>-0.12</td>
<td>-0.06</td>
<td>-0.02</td>
<td>0.03</td>
<td>-0.10</td>
<td>-0.15</td>
</tr>
</tbody>
</table>

Note. TGT=Teacher-Generated content Task; N=Number; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.

* p<0.05. ** p<0.01.

5.5.2 In the LGT

Table 21 presents the correlations between the written performance measures and flow and anti-flow measures in the LGT while controlling for students’ language proficiency. It can be found that fluency (TW), MLT and MLC (indicating syntactic complexity), and accuracy (indicated by %EFT, %EFC, and WCR) seemed to be independent of the flow and anti-flow measures. However, lexical diversity indicated by D-value had a mild positive correlation with anxiety (r=0.50, p<0.05), indicating that the more anxious the students felt, they were likely to use more diverse words in their writings. Meanwhile, MNCT was found strongly positively correlated with interest (r=0.59, p<0.01), showing that the students wrote more simple sentences or more subordinations in complex sentences when they felt interested in the task.
Table 21

Correlations between the written performance measures and flow and anti-flow measures in the LGT controlling for language proficiency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Interest</th>
<th>Attention</th>
<th>Control</th>
<th>Csbalance</th>
<th>Boredom</th>
<th>Apathy</th>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td>0.02</td>
<td>0.18</td>
<td>0.21</td>
<td>0.25</td>
<td>0.05</td>
<td>-0.04</td>
<td>0.20</td>
</tr>
<tr>
<td>D-value</td>
<td>0.07</td>
<td>-0.04</td>
<td>0.26</td>
<td>0.14</td>
<td>0.03</td>
<td>-0.03</td>
<td>0.50*</td>
</tr>
<tr>
<td>P_Lex</td>
<td>-0.29</td>
<td>-0.39</td>
<td>-0.40</td>
<td>-0.30</td>
<td>0.25</td>
<td>-0.35</td>
<td>0.36</td>
</tr>
<tr>
<td>MLT</td>
<td>0.22</td>
<td>-0.09</td>
<td>0.14</td>
<td>0.26</td>
<td>0.6</td>
<td>-0.16</td>
<td>0.35</td>
</tr>
<tr>
<td>MNCT</td>
<td>0.59**</td>
<td>0.38</td>
<td>0.46</td>
<td>0.39</td>
<td>-0.38</td>
<td>-0.32</td>
<td>-0.08</td>
</tr>
<tr>
<td>MLC</td>
<td>0.11</td>
<td>-0.15</td>
<td>0.01</td>
<td>0.17</td>
<td>0.16</td>
<td>-0.13</td>
<td>0.31</td>
</tr>
<tr>
<td>%EFC</td>
<td>-0.33</td>
<td>-0.24</td>
<td>-0.17</td>
<td>-0.28</td>
<td>0.25</td>
<td>0.29</td>
<td>0.18</td>
</tr>
<tr>
<td>%EFT</td>
<td>0.17</td>
<td>-0.06</td>
<td>0.01</td>
<td>-0.05</td>
<td>-0.06</td>
<td>-0.02</td>
<td>0.33</td>
</tr>
<tr>
<td>WCR</td>
<td>-0.33</td>
<td>-0.24</td>
<td>-0.20</td>
<td>-0.22</td>
<td>0.29</td>
<td>0.28</td>
<td>0.17</td>
</tr>
</tbody>
</table>

Note. LGT=Learner-Generated content Task; N=Number; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.

* \( p<0.05 \). ** \( p<0.01 \).

5.5.3 In the CT

Table 22 presents the correlations between the written performance measures and flow and anti-flow measures in the CT. It seems that most of the task performance measures had no correlation with the flow and anti-flow measures. Specifically, fluency, lexical variety (P_Lex), and MLT and MLC (syntactic complexity) seemed to stand independent of the flow and anti-flow measures. Nevertheless, lexical complexity indicated by D-value strongly was found positively correlated with interest (\( r=0.58 \), \( p<0.01 \)) and mildly negatively correlated with boredom (\( r=-0.51 \), \( p<0.05 \)), showing that the words that the students used were more diverse when they felt more interested in and less bored with the task. Meanwhile, MNCT that referred to syntactic complexity negatively correlated with control(\( r=-0.50 \), \( p<0.05 \)), indicating that more control over the task was likely to result in lower syntactic complexity. Finally, in this CT, accuracy produced by the CSC did not seem to be correlated with their flow and anti-flow experiences.
Table 22

Correlations between the written performance measures and flow and anti-flow measures in the CT controlling for language proficiency

<table>
<thead>
<tr>
<th>Measures</th>
<th>Interest</th>
<th>Attention</th>
<th>Control</th>
<th>Csbalance</th>
<th>Boredom</th>
<th>Apathy</th>
<th>Anxiety</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td>0.34</td>
<td>-0.15</td>
<td>0.13</td>
<td>0.39</td>
<td>0.15</td>
<td>0.24</td>
<td>0.23</td>
</tr>
<tr>
<td>D-value</td>
<td>0.58**</td>
<td>0.37</td>
<td>0.07</td>
<td>0.37</td>
<td>-0.51*</td>
<td>-0.14</td>
<td>-0.24</td>
</tr>
<tr>
<td>P_Lex</td>
<td>0.38</td>
<td>-0.09</td>
<td>0.05</td>
<td>0.16</td>
<td>-0.41</td>
<td>-0.27</td>
<td>0.16</td>
</tr>
<tr>
<td>MLT</td>
<td>-0.26</td>
<td>-0.29</td>
<td>-0.35</td>
<td>-0.32</td>
<td>0.34</td>
<td>0.21</td>
<td>-0.02</td>
</tr>
<tr>
<td>MNCT</td>
<td>-0.21</td>
<td>-0.20</td>
<td>-0.50*</td>
<td>-0.44</td>
<td>0.28</td>
<td>0.22</td>
<td>0.07</td>
</tr>
<tr>
<td>MLC</td>
<td>-0.16</td>
<td>-0.19</td>
<td>-0.04</td>
<td>-0.01</td>
<td>0.21</td>
<td>0.08</td>
<td>-0.04</td>
</tr>
<tr>
<td>%EFC</td>
<td>-0.38</td>
<td>-0.14</td>
<td>0.08</td>
<td>-0.35</td>
<td>0.29</td>
<td>0.02</td>
<td>0.11</td>
</tr>
<tr>
<td>%EFT</td>
<td>-0.41</td>
<td>-0.14</td>
<td>0.09</td>
<td>-0.51</td>
<td>0.15</td>
<td>-0.17</td>
<td>-0.04</td>
</tr>
<tr>
<td>WCR</td>
<td>-0.33</td>
<td>-0.05</td>
<td>0.12</td>
<td>-0.29</td>
<td>0.26</td>
<td>0.01</td>
<td>0.09</td>
</tr>
</tbody>
</table>

Note. CT=Creative Task; N=Number; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.

5.6 Discussion on the correlations between the flow and anti-flow experiences and the written performances of the CSC in different task conditions while controlling for language proficiency

This part of the dissertation discusses the results that were presented in section 5.5. The correlations between the flow and anti-flow and the written performance measures of the CSC in each task condition are analyzed in detail in the following.

5.6.1 In the TGT

Regarding fluency, it was found that more challenge-skill balance might lead to higher fluency. The reason behind this might be that when the students felt mastery and control over the task and thought that they could meet the challenges that the task required, they were more likely to reach the flow state when doing the task (Csikszentmihalyi, 1975; Egbert, 2003; Lo & Hyland, 2007) and possibly become more willing to do the task (Fan, 2014), as a result of which, they produced better output, here, in terms of fluency.

Besides, less apathy and less anxiety seemed to result in higher lexical diversity. This task asked the students to write the story depicted by the pictures, offering the students no control over the content, which might cause no pressure on generating
content. However, the predetermined content might also deprive their excitement to some degree. Less apathy could be beneficial for their writing since they could make more effort to produce an interesting story by using more diverse words. As for anxiety, less anxious students were more likely to make full use of their knowledge since they could free up their minds and devote more task-intensive effort to the task (Ainley et al., 2005).

However, regarding the syntactic complexity and accuracy, it seems that the students’ flow and anti-flow experiences did not make any difference in these aspects. This might be because they had enough language competence to complete the task, regardless of their subjective feelings. Another explanation for this might be that they avoided risk using difficult language.

5.6.2 In the LGT

For the LGT, fluency seemed not to be affected by the flow and anti-flow measures. This task required the students to write something memorable that they had experienced. It was different from the TGT in the sense that in this task, the students could decide how detailed and elaborated the story would be. However, although the LGT gave them control over the content to write, it might also excuse them from pushing themselves harder to write longer texts, regardless of their flow and anti-flow state at the time of writing.

Concerning lexical complexity, more anxiety was likely to result in a higher lexical diversity. This again demonstrated the facilitating effect of lower-level anxiety (Dörnyei, 2005; Gable & Harmon-Jones, 2010; Hewitt & Stephenson, 2012). Besides, syntactic complexity was found strongly positively correlated with interest. This finding was reasonable since interest that the students had in the task could enhance their motivation and engagement in performing the task (Valiente et al., 2012), therefore leading to better performance that was reflected by greater syntactic complexity in this task.

Finally, accuracy produced by the CSC seemed to be independent from the influence of their flow and anti-flow experiences. This might be caused by their intermediate language proficiency or their avoidance of risk-taking in using more difficult language.

5.6.3 In the CT

As for the CT, fluency produced by Chinese students in China seemed to be independent of the flow and anti-flow measures. This might be caused by the fact that
this type of task was unfamiliar to them, in which case they might just focus on completing it.

Concerning lexical complexity, lexical diversity was found to be strongly correlated with interest positively and mildly correlated with boredom negatively. This finding is not unusual and supports Maehr’s (1984) personal investment theory. The students were more willing to devote more effort and become more engaged in performing the task when they found it interesting and meaningful (Valiente et al., 2012). This task contained constraint, asking the students to create an interesting story with the given words being integrated, seemed to help the students stretch and explore beyond their existing knowledge and get them more engaged cognitively, therefore leading to the emergence of complex language at the lexical level (Tin, 2011). Besides, positive emotion like interest might broaden learners’ thoughts and encourage them to seek solutions to the problem (Bolitho et al., 2003; Fredrickson, 2004; Fredrickson, 1998; MacIntyre & Gregersen, 2012). In other words, when performing the task, the students’ positive emotions such as their interest in the task seemed to enhance their cognitive engagement (Pu, 2017; Ren, 2005; Zhang 2017), thus making a difference in their language complexity at the lexical level. As for syntactic complexity, the students seemed to write complex sentences with more subordinations included when they sensed less control over the task. A possible explanation for this might be that less control over the task content encouraged the students to take the task more seriously, which caused more conscious attention to language form, and thus, brought about more complex language, which might be partially attributed to their good language proficiency. This also infers the facilitating effect of low-level negative emotion on task performance (Gable & Harmon-Jones, 2010).

As for accuracy in this task, just like what happened in the TGT and the LGT, it seemed not to be correlated with the flow and anti-flow measures. This might be due to their good language proficiency or their not taking risks to use more difficult language.

5.7 Summary: comparison between the CSH and the CSC

Having analyzed (1) the effects of task conditions on the written performances of the CSH and the CSC respectively, (2) the effects of task conditions on the flow and anti-flow experiences of the CSH and the CSC respectively, and (3) the correlations between the students’ written performances and their flow and anti-flow experiences in each task condition for the CSH and the CSC respectively, this section of the dissertation makes a brief comparison between the CSH and the CSC with regard to the
same three aspects to see the differences in the trends observable, and to exam if the English learning contexts have a moderating role in any of these three aspects.

5.7.1 Results on the differences between the CSH and the CSC concerning the relationship between task conditions and their written performances

First, between TGT and the LGT, the differences between the two groups of Chinese students in the relationship between task conditions and the written performance measures can be summarized as: (1) a significant difference appeared in the written performances for the CSC between these two tasks. Specifically, the CSC produced higher fluency in the TGT than in the LGT, while no difference occurred with regard to fluency for the CSH; (2) for both CSH and CSC, there was a significant difference in syntactic complexity indicated by MNCT, with a greater MNCT being produced in the LG condition than in the TG condition; and (3) for the CSC, a significant difference in terms of accuracy indicated by %EFT and WCR also appeared between the TGT and LGT, with a greater %EFT and WCR in the LG condition. However, for the CSH, tasks that differed in learner agency did not seem to be associated with their accuracy.

Second, between the LGT and the CT, for both CSH and CSC, a significant difference concerning MNCT occurred, with a greater MNCT being produced in the LGT, showing that the two groups of students both produced greater syntactic complexity in the task characterized by more learner agency as opposed to more potential for creativity.

Third, between the TGT and the CT, for the CSH, there was a significant difference in P_Lex, with a higher P_Lex in the TGT, indicating that they used more varied and difficult vocabulary in the teacher-generated content task condition rather than a task that had more potential to elicit creative language use. However, no significant difference with this respect was found for the CSC between these two tasks.

5.7.2 Discussion on the differences between the CSH and the CSC concerning the relationship between task conditions and their written performances

First of all, between the TGT and the LGT, the results indicated that the learner agency exerted an effect on written performance measures of fluency, complexity (in the sense of MNCT), and accuracy (in the sense of %EFT and WCR) for the CSC, while only on written performance measure of syntactic complexity (in the sense of MNCT) for the CSH. Specifically, the CSC produced lower fluency, but greater syntactic complexity and accuracy in the LGT, indicating a trade-off effect caused by learner
agency provided by the tasks; and the CSH produced greater syntactic complexity in the LGT, indicating a facilitative effect of learner agency on syntactic complexity for the CSH. Firstly, regarding the fact that a difference in fluency between the two task conditions occurred for the CSC while it did not show for the CSH, the reason for this might be that, for the CSC, they were more familiar and experienced with the TGT and might prefer the TGT than the LGT due to a lot of practice of this type of writing in formal English writing instruction since it has been the most used writing form in college entrance exams in most cities in China (Zhang, 2012). It seems that they were good at this kind of writing. For example, it might be the case that they tried to be as detailed as possible in describing the contents presented by the pictures. However, when doing the LGT, they might have had difficulty automatically retrieving language from their linguistic reservoir. This might be explained from the perspective of their English learning context and their English learning experiences. For instance, the CSC have few opportunities for producing output in the English language, such as talking with foreign friends or classmates from other countries, which, nevertheless, are more common for the CSH (see Table 6 for the description of the out-of-school English learning activities of the CSH and the CSC). Besides, they seem to have no output and small amount of input of English outside school (see Table 6 for the description of the out-of-school English learning activities of the CSH and the CSC). What is more, in China, the teaching of English writing is paid little attention to in English classes, and the students are rarely required to practice English writing (Chen, 2018; Liu, 2007; Yuan, 2013). When it comes to narrative writing, usually the content is half teacher-controlled and half learner-controlled since the students are required to include and elaborate on the information that is provided. This type of task is similar to picture narration except for the fact that the prompt is given in the form of text, rather than pictures; thus, it seems similar to the teacher-controlled content condition. In other words, the personal story narration with total control on the learners’ part is still unfamiliar to the CSC. In summary, for the CSC, the lower fluency in the LGT, from a contextual perspective, might be due to a lack of English use in daily life and a lack of writing practice of personal story narration no matter in daily life or in English writing classes, which might require more effort, on the one hand, to generate content and encode it into English, and on the other, to automatically retrieve to the English language reservoir to find the right language forms, both of which added a lot of cognitive burden on the students as well required great attentional resources and effort (Kellogg, 1996; Skehan, 1998). On the contrary, to the best of my knowledge, reading in English and English writing is stressed in Hungary, especially for those students who go to international schools (Bennet, 2019).
Students have to write about different topics with different prompts as homework, including personal experience narrations or pictures descriptions. Therefore, the tasks used in the TGT and the LGT probably did not seem very different to them in terms of the content-generating and language retrieving. They just completed the task as required. Secondly, both groups of Chinese students produced greater syntactic complexity reflected by using more subordinations or sentences of simple syntactic structures in the LG condition than in the TG condition. Interpreted from Maehr’s (1984) personal investment theory, this might be attributed to more control over the content to write in the LGT which might foster their cognitive engagement as well as their motivate to write. Nevertheless, in the TG condition, they had no control over the content because it was already presented, making them signal that all was going well and they did not need to invest much effort in the task (Schwarz & Clore, 1996); therefore, they might just translate the information depicted by the pictures mechanically instead of working extra time and effort on linguistic complexity. It seems true for both groups of Chinese students that the LG condition led to greater complexity than the TG condition, and it is not likely to be mediated by their English learning contexts and their previous English learning or writing experiences. Thirdly, the CSC produced a higher accuracy in TGT than in the LGT, while for the CSH, task conditions of the TGT and the LGT did not have a major effect on their accuracy. In other words, a facilitative learner agency effect on accuracy was present in the CSC but not in the CSH. This might be caused by a difference between the CSH and the CSC in cognitive load with conceptualization and language encoding, which might originate from their English learning experiences in the different English study contexts. Specifically, for the CSC, the personal story narration might neither be practiced frequently in their daily English output nor at their English writing classes, which possibly gave rise to the possibility that they needed more attentional resources and effort in expressing their ideas into proper English than the CSH.

Besides, between the LGT and the CT, for both CSH and CSC, the LGT elicited greater syntactic complexity shown by more subordinations used, indicating a facilitative effect of total control over the tasks as opposed to the half-control and half-constrained task condition on syntactic complexity of the written output for both groups of Chinese students. A possible account for this finding might be that the LGT might reduce learners’ cognitive load for content conceptualization (Skehan, 1998), on the one hand, and make them emotionally more engaged, on the other (Liu, 2006; Maehr, 1984; Ren, 2017; Reschly et al., 2008). This suggests that Chinese students, regardless of their
English learning contexts and their English learning experiences, were likely to produce greater syntactic complexity in the LGT than in the CT.

Finally, between the TGT and the CT, the results revealed that the TG condition had a facilitative effect on the written performance measure of lexical complexity for the CSH, while it had no significant effect on it for the CSC. A possible explanation for the occurrence of the difference in lexical variety between the two groups of students might be due to the different amount of engagement in doing the two tasks. To be more specific, the CSH might be more cognitively engaged in the TGT than in the CT, while the writing engagement of the CSC in the two writing tasks might not be so much different. This explanation can find support in the difference between the two groups of students in their flow experiences in the two tasks. The CSH felt higher challenge-skill balance, more apathy, and more anxiety in the TGT than in the CT, and the difference was statistically significant. A sense of confidence in completing the task as well as the low-level negative emotions seemed to foster their engagement, especially cognitive engagement in doing the task (Gable & Harmon-Jones, 2010), thus leading to better performance, here in terms of lexical variety. However, for the CSC, no significant different in their flow experiences, or in another word, their engagement in doing the two tasks can be detected, inferring that the amount of engagement in doing the two tasks were not so much different, thus leading to no significant difference in task performance.

5.7.3 Results on the differences between the CSH and the CSC concerning the relationship between task conditions and their flow and anti-flow experiences

English learning contexts seemed to play a role in manipulating the effects of task conditions on the flow and anti-flow experiences of Chinese students in their writing process. For the CSH, a significant difference appeared in terms of challenge and skill balance between the TGT and the CT. The CSH felt a higher challenge-skill balance in the CT than in the TGT. Besides, a significant difference concerning apathy and anxiety also occurred between the two tasks, with the students having experienced more apathy and anxiety in the TGT, suggesting that they felt less apathy and anxiety in the CT. However, for the CSC, between the TGT and CT, no significant difference in terms of challenge-skill balance, apathy and anxiety were found, even though there was a tendency showing that the CT was likely to be both more flow and anti-flow generating for the CSC. Moreover, for both groups of Chinese students, no significant difference concerning flow and anti-flow experiences was found between the TGT and LGT, and between the LGT and the CT.
5.7.4 Discussion on the differences between the CSH and the CSC concerning the relationship between task conditions and their flow and anti-flow experiences

First of all, between the TGT and the LGT which differed in learner agency, it seems that learner agency had no significant effect on the students’ flow and anti-flow experiences, no matter which cultural backgrounds the students were from.

Besides, between the TGT and the CT, the CSH sensed significantly higher challenge-skill balance, less apathy, and less anxiety in the CT. However, for the CSC, no significant difference with regard to the flow and anti-flow aspects was found. This finding showed that the two groups of the students’ characteristics, such as their English learning experiences and English study contexts might play a role in their flow and anti-flow experiences when performing the TGT and the LGT. In other words, whether a task that offered them possibility for creativity, as opposed to a task that was completely teacher-controlled, affected their flow and anti-flow experiences was mediated by learners’ characteristics and their English learning contexts.

What is more, if we look at the mean values of all the anti-flow variables, it can be found that, for the CSH, the CT seemed to provoke the least anti-flow while for the CSC, it was just the opposite, with the CT provoking the most anti-flow. Besides, in terms of the students’ flow experiences, for both groups of Chinese students, the TG condition seemed to be the least flow-inducing. Even though these differences regarding the flow and anti-flow measures for the two groups of students respectively were not statistically significant, they still showed a tendency that students with different characteristics and from different cultural backgrounds might have different emotional responses or affective experiences when performing tasks that differed in learner agency and the potential for creativity.

5.7.5 Results on the differences between the CSH and the CSC concerning the correlations between their flow and anti-flow experiences and their written performances in different task conditions

The section compares the CSH and the CSC concerning the correlations between their flow and anti-flow experiences and their written performances in different task conditions (as shown by Table 23, Table 24, and Table 25). What needs mentioning is that the symbols “+”, “++”, and “-” mean that the two referred variables were found mildly positively, strongly positively, and mildly negatively correlated in the task.
Table 23

*Differences between the CSH and the CSC regarding correlations between the flow and anti-flow measures and the written performance measures in the TGT*

<table>
<thead>
<tr>
<th>CSH</th>
<th>CSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW &amp; Apathy -</td>
<td>TW &amp; Csbalance +</td>
</tr>
<tr>
<td>D-value &amp; Interest -</td>
<td>D-Value &amp; Apathy -</td>
</tr>
<tr>
<td>D-value &amp; Boredom +</td>
<td>D-Value &amp; Anxiety -</td>
</tr>
<tr>
<td>D-value &amp; Anxiety +</td>
<td></td>
</tr>
<tr>
<td>P_Lex &amp; Anxiety +</td>
<td></td>
</tr>
<tr>
<td>MLC &amp; Control +</td>
<td></td>
</tr>
<tr>
<td>%EFC &amp; Anxiety -</td>
<td></td>
</tr>
<tr>
<td>%EFT &amp; Attention +</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* CSH=Chinese Students studying in Hungary; CSC=Chinese Students studying in China; TGT=Teacher-Generated Content Task; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.

First, in the TGT (see Table 23), for the CSH, fluency (TW), lexical diversity (D-value), lexical variety (P_Lex), syntactic complexity (MLC), and accuracy (%EFT and %EFC), all seemed correlated with some of the flow and anti-flow measures. It seems that the written performances of the CSH and their flow and anti-flow experiences interacted more actively in this TG condition. However, in this task, for the CSC, only fluency (TW) and lexical diversity (D-value) correlated with some of the flow and anti-flow measures. To be more specific, for the CSH, fluency negatively correlated with apathy. However, for the CSC, it was positively correlated with the challenge-skill balance. In terms of lexical complexity, for the CSH, first of all, lexical diversity (D-value) was found to be weakly correlated with interest negatively and with boredom and anxiety positively. Besides, for the CSH, there was a weak positive correlation between lexical variety (P_Lex) and anxiety. For the CSC, only lexical diversity (D-value) negatively correlated with apathy and anxiety. What is more, for the CSH, syntactic complexity (MLC) correlated with control positively. In terms of accuracy, for the CSH, %EFT was found to correlate with attention positively and %EFC correlated with anxiety negatively. For the CSC, however, no correlation was found between the flow and anti-flow measures and syntactic complexity as well as accuracy.
Table 24

*Differences between the CSH and the CSC regarding correlations between the flow and anti-flow measures and the written performance measures in the LGT*

<table>
<thead>
<tr>
<th>CSH</th>
<th>CSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW &amp; Boredom -</td>
<td>D-Value &amp; Anxiety -</td>
</tr>
<tr>
<td>MLC &amp; Anxiety +</td>
<td>MNCT &amp; Interest ++</td>
</tr>
<tr>
<td>%EFC &amp; Anxiety -</td>
<td></td>
</tr>
</tbody>
</table>

*Note.* CSH=Chinese Students studying in Hungary; CSC=Chinese Students studying in China; LGT=Learner-Generated Content Task; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.

In the LGT (see Table 24), for both groups of Chinese students, their written performances did seem to actively relate to their flow and anti-flow experiences. Concerning fluency, for the CSH, it was negatively correlated with boredom. However, for the CSC, fluency seemed to be independent of the flow and anti-flow measures. Besides, for the CSH, no correlation between the flow and anti-flow measures and lexical complexity was observed. However, for the CSC, lexical diversity (D-value) was negatively correlated with anxiety. In terms of syntactic complexity, for the CSH, it (MLC) slightly positively correlated with anxiety, while for the CSC, syntactic complexity (MNCT) was strongly positively correlated with interest. In terms of accuracy, for the CSH, it (%EFC) was negatively correlated with anxiety, while for the CSC, no correlation between the flow and anti-flow measures and accuracy was found.
Table 25

*Differences between the CSH and the CSC regarding correlations between the flow and anti-flow measures and the written performance measures in the CT*

<table>
<thead>
<tr>
<th>CSH</th>
<th>CSC</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW &amp; Attention +</td>
<td>D-Value &amp; Interest ++</td>
</tr>
<tr>
<td>D-value &amp; Boredom -</td>
<td>D-Value &amp; Boredom -</td>
</tr>
<tr>
<td>P_Lex &amp; Csbalance +</td>
<td>MNCT &amp; Control -</td>
</tr>
</tbody>
</table>

*Note. CSH=Chinese Students studying in Hungary; CSC=Chinese Students studying in China; CT=Creative Task; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio.*

In the CT (see Table 25), for the CSH, fluency, lexical complexity, and accuracy seemed to be correlated with some of the flow and anti-flow measures, while syntactic complexity did not seem to correlate with the flow and anti-flow measures. In specific, fluency was slightly positively correlated with attention, while for the CSC, fluency seemed independent of the flow and anti-flow measures. In terms of lexical complexity, for the CSH, lexical diversity (D-value) was negatively correlated with boredom and lexical variety (P_Lex) positively correlated with challenge-skill balance. However, for the CSC, it was found that only lexical diversity (D-value) was strongly positively correlated with interest and mildly negatively correlated with boredom, with no correlation between flow and anti-flow measures and lexical variety (P_Lex) was found. Besides, for the CSC, syntactic complexity (MNCT) seemed correlated with control negatively. Moreover, for both groups of the Chinese students, no correlation between the flow and anti-flow measures and accuracy was found in this task.

### 5.7.6 Discussion on differences between the CSH and the CSC concerning the correlations between their flow and anti-flow experiences and their written performances in different task conditions

Firstly, in the TG condition, for the CSH, the correlations between their flow and anti-flow experiences and their written performances were more complicated than those for the CSC. It seems that their written performances were more sensitive to their flow and anti-flow experiences. First of all, the way the fluency interacted with the students’
flow and anti-flow experiences between the CSH and the CSC on this task was different, with it being negatively correlated with apathy for the CSH while positively correlated with challenge-skill balance for the CSC, inferring an influence possible caused by their English learning experiences and English study contexts. With regard to complexity, for the CSH, lexical diversity was positively correlated with negative emotions like boredom and anxiety and negatively correlated with interest. However, for the CSC, lexical diversity was negatively correlated with negative emotions like apathy and anxiety. Therefore, it is reasonable to assume that the students’ English learning experiences and study contexts might have a moderating role in these correlations between lexical complexity and their flow and anti-flow experiences when performing the task. Another interesting finding was that %EFC and %EFT that reflected accuracy were negatively correlated with anxiety and positivity correlated with attention respectively for the CSH, while they were shown to be irrelevant to the flow and anti-flow experiences for the CSC, again showing a role of the students’ English learning experiences and study contexts on the relationships between accuracy and the flow and anti-flow measures.

Secondly, in the LGT, the correlations between the written performance measures and flow and anti-flow measures seemed not so complicated for either CSH or CSC. Specifically, for the CSH, fluency, complexity, and accuracy seemed to be correlated with their anti-flow experiences, while, for the CSC, only complexity correlated with their flow and anti-flow experiences. Here, I would like to claim that their study contexts and their English learning experiences might account for these differences in the same task condition.

In the CT, generally speaking, the written performances of both groups of Chinese students and their flow and anti-flow experiences did not seem to have a lot of relationships. For the CSH, fluency and lexical complexity seemed correlated with their flow and anti-flow experiences, while for the CSH, only complexity of their written texts was correlated with their flow and anti-flow experiences. Again, the way that the flow and anti-flow experiences interacted with the written performances by the CSH and the CSC seemed to be different.

In summary, it seems that, in a specific task condition, not all flow elements were associated with the written performances in the same ways to the same degrees. Besides, the reasons behind these associations were not likely the same but different with the students who came from different contextual English learning backgrounds.
Chapter 6: Comparison between the CSH and the CSC concerning their Written Performances and their Flow and Anti-flow Experiences in Different Task Conditions

6.1 Introduction

This section of the dissertation presents the results of RQ 4 that addressed the differences between the two groups of Chinese students’ written performances and their flow and anti-flow experiences in each task with the attempt to examine the role of the students’ English learning experiences and their English learning contexts in these two aspects.

Before tapping to the main issue of the role of study contexts, two things need to be clarified. On the one hand, even though it seems reasonable to combine the two groups’ results (task performances and flow experiences) to investigate the general features of the Chinese intermediate English learners’ task performances and their flow and anti-flow experiences across the different tasks, there is no point in doing so since the two groups of students were very different in their English learning experiences and their prior experiences with the trainings of the three different tasks. However, for each group, their English learning experiences and their prior experiences with the trainings of the three writing tasks are very similar and homogeneous. In addition, it was found that the two groups were different in terms of their written performances and their flow experiences across different tasks, showing the role of English study contexts in their written performances and flow and anti-flow experiences. Therefore, it is reasonable to only compare their task performances and flow and anti-flow experiences across the different task conditions to examine the role of English study contexts in the above two aspects.

On the other hand, this section is not a repetition of section 5.8. To illustrate, section 5.8 was a summary of the differences that task conditions had on two groups of Chinese students’ written performances and their flow and anti-flow experiences across the different tasks, and the correlations between their written performances and their flow and anti-flow experiences in each task. In other words, it addressed the role of the task conditions and focused on the relationships of the ask conditions on each group of students’ task performance and flow and anti-flow experiences. Therefore, the three tasks had to be taken into consideration at the same time since I was comparing every task performance and flow and anti-flow measures across the task conditions as well as examining the general trends in each task condition with regard to the correlations between the flow and anti-flow and the written performance measures and compare the
trends in these correlations in each task between the CSH and the CSC. Nevertheless, this section looked at each task condition separately and exclusively and addressed the role that the students’ English learning experiences and learning contexts played in their written performances and their flow and anti-flow experiences in a certain and specific task condition (i.e., the TG condition, the LG condition, the CT condition). In this sense, given that the task condition was the same for both groups of Chinese students, it is safe to assume that the differences in the mentioned two aspects (written performances and flow and anti-flow experiences) possibly resulted from their English learning experiences and English learning contexts. However, the shared part by the section 5.8 and this section was that the differences between the CSH and the CSC in the written performances and the flow and anti-flow experiences, whether across different task conditions covered in section 5.8 or in a specific task condition that was talked about in this section were explained from the contextual perspective.

To compare the two groups of Chinese students in terms of their written performances and their flow and anti-flow experiences respectively in each task condition, independent samples t-tests were employed.

6.2 Differences between the CSH and the CSC concerning their written performances and their flow and anti-flow experiences in the TGT

This section presents the results concerning the differences between the CSH and the CSC with regard to their written performances and their flow and anti-flow experiences in the TGT and also discusses them.

6.2.1 Results

The TGT was considered as offering the least control for the learners and had no potential to elicit creativity since the contents were predetermined and the storyline was given. Table 26 presents the descriptive statistics of written performance measures and flow and anti-flow measures in the TGT of the CSH and the CSC respectively. With regard to the written performances in the TG condition, the t-test showed no significant difference between the two groups of Chinese students in terms of fluency (TW), complexity (D-value, P_Lex, MLT, MNCT, and MLC) and accuracy (%EFC, %EFT, and WCR) (p>0.05).

As for the flow and anti-flow experienced by the students in this task, independent samples t-test showed a significant difference between the two groups of Chinese students in terms of control (p<0.05) and anxiety (p<0.05). The CSH sensed more control (Mean=3.79 for the CSH and Mean=3.26 for the CSC respectively) and felt less
anxious (Mean=2.08 for the CSH and Mean=2.58 for the CSC respectively) than the CSC when doing this task.

Table 26

Descriptive statistics of students’ written performance measures and flow and anti-flow measures in the TGT

<table>
<thead>
<tr>
<th>Measure</th>
<th>CS H</th>
<th>CS C</th>
<th>CSH Mean</th>
<th>SD</th>
<th>CSC Mean</th>
<th>SD</th>
<th>Difference</th>
<th>T</th>
<th>df</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td>20</td>
<td>20</td>
<td>194.1</td>
<td>57.8</td>
<td>48.4</td>
<td>28.10</td>
<td></td>
<td>1.67</td>
<td>38</td>
<td>0.10</td>
</tr>
<tr>
<td>D value</td>
<td>20</td>
<td>20</td>
<td>64.70</td>
<td>12.1</td>
<td>14.4</td>
<td>0.20</td>
<td></td>
<td>0.48</td>
<td>38</td>
<td>0.96</td>
</tr>
<tr>
<td>P_Lex</td>
<td>20</td>
<td>20</td>
<td>1.08</td>
<td>0.34</td>
<td>0.35</td>
<td>0.20</td>
<td></td>
<td>1.83</td>
<td>38</td>
<td>0.08</td>
</tr>
<tr>
<td>MLT</td>
<td>20</td>
<td>20</td>
<td>8.94</td>
<td>1.51</td>
<td>1.68</td>
<td>0.16</td>
<td></td>
<td>0.32</td>
<td>38</td>
<td>0.75</td>
</tr>
<tr>
<td>MNCT</td>
<td>20</td>
<td>20</td>
<td>1.23</td>
<td>0.13</td>
<td>0.13</td>
<td>-0.56</td>
<td></td>
<td>-1.3</td>
<td>38</td>
<td>0.19</td>
</tr>
<tr>
<td>MLC</td>
<td>20</td>
<td>20</td>
<td>7.31</td>
<td>6.84</td>
<td>1.14</td>
<td>0.46</td>
<td></td>
<td>1.26</td>
<td>38</td>
<td>0.22</td>
</tr>
<tr>
<td>%EFC</td>
<td>20</td>
<td>20</td>
<td>68.25</td>
<td>21.8</td>
<td>20.4</td>
<td>11.66</td>
<td></td>
<td>1.74</td>
<td>38</td>
<td>0.09</td>
</tr>
<tr>
<td>%EFT</td>
<td>20</td>
<td>20</td>
<td>62.48</td>
<td>25.5</td>
<td>20.1</td>
<td>9.25</td>
<td></td>
<td>1.27</td>
<td>38</td>
<td>0.21</td>
</tr>
<tr>
<td>WCR</td>
<td>20</td>
<td>20</td>
<td>0.88</td>
<td>0.09</td>
<td>0.01</td>
<td>0.56</td>
<td></td>
<td>1.89</td>
<td>38</td>
<td>0.07</td>
</tr>
<tr>
<td>Attention</td>
<td>20</td>
<td>20</td>
<td>3.36</td>
<td>0.63</td>
<td>0.81</td>
<td>0.28</td>
<td></td>
<td>1.20</td>
<td>38</td>
<td>0.24</td>
</tr>
<tr>
<td>Control</td>
<td>20</td>
<td>20</td>
<td>3.79</td>
<td>0.69</td>
<td>0.87</td>
<td>0.53</td>
<td></td>
<td>2.14</td>
<td>38</td>
<td>0.04</td>
</tr>
<tr>
<td>Csbalance</td>
<td>20</td>
<td>20</td>
<td>3.58</td>
<td>0.61</td>
<td>0.81</td>
<td>0.36</td>
<td></td>
<td>1.59</td>
<td>38</td>
<td>0.12</td>
</tr>
<tr>
<td>Boredom</td>
<td>20</td>
<td>20</td>
<td>2.78</td>
<td>0.67</td>
<td>0.96</td>
<td>0.21</td>
<td></td>
<td>0.82</td>
<td>38</td>
<td>0.42</td>
</tr>
<tr>
<td>Apathy</td>
<td>20</td>
<td>20</td>
<td>2.28</td>
<td>0.56</td>
<td>0.89</td>
<td>-0.21</td>
<td></td>
<td>-0.9</td>
<td>38</td>
<td>0.37</td>
</tr>
<tr>
<td>Anxiety</td>
<td>20</td>
<td>20</td>
<td>2.08</td>
<td>0.58</td>
<td>0.75</td>
<td>-0.50</td>
<td></td>
<td>-2.3</td>
<td>38</td>
<td>0.02</td>
</tr>
</tbody>
</table>

Note. TGT=Teacher-Generated content task; N=Number; SD=Standard Deviation; CSH=Chinese Students Studying in Hungary; CSC=Chinese Students Studying in China;
6.2.2 Discussion

In my samples, the t-test showed no significant difference between the two groups of Chinese students in terms of fluency, complexity, and accuracy produced in this task. This might be caused by the task characteristic in the sense that the task required the students to narrate a ready-made story about a family going to the park on weekends and the boy falling into the lake when playing with the ducks and then being saved by a young man. Since the story had a tight and clear storyline and story structure, the students might just transcribed the information presented by the pictures. Moreover, the students were not likely to expand the story and add more information to it during writing even though it was made clear in the task prompt that adding more details was allowed. In this case, it seems that the students did not push themselves to use more complex language to write the story longer or more complex. Therefore, we can assume that the English learning experiences and English study contexts did not make a difference to the Chinese students’ written performances on tasks operated on the TG condition.

Besides, the two groups of Chinese students were found to have experienced different levels of flow and anti-flow manifested by control and anxiety sensed when doing the task. The CSH were in more control over the task and less anxious during task performance than the CSC, indicating that the CSH were more engaged when doing this task. This is assumed to be caused by their English learning environment and learning experiences. However, this finding is hard to explain in the present study and a qualitative study might be able to explain this finding.

6.3 Differences between the CSH and the CSC concerning their written performances and their flow and anti-flow experiences in the LGT

This section presents the results concerning the differences between the CSH and the CSC with regard to their written performances and their flow and anti-flow experiences in the LGT and also discusses them.
6.3.1 Results

Table 27 presents the results of the descriptive statistics of task performance measures and flow and anti-flow measures of the two groups of students in the LGT. Independent samples t-test showed that there was a statistically significant difference in fluency indicated by TW (p<0.05) at the group level. The CSH wrote more than the CSC in this task (Mean=198.65 for the CSH and Mean=145.20 for the CSC respectively).

As for the flow and anti-flow experienced by the two groups of students in this task, the t-test showed that there was a significant difference in the flow experience indicated by control (p<0.01) and challenge and skill balance (p<0.05) between the two groups of Chinese students. The CSH felt more control over the task (Mean=4.13 for the CSH and Mean=3.45 for the CSC respectively) and a greater balance between challenge and skill (Mean=3.84 for the CSH and Mean=3.36 for the CSC). In terms of the anti-flow variables, the two groups of students were significantly different in anxiety (p<0.05). The CSH seemed to feel less anxious than the CSC when doing the task (Mean=2.05 for the CSH and Mean=2.54 for the CSC respectively).

Table 27

Descriptive statistics of students’ written performance measures and flow and anti-flow measures in the LGT

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>Mean</th>
<th>t</th>
<th>df</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS</td>
<td>H</td>
<td>CS</td>
<td>H</td>
<td>C</td>
<td>C</td>
<td></td>
</tr>
<tr>
<td>TW</td>
<td>20</td>
<td>198.6</td>
<td>58.8</td>
<td>145.2</td>
<td>40.0</td>
<td>53.45</td>
<td>3.36</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0</td>
<td>0</td>
<td>7</td>
<td>0</td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>D value</td>
<td>20</td>
<td>65.15</td>
<td>19.0</td>
<td>60.90</td>
<td>13.8</td>
<td>4.25</td>
<td>0.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td>6</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>P_Lex</td>
<td>20</td>
<td>0.87</td>
<td>0.43</td>
<td>0.88</td>
<td>0.44</td>
<td>-0.01</td>
<td>-0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>MLT</td>
<td>20</td>
<td>9.29</td>
<td>1.84</td>
<td>8.51</td>
<td>1.56</td>
<td>0.78</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>MNCT</td>
<td>20</td>
<td>2.43</td>
<td>2.38</td>
<td>1.84</td>
<td>1.56</td>
<td>0.78</td>
<td>1.45</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>MLC</td>
<td>20</td>
<td>6.87</td>
<td>6.87</td>
<td>1.28</td>
<td>1.10</td>
<td>0.12</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>%EFC</td>
<td>20</td>
<td>68.17</td>
<td>18.3</td>
<td>68.23</td>
<td>18.1</td>
<td>-0.07</td>
<td>-0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38</td>
</tr>
<tr>
<td>%EFT</td>
<td>20</td>
<td>65.66</td>
<td>19.9</td>
<td>68.17</td>
<td>18.3</td>
<td>2.51</td>
<td>-0.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>38</td>
</tr>
<tr>
<td></td>
<td>N</td>
<td>20</td>
<td>20</td>
<td>0.87</td>
<td>0.87</td>
<td>0.11</td>
<td>0.08</td>
</tr>
<tr>
<td>-------</td>
<td>---</td>
<td>-----</td>
<td>-----</td>
<td>------</td>
<td>------</td>
<td>------</td>
<td>------</td>
</tr>
<tr>
<td>Interest</td>
<td>20</td>
<td>20</td>
<td>3.12</td>
<td>3.08</td>
<td>0.82</td>
<td>0.73</td>
<td>0.04</td>
</tr>
<tr>
<td>Attention</td>
<td>20</td>
<td>20</td>
<td>3.53</td>
<td>3.13</td>
<td>0.74</td>
<td>0.75</td>
<td>0.40</td>
</tr>
<tr>
<td>Control</td>
<td>20</td>
<td>20</td>
<td>4.13</td>
<td>3.45</td>
<td>0.50</td>
<td>0.76</td>
<td>0.68</td>
</tr>
<tr>
<td>Csbalance</td>
<td>20</td>
<td>20</td>
<td>3.84</td>
<td>3.36</td>
<td>0.46</td>
<td>0.78</td>
<td>0.48</td>
</tr>
<tr>
<td>Boredom</td>
<td>20</td>
<td>20</td>
<td>2.45</td>
<td>2.56</td>
<td>0.88</td>
<td>0.86</td>
<td>-0.11</td>
</tr>
<tr>
<td>Apathy</td>
<td>20</td>
<td>20</td>
<td>2.18</td>
<td>2.59</td>
<td>0.68</td>
<td>0.74</td>
<td>-0.41</td>
</tr>
<tr>
<td>Anxiety</td>
<td>20</td>
<td>20</td>
<td>2.05</td>
<td>2.54</td>
<td>0.67</td>
<td>0.63</td>
<td>-0.49</td>
</tr>
</tbody>
</table>

**Note.** TGT=Teacher-Generated content task; N=Number; SD=Standard Deviation; CSH=Chinese Students Studying in Hungary; CSC=Chinese Students Studying in China; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio; Csbalance=Challenge-skill balance.

### 6.3.2 Discussion

For this task, in terms of task performance, the two groups of participants differed significantly in fluency, with the CSH having written more. There might be two explanations for this result. On the one hand, there was a difference in writing engagement that might lead to the difference in fluency between the two groups of Chinese students. To be more specific, the CSH were more engaged when doing the ask as manifested by the greater control over the task, the greater challenge and skill balance, and less anxiety they sensed when doing the task (see section 6.3.1, Table 27). The students were likely to perform better when they were in flow (Egbert, 2003; Maehr, 1984). Therefore, the difference seemed to show that the CSH benefited more from the LG task condition than the CSC did; on the other hand, this outcome might be interpreted from a contextual perspective. For the CSH, they were studying in a more open and international context, and some of the participants of the present study went to British, American, or Bilingual schools, and most of them often engage themselves with diverse kinds of out-of-school English learning activities; therefore, they have more opportunities to be exposed to an English-speaking environment or more opportunities...
for English input and output. As suggested by Myles (2002), higher performance requires deliberately routine linguistic knowledge input which can be automatized in a fast-errorless production of L2 speech through practice in speech production. Foreign language learners, through studying abroad, gain opportunities for such automatization usually in meaningful interactions with native speakers and massive exposure to L2 input (Dekeyser, 2007; Segalowitz et al., 2004). This effect caused by the English learning context on speech might also exist in writing, since, for teenage students, narrative writing can be considered as a written form of the personal experiences-sharing in daily conversations. Besides, this explanation that the study contexts might cause the difference in fluency produced by the two groups of students can also be supported by Maehr’s (1984) personal investment theory. As is suggested by Maehr (1984), socio-cultural context and prior personal experience are important factors that determine the meaningfulness of educational tasks and situations for a learner, and the degree of meaningfulness that a task or a situation has for a learner will, in turn, determine learners’ willingness to invest his or her personal resources (i.e., time, talent and energy) into completing the task. Interpreted in this way, the CSC might not find this genre of English writing very meaningful because the social solidarity goals or extrinsic goal for learning English in the larger educational context for them is to perform well in the College Entrance Exam, and they might not be very engaged in completing this task. In fact, from the mean values for anti-flow variables, they seemed to have experienced greater anti-flow (i.e., boredom, anxiety, and apathy) and felt less control and a lower balance between challenge and skill than the CSH did when doing this LGT (see Table 27), which provided some support for this explanation.

In terms of flow and anti-flow, it was found that the two groups were significantly different in control and challenge-skill balance during task performance, with the CSH feeling more control and a greater balance between challenge and skill than the CSC. Besides, the CSH also felt less anxiety that the CSC. To explain this, first, it is essential to note that, in the first place, control in foreign language learning classroom refers to the control over the learning situation in oral interaction (Egbert, 2003). In the present study, control primarily referred to the control over the contents to write and the overall control over the task-related factors (i.e., the time limit for writing). In the present study, the significant difference in control between the CSH and the CSC might have been caused by their English learning experiences and study contexts. First, the LGT, which had no storyline and offered the participants total freedom to generate the contents and story structure, might be less familiar and practiced by the CSC. Even though, according to my personal experiences, these kinds of writing tasks are frequently practiced in
Chinese language, this cannot guarantee that the students can also do it well in English (Kern, 2000). Writing ability is not a naturally acquired skill; instead, it is usually learned or culturally transmitted as a set of practices in formal instructional settings or other environments (Myles, 2002). However, in high schools in China, English writing is not paid enough attention to and not taught and practiced systematically (see section 2.2 for the problems of writing instruction in China), which might result in the feeling of less control over the task and a lower balance between task challenges and writing skills. When the students did not sense the control over the task and felt that they could complete the task, they would get anxious and might lose their interest to do the task, which was likely to decrease their task engagement. However, for the CSH, for most of them, producing output in English was relatively common or even daily-based. Therefore, they might not get anxious when doing this kind of task. In this sense, it can be assumed that their English teaching and learning environment might play a role in manipulating the students’ engagement in doing the task.

6.4 Differences between the CSH and the CSC concerning their written performances and their flow and anti-flow experiences in the CT

This section presents the results concerning the differences between the CSH and the CSC with regard to their written performances and their flow and anti-flow experiences in the CT and also discusses them.

6.4.1 Results

Table 28 presents the results of the descriptive statistics of task performance measures and flow and anti-flow measures of the two groups of students in the CT. Independent samples t-test showed that there was a significant difference in fluency (indicated by TW) (p<0.01) in the CT between the CSH and the CSC, with the CSH writing more than the CSC (Mean=242.40 for the CSH and Mean=157.25 for the CSC respectively). What is interesting about the CSH is that the standard deviation value (SD=111.78) for their fluency in this task was very high, which means that there was a great individual variation in fluency they produced.

Regarding flow and anti-flow, independent samples t-test showed that there was a significant difference in Chinese students’ flow experience in terms of control (p<0.01) and challenge-skill balance (p<0.01). The CSH felt more control than CSC during the task performance (Mean=4.25 for the CSH and Mean=3.36 for the CSC respectively). Besides, the CSH experienced a better challenge-skill balance than the CSC (Mean=3.95 for the CSH and Mean=3.28 for the CSC respectively). In terms of the anti-flow
variables, there were significant differences between the two groups of Chinese students regarding apathy (p<0.01) and anxiety (p<0.01). The CSH experienced less apathy (Mean=1.90 for the CSH and Mean=2.68 for the CSC respectively) and less anxiety (Mean=1.75 for the CSH and Mean=2.72 for the CSC respectively) than the CSC.

Table 28

Descriptive statistics of students’ written performance measures and flow and anti-flow measures in the CT

<table>
<thead>
<tr>
<th>Measure</th>
<th>N</th>
<th>CS H</th>
<th>CS C</th>
<th>CSH</th>
<th>CSC</th>
<th>Mean Difference</th>
<th>t</th>
<th>df</th>
<th>Sig (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TW</td>
<td>20</td>
<td>20</td>
<td>242.4</td>
<td>157.2</td>
<td>111.7</td>
<td>59.2</td>
<td>85.15</td>
<td>0.30</td>
<td>38</td>
</tr>
<tr>
<td>D value</td>
<td>20</td>
<td>20</td>
<td>59.55</td>
<td>59.50</td>
<td>14.88</td>
<td>14.9</td>
<td>0.05</td>
<td>0.01</td>
<td>38</td>
</tr>
<tr>
<td>P_Lex</td>
<td>20</td>
<td>20</td>
<td>0.77</td>
<td>0.73</td>
<td>0.40</td>
<td>0.30</td>
<td>0.04</td>
<td>0.35</td>
<td>38</td>
</tr>
<tr>
<td>MLT</td>
<td>20</td>
<td>20</td>
<td>8.97</td>
<td>8.23</td>
<td>2.35</td>
<td>1.36</td>
<td>0.74</td>
<td>1.22</td>
<td>38</td>
</tr>
<tr>
<td>MNCT</td>
<td>20</td>
<td>20</td>
<td>1.27</td>
<td>1.29</td>
<td>0.19</td>
<td>0.21</td>
<td>-0.02</td>
<td>-0.2</td>
<td>38</td>
</tr>
<tr>
<td>MLC</td>
<td>20</td>
<td>20</td>
<td>7.09</td>
<td>6.47</td>
<td>1.73</td>
<td>1.18</td>
<td>0.62</td>
<td>1.32</td>
<td>38</td>
</tr>
<tr>
<td>%EFC</td>
<td>20</td>
<td>20</td>
<td>58.93</td>
<td>67.67</td>
<td>1.73</td>
<td>1.18</td>
<td>-8.74</td>
<td>-1.3</td>
<td>38</td>
</tr>
<tr>
<td>%EFT</td>
<td>20</td>
<td>20</td>
<td>58.93</td>
<td>60.40</td>
<td>24.23</td>
<td>17.1</td>
<td>-1.48</td>
<td>-0.2</td>
<td>38</td>
</tr>
<tr>
<td>WCR</td>
<td>20</td>
<td>20</td>
<td>0.87</td>
<td>0.87</td>
<td>0.10</td>
<td>0.07</td>
<td>-0.00</td>
<td>-0.0</td>
<td>38</td>
</tr>
<tr>
<td>Interest</td>
<td>20</td>
<td>20</td>
<td>3.32</td>
<td>3.01</td>
<td>0.83</td>
<td>0.87</td>
<td>0.31</td>
<td>1.16</td>
<td>38</td>
</tr>
<tr>
<td>Attentin</td>
<td>20</td>
<td>20</td>
<td>3.56</td>
<td>3.29</td>
<td>0.69</td>
<td>0.68</td>
<td>0.28</td>
<td>1.28</td>
<td>38</td>
</tr>
<tr>
<td>Control</td>
<td>20</td>
<td>20</td>
<td>4.25</td>
<td>3.36</td>
<td>0.48</td>
<td>0.56</td>
<td>0.89</td>
<td>5.39</td>
<td>38</td>
</tr>
<tr>
<td>Csbalance</td>
<td>20</td>
<td>20</td>
<td>3.95</td>
<td>3.28</td>
<td>0.48</td>
<td>0.60</td>
<td>0.68</td>
<td>3.93</td>
<td>38</td>
</tr>
<tr>
<td>Boredom</td>
<td>20</td>
<td>20</td>
<td>2.44</td>
<td>2.73</td>
<td>0.69</td>
<td>0.83</td>
<td>-0.29</td>
<td>-1.2</td>
<td>38</td>
</tr>
<tr>
<td>Apathy</td>
<td>20</td>
<td>20</td>
<td>1.90</td>
<td>2.68</td>
<td>0.61</td>
<td>0.69</td>
<td>-0.78</td>
<td>-3.7</td>
<td>38</td>
</tr>
</tbody>
</table>
Anxiety 20 20 1.75 2.72 0.62 0.63 -0.97 -4.9 38 0.00

Note. TGT=Teacher-Generated content task; N=Number; SD=Standard Deviation; CSH=Chinese Students Studying in Hungary; CSC=Chinese Students Studying in China; TW=Total Words; MLT=Mean Length of T-unit; MNCT=Mean Number of Clauses per T-unit; MLC=Mean Length of Clause; %EFT=%Error-Free T-units; %EFC=%Error-Free Clauses; WCR=Weighted Clause Ratio; Csbalance=Challenge-skill balance.

6.4.2 Discussion

For the CT, in terms of the students’ written performances, the two groups exhibited a significant difference in fluency, with the CSH writing more than the CSC. A possible explanation for this could be that some students in this group enjoyed doing this task much more than others, and they seemed to become extraordinarily excited and behaviorally more engaged in doing the task. This within-group difference might cause the difference between the groups in fluency. Another explanation might be that the two groups of the students experienced different amounts of flow while doing the task. Specifically, the CSH felt more control and a greater balance between challenge and skill, less apathy, and less anxiety than the CSC during the task performance, showing that the CSH were more engaged in doing this task, and thus profited the fluency produced on this task. As proved by a few studies, flow experiences enhance task engagement and language learning (e.g., Egbert, 2003; Lambert & Minn, 2007; Philp & Duchesne, 2016).

As for the differences in the flow and anti-flow experienced by the two groups of Chinese students, it might be related to the task characteristics of the CT as well as their previous learning experiences with the CT. To be more specific, comparatively speaking, the CSH might have been more familiar with this kind of task than the CSC, and they might have the skills that matched the task challenges since creative writing is part of high school English curricula in America high schools and they have been trained in creative writing (Bennet, 2019). However, for the CSC, the genre of argumentation and half teacher-generated content narration are frequently practiced in high schools in China while other forms of writings are often paid less attention to and rarely practiced (Chen, 2015; Zhang, 2016). This might be the reason why the CSC felt anxious about the task. Moreover, since this kind of task was somewhat new to them, they might have
felt it hard and too challenging, which might have caused them more apathy and less motivation to do the task (Noels, 2001; Williams & Burden, 1997).

Chapter 7: Conclusions and Pedagogical Implications

7.1 Introduction

Since the 1980s, Task-based language learning and teaching (TBLT) has become popular and the framework has been widely applied to the research on L2 learning and instruction. Given the important role of task in the L2 or foreign language instruction, task-based language learning research primarily focuses on task design and task implementation in language learning activities and classes. Besides, recent studies have started to pay attention to the role of learner engagement in task performance and language learning in a long run (e.g., Lambert & Minn, 2007; Lambert et al., 2017). The present study has extended past research to reveal the effects of task design in improving learners’ engagement in writing as well as their written performances.

The tasks used in this study were characterized by different levels of learner agency/control and different potential for creativity, both of which were expected to lead to student engagement to perform the tasks. The three narrative tasks operated on different task conditions, namely the teacher-generated content (TG) condition, also called the teacher-generated task (TGT), the learner-generated content (LG) condition, also called the learner-generated task (LGT), and the creative task (CT). Among the three tasks, the TGT had the least control on the learners’ part while the LGT had the most. Besides, concerning the potential for creativity, the CT had more potential to elicit creative and complex language use than the TGT and LGT. The present study aimed to explore the relationship between task conditions and the Chinese teenage students’ written performances measured by fluency, complexity, and accuracy of the written texts, and their writing engagement measured by their flow and anti-flow experiences during the writing process, as well as correlations between their written performances and their flow and anti-flow experiences in each task condition. Besides, since language learning context has been proved to affect language learning, the present study also attempted to find its role in Chinese students’ written performances and their flow and anti-flow experiences during the writing process.

This chapter of the dissertation provides the conclusion of the study. It includes a summary of the findings that answer the RQs posed in section 1.2, a discussion of the limitations of the study, a presentation of the pedagogical implications of the findings, and the suggestions for future research.
7.2 Summary of findings

The first research question referred to the relationship between task conditions and the written performances of the Chinese students studying in Hungary (the CSH) and Chinese students studying in China (the CSC) with intermediate English proficiency in terms of fluency, complexity, and accuracy. The study revealed within-learner variations in terms of written performances across different tasks. For the CSH, the study showed that: (1) between the TGT and the LGT, the students produced greater syntactic complexity in the LGT by producing longer simple sentences or longer subordinations in complex sentences, showing that more control in the task was likely to be beneficial for the students’ written performance in terms of syntactic complexity; (2) between the TGT and the CT, the words that the students used were more rare and difficult in the TGT, indicating that the task with more potential for creativity seemed to bring about less lexical complexity than the task that was operated on the teacher-generated content condition and had no potential for creativity; and (3) between the LGT and the CT, the sentences produced in the LGT were more complex with more subordinations than those produced in the CT, showing that greater syntactic complexity was more likely to occur in the task with more learner agency than the task that had more potential for creativity but less control on the learners’ part. For the CSC, it was found that: (1) between the TGT and the LGT, the students produced greater syntactic complexity and accuracy in the LGT. However, the students wrote more words in the TGT than in the LGT, showing the trade-off effects on task performance in the two task conditions. These findings showed that the task offering the students more control seemed to result in greater syntactic complexity and lower fluency than the task that provided no control for the students; (2) between the TGT and the CT, no significant difference was found in the written performances; and (3) between the LGT and the CT, the students produced greater syntactic complexity in the LGT, again, showing that, for the CSC, the task that had more potential to elicit creativity might not necessarily result in greater complexity than the task that was totally learner-controlled.

The second research question addressed the relationship between task conditions and the flow and anti-flow experiences of the CSH and the CSC during their writing process. For the CSH, the study found that: (1) between the TGT and the LGT, no significant difference in the flow and anti-flow experiences of the students was observed; (2) between the TGT and the CT, the students sensed a higher challenge and skill balance, more apathy, and more anxiety in the TGT, indicating that the task with more potential for creativity was likely to lead to lower challenge-skill balance, less apathy, and less anxiety than the TGT with no potential for creativity; and (3) between the LGT
and the CT, the students seemed to go through less anxiety in the CT, indicating that, for the CSH, the task with more potential for creativity seemed to result in less anxiety than the task with more learner agency but no potential to elicit creativity. For the CSC, no significant difference regarding their flow and anti-flow experiences across the different tasks was found. However, the mean scores for the flow and anti-flow measures tended to indicate that, among the three tasks, the TGT seemed to be the least flow-provoking since the students seemed to have the lowest interest, least attention, lowest control, and lowest challenge-skill balance in the TGT compared to those in the LGT and the CT. Besides, between the TGT and the LGT, the LGT seemed to be more flow-generating, showing that having more control over the contents to write and the opportunity to generate their own ideas were likely to bring about more flow experience and enhance the students’ writing engagement. However, this difference was not statistically significant, which might be due to the small size of the sample. A similar study with a larger sample size might find some meaningful findings on this aspect.

The third research question referred to the correlations between the written performance measures and flow and anti-flow measures in each task for the CSH and the CSC respectively with the effects of students’ language proficiency being excluded. It was found that, for the CSH, in the TGT, fluency was found to be negatively correlated with apathy. Besides, lexical diversity was found to be negatively correlated with interest and positively correlated with boredom and anxiety, which was unexpected. Besides, lexical variety seemed to be positively correlated with anxiety. Moreover, more control seemed to lead to greater syntactic complexity shown by longer simple sentences or longer subordinations in complex sentences. What is more, accuracy was found to be positively correlated with attention and negatively correlated with anxiety. In the LGT, boredom was found to be negatively correlated with fluency. Besides, in terms of complexity, more anxiety seemed to contribute to higher syntactic complexity, which was reflected in longer sentences of simple structures or longer subordinations in complex sentences produced by the students. Regarding accuracy, anxiety was found to be negatively correlated with the ratio of error-free clauses. In the CT, first of all, more attention seemed to result in higher fluency. Besides, lexical diversity was found to be negatively correlated with boredom, and lexical variety was found to be positively correlated with challenge-skill balance. Moreover, accuracy shown by the ratio of error-free clauses seemed mildly correlated with anxiety negatively. For the CSC, the findings showed that, in the TGT, first of all, a better challenge-skill balance seemed to lead to higher fluency. Besides, lexical diversity was found to be negatively correlated with apathy and anxiety. However, there seemed no relationship between the flow and
anti-flow measures and the syntactic complexity and the accuracy produced in this task. In the LGT, for the CSC, fluency did not seem to be associated with the students’ flow and anti-flow experiences. As for lexical complexity, less anxiety seemed to lead to greater lexical diversity. Besides, the study found that interest was strongly positively correlated with syntactic complexity indicated by longer simple sentences or longer subordinations included in complex sentences produced by the students. In addition, no correlation was detected between accuracy and the flow and anti-flow measures. In the CT, first of all, no correlation was found between fluency and the flow and anti-flow measures. In terms of complexity, more interest and less boredom seemed to benefit lexical diversity. Besides, what is interesting is that syntactic complexity was found to be negatively correlated with control. Finally, no correlation was found between accuracy and the flow and anti-flow experiences that the students were going through during task performance.

What needs to be mentioned is that a brief comparison was made between the CSH and the CSC concerning the differences in: (1) the relationship between task conditions and their written performances, (2) the relationship between the task condition and their flow and anti-flow experiences, and (3) the correlations between their written performances and their flow and anti-flow experiences in each task. It was found that the task conditions seemed to relate to the linguistic features (i.e., fluency, complexity, and accuracy) and the flow and anti-flow experiences of the two groups of participants in different ways. Besides, it was also found that the ways that the written performance measures and the flow and anti-flow measures interacted in each task condition were different for the CSH and the CSC respectively. These might be caused by, not only the task conditions, but also the characteristics of the participants, specifically, their English learning experiences and their English study contexts.

The fourth research question was to investigate the role that the students’ English learning experiences and their English learning contexts might have played in the written performances and flow and anti-flow experiences of the CSH and the CSC in each task condition. In the TGT, no significant difference regarding the written performances at the group level was found, showing that English learning experiences and study contexts might have no moderating role in their written performance in the TGT. In terms of the flow and anti-flow experiences, the CSH sensed more control over this task and felt less anxious than the CSC, showing a difference in the sense of control over this task. In the LGT, the CSH produced higher fluency than the CSC. In terms of the flow and anti-flow measures, the CSH felt more control over the task and a greater balance between challenge and skill than the CSC. Besides, the CSH went through less
anti-flow in terms of apathy and anxiety than the CSC when performing the LGT. These findings showed that the CSH seemed to be more engaged behaviorally and emotionally when doing the LGT than the CSC. In the CT, first of all, the CSH produced higher fluency than the CSC. Besides, an interesting finding was that there was a large individual variation regarding fluency produced by the CSH, indicating that some of them were likely to become much more engaged in the task than the others. In terms of the flow and anti-flow experiences, the CSH felt more control and a greater challenge-skill balance, and less apathy and less anxiety during the task performance than the CSC, indicating that the CSH were more likely to be in flow and more engaged when doing the CT than the CSC. All in all, the findings showed that the CSH seemed to be more engaged than the CSC when performing each task, mostly shown by sensing more control over the task and a better challenge-skill balance, as well as going through less apathy and less anxiety. It can be inferred that, for the participants in the present study, it seems that their English learning experiences and English study contexts had played a role in their written performances and their flow and anti-flow experiences in different task conditions.

7.3 Limitations

The current dissertation attempted to explore the Chinese teenage students’ written performances and their flow and anti-flow experiences during the writing process by utilizing a TBLT approach, with a hope to shed some light on the understanding of the relationship between the task conditions, namely the TG condition, the LG condition, and the CT condition and the students’ written performances and their flow and anti-flow experiences, as well as to investigate the correlations between the written performance measures and flow and anti-flow measures with their language proficiency controlled in each task condition. Besides, this study also aimed to examine the role of study contexts in the two researched aspects. However, some limitations of the current study should be pointed out and taken into consideration when the results are being interpreted.

The first limitation is the sample size. Each group included 20 participants. In the present study, the strict requirements for choosing the participants made it impossible to have a larger sample. Nevertheless, some tendencies, for example, regarding the differences in some aspects of task performances and flow and anti-flow experiences between different tasks, might have reached statistical significance if the study had been conducted with a larger sample size. Therefore, this limitation has to be kept in mind when generalizing based on the outcomes of the study.
Another limitation of the current study is that it is quantitative only in nature. This might weaken its power in explaining some findings, especially the findings concerning the correlations between the flow and anti-flow measures and the written performance measures in each task. For example, the present quantitative study could only speculate about but could not explain the positive correlation between apathy on fluency in the TGT for the CSH.

The last limitation of the present study concerns the contextual factors, referring to the circumstances in which the writing happened. A small portion of the writing samples of the students in Hungary was collected outside the classroom without my presence while others were collected in the classroom by myself. Although clear and strict requirements were given before the administration of the tasks that they were not allowed to use particular tools to assist them and were told that the writings they would produce would only be for research use so as to unload their pressure, it is impossible to know for sure that they followed the requirements and did not use particular tools to assist them. Therefore, it might be possible that the circumstances affected their written performances and their flow and anti-flow state during their task performance.

7.4 Pedagogical implications and directions for future research

Language learning is a broad concept. For instance, it involves the learners, the teachers and the textbooks and the learning materials, and they cannot be considered separately. The research findings of the present study provide more practical implications for writing instruction in high schools in China, which consists of the teachers, the students, and the writing activities and writing materials. In the practices of writing instruction, these three aspects should work in an integrative and cooperative way to benefit the instruction and learning of English writing in high schools in China. Based on the findings of the present study, three implications for language pedagogy are offered as the following:

First of all, since one of the main problems with writing instruction in high schools in China is that the students are not interested in writing and the writing activities and writing contents seem to be considered not well designed and well organized, the present research might provide some understanding for learners’ interest in writing and offer the English writing teachers some guidance for writing task design. Teachers need to recognize that tasks operated on different conditions are likely to elicit written texts of different linguistic features in terms of fluency, complexity, and accuracy. For example, for Chinese students in China, the learner-generated content condition is more likely to elicit written products of greater linguistic complexity and accuracy than the
teacher-generated content condition. Besides, no matter whether the Chinese students study in China or Hungary, it seems that the learner-generated content task, as opposed to the creative task, is likely to stimulate greater syntactic complexity in their written texts. Therefore, when teachers are assessing students’ writing capability in terms of specific language features (i.e., fluency, vocabulary, grammar), it might be important to design and use different tasks. Besides, teachers or language practitioners can design and use specific tasks on the basis of their aims of writing instruction.

In addition, it has been found that tasks with different features (i.e., learner agency, potential for creative language use) seem to relate to the writing experiences of flow and anti-flow which are likely to correlate with written performances; therefore, writing tasks can be designed with certain task features and implemented in task conditions to elicit and enhance flow experiences and inhibit anti-flow experiences, with the expectation to foster written language performances and language development in the long run. For example, for Chinese students in Hungary, from the original data, the creative task seemed to provoke the most flow and the least anti-flow while the teacher-generated content task worked oppositely.

What is more, the students’ characteristics concerning their English learning experiences and their English study contexts should be taken into consideration when teachers are designing tasks, because, for the students studying in different contexts, the flow and anti-flow experiences seem to associate with the writing engagement and the written performances in different ways. For example, for Chinese students in Hungary, in the creative task, a better challenge-skill balance is more beneficial for lexical complexity of the written texts, implying that the topic and constraint provided for creating a story should be carefully chosen so as to provide a task operated on this condition with a favorable challenge and skill balance for learners. To maximally engage the students in order to produce optimal performance and enhance language learning, teachers or language practitioners could gauge the students’ affective appraisals of the writing tasks and make adjustments in tasks and task choices in order to make them possibly more engaging for learners.

In light of the results and limitations of the current study, future research is recommended to be conducted on task-based research on student engagement in writing (measured by flow in the current study) on larger sample size, especially when it comes to a comparative study with the participants from different cultural or contextual backgrounds. Besides, the writing should happen under the same circumstances if the students’ flow and anti-flow experiences in the tasks are under investigation. Furthermore, conducting mixed method research would be more profitable and fruitful.
in order to get a better understanding of the relationship between task conditions and learners’ flow and anti-flow experiences and the correlations between them. In addition, a longitudinal study is desirable to find out the mechanisms that are behind and thus might maintain these associations between the flow and anti-flow measures and written performance measures. Finally, since L2 or foreign language study contexts seem to play a role in the Chinese students’ written performances and their flow and anti-flow experiences, well-designed and systematic research might be conducted on the moderating role of cultural and contextual factors in learners’ written task performances and their flow and anti-flow experiences during the writing process as well as the relationships of the two aspects.
References


Cao, Q. Ch. (2011). Chuzhong yingyu xiezuo nengli de xianzhuang fenxi [An analysis of the current situation of English writing ability in junior middle school and teaching strategies]. Foreign Language Teaching and Research, 1.


Cui, B. L. (2016). Chuzhong yingyu xiezuo xianzhuang fengxi ji jiaoxue celve anjiu [A study on analysis of current situation of Junior High school English writing teaching and teaching strategies]. *Course Education Research.*


He, L. L. (2013). Qiantang xinkebiao gaozhong yingyuzhong yingyu xiezuo jiaoxue cunzai wenti ji duice [A study on the existing problems in English writing teaching at Senior High school according to the new curriculum standards and the suggested solutions]. *Education for Chinese After-school, 12*, 77-77.


Michel, M., Kuiken, F., & Vedder, I. (2012). Task complexity and task condition in interaction. *Special Issue on task complexity and Instructed SLA. IRAL."


Wang, J. L. (2012). Jiyu huati de cihui lianxiang nengli yu yingyu xiezuo chengji de xiangguanxing[A study on correlation between a tapic-based word association ability and English writing proficiency]. *Journal of Tianjin Foreign Studies University, 19*(6).


1. A T-Unit is defined an independent clause and all its dependent clauses.

2. Count run-on sentences and comma splices as two T-units with an error in the first T-unit. If several comma-splices occur in a row, count only the last as error free.

   E.g. My school was in Saudi Arabia, it was the best school there.

   1T-unit 1T-unit
   1 error error-free

3. For sentence fragments, if the verb or copula is missing, count the sentence as 1T-unit with an error. If an NP/Prep.P/subordinate/main clause of a complex sentence is standing alone, attach it to the preceding or following T-unit/sentence as appropriate and count as an error (in the subordinate/main clause itself).

   E.g. As night fall, I didn’t think about escape. Like a puppet, quietly waiting for destiny.
   E.g. But the second day is very slow. cause the second day happened something.
   E.g. Many people don’t remember many things happen in the past, but if it is memorable. It can be remembered forever.

4. When there is a grammatical subject deletion in a coordinate clause, count the entire sentence as 1 T-unit.

   E.g. First, we went to our school and then went out with our friends.

5. Count both “so” and “but” as coordinating conjunctions. Count “so that” as a subordinating conjunction unless “so” is obviously meant.

6. Do not count tag-questions as separate T-units.

7. Subordinate clauses of a complex sentence are not counted as T-unit as in: I believe that A and (that) B = 1 T-unit.

8. Count T-units in parentheses as individual T-units.

9. In a sentence that has a subject with only an auxiliary verb, do not count that subject and verb as a separate clause (or as a separate T-unit).

   E.g. John likes to ski and Mary does too; John likes to ski, doesn’t he?; John is happy and Mary is too.

10. An imperative clause which does not require a subject is considered as a T-Unite.

11. Sentence started with coordinators such as and, but, so, for are counted as 1 T-unit with no error.

   E.g. This is my home. But the painful wound on the …told me that it’s ….

   1T-unit 1 T-unit
12. A complete exclamation with its subject and predicate is counted as a T-unit while the one with grammatical subject and predicate deletion is not counted as a T-unit.
   E.g. What a sunny day it is! (1 T)
   E.g. What a sunny day! (not a T-unit)
13. Direct quotes are treated individually while counting T-units and clauses except that the quotes made of just one word or several non-verb words (verb/verb phrases are considered imperatives) are attached to the preceding or following T-units/sentences as appropriate.
14. The prompts given by the researcher in the writing tasks are not included in T-unit counting, clause counting or word counting.
Appendix B-Clause guidelines (Based on Polio, 1997, and Lee, 2009)

1. Count run-on sentences and comma splices as two clauses with the last error-free.
   E.g. My school was in Saudi Arabia, it was the best school there.
       1 error error-free

2. For sentence fragments, if the verb or copula is missing, count the sentence as 1 clause with an error. If an NP/Prep.P is standing alone, attach it to the preceding or following sentence as appropriate and count as 1 clause with an error.
   E.g. They very angry with me. (1 clause with an error)
   E.g. As night fell, I didn’t think about escape. Like a puppet, quietly waiting for destiny.
       1 error-free clause 1 clause with an error

3. When a dependent clause is connected to the main clause with a subordinator/a relative pronoun/a nominal conjunction/a comparative conjunction, it is counted as a clause.
   E.g. The more I like the class, the more I enjoyed it.
   E.g. If the subordinate stands independently with a period and the main clause stand alone, no matter which comes first, count them 2 clauses with an error in the first.
   E.g. But the second day was very hard. Because something terrible happened so. (2 clauses with an error in the first)

4. Only an imperative does not require a subject to be considered a clause.

5. In a sentence that has a subject with only an auxiliary verb, do not count that subject and verb as a separate clause (or as a separate T-unit).
   E.g. John likes to ski and Mary does too; John likes to ski, doesn’t he? John is happy and Mary is too.

6. When a coordinator is connecting two or more than two verb phrases, they are not counted as clauses.

7. A reduced adverbial clause is not counted as a clause.
   E.g. By using reliable result, when intaking those medical, there are less chances to cause problem. (1 clause)
   E.g. You will look further rather than being shortsighted when doing investment. (1 clause)
Appendix C - Error guidelines (Based on Polio & Shea, 2014, and Arsanyilmaz & Pedersen, 2010)

Errors that are not considered errors in this research

1. Comma used between two or the last two parallel elements in a sentence instead of “and” or “or” is not counted as an error.
   E.g. The survivors and I decided to build woody house, go hunting.
   E.g. We ate fruits, cheese, meat, ice-cream.

2. Spelling errors (including word changes like “their/there”, “holped/hoped”) are not counted.
   E.g. I am a very unusual poeple.

3. Missing commas between clauses or after prepositional phrases are not counted as errors.
   E.g. He still remembered the person’s name in that meeting which was unbelievable.
   E.g. Hotel costs 200 yuan, per day.

4. Failing to use capitalization at the beginning of sentences is not counted as an error.
   E.g. that was really unforgettable.

5. Abbreviation and ellipsis are not counted as errors (e.g., sth., sb., etc.).
Appendix D-Weighted Clause Ratio (WCR) (Based on Wigglesworth & Foster, 2008)

Weighted Clause Ratio (WCR)

According to Wigglesworth and Foster (2008), clauses can be categorized into the following four based on the types of errors they contain:

1. Entirely accurate
2. Level 1: The clause has only minor errors (e.g., in morphosyntax) that do not compromise meaning.
3. Level 2: The clause contains serious errors (e.g., verb tense, word choice, or word order), but the meaning is recoverable, though not always obvious.
4. Level 3: The clause has very serious errors that make the intended meaning far from obvious and only partly recoverable.

E.g. I first feel about something that help other people.
E.g. That’s my first drive the boat.

However, the above categorization is not so clear in categorizing level 1, level 2, and level 3 clauses. Therefore, on the basis of the definitions given to the four types of clauses (Wigglesworth & Foster, 2008) and my data samplings, I worked out the specific error guidelines, namely level 1 error, level 2, and level 3 errors for grouping clauses into level 1, level 2, and level 3 clauses. Besides, what is important to point out is that while grouping clauses, if a clause contains an level 2 error or several level 2 errors that makes or make it a level 2 clause, or if the clause has a mix of level 2 and level 1 errors, the clause is considered as a level 2 clause (the most seriously incorrect one). If a clause contains one or several level 1 errors that makes or make it a level 1 clause, no matter how many errors there are, the clause is considered a level 1 clause.

E.g. Because of my careless, I dropped in the pool. (level 2 clause)
   level 1 error  level 2 error
E.g. I was too scary to swimming. (level 1 clause)
   level 1 error  level 1 error

Finally, a sentence, even though it makes little or no sense, is counted as 1 clause (1 T-unit) if it has a subject and a verb as predicate, and the sentence is considered as having level 3 errors and counted as level 3 clause.

E.g. That’s my first drive the boat.

Level 1 errors that make level 1 clauses
1. Wrong use of word form (e.g., adjective for noun, adjective for adverb)
E.g. My parents and I were very thank for him.
E.g. I was too bored to sleepy.
E.g. You know you were almost died.
E.g. I didn’t want to explain something.

2. Overgeneralizations of -ed/-es/-ing/-s/-er/-est/etc. markers that build
tense/aspect/voice/participle/the plural form of noun/the third person
single/comparative/superlative formation etc..
E.g. I didn’t noticed that I have hurt her very hard.
E.g. There are many childs in the park.
E.g. He flyed to the castle.
E.g. It is the beautifulest city in the world.
E.g. But four minutes later, the plane was shaked strongly.

3. Comma splices in run-on sentences
E.g. I saw a mother duck with her baby duck, the baby is very lovely.

4. Sentence fragment (e.g., an standing-alone noun phrase/preposition phrase/a main or a
subordinate clause of a complex sentence)
E.g. In a happily family, the boy points at the place out of the windows and talk to his
parents. That he wants to play out.
E.g. But the second day was very hard. Because something terrible happened.

5. Wrong, extra, or missing article
E.g. This uncle has reason.

6. Wrong possessive and genitive uses
E.g. My brothers house. My university’s friend.

7. Gerund/infinitive errors
E.g. I did my best to shouting for help.
E.g. And they considered to remove the rock and get the carpet.

8. Preposition errors (missing, extra, wrong)
E.g. After several hours, we arrived the gate of the park.
E.g. I was very thankful of him.
E.g. In all directions, the audience were applauded.

9. Wrong pronoun or possessive determiner (including reflexive and it/there)
E.g. I wanted to find a museum because my friend invited my to join a ceremony.
E.g. My doctor want to stay in the hospital for a week. (It should be “My doctor
wanted me to stay in the hospital for a week”.)
E.g. There were blood red eyes. Just like the murals, it was the eyes of a dragon.
E.g. The island is my.
10. Quantifier–noun agreement (much/many, this/these) or other quantifier problems (a few/few), not including singular plural
   E.g. We had relax for few hours. (It should be “We had relaxation for a few hours”.)
11. Word order problem in complex sentences
   E.g. We were talking about what should we do next week.
   E.g. But although I runned how fast, how hard I was still on the same place.
12. The missing comma in an infinite attributive clause
   E.g. We cannot have a normal Saturday and Sunday like everyone else which is also really bad.
13. Others where the meaning is recoverable but with (some level 1 errors) wrong grammar.
   E.g. It teach a thing that you shouldn’t because out of curiosity to do something you can’t. (The clause “you can’t” should be “you can’t do”. It has a level 1 error. However, the meaning is recoverable, so it is considered as level 1 clause.)

**Level 2 errors that make level 2 clauses**

14. A sentence misses predicate or the linking verb or mixes them together
   E.g. My parents scared when they saw me in the water.
   E.g. The boy in the mirror as pretty as a doll.
   E.g. I didn’t a lucky girl.
15. Failing to use or not using tense/aspect/voice/participle/the plural form of noun/the third person single/comparative/superlative/etc. formation properly
   Note: Tense/reference errors are counted within the context of the preceding discourse.
   E.g. I was so scared that I got hurt and I lose a little blood.
   E.g. And then a trip of magic is began.
   E.g. A dragon was like snakes squirming in sand.
   E.g. There were many thing about the island history.
   E.g. The trees were falled down.
   E.g. Unlucky things were always came quietly.
   E.g. My grade was down and down.
16. Negation problem
   E.g. So, my father couldn’t study no more.
   E.g. But this was over yet. They had a whole new challenge in front of them surviving.
17. Problem with relative clause formation including wrong relative pronoun and reduced relative clause (use of infinitive instead of participle)
   E.g. The day is a good day, which the weather is warm and the sun is shining.
   E.g. Then, I saw a river which water was clean and filled with fishes.

18. Wrong lexical items
   E.g. We heard the voice when the car was behind. (The word “voice” should be used instead of the word “sound”.)
   E.g. Suddenly, I dropped into the water. (The word “fell” should be used instead of the word “dropped”.)
   E.g. I feel so happy when I memory the trip. (The word “recall” should be used of the word “memory”.)

19. Serious word order problems but the sentence makes complete sense if re-ordered
   E.g. The child who helped others though I only met once, his face was deeply reflected in my mind.
   E.g. I still remember the day when take place something.

**Level 3 errors makes level 3 clauses**

20. Incomplete sentences or sentences make little or even no sense with the context being taken into consideration
   E.g. I first feel about something that help other people.
   E.g. There is a thing never forget.
   E.g. I try my best loudly help.
   E.g. My an early age very like animals.
Appendix E-Flow questionnaire (From Czimmermann & Piniel, 2016)

The English version

Name:
Dear students,

I kindly request you to recall the writing task that you just did and how you felt during the task. Please put a cross (X) in the box which shows the best how much the statement is true in your case (the five circles from the right to the left stand for, Totally disagree, Disagree, Neither disagree nor agree, Agree, and Totally agree respectively). There are no right or wrong answers, and I am just interested in your true experiences and feelings during the writing process.

Thank you for your help!

1. This task made me curious. 
2. During this task, I could make my own decisions about what to write.
3. When doing this task, I was completely absorbed in what I was doing.
4. I knew exactly what I had to do.
5. This task was uninteresting for me.
6. During this task, I was in control of what I was doing.
7. I got bored during the task.
8. While doing this task, I was thinking about other things.
9. The task made me upset.
10. I continuously felt that things were going smoothly.
11. I would do this task again even if it were not required.
12. I felt that I had no control over what was happening during the task.
13. The requirement of the task made me feel tense.
14. When doing this task, I noticed that I was distracted.
15. This task was an exciting challenge for me.
16. This task was fun. 

17. During the task, I could make decisions about how to write it.

18. My attention wandered during the task.

19. This task was tiring.

20. I felt worried during the task.

21. I did not have anything to do during the task.

22. I felt that I could meet the requirements.

23. The task was unexciting for me.

24. What I had to do was beyond me.

25. The task was not attractive to me.

26. I felt frustrated during the task.

27. I didn’t feel like doing the task.

28. I was focused during the task.

29. This task was boring.

30. I could not decide how to deal with the task.

31. During the task, I did not feel confident about finishing it.

The Chinese version

姓名：

亲爱的同学，你好！

请回想一下你刚才所做的写作任务，在写作过程中你有什么感受？下面的各种说法能在何种程度上反映你的这种感受？每句话后面有个衡量线，线上有 5 个圆圈，请根据你自己的情况，在每句话后面衡量线上相应的圈里画“x”（从左到右的五个圈分别代表：完全反对，反对，既不反对也不同意，同意，完全同意）。你的回答没有正确和错误之分，我只希望了解到你写作过程中的真实体验和感受。单项选择。

谢谢你的配合！
1. 这项写作任务激发了我的好奇心。
2. 写作过程中，我可以自己决定写什么内容。
3. 写作时，我完全沉浸其中。
4. 我很清楚地知道我该做什么。
5. 这项写作任务很有意思。
6. 写作过程中，我能控制我正在做的事情。
7. 写作过程中，我感到无聊。
8. 写作过程中，我在想别的事情。
9. 这项写作任务让我感到心烦。
10. 我一直觉得任务进行得很顺利。
11. 即使不被要求做这项写作任务，我也会去做。
12. 写作过程中，我感觉自己无法控制正在进行的事情。
13. 这项写作任务的要求让我感到紧张。
14. 在写作过程中，我注意到了外界干扰。
15. 这项写作任务是一项激动人心的挑战，有能力完成这项任务，我很开心。
16. 这项写作任务很好玩，新鲜。
17. 写作过程中，要怎么写，我可以自己做决定。
18. 写作过程中，我注意力不集中，跑神了。
19. 这项写作任务使我感觉到累。
20. 写作过程中，我感到紧张。
21. 写作过程中，我无事可做。
22. 我觉得我可以达到任务的要求。
23. 这项写作任务很无趣。
24. 我要做的事情超出了我的能力范围。
25. 这项写作任务无法吸引我。
26. 在写作过程中，我感到沮丧。  ○○○○○○○
27. 我不想做这项写作任务。  ○○○○○○○
28. 写作过程中，我注意力集中。  ○○○○○○○
29. 这项写作任务太无聊了。  ○○○○○○○
30. 我不能决定应该如何处理这项写作任务。 ○○○○○○○
31. 写作过程中，我觉得没有信心完成这项任务。 ○○○○○○○
Appendix F-Three written narrative tasks

The English version

The Teacher-Generated Content Task (The TGT)

Observe the following six pictures carefully and write a story about what happened. You have 30 minutes. You can write as much as you like, but write minimum 100 words.

A memorable event

This is something that I will never forget....
The Learner-Generated Content Task (The LGT)

Please, write a story about a memorable event that happened to you. The beginning sentence is given. You have 30 minutes. You can write as much as you like, but minimum 100 words.

A memorable event

This is something that I will never forget.…

The Creative Task (The CT)

Create a story, using at least 6 of the following 10 words. You have 30 minutes. You can write as much as you like, but minimum 100 words.

单词: desert (沙漠), island (岛), museum (博物馆), plane (飞机), blood (血), ceremony (典礼, 仪式), storm (暴风雨), dragon (龙), bomb (炸弹), flying carpet (飞行地毯)

A memorable event

This is something that I will never forget…. 
要求：请仔细观察下面6幅图，根据图的内容写一个故事，题目为“一件难忘的事”。所写故事要包括图片所有信息，但也可添加你想要的细节。写作时间为30分钟。你可以尽量多写，但不少于100字。

A memorable event

This is something that I will never forget…. 
The Learner-Generated Content Task (The LGT)
要求：请写一则故事，关于发生在你身上的“一件难忘的事”。写作时间为 30 分钟。
你可以尽量多写，但不少于 100 字。故事开头第一句话已给出，不计入总词数。

A memorable event
This is something that I will never forget.…

The Creative Task (The CT)
要求：下面给出了 10 个单词，请至少使用其中的任意 6 个单词来创作一个有趣的故事，题目为“一件难忘的事”。写作时间为 30 分钟。你可以尽量多写，但不少于 100 字。
单词：desert （沙漠）, island （岛）, museum （博物馆）, plane （飞机）, blood （血）, ceremony （典礼, 仪式）, storm （暴风雨）, dragon （龙）, bomb （炸弹）, flying carpet （飞行地毯）

A memorable event
This is something that I will never forget.…
Appendix G-Bio-data questionnaire

The English version

1. Your age is ______.
2. Your gender is ______.
3. Which school do you go to? ______ In which grade? ______
4. In which language are your English classes taught at your current school? ______
   A. Completely English
   B. Mostly Chinese
   C. Interchange between English and Chinese
   D. Mostly Hungarian
   E. Interchange between English and Hungarian
   F. Others _____________________
5. When did you start learning English? ______
   A. Kindergarten
   B. Primary school from 1th grade
   C. Primary school from 2nd grade
   D. Primary school from 3rd grade
   E. Junior high school
   F. Others ______
6. How many years have you studied in Hungary? (only for those who are studying in Hungary) ______ years
7. In what ways do you have contact with English outside school? (multiple choices) __
   A. Talking with foreign friends (or relatives)
   B. Talking with classmates from other countries
   C. Reading English novels or magazines
   D. Reading English websites
   E. Learning English from English learning Apps
   F. Watching English films, plays, or TV series
   G. Playing English video games
   H. Listening to English songs
   I. Taking English lessons outside school (i.e., having an English tutor, taking English
lessons in an educational institute, etc.)

J. Others ______________

8. Do you speak other languages besides Chinese? Yes or no? if yes, what are they? _

9. Do you like writing? ______
   A. Yes B. No

10. How do you evaluate your writing proficiency? __________
   A. Very good B. Good C. Average D. Bad E. Very bad

Thank you very much for your help!

The Chinese version

请根据个人情况，回答下列问题或选择恰当的答案。

1. 你的年龄是 ______ 岁。

2. 你的性别是 _____。
   A. 男 B. 女

3. 你在哪所学校上学？ __________ 年级？_____

4. 你目前所在学校英语课的授课语言是？
   A. 全部英语
   B. 大部分是汉语
   C. 英语和汉语交替使用
   D. 大部分是匈牙利语
   E. 英语和匈牙利语交替使用
   F. 其他，请注明 ______

5. 你从什么时候开始学习英语？
   A. 幼儿园
   B. 小学一年级
   C. 小学二年级
   D. 小学三年级
   E. 初中 (七年级)
F. 其他，请注明 __________

6. 你在匈牙利学习多久了？（仅在匈牙利的学生做答）_________年

7. 学校外，你与英语接触的方式是？（可以多选） ______________
   A. 和外国朋友或亲人交流
   B. 和来自其他国家的同学交流
   C. 读英语小说、杂志
   D. 浏览英语网站
   E. 英语学习应用程序
   F. 看欧美电影、戏剧或电视剧
   G. 玩以英语为解说的电子游戏
   H. 听英文歌曲
   I. 上课外英语课（比如，上英语家教课、去辅导机构或学校上英语课）
   J. 其他，请注明 __________

8. 除汉语外，你说其他语言吗？________
   A. 是，说其他语言，请注明 __________ B. 不说

9. 你喜欢写作吗？________
   A. 喜欢 B. 不喜欢

10. 你认为你的写作水平是——？
    A. 很高 B. 高 C. 中等 D. 低 E. 很低

非常感谢！