

An exploration of the relationships among task repetition,
formulaic language, and perceived fluency in Chinese L1
EAP students' oral presentations

by

Yana Lysiak

A thesis submitted to the Faculty of Graduate and Postdoctoral
Affairs in partial fulfillment of the requirements for the degree of

Master of Arts

in

Applied Linguistics and Discourse Studies

Carleton University
Ottawa, Ontario

© 2019, Yana Lysiak

Abstract

This study explored the impact of the task repetition (TR) – “repetitions of the same or slightly altered tasks” (Bygate & Samuda, 2005, p. 43) – on the development of formulaic language (FL) – prefabricated sequences that are “stored and retrieved whole from memory” (Wray, 2002, p. 9) – and perceived L2 speech fluency gains. While there is a consensus that FL facilitates fluent speech (Wood, 2015) and that TR can lead to L2 fluency development (Bygate, 1996), it is not clear whether TR alone can yield gains in the FL use and fluency. This is especially the case with investigations in the EAP context that tend to focus on the development of writing and grammar abilities, not on speaking skills (Barnard & Scampton, 2008).

To investigate the effectiveness of TR in the development of FL use and perceived fluency gain, two versions of oral presentations delivered by 10 Chinese L1 EAP students were studied for instances of FL use and rated for fluency. Students repeated the task twice, four weeks apart. While the FL analysis was done using the Wray and Namba’s (2003) checklist, perceived fluency judgments (e.g., speech rate, comprehensibility, pauses and hesitations, language proficiency) were provided by three independent raters who assessed a representative portion of the presentations. The results indicate that TR led to gains in fluency and promoted an increased use of and variety in FL.

Keywords: task repetition, formulaic language, perceived fluency, EAP, oral presentations

Acknowledgments

I am proud that I have challenged myself both professionally and personally when I applied for the MA in ALDS program at Carleton University and chose to pursue the research option. This has turned out to be one of the most rewarding experiences of my life.

I feel honored to have had Dr. Eva Kartchava and Dr. David Wood as my thesis supervisors. Their contribution to this project has been immeasurable. I'd like to thank Dr. Eva Kartchava who has been inspiring me ever since I met her, who motivated me to write this project and shared her research data with me. I am thankful for all valuable suggestions and feedback she gave me, for everything she has taught me, and for her time, which she always gave generously to this work. I would like to thank Dr. David Wood who provided continuous support to this project, shared his expertise and inspired me with his research. I appreciate his valuable critique, guidance and help with choosing appropriate procedures to analyze the data. I'd like to thank all professors who have helped me learn. Thank you to Dr. Michael Rodgers, Dr. Geoff Pinchbeck and to Dr. Jaffer Scheyholislami for helping me shape my knowledge and ideas about teaching. A huge thanks go to Dr. Don Myles for his cheerful encouragement and valuable suggestions regarding this project. I am also grateful for Joan Grant who was always there for me and always answered all my questions.

I would like to acknowledge my ALDS peers for their wonderful collaboration, great support and for their kind assistance with the judgment of my data. A huge thank you go to Alex, Gillian, Betul, Alexis, Marissa, Melissa, Laura, Yan, and Christine. I also have to extend my huge thanks to my friends both in Canada and Ukraine, who kept supporting and cheering me on during this whole time.

Finally, I owe the greatest debt of gratitude to my husband and my family to whom I dedicate this thesis. They have been extremely supportive and encouraging throughout this whole time and I think that without their unconditional love and belief in me I would not have completed this project. Thank you for praying for me and for encouraging me to keep moving forward.

Table of Contents

Abstract	ii
Acknowledgments	iii
Table of contents	v
List of Tables	ix
List of Figures	x
List of Appendices	xi
Chapter 1: Introduction	1
1.1. Research gap and guiding questions.....	1
Chapter 2: Literature Review	6
2.1. Overview.....	6
2.2. Formulaic language.....	6
2.3. Fluency and formulaic language.....	11
2.3.1. Speech fluency and its potential indicators.....	11
2.3.2. The role of formulaic language in fluency.....	17
2.4. Focused instruction on formulaic language.....	19
2.5. The role of task repetition in L2 performance.....	27
2.6. Conclusion.....	33
Chapter 3: Methodology	35
3.1 Overview.....	35
3.2. EAP context.....	35
3.3. Participants.....	38

3.4. Data collection.....	40
3.5. Data analysis.....	42
3.5.1. Formulaic language analysis.....	43
3.5.1.1. Formulaic language – Assessment of Inter-rater reliability.....	47
3.5.1.2. FS classification.....	50
3.5.2. Fluency judgment.....	51
3.5.2.1. Fluency – Assessment of Inter-rater reliability.....	54
3.5.3. Combining formulaic language analysis and fluency judgment.....	56
Chapter 4: Results.....	58
4.1. Overview.....	58
4.2. Research Question 1.....	59
4.2.1. Formulaic language analysis.....	59
4.2.2. Analysis of FS functions.....	65
4.2.3. Summary of the results for question 1.....	70
4.3. Research Question 2.....	70
4.3.1. Findings from the fluency judgment.....	71
4.3.2. Summary of the results for question 2.....	79
4.4. Research Question 3.....	80
4.4.1. Comparison of the FS and fluency ratings.....	80
4.4.2. Summary of the results for question 3.....	99
Chapter 5: Discussion and conclusion.....	100
5.1. Overview.....	100
5.2. Research question 1.....	100

5.2.1. The common FS used by Chinese L1 EAP students.....	101
5.2.2. The increased total and variety of formulaic language in the presentations.....	102
5.2.3. FS functions in the presentations.....	106
5.2.4. Summary of the discussion of research question 1.....	110
5.3. Discussion of research question 2.....	111
5.3.1. Speech rate.....	111
5.3.2. Comprehensibility.....	115
5.3.3. Pausing phenomenon in the participants' presentations.....	116
5.3.4. Language proficiency.....	119
5.3.5. Summary of the discussion of research question 2.....	121
5.4. Discussion of research question 3.....	122
5.4.1. The role of task repetition in formulaic language and fluency development.....	122
5.4.1.1. Task repetition and formulaic language.....	125
5.4.1.2. Task repetition and perceived fluency.....	126
5.4.2. The role of formulaic language in speech fluency.....	127
5.4.3. Summary of the discussion of research question 3.....	128
5.5. Pedagogical implications.....	129
5.6. Limitations and future suggestions.....	131
References.....	135
Appendices.....	151
Appendix A. Secondary data clearance.....	151

Appendix B. Evaluation rubric from the task repetition in class.....	153
Appendix C. Wray and Namba's (2003) checklist.....	154
Appendix D. Confidentiality agreement.....	156
Appendix E. The full list of FS per group.....	157
Appendix F. The results of the formulaic language analysis.....	161
Appendix G. FS used by students who showed some progress.....	162

List of Tables

Table 1. <i>Formulaic language analysis example</i>	46
Table 2. <i>Fluency scale and illustrative example</i>	53
Table 3. <i>FS group results</i>	63
Table 4. <i>The percentage of the increased total and newly occurring FS types</i>	64
Table 5. <i>FS functions</i>	66
Table 6. <i>Repeated FS and their functions in participants' presentations</i>	69
Table 7. <i>Fluency ratings</i>	72
Table 8. <i>FL analysis results for the 6 improved participants</i>	81
Table 9. <i>FL analysis results for the 6 improved participants</i>	82

List of Figures

<i>Figure 1.</i> The percentage of the functional taxonomy of FS.....	68
<i>Figure 2.</i> The percentage of the fluency scores per group.....	73

List of Appendices

Appendix A. Secondary data clearance.....	152
Appendix B. Evaluation rubric from the task repetition in class.....	153
Appendix C. Wray and Namba's (2003) checklist.....	154
Appendix D. Confidentiality agreement.....	156
Appendix E. The full list of FS per group.....	157
Appendix F. The results of the formulaic language analysis.....	161
Appendix G. FS used by students who showed some progress.....	162

Chapter 1: Introduction

1.1. Research gap and guiding questions

Over the past decades, interest in the phenomenon of formulaic language, referred to as formulaic sequences (FS) which appear to be prefabricated and recalled from memory as a single item (Wray, 2002), has grown. Sequences are considered formulaic when they contain at least two words (Coulmas, 1979) and share “a single meaning or function” in discourse (Wood, 2015, p. 3). FS may be very diverse as the following examples indicate: *by and large, quality time, long story short, get along, on the other hand, how is it going*. With the development of different corpus analysis tools researchers tried to identify an approximate proportion of formulaic language in discourse or a given register and discovered that it is quite common in written language and especially frequent in spoken language (Altenberg, 1998; Erman & Warren, 2000). In fact, according to Pawley and Syder (1983), memorized sequences form a big part of everyday conversations. Moreover, the research which explored the use of lexical bundles (FS which are identified by the means of corpus and not necessarily retrieved as wholes) in academic registers revealed that FS serve important functions and appear to be building blocks of discourse (Biber, Conrad, & Cortes, 2004).

Interestingly, the research which looked into how beneficial mastery of using formulaic language might be for second language (L2) learners revealed that those who apply FS in their speech tend to sound more fluent and nativelike (Boers, Eyckmans, Kappel, Stengers, & Demecheller, 2006; McGuire & Larson-Hall, 2017; Pawley & Syder, 1983; Thomson, 2017; Wood, 2006, 2009, 2010). Fluency is usually defined as a “nativelike” ability in the L2 and a function of “temporal variables of speech” such as: (1) rate of speech, (2) length of runs and (3) pauses/hesitations (Wood, 2015, p. 86). On the other hand, Segalowitz (2010) chose a cognitive

science-based approach to define fluency and views it as a multidimensional construct beyond the audible aspects of language. Thus, he defines “perceived fluency” as a listener’s impression about a speaker’s cognitive fluency based on the speaker’s utterance fluency (p. 165).

Interestingly, the studies which investigate the complexity of interrelationship between formulaic language and fluency are rather scarce and there is call for more research which would further explore the link between these two phenomena. For instance, some of the possible reasons why using formulaic language might facilitate L2 fluency is because of its holistic nature and the fact that it seems to be prefabricated (Wray, 2002). Consequently, relying on FS might ease mental processing and lead to fewer hesitations and long pauses in speech (Skehan, 1998). Furthermore, using ready-made stretches of language might provide an opportunity for sounding more proficient and error-free in L2 because these sequences tend to be above an actual proficiency level of a speaker (Boers et. al, 2006). Therefore, formulaic language seems to be a very important part of spoken language production and may be key to achieving L2 speech fluency (Wood, 2015, p. 163).

Considering how abundant FS are in the language and the important role they play in spoken language production, researchers looked into different ways on how to incorporate formulaic language into second language teaching and foster its acquisition (Nattinger & DeCarrico, 1992; Lewis, 1997; Willis, 1990). Lewis’ (1997) Lexical Approach might be one of the most comprehensive efforts to date to teach formulaic language in the L2 classroom. The approach is based on the idea that FL instruction should be in the form of awareness-raising activities and emphasize noticing. Pedagogical intervention studies by Boers et al. (2006) and Stengers, Boers, Housen and Eyckmans (2010) attempted to implement Lewis’ approach (1997) but did not indicate significant improvement in the FS use. It may be argued that even though

awareness raising benefits learners and could be useful for the instruction of FS in the L2 classroom, noticing activities do not necessarily guarantee active use of FS and more repetition should be encouraged (Boers et al., 2006). One way to practice FS in L2 classroom could be through task repetition, which is reported to be an effective strategy within a task-based language (TBLT) approach assisting L2 learners to improve their oral performance (Bygate, 1996). As highlighted by previous research, task repetition has revealed its impact on L2 complexity, accuracy, and fluency (Ahmadian & Tavakoli, 2011; Bygate, 1996, 2001; Gass, Mackey, Fernandez, & Alvarres-Torez, 1991; Lynch & Maclean, 2001). Despite positive effects of task repetition on spoken language production, this researcher has not been able to find any study which would investigate the role of task repetition in the acquisition and use of formulaic language and fluency development.

Focused instruction, which dominates research on formulaic language teaching and acquisition, seems to benefit L2 learners (Boers et. al, 2006; Stengers et al., 2010; Peters, 2012; Webb, Newton, & Chung, 2013; Wood, 2009; Wray & Fitzpatrick, 2008) but in some L2 classrooms a teacher may not always have enough class time to address formulaic language explicitly or to provide sufficient oral practice of FS through different activities. This is especially the case with investigations in the EAP context where instructors tend to focus on the development of writing and grammar skills and spend limited time on speaking skills, rarely addressing formulaic language in both written and spoken language (Barnard & Scampton, 2008; Evans & Green, 2007). Therefore, the present study, which was carried out in an EAP classroom at a Canadian university, aimed to explore an effect of task repetition on formulaic language use and fluency development without any direct/explicit pedagogical intervention. In other words,

the purpose of this thesis was to determine whether repeating a task may be enough to promote FS use as well as fluency gains.

Thus, the research questions of this study are as follows:

1. What are the common formulaic sequences (FS) used by Chinese L1 EAP students and what functions did they serve in the students' presentations?
2. What are the fluency profiles of Chinese L1 EAP students in their presentations?
3. Does task repetition bring on FS and fluency gains (independently and/or together) in the presentations of Chinese L1 EAP students?

To explore the effect of task repetition on the development of formulaic language use and perceived fluency gain, two versions of oral presentations delivered by 10 Chinese L1 EAP students were analyzed for the FS use and rated for fluency. The formulaic language analysis was done using the Wray and Namba's (2003) checklist and facilitated by the native speaker judgment of two independent raters. The identified FS were organized according to the functions they served in the discourse using the functional taxonomy of the lexical bundles conducted by Biber et al. (2004). Perceived fluency scores were assigned by three independent judges. Overall, task repetition showed a positive effect on both formulaic language use and perceived fluency of the participants.

The remaining chapters are organized in the following way. Chapter 2 starts with an overview of the theoretical background of formulaic language and L2 fluency followed by the pedagogical discussion of how formulaic language may be taught and its acquisition fostered. Then, positive effects of task repetition on L2 oral production found in the previous research are discussed. Chapter 3 examines the EAP context where the study was carried out, profiles the participants whose oral presentations made up the data of the present study, and describes the

task they completed. Next, it details the data analyses undertaken. Chapter 4 reports on the results of the analysis of formulaic language use and perceived fluency of the on both iterations of the task. Finally, Chapter 5 discusses the findings of this investigation in light of the previous research on formulaic language, L2 fluency, and task repetition. The chapter ends with a discussion of pedagogical implications, limitations, and suggestions for future research.

Chapter 2: Literature review

2.1. Overview

This chapter provides an overview of the theoretical and empirical literature which has explored formulaic language in terms of its definition, mental processing and acquisition, as well as its diverse functions in discourse. Following this, research on oral fluency and the role that formulaic language might play in its development are reviewed. Next, research on formulaic language teaching and practice is detailed. Lastly, the role of task repetition in second language (L2) production is discussed. The chapter ends with the research questions which guided the present study.

2.2. Formulaic language

Since formulaic language is often referred to as formulaic sequences (FS) – sequences which appear to be prefabricated and recalled from memory as wholes (Wray, 2002) – these terms will be used interchangeably in this paper. Formulaic language has garnered interest among researchers from various fields over the past few decades, including literary studies, educational psychology, philosophy, neurology, sociology, and language teaching methodologies (Wood, 2015, pp. 4-8). Since the 1970s, research on phraseology (the study of formulaic language) has been rapidly developing with the extensive work of lexicographers and the development of technology and corpus analysis tools, and remains a popular inquiry (Pawley, 2007). There is no definite figure as to what portion of the English language can be classified as formulaic since the density of formulaic sequences may vary in different discourse genres and situational contexts. However, according to some corpora estimates, formulaic language constitutes between 50% and 80% of words in the English language (Altenberg, 1998; Erman &

Warren, 2000). Considering how abundant formulaic language is in the English language, it can be argued that it is a “fundamental aspect of language and communication” (Wood, 2015, p. 3). Wray and Perkins (2000, p. 3) compiled a list of all existing terms which have been used to define formulaic language and came up with at least 40 different definitions: chunks, clichés, collocations, fixed expressions, ready-made expressions, formulas, frozen phrases, etc. By exploring the definitions of formulaic language, they illustrate the main categories of formulaic language (Wray & Perkins, 2000). These include: collocations, idioms, lexical phrases, lexical bundles, metaphors, proverbs, phrasal verbs, n-grams, concgrams, and compounds (Wood, 2015, pp. 37-51). As can be seen from the list of the formulaic language categories, the range is quite impressive. As was mentioned earlier, the most established definition of formulaic language/FS in the literature, and the definition adopted in this study, is that of Wray (2002) because it seems to clearly acknowledge the complex nature of this phenomenon:

A sequence, continuous or discontinuous, of words or other meaning elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar. (p. 9).

Common examples of formulaic language in English include the following sequences: *break a leg, head over heels, come up with, good job, have a nice day* etc. These examples suggest that while FS can be rather diverse, certain common features unite them under the category of formulaic language. Firstly, they have at least two or more words which carry one meaning and serve a particular function in discourse (Coulmas, 1979), and secondly, they appear to be prefabricated or stored and retrieved from memory as a whole unit (Wray, 2002). This second characteristic has been particularly challenging for researchers to establish and there is

still more empirical evidence required in this area (Wood, 2015). Conversely, not all FS seem to be holistically processed. For instance, research on a broad category of formulaic language referred to as lexical bundles, which are identified by means of corpus analysis software programs, indicates that their processing seems to be more controlled and conscious (Wood, 2015, p. 162). Common lexical bundles include: *what does that mean, at the same time, as can be seen, for example, in other words*, etc. Moreover, while certain FS may be processed holistically by some individuals, this does not necessarily mean that they will be processed in the same manner by others. For example, when native speakers decide whether a sequence in a text is formulaic or not, it is often based on their intuition, and there is usually some variation in their answers (Eyckmans, Boers, & Stengers, 2007). Schmitt and Underwood (2004) in their psycholinguistic experiments also revealed that a similar variation occurs between native speakers and L2 learners when it comes to agreeing on what is formulaic and what is not.

Research arguing that formulaic language might be holistically processed often considers those sequences which are easily recognized as formulaic and usually have unitary meaning (Gibbs & Gonzales, 1985). For example, idioms (e.g., *apple of my eye, safe and sound, through thick and thin*) as opposed to novel phrases with nonfigurative meaning seem to be stored and retrieved from memory as a whole because they are processed more quickly by native speakers (Underwood, Schmitt, & Galpin, 2004). However, when it comes to non-native speakers, they tend to spend more time on processing idioms in texts due to their figurative meaning and less time on words with literal meaning. This was evidenced in a study by Siyanova-Chanturia, Conklin, and Schmitt (2011) which tracked participants' eye movement while they read texts. Nonetheless, both native and non-native speakers, especially high-proficiency language learners, seem to process high frequency FS in input faster and therefore, possibly holistically (Bod, 2000;

Siyanova-Chanturia et al., 2011). Thus, it can be argued that frequency and automatization could play a role in holistic processing (Logan, 1988). The idea that FS could be automatized through repeated exposure in the input is particularly attractive for educational linguists. It implies that certain FS may eventually become holistic for L2 learners and be retrieved from memory as wholes automatically through exposure and practice (Wood, 2006; Wood & Namba, 2013). By processing FS holistically, L2 learners can produce stretches of discourse more easily without constructing utterances word-by-word (Skehan, 1998). This saves them time and eases mental processing. Additionally, an investigation of L2 development and formulaic language in a keynote study by Schmidt (1983) illustrated that a native Japanese-speaking participant, Wes, relied on prefabricated sequences in his English speech and kept increasing their number and variety through exposure to them in the input as well as practice. Schmidt (1983) suggested that Wes used chunks as a communicative strategy in order to avoid building full sentences himself and to avoid making more mistakes in communication. Even though Wes did not attempt to break down and analyze FS, it seems that L2 learners should be encouraged to do so as it may facilitate their acquisition of the target language. Boers et al. (2006) argue that even though most of FS might be processed more or less holistically, there is a big portion of FS which can also be consciously analyzed or even composed.

Another important characteristic of FS is that they serve certain functions in discourse. It is possible that by recognizing and remembering a function or meaning of FS as a whole L2 learners can eventually process that FS holistically (Wood, 2015, p. 58). For example, Nattinger and DeCarrico (1992, p. 60) offered a long and detailed function-based classification of formulaic language with a large number of sub-categories:

- social interactions (e.g., *I didn't catch your name*);

- necessary topics (e.g., *my name is*);
- discourse devices (e.g., *my point is that*).

On the other hand, Wray (2000, p. 97) proposed a different broad classification and identified four major functions of formulaic language which aid speech production:

- enabling short-cuts (e.g., *My capacity as ___*);
- time-byers (e.g., *If you want my opinion*);
- manipulation of information (e.g., *Shall I compare thee to a summer's day?*);
- signaling the organization of spoken discourse (e.g., *I wonder if you'd mind*).

These function-based classifications of formulaic language illustrate that FS is not simply a linguistic unit but rather a “tool that can be put to many uses” (Wray & Perkins, 2000, p. 9), and therefore, sometimes, it is difficult to separate the form and function of FS (Wray, 2002). Interestingly, Moon (1998) argued that FS are usually genre specific, and ideally, research on FS functions needs to be conducted within one register rather than by encompassing all categories of FS and creating abstract models. For instance, Biber et al. (2004) explored the frequency of functions of lexical bundles specifically used in university teaching and textbooks. They distinguished the following discourse functions for lexical bundles:

- stance expressions (e.g., *I don't know if*);
- discourse organizers (e.g., *you know I mean*);
- referential expressions (e.g., *those of you who*);
- special conversational functions (e.g., *thank you very much*).

The present study, which was conducted in a university setting, adopted the functional taxonomy of FS proposed by Biber et al. (2004) because it was relevant for the context and language being analyzed; the taxonomy is also well-established in the field and widely used by

second language acquisition (SLA) researchers. Wood (2006) also attempted to categorize FS by functions specific to the genre of spontaneous spoken narrative retellings of 11 English as a second language (ESL) learners at a Canadian university. Wood explored the role of FS functions in the development of oral L2 fluency. His findings illustrated that this group of learners used FS to serve five functional categories which facilitated more fluent speech over time. The functions included: use of multiple FS to extend the length of runs of speech (i.e., the stretches of uninterrupted utterance between pauses); repetition of FS in a run of speech; reliance on one FS or filler repeatedly; use of self-talk and fillers; use of FS as rhetorical devices (p. 24).

Overall, formulaic language is a complex phenomenon which constitutes a large portion of the English language and has attracted interest of researchers from various fields. FS serve important functions in discourse and appear to be fundamental to spoken language (Pawley & Syder, 1983; Wood, 2006, 2009). Moreover, previous research has emphasized that relying on FS which are retrieved holistically may help L2 learners to “buy” processing time while speaking (Skehan, 1998, p. 40). This leads to fewer pauses and more natural speech (Granger, 1998; Pawley & Syder, 1983; Wray, 2002). The next section discusses the possible connection between formulaic language and L2 fluency found in the literature.

2.3. Fluency and formulaic language

2.3.1 Speech fluency and its potential indicators

For the purpose of this study, it is important to discuss such a complex phenomenon as oral fluency and its potential markers in order to examine what role formulaic language might play in its development. Studies on L2 fluency have increased rapidly ever since the 1990s. Fluency has become a focus of the language testing literature which addressed the issues of how

to operationalize measures of fluency for assessment purposes (Bachman, 1990; Fulcher, 1996; Lennon, 1990). Even though the definition of fluency has been a matter of debate for several decades, in the broad sense, the term is usually used to describe various qualities of speech (Chambers, 1997). According to Fillmore (1979), fluency is a multi-faceted phenomenon not only in L2 but also in the first language (L1) where it includes such features as pausing, complexity, coherence, and creativity. When L2 speakers reach a higher level of proficiency, their performance is usually compared to that of native speakers' (Chambers, 1997). Pawley and Syder (1983) provided one of the first definitions of fluency which is still commonly used in the field. They understood fluency as "the native speaker's ability to produce fluent stretches of discourse" (p. 191). Interestingly, Segalowitz (2010) compared different meanings of the word "fluency" in different languages and came to the conclusion that its dominant metaphor is a "language in motion". Therefore, fluency is often referred to as "flow", "smoothness" or "fluidity" of speech (p. 4). In fact, according to Kormos' (2006) examination of different measures of fluency, fluidity is considered a predominant feature of fluency.

Despite several decades of work, researchers have not discovered the most objective measures of fluency (Koponen & Riggenbach, 2000). However, analysis of key temporal variables of speech (i.e., speech rate, mean length of runs, and pausing), which are usually linked to fluency, has become a focus of many studies (Grosjean, 1980; Dechert & Raupach, 1987; Raupach, 1980; Sajavaara, 1987). Overall, previous research on temporal measures illustrated that speech rate and mean length of runs are the best predictors of fluency (Kormos & Denes, 2004; Riggenbach, 1991; Towell, Hawkins, & Bazergui, 1996). Speech rate is "the number of syllables articulated per minute" and the mean length of runs is "the average number of syllables produced in utterances between pauses of 0.25 seconds and above" (Kormos & Denes, 2004, p.

148). Riggnebach's (1991) study with Chinese students revealed that the participants who spoke faster were rated as more fluent by the judges. Kormos and Denes (2004) explored perceptions of L2 fluency with 16 Hungarian students learning English. The raters who judged the participants' speech samples were both native and non-native speaking English teachers. This study is particularly interesting because teachers are the most frequent judges of L2 fluency. The results indicated that for all the judges the speech rate, the mean length of utterance, phonation time ratio (i.e., "the percentage of time spent speaking as a percentage proportion of the time taken to produce the speech sample"), and the number of stressed words produced per minute were the best predictors of fluency scores (p. 148). Freed (1995) observed fluency development among American students learning French and concluded that the most noticeable change in the participants' fluency during their stay in France was increased speech rate. Additionally, Freed (1995) identified a tendency toward longer runs in the speech of the participants over time. In another longitudinal study, Lennon (1990) found that the mean length of runs between pauses in the speech of his participants also increased over a period of 23 weeks.

Pausing is a complex phenomenon and is usually used interchangeably with the term "hesitation" (e.g., *um, you know*) (Wood, 2015, p. 87). Usually, research on pauses and hesitations measures their duration, frequency, and syntactic location (Dechert, 1980; Freed, 1995; Riggnebach, 1991). The studies which aimed to compare the use of pauses between native speakers of the target language and L2 speakers indicated that unlike L1 speakers who pause at the end of clauses, L2 learners tend to pause in the middle of the clause interrupting the flow (Deschamps, 1980; Freed, 1995; Lennon, 1984).

In general, objective measures of the rate of speech tend to correlate with human judges' perceptions of fluency (Freed, 1995; Riggnebach, 1991; Towel et al., 1996). On the other hand,

the objective judgment of pausing phenomenon might not always coincide with human judgments. Extensive use of filled pauses/fillers (e.g., *ums*, *uhs*) should not always be interpreted as a sign of dysfluency, but rather as an important communicative strategy (Fox Tree, 2001; Clark & Fox Tree, 2002). For example, native speakers may use fillers in their speech to draw attention of the interlocutor to something which is upcoming in speech (e.g., rephrasing, errors, an intention to speak) (Clark & Fox Tree, 2002). Therefore, sometimes, judges do not consider the use of silent pauses or filled pauses distracting while analyzing L2 fluency (Derwing, Rossiter, Munro & Thomson, 2004; Kormos & Denes, 2004). Another important feature which may impact judges' perception of someone's fluency is accentedness (i.e., a combination of different speech sounds different from a local variety) and comprehensibility of L2 speakers (Munro & Derwing, 2001). Moderate speech rate is believed to be the most appropriate for optimal comprehensibility (Munro & Derwing, 2001).

It can be argued that speech production is tied to speech perception. Even though the studies which focused on perceptions of fluency have been rather scarce and those which did had few subjects, holistic judgment can provide more insights into someone's fluency than the use of analytic measurement of temporal variables on its own (Freed, 1995; Lennon, 1990; Riggenbach, 1991; Towell et al., 1996). For example, Lennon (1990) argued that fluency is a performance phenomenon and therefore, it is "an impression on the listener's part that the psycholinguistic processes of speech planning and speech production are functioning easily and efficiently" (p. 391). Lennon (1990) examined different quantifiable features of oral performance in order to determine which of them might be the best indicators of fluency. His findings indicated that four German participants learning English in the UK for six months managed to increase their speech rate over time and decrease the number of pauses. He also reported that the improvement in

speech rate was a reflection of pause time differences rather than articulation speed. Ultimately, Lennon (1990) suggested that speech rate or pausing individually were not critical indicators of fluency. Instead, it seemed that the operation of higher level mental processes (i.e., planning and production processes) was a more significant indicator of fluency. The idea that fluency might be a listener's impression of a speaker's speech is supported by Segalowitz (2010) who proposed his definition of fluency from a cognitive perspective. He described fluency as a multidimensional phenomenon which can be divided into three senses: cognitive fluency, utterance fluency, and perceived fluency (pp. 48–49). According to Segalowitz (2010), cognitive fluency is the ability to efficiently mobilize mechanisms responsible for planning an utterance, for producing it and for integrating interactive cognitive processes. It is important to note that this type of fluency is the one which a speaker possesses. Utterance fluency refers to the temporal characteristics of speech (e.g., speech rate; stressed words per minute) which serve as properties of the utterance. Finally, perceived fluency has to do with “the inferences listeners make about a speaker's cognitive fluency based on their own perception of utterance fluency” (p. 48). Namely, the judgment about a speaker's fluency is based on impressions which are drawn from listening to speech samples. Segalowitz's (2010) definition of perceived fluency was used in the present study because fluency judgments were made by human raters.

Interestingly, research on temporal variables of speech has attracted psycholinguists because of the potential for explicating models of speech processing (de Bot, 1992; Levelt, 1989). A most cited model often applied in L2 studies is Levelt's (1989) speech production model. Levelt (1989) observed that speakers go through three different stages in order to deliver their message: conceptualization, formulation, and articulation. Conceptualization is the stage in which the speaker conceives a message; formulation is the next stage where the speaker selects

appropriate lexical and syntactic elements to produce the message; finally, articulation stage is when a “phonetic plan” is converted into overt speech (de Bot, 1992, p.1). Levelt’s model may be applicable to L2 fluency research in terms of discovering how formulation, which became proceduralized (i.e., unconscious and automatic), might be represented by temporal measures. For example, Towell et al. (1996) used Levelt’s model in their study on fluency development. They focused on the temporal aspects of L2 fluency exploring them longitudinally and examining whether the oral variables can become proceduralized. They measured fluency of 12 French participants studying English whose task was to tell a story. Towell et al. (1996) reported that there were significant increases in the articulation rate as well as improvements in the mean syllables per run. These results could be explained by activating Levelt’s (1989) speech production model. It should be noted that besides performing objective fluency measurements, Towell et al. (1996) also conducted a detailed qualitative analysis by choosing two subjects’ oral samples and examining which aspects of their speech might have become proceduralized. They revealed that such aspects of speech as formulaic language became proceduralized. One of the speakers extensively used FS, and as a result, had longer runs and increased speech rate. In contrast, another speaker who did not attempt to use formulaic language seemed to be “laboriously putting together sentences very much under the control of interpretive mechanisms” (p. 109). Finally, Towell et al. (1996) concluded that increases in fluency observed by the listeners reflect not only improvements in the use of temporal variables of speech, but also complex underlying processes such as proceduralization.

In summary, fluency seems to be a very complex phenomenon, and it is challenging for researchers to find standardized methods for analyzing and measuring it in speech (Segalowitz, 2010). Therefore, the idea of perceived fluency seems attractive because it suggests that all

temporal variables may interfere with one another and should be taken into consideration when assessing fluency (Towell et al., 1996). Moreover, while some variables (e.g., filled pauses) may seem dysfluent by objective measures, they can also be interpreted, as a communicative strategy or a discourse function by human raters (Clark & Fox Tree, 2002; Fox Tree, 2001). Finally, it appears that formulaic language may play an important role in the increased speech rate and mean length of runs (Towell et al., 1996).

2.3.2. The role of formulaic language in fluency

As previously mentioned, FS serve important functions in discourse and may be stored and retrieved from memory as wholes, saving processing time during spontaneous conversation and facilitating fluent speech (Coulmas, 1979; Skehan, 1998). However, there is a limited body of research exploring the link between formulaic language and L2 fluency.

Coulmas (1979) was one of the first researchers who tried to explain the role of formulaic language in oral production. He argued that instead of always producing sentences according to grammatical and lexical systems, native speakers use standardized and high-frequency sequences because they come to mind in particular social situations and contexts. Since FS are shared by a large group of speakers, they are highly predictable in certain situations and therefore, anticipated by the interlocutor (Coulmas, 1979). Later, in a landmark paper, Pawley and Syder (1983) drew on Coulmas' (1979) ideas to point to a possible connection between formulaic language and fluency. They presented an analysis of oral production of two native speakers and revealed that memorized sequences (i.e., stored in memory and recalled precisely without any deviation from form) are a big part of everyday conversations. According to Pawley and Syder (1983), an L2 speaker using FS, says what an interlocutor expects to hear. As a result, their speech sounds more nativelike and natural. Boers et al. (2006) in their study explained how using

FS can benefit L2 learners. The authors explained that when L2 speakers master the “...”idiomatic” dimension of natural language”, they sound more nativelike (p. 246). Also, using FS helps L2 learners to sound more error-free. Finally, FS tend to surpass an L2 speaker’s proficiency level, which creates the impression that the L2 speaker’s command of the target language is more advanced than it really is. Furthermore, Boers et al. (2006) argue that L2 learners treat FS as “zones of safety” and rely on them in spontaneous conversations (p. 247). In other words, L2 learners can use FS as a means to bridge gaps in their linguistic knowledge, which, in turn, leads to fewer hesitations in speech. Since FS may be holistic, using them eliminates internal pauses and, therefore, leads to longer runs of speech, and the latter appears to be the key indicator of fluency (Wood, 2015, p. 88). Extensive use of FS is usually found in the speech of fluent L2 speakers. This is evidenced by Ejzenberg’s (2000) study comparing the use of formulaic language by fluent and non-fluent L2 speakers. She found that fluent speakers relied on FS in speech and used them more than non-fluent speakers who had less FS and struggled with applying them appropriately.

A key study which investigated the role of formulaic language in L2 fluency was that by Wood (2010), where drawing on his earlier research (Wood, 2006), he pointed out possible benefits of FS functions for developing fluency in spontaneous communication. Eleven L2 participants in an intensive English program abroad met once a month over a period of six months and watched short silent movies. Their spontaneous retellings of what they saw in the short movie were collected afterwards. In the end, Wood’s participants showed significant gains in fluency over the six months. Repetition of FS helped the participants to lengthen speech runs and avoid pauses. More importantly, the use of FS as self-talk and fillers (e.g., *I guess, I don’t know*) also allowed them to produce longer runs. Finally, the use of FS as rhetorical devices

helped the participants improve their storytelling, avoid pauses, and extend mean length of runs. This mixed method study provided important findings about the connection between formulaic language and fluency, and illustrated that improved spontaneous L2 speech, as well as more frequent use of FS, can be a result of living and studying in an ESL environment.

Given the potential benefits for L2 oral production when using FS, many applied and educational linguists have emphasized the importance of addressing formulaic language in the L2 classroom (Boers et. al., 2006; Foster, 2001; Lewis, 1993; Nattinger & DeCarrico, 1992; Wood, 2009; Wray, 2002). The following discussion addresses research undertaken in the field of teaching formulaic language and illustrates how the acquisition of FS may affect L2 fluency.

2.4. Focused instruction on formulaic language

Formulaic language instruction has been overlooked in SLA research for many years due to the difficulties inherent in teaching the idiomatic and special structures of FS (Wray, 2000). One very specific aspect of FS is that they may be semantically irregular (e.g., *beat around the bush, by and large*) which makes them rather challenging to acquire for L2 learners who tend to fail to interpret the meaning correctly (Boers, Eychmans, & Stengers, 2007; Wray & Perkins, 2000). Even though teaching and mastering formulaic language may be rather challenging, FS are nonetheless believed to be more or less prefabricated and ready to use, and, if memorized, can save a lot of time and effort in L2 production (Wray, 2000). Previous research on formulaic language acquisition focused on the most effective ways of learning and teaching FS in L2 classroom and exploring how instruction of FS may facilitate more proficient and fluent oral production. Moreover, there is a large body of research which studied effectiveness of using formulaic language in written language, especially in academic discourse (Ellis, Simpson-Vlach,

& Maynard, 2008; Jones & Haywood, 2004). For the purpose of this study, research on formulaic language in writing will not be discussed but instead the focus is on the literature which explored ways of teaching formulaic language in order to facilitate oral communication.

Nattinger and DeCarrico (1992) recommended engaging L2 learners in practice and drills in order to develop their fluency in using FS. They also emphasized that students should analyze the use of FS on the sentence-level as well as explore the role of FS in producing the whole text, and used model texts and texts written by students for practice. Willis (1990) created methodologies based on lexis with a focus on formulaic language and involved L2 learners in analysis of the form and meaning of certain FS as a part of explicit instruction. Even though this type of instruction may seem somewhat artificial, as argued by Wray (2000), an analytical approach to teaching FS may help L2 learners understand what a particular FS means and how it can be used in a particular context.

Several pedagogical intervention studies investigated teaching strategies on how to instruct and/or facilitate acquisition of formulaic language. The commonality among them is that the teaching approach involved focused instruction and raising awareness of formulaic language (Boers et al., 2006; Peters, 2012; Stengers et al., 2011; Webb et al., 2013). Lewis (1993) was one of the first advocates of teaching formulaic language in the L2 classroom and developed Lexical Approach suggesting that lexis, rather than grammar, should be the focus of the syllabus. The author created different pedagogical techniques in the form of awareness-raising activities aiming to optimize acquisition of FS. According to Lewis's approach (1993), teachers should systematically make L2 learners notice recurring FS in the authentic input through different activities. He also argued that activities which raise the awareness of FS may assist in individual learning.

Interestingly, Boers et al. (2006) decided to implement Lewis's (1993) Lexical Approach in order to test whether "chunk-noticing" technique could facilitate production of target FS in the post-task (p. 248). Thirty-two adult L2 learners in an as a foreign language (EFL) program in Belgium, with an advanced level of proficiency in English, were selected for this small-scale experiment over an eight month period. Both experimental and control groups had the same course materials and authentic language input (audio, video and textual). The students were provided with transcripts of recordings used for listening tasks. The experimental group engaged in noticing activities and was to highlight high-frequency FS in input. The control group, in turn, did not follow the Lexical Approach and only focused on some grammar patterns in listening and reading materials, but were asked to highlight "interesting" individual words they found in the texts (p. 251). The experimental group processed FS in the input twice: first, focusing on meaning, and then, on the input itself. The attention of the experimental group was called by the teacher to the exploration of co-text, "the environment in which words appeared" (p. 251). As part of an oral proficiency assessment at the end of the course, the participants needed to retell what they understood from a given text (i.e., short article copied from a newspaper or a magazine), and then to perform spontaneous speech on some unprepared, yet familiar topic. The assessment was done by two different judges who listened to a recording or had a face-to-face interview with the participants. Boers et al. (2006) were interested in determining whether students could be perceived as notably more proficient because of the use of FS in their speech. The results seemed to suggest that this was the case. The experimental students received higher scores because of their enhanced awareness of formulaic language. However, the difference between the control group and the experimental group in terms of the use of FS was not significant. A limitation of the findings in the study of Boers et al. (2006) is that the experimental

students who used more FS in their retellings transferred some FS from the texts and repeated them word for word. This limitation was taken into consideration in another study by Stengers et al. (2010) where the authors repeated the procedure from the study conducted by Boers et al. (2006) but prepared a final text in students' L1 in order to avoid transferring of FS word for word. Similarly, there was no significant difference in the use of formulaic language between the two groups of participants. Considering these results, it may be argued that activities which raise awareness of formulaic language can be useful for L2 learners, but it seems that noticing FS in input does not necessarily guarantee their active use or acquisition (Wood, 2015). Boers et al (2006) suggested that, in addition to noticing activities, L2 learners need to be involved in more practice with FS.

Webb et al. (2013) tested an input flooding technique with their 161 Taiwanese students of English at three different universities in order to see what effect it could have on incidental learning of formulaic language (collocations, in particular). This technique implies using a modified text which would contain a certain amount of target FS. There were four experimental groups, each of them read a modified graded reader while listening to a recording of the story which contained 18 target FS, encountered different number of times (i.e., 1, 5, 10, 15). The control group did not read or listen to any version of the graded reader but had the same pre-test and post-tests as the experimental groups. The pretest measured the participants' receptive knowledge of the target FS form while four post-tests measured productive and receptive knowledge of form and meaning of the target FS. The findings of Webb's et al. (2013) research confirmed that frequent exposure to target FS within a stretch of discourse may facilitate incidental learning of FS. In fact, when L2 learners encounter a certain FS 15 times within a modified graded reader text, they can acquire that FS.

Even though previous research emphasized that increasing exposure to and awareness of target FS might be a way of teaching formulaic language in the L2 classroom, it seems that teachers should also provide students with more opportunities to practice those FS and, possibly, even encourage their memorization in preparation for future conversations. Verbatim memorization of complex forms (i.e., formulaic language) may be a necessary prerequisite for language development in real communication and it may interact with analytical processes in language learning (i.e., analysis of rules) (Johnstone, 1988; Nattinger & DeCarrico, 1992). Wray (2008) explained that even though lower proficiency students may benefit more from memorizing FS, reproducing memorized material as a learning method may be effective for advanced learners as well. Memorization produces the ability to use FS (Yu, 2009). For example, Wray and Fitzpatrick (2008) reported on a study where they provided positive results of pushing learners to memorize and use utterances which were beyond their current competence. They suggested that a number of FS should be memorized by L2 learners so that they could be transferred later into real life encounters with native speakers. In their study, practice and rehearsal of FS made the participants sound more nativelike and feel more confident during interactions with native speakers about their language. It seems that memorization may be effective when L2 learners have an understanding of the material being practiced. Similarly, Gatbonton and Segalowitz (2005) in their teaching methodology, encouraged genuine repetition and automatization with implications for dealing with FS, arguing that a task should create a genuine need for L2 learners to use FS. In order to ensure holistic and automatic retrieval of FS from memory, then, it seems that L2 learners should be encouraged to process FS on a deeper level, to practice and to automatize them. For automatization to occur, L2 learners should be

engaged in the production of long stretches of discourse such as presentations, dramatizations, and role-plays (Wood, 2015, p. 145).

Interestingly, Wood (2009) in his case study with a Japanese L2 speaker of English, also implemented Lewis's (1993) Lexical Approach to explore the development of FS and fluency due to focused instruction. However, he provided his participant with some additional practice of target FS. Wood's participant was engaged in a fluency workshop which consisted of three blocks of six hours each, with the focus on formulaic language. The fluency workshop contained four stages: the input stage, the automatization stage, the practice and production stage, and finally the free-talk stage. As part of the automatization stage, the participant participated in different meaning-focused and time-pressured activities. One of them was a shadowing technique, which is usually used in pronunciation teaching, where the students are supposed to shadow "the recording with the transcript in the language laboratory at least eight times" (Wood, 2009, p. 48). Then, a dictogloss – an activity during which a teacher reads a text aloud and gets students to note key ideas from it to then reconstruct the text together (Wajnryb, 1990) – was among these. Another activity was a mingle jigsaw where the learners are given a slip of paper with a key FS written on it, so that they can share their assigned FS with peers in a chat circle and listen to the FS produced by their fellow students in return (Wood, 1998). The participant's narratives were collected before and after the fluency workshop. Her speech was rated as more fluent in the second sample, with a larger number of FS compared to the first time. Moreover, the second narrative employed a greater variety and complexity of FS. Even though the results of this study are limited to one participant, they seem to have persuasive implications for teaching and practicing formulaic language and fluency in the L2 classroom.

Thomson (2017) reported on a quasi-experimental, six-week study with Japanese participants in an EFL environment, which explored the benefits of a focused instruction on formulaic language and fluency, and used the fluency workshop proposed by Wood (2009). In order to test the effectiveness of the fluency workshop, Thomson (2017) divided her participants into experimental (n = 44) and control (n = 29) groups. Students from the control group were from a general EFL class where they were not exposed to target FS and did not have activities which the experimental group had. The students in the experimental group were involved in the following activities as a part of the fluency workshop: listen with gist questions, marking pauses, phrase instruction, shadowing, dictogloss, mingle jigsaw and different kinds of role-plays. Three model dialogues with ten target FS each were created for this study and used in a pre-test and a post-test. Even though Thomson (2017) used the fluency workshop in her study, there were a number of differences between her and Wood's (2009) research. For example, Thomson's quasi-experimental study was undertaken in a small Japanese city with limited exposure to English; whereas Wood's case study was conducted in an ESL context, in Canada. Another difference was that the proficiency of Wood's participant was higher than that of Thomson's participants. The results of Thomson's (2017) study indicated that, the fluency workshop turned out to be effective for the experimental group where it showed development of productive knowledge by improving recall of target FS and also, increased the speech rate in the post-test, compared to the pre-test results. Thus, her findings illustrated a positive correlation between fluency and the use of FS. However, speaking fluency and the number of FS in the experimental group did not increase significantly on the post-test when compared to the control group who did not have activities from the fluency workshop. Thomson (2017) suggested that the reason for this discrepancy might have been the limited sample size of the spoken dialogues (i.e., 3).

In contrast, a recent study by McGuire and Larson Hall (2017) which explored the effects of explicit instruction on formulaic language, revealed positive results for the use of FS and fluency after noticing and practicing FS. They reported on a five-week study conducted in the USA with L2 students, who were divided into control (n =8) and treatment (n =11) groups and were taught using task-based teaching techniques. The control group practiced speaking and listening using authentic materials in English; the treatment group did the same but also focused on noticing FS in the input. Unlike previous studies which investigated the role of formulaic language in fluency using a monologue as the final task, McGuire and Larson Hall (2017) chose a dialogic task on the same topic in both a pre-test and a post-test. Objective fluency measures (i.e., speech rate and mean length of runs) were supplemented with subjective native-speaker judgments. Overall, the students from the treatment group significantly increased their FS use as well as the speech rate and mean length of runs, outperforming the control group on most measures. These findings suggest that explicit teaching of FS can help students use more FS utterances and that this use might lead to students becoming more fluent in spontaneous speech. It should be noted though that this study might have had such positive results in the use of FS and fluency not only due to the pedagogical intervention, but also because the participants were all living in the United States and had significant exposure to the L2. In addition, the students had mid-intermediate to advanced levels of L2 proficiency. This may also suggest, then, that when L2 students have a higher level of proficiency, they are likely to respond more positively to focused formulaic language instruction. This idea is supported by research which showed that only advanced L2 learners demonstrate a nativelike use of formulaic language (Laufer & Waldman, 2011).

Since practice and automatization seem to be key elements leading to FS acquisition, it can be argued that task repetition, a popular classroom strategy (Bygate, 2001), may be an effective pedagogical tool for promoting acquisition and active use of formulaic language. The purpose of the next section is to discuss the technique of task repetition and its benefits to L2 oral production.

2.5. The role of task repetition in L2 performance

Over the past two decades, task repetition has attracted the attention of many researchers. Studies on task repetition have focused on different task design features, operationalizations of the task, and the effects that task repetition has on the quality of L2 oral production in terms of fluency, complexity, accuracy, and lexis (Ahmadian & Tavakoli, 2011; Bygate, 1996; Gass et al., 1999; Lynch & Maclean, 2001). Task repetition is defined as “repetitions of the same or slightly altered tasks – whether whole tasks or parts of a task” (Bygate & Samuda, 2005, p. 43).

Accordingly, some studies have investigated the extent to which task repetition influences L2 performance in a new task (Bygate, 2001; De Jong & Perfetti, 2011; Gass et al., 1999; van de Guchte, Braaksma, Rijlaarsdam, & Bimmel, 2016) while some studies explored the benefits of repeating the same task or a slightly altered task (Bygate, 1996; Lambert, Kormos, & Minn, 2017; Lynch & Maclean, 2001). In this study, task repetition refers to repeating the same task.

Task repetition is a widely used classroom strategy within a TBLT methodology which is often discussed as an implementation procedure for oral tasks (Qui & Lo, 2017). Over the last thirty years, TBLT has been an area of interest for many SLA researchers, teachers, and curriculum designers (Ellis, 2003; Skehan, 1998; Willis, 1996). TBLT is an innovative language teaching approach which fosters authenticity in the L2 classroom by including real-world tasks

and usually following the three-phase procedure of pre-task, on-task, and post-task (Bygate, 2016). Ellis (2003) explained that some of the main characteristics of the task is that it should focus on meaning and be outcome-oriented. Even though TBLT does not include explicit grammar instruction, there is a way to focus on form during meaningful communication. For instance, it can be done through task repetition after a form-focused feedback on some target grammatical structures during a main task, according to van de Guchte et al. (2016). The results of their experimental study with 48 Dutch students learning German as a foreign language indicated that task repetition, as a post-task activity, could promote accurate use of new grammatical structures. However, van de Guchte et al. (2016) reported that students who performed the task more accurately showed less fluency. This can be explained by the fact that the participants primarily focused on retrieval and application of the new grammatical structures and on making fewer errors (van de Guchte et al., 2016). Similarly, Skehan (1998) argued that attending to one dimension of performance (e.g., accuracy and focus on form) may have some detrimental effects on other aspects of performance (e.g., fluency and focus on meaning).

The majority of aforementioned studies on task repetition employed Levelt's (1989) speech production model as a theoretical framework. This model has already been discussed in the chapter in regard to fluency research on temporal variables (Towell et al., 1996). As previously mentioned, three processing stages posited in Levelt's (1989) model include: conceptualization, formulation, and articulation. Conceptualization is believed to be a conscious and cognitively demanding process in both L1 and L2, when a speaker generates ideas and creates content for a message during a task. Accordingly, during second enactment of the task, L2 speakers rely on previously conceptualized task content and activate recently used linguistic constructions to express their message (Bygate, 2001). This results in better L2 performance

because L2 speakers spend more attention on the language being used. Thus, task repetition “might reduce the attentional demands on learners to conceptualize, encode and monitor their message simultaneously” (Lambert et al., 2017, p. 8).

One of the earliest attempts to study task repetition was Bygate’s (1996) study where he investigated the effects of exact repetition of a task on language production in two performances of one participant who watched a cartoon video and was asked to retell it. The results of this study supported the hypothesis that during the initial attempt the participant was concerned with content planning and was under time pressure to use linguistic resources (i.e., lexical and syntactic elements) to communicate. During the second attempt, because the participant was already familiar with the content, she paid more attention to the formulation of the intended meaning, producing output of higher fluency, complexity (i.e., a large number of subordinate clauses, lexical richness) and accuracy (i.e., the proportion of error-free clauses to all clauses) (Bygate, 1996). Another interesting study was conducted by Gass et al. (1999) who explored whether task repetition may lead to more sophisticated language use. One hundred and three native speakers of English learning Spanish as a foreign language were asked to watch fragments of Mister Bean videos with no audio in a language laboratory; while watching, they needed to narrate and record their own version of what was happening in the story. Two experimental groups watched the video and retold it four times, while a control group performed the task only twice. Interestingly, one experimental group had the same content for task repetitions while another one had new content each time. Overall, task repetition led to greater proficiency, accuracy and lexical sophistication. Gass et al. (1999) argued that “freeing up attention to meaning allows learners to gain greater control over their linguistic knowledge” (p. 573). However, the results did not transfer to a new context.

Conversely, a growing body of research is now suggesting that three or more iterations of a task may be necessary in order to bring about changes to accuracy (Boers, 2014; Hunter, 2017). Lambert et al. (2017) conducted a study which involved six task repetition performances with different interlocutors, and explored their effects on short-term L2 fluency. They reported that their participants' speech rate improved markedly over the first three performances and then more gradually until the fifth performance. While the speech rate increased with the first repetition, self-repairs and reformulations decreased only in the fifth and sixth performances. Lambert et al. (2017) argued that this happened later rather than sooner because only by this point were the students able to optimize their conceptualization and encoding processes resulting in fewer self-corrections. These findings illustrate different stages of L2 production and suggest that when learners are given more opportunities to further repeat a task, additional changes may be observed in terms of fluency and complexity. Even though it seems that task repetition may be less interesting or engaging, compared to other methods, recent studies report that L2 learners actually find it very facilitative and not redundant (Lambert et al., 2017; Lynch & Maclean, 2000). For instance, Qui and Lo (2017) suggested that if teachers are concerned that students may find repetition of the same task boring, they could let them repeat this procedure with different interlocutors. Furthermore, as evidenced by the results of their study, L2 learners are behaviorally and cognitively more engaged in tasks with familiar topics and report more enthusiasm, confidence, and interest when performing them. Their Chinese EFL learners also had fewer self-repairs in tasks on a familiar topic because they were more familiar with the content and required "fewer attentional resources to conceptualization" (Qui & Lo, 2017, p. 693). Finally, Qui and Lo (2017) also found that when their participants repeated a task on an unfamiliar topic, they felt highly motivated to complete it and were more involved in monitoring

their language. Overall, it seems that teachers may choose a topic for task repetition based on the goal they have for their students as tasks on familiar or unfamiliar topics reap their own benefits in students' oral production.

Interestingly, Bui and Huang (2018) explored the effects of content familiarity on fluency and revealed that it had a positive influence on speech rate, phonation time (i.e., the ratio of voicing time to the total time of utterance), breakdowns (i.e., hesitations, pauses) in the middle of the clause, and repetitions. In contrast, their findings illustrated that content familiarity was not as helpful in reducing pauses, self-repairs or increasing mean length of runs. Bui and Huang (2018) also compared the effects of content familiarity and pre-task planning on L2 fluency during task repetition and found that pre-task planning had a stronger impact on fluency measures. In terms of pedagogy, it is important to provide L2 learners with some help in performing tasks in which planning may represent a facilitative aspect.

While content familiarity may be an implicit type of planning, Ellis (2009b) distinguished three explicit types of planning: pre-task planning, rehearsal, and within-task planning or on-line planning. It should be noted that, Ellis (2009b) divided pre-task planning into strategic planning (i.e., "planning what content to express and what language to use but without opportunity to rehearse the complete task) and rehearsal (i.e., "an opportunity to perform the complete task before performing it a second time") (p. 474). Thus, pre-task planning gets learners ready for the task by offering preparation time and helps to compensate for lack of content familiarity (Bui & Huang, 2018). In task repetition, the first performance serves as a preparation (pre-task activity) or rehearsal before further enactments of the task (Ellis, 2009b). In general, pre-task planning seems to assist fluency the most while on-line planning benefits accuracy; however, both types of planning seem to have a positive effect on complexity (Yuan & Ellis, 2003). Furthermore,

Ellis (2009b) argued that rehearsal alone can have a positive effect on L2 performance but when it is followed by some treatment (e.g., corrective feedback) it has more effect on all three aspects: fluency, complexity and accuracy. For example, in Sheppard's (2006) study, Japanese participants were provided with feedback on their initial performance of an oral narrative task, before task repetition. This led to gains in fluency, complexity (i.e., lexical complexity), and markedly in accuracy.

Studies on task repetition also explored different spacing between the original and repeated performance: immediate (Lynch & Maclean, 2001; De Jong & Perfetti, 2011), after a few days (Bygate, 1996), after a few weeks (Fukuta, 2016), and finally, after a few months (Bygate, 2001). For example, Bygate (2001) who compared task repetition performances of 48 participants within a ten-week interval revealed that even after this period, the students showed increases in fluency and complexity; however, there were no significant changes in accuracy. In contrast, Lynch and Maclean (2000, 2001) conducted a task repetition study with an immediate interval of three minutes, where 14 oncology students learning English for specific purposes were required to present their research six times in a poster presentation carousel. Their findings indicated that the participants' speech became more fluent and accurate after several repetitions. Bui, Ahmadian, and Hunter (2019) in their recent study on task repetition, explored five performance conditions with different time intervals between the tasks. They found that a one-week interval had the strongest effects on structural complexity and repair fluency (i.e., related to self-correction, reformulations, repetitions), whereas immediate repetition benefited speed fluency (i.e., related to speech rate), as defined by Skehan (2003). Interestingly, the results indicated that fluency increased due to task repetition regardless of spacing between attempts. On the other hand, results associated with accuracy and complexity are not as predictable and are

mostly observed in task repetition with a short interval between performances (Bui et al., 2019). Overall, task repetition is said to be particularly beneficial for fluency and complexity (Bygate, 2001; Gass et al., 1999) and may be used as a focus on form for improving accuracy in L2 classrooms (van de Guchte et al., 2016).

2.6. Conclusion

In sum, the goal of this literature review was to discuss the importance of such complex phenomena as formulaic language and fluency in L2 performance, to illustrate a potential link between them, as well as to suggest the use of task repetition as a way of improving fluency and possibly, promoting formulaic language in the L2 classroom. Even though focused instruction on formulaic language, which encourages awareness and noticing of FS, may be one of the most popular ways to instruct FS, it does not always lead to acquisition and therefore, should include more practice, repetition and automatization in order to be effective (Boers et al., 2006, Wood, 2015, p. 141). As highlighted in this chapter, task repetition has a positive effect on spoken language production and fluency (Ahmadian & Tavakoli, 2011; Bygate, 1996; Gass et al., 1999; Lynch & Maclean, 2001), but there is no study to date which tests the role of task repetition on formulaic language use. Therefore, the purpose of the present study is to explore whether task repetition can be an effective practice for promoting the use of FS and fluency without any explicit instruction on formulaic language. Furthermore, the present study was conducted in an English for academic purposes (EAP) classroom, where instructors usually focus on writing and reading rather than on speaking, and rarely address formulaic language explicitly (Barnard & Scampton, 2008; Evans & Green, 2007). Thus, it is important to discover whether task repetition,

as a meaningful speaking practice, can be a way to encourage the use of FS and improve L2 students' oral production, as a result. The following research questions guided the investigation:

4. What are the common formulaic sequences (FS) used by Chinese L1 EAP students and what functions did they serve in the students' presentations?
5. What are the fluency profiles of Chinese L1 EAP students in their presentations?
6. Does task repetition bring on FS and fluency gains (independently and/or together) in the presentations of Chinese L1 EAP students?

Chapter 3: Methodology

3.1. Overview

This chapter elaborates on the methodological processes adopted in the present study. The small sample (n=10) provided an opportunity to present the findings in rich contextualized detail (Dörnyei, 2007). The research employed a time-series design, which is frequently used with a small group of participants and involves repeated observations (Mackey & Gass, 2015). The chapter starts with a brief examination of the English for Academic Purposes (EAP) programs across Canada and describes the program from which the participants for this study were recruited. Then, the information about the participants and the task they engaged in is provided. Analyses of formulaic language and fluency are described in the data analysis section. Since this research is qualitative in design, the potential for researcher bias is high; to counteract subjectivity in the analysis, inter-rater reliability and intra-rater reliability measures, as well as rater training employed are presented at the end.

3.2. EAP context

The current study examines the effects of task repetition on the development of formulaic language and perceived fluency gains among Chinese L1 speakers attending a university-level EAP program in Eastern Ontario, Canada. EAP programs usually serve as English language support programs, which prepare students to transition to their university studies with English being the medium of instruction (Cheng & Fox, 2008). These programs are very popular across Canada (and other English-speaking countries) due to the increasing numbers of L2 students studying at Canadian universities (Fox, Cheng & Zumbo, 2014). The population of these classes is usually international students and other language learners required to demonstrate their language

proficiency to be able to follow content-based English-medium instruction. The EAP classes provide students with the necessary knowledge of the basic study techniques specific for the curriculum of the university, promote the development of their academic language use, and teach them analytical and critical thinking skills (Cheng & Fox, 2008). As explained by Cheng, Myles and Curtis (2004), the activities in EAP classes usually “simulate academic work; for example, reading academic texts and taking notes; writing research reports or essays; making formal seminar presentations” (as cited in Fox et. al, 2014, p. 58). It is important to note that even though in the past EAP programs used to be referred to as English as a Second Language (ESL) programs at some universities, they are different (Fox et. al, 2014) since ESL programs are usually oriented towards the development of the communicative competence and involve students in role-plays and group work, emphasizing a sense of integrating and belonging to the new society (Savignon, 2002). Furthermore, EAP programs differ in their methodology and approach, and surveys conducted across Canada did not demonstrate any consistency among them (Berman, 2002).

Berman and Cheng (2001) conducted a Canadian needs assessment survey in order to compare the questionnaire answers by L2 and native English-speaking undergraduate and graduate students. The survey revealed that EAP students consider giving presentations and writing essays and reports the hardest language skills to master. Similarly, native speakers at the graduate level reported participating in discussions and delivering presentations difficult. Usually, oral presentations are a standard requirement for students in EAP classes. As a rule, they need to deliver an oral presentation at the end of the term, where they have a chance to practice their presentation skills and public speaking, which are highly useful for their study at the university. With a limited amount of contact hours (e.g., six hours a week during one term in the ESLA program described in the present study – see below), teachers struggle to focus on four language skills equally,

choosing to favour some skills over the others. This was also the case in the ESLA program that favoured practice of reading and writing skills over listening and speaking, similarly to many other EAP programs (Ferris & Tagg, 1996a). And this is despite the fact that presenting in English or participating in discussions during lectures and seminars is an essential part of university success (Ferris & Tagg, 1996b). In fact, since oral presentations are normally scheduled for the end of the term, students do not have a chance to get detailed feedback about their performance and areas to improve on in the future. Instead, they usually get a grade along with a few very general comments and no chance to deliver another presentation, where they could correct the mistakes they had made during their first attempt. Clearly, one presentation is not enough for students to practice their speaking skills. Moreover, international students from Asian countries, such as China, Japan and Korea, who are expected to be silent and passive learners in their educational system back home, are usually known to be silent in class and might experience speaking anxiety, and for those reasons, need more oral practice (Kim, 2006; Woodrow, 2006). Therefore, despite completing an EAP program, some students might still remain anxious about their conversational ability in English because of the lack of sufficient oral practice. This sentiment, unfortunately, is not surprising since EAP programs, normally, fail to supply students with enough feedback on and practice of their speaking skills (Barnard & Scampton, 2008; Evans & Green, 2007).

The EAP program from which the participants of the present study were recruited is called ESLA (English as a Second Language) and provides instruction at three levels of proficiency: introductory (ESLA 1300), intermediate (ESLA 1500), and advanced (ESLA 1900) (“Credit ESL”, n.d.). The ESLA students are required to take the courses when their English proficiency scores on such standardized tests as IELTS, CAEL, or TOEFL are lower than the university admission minimum (i.e., IELTS=6.5; TOEFL=86; CAEL=70 (“Credit ESL”, n.d.). Since EAP students

usually come from a variety of disciplinary backgrounds, it can be quite challenging for programs to organize the curriculum and materials for these courses. The ESLA program follows a Theme-based Language Instruction (TBLI), which implies that the instructors create an overarching theme for the ESLA course that they believe to be beneficial for all students regardless of their major (Myers, 1996). They choose one or more mini-topics which would meet the interests of varied students and could be connected to the students' majors. Instructors at the ESLA program usually try to integrate four language skills in their classes. However, the main focus is mostly on writing rather than on speaking. The focus on speaking might vary depending on the needs analysis which ESLA instructors do at the start of the term.

This study presents a unique case where an EAP instructor created an assignment that had the students deliver two oral presentations on a topic of their choice, once at the beginning of the term and then again, at the end of the term. Additionally, the students had a chance to reflect and receive teacher's feedback on their general linguistic performance as well as on their presentation skills after each presentation (see section 3.4. for a detailed description of the task completed by the participants of the present study). Overall, the participants viewed the task repetition experience positively and confirmed that they had benefitted from it. According to an EAP instructor, most of the students had never recorded themselves in English before and therefore, found the task repetition and the need to reflect on the presentation very useful for evaluating their own progress.

3.3. Participants

The data were drawn from the research conducted by Kartchava and Nassaji (2019) with participants (n=26) recruited from the ESLA program described above. For this project, however,

a secondary data clearance was obtained (see Appendix A for the certificate). Ten randomly selected student files made up the data for the present study. All the students attending the ESLA course were L1 Mandarin speakers, which is a rare case since, usually, EAP classes in the Canadian context involve students of different backgrounds. This provided a unique opportunity to examine this particular group of learners in detail. The students (mean age: 21) were pursuing different majors (e.g., engineering, business, computer science). At the time of the study, they were attending an intermediate ESLA 1500 level course. Most of the students in this group were repeating the course because they had not gotten the required grade of C+ or higher, and as such, could not advance to the next level in the program (i.e., the ESLA 1900 level).

Although their out-of-class exposure to English (via media or Internet) was not directly investigated for the purposes of this study, the EAP instructor responsible for the class explained that the majority of his students were newcomers to Canada and were new to the rules of the multicultural society. They spent their free time mostly with their Chinese speaking friends or relatives, forming a primarily L1 focused community that rarely engaged in the practice of English with native speakers or other non-Chinese speaking peers outside the classroom (Will, 2016). Moreover, during in-class group work the participants were inclined to use L1 instead of English. This might be explained by the natural desire to use L1 to cope with tasks faster and easier; it can also be a cultural transfer from the education system in China where teachers may allow the use of Chinese in the English classroom as a tool to facilitate such tasks as ensuring general comprehension or focusing on explanations of specific grammar, vocabulary, and difficult concepts (Jan, Li & Lin, 2014).

In terms of the participants' exposure to formulaic language, it can be claimed that because this group was not fully integrated into the Canadian society, their exposure to formulaic language

was quite limited. An examination of the first version of each presentation revealed that the amount of formulaic language produced was rather small. The EAP instructor's assessment of the students' L2 performance also indicated that their speech did not contain a large amount or variety of formulaic sequences (FS). Moreover, having examined the publicly available information on the kinds of language tested by IELTS, TOEFL or CAEL, it appears that none of the tests explicitly claim measuring the knowledge of and proficiency in formulaic language ("CAEL", n.d., "ETS TOEFL", n.d., & "IELTS Official Test Centre", n.d.), but see IELTS' references "uncommon lexical items" and "idiomatic vocabulary" ("IELTS Scoring guide", n.d). Therefore, it can be assumed that the participants in this study had not focused on the development of formulaic language as part of their test preparation. It can also be suggested that most of the students might have not been explicitly taught about the nature of FS during their studies back home at all. The assumption is strengthened further by the complicated nature of formulaic language, which makes it challenging to integrate into language pedagogy most of the times, and which might discourage some teachers from trying to focus on formulaic language explicitly in their classroom (Meunier, 2012).

3.4. Data collection

The data came from two required presentations the students produced during the ESLA course. The learners were asked to prepare and deliver a 5-minute oral presentation on a topic of their choice; the topic, however, had to be confirmed beforehand with their instructor. They were required to choose a familiar topic on which to present in order to avoid memorizing new content, focusing instead on improving their academic presentation skills. For the presentation, the students were required to use slides and were allowed to use notes albeit they were strongly encouraged not

to read their presentations verbatim. Their oral presentations were audio-recorded. Using an evaluation rubric (Appendix B), the teacher supplied feedback on the students' presentation skills and spoken English but did not comment on their formulaic language use. Then, at home, the students were asked to listen to their recording, analyze the teacher's feedback, and assess the presentation quality in terms of format and form. They were required to use the same rubric that their teacher used for evaluating their presentation in class. The five-trait rubric focused on the following areas: voice quality (volume and pace), pronunciation, intonation, word and sentence stress, eye contact and presentation style, organization, content, and the appropriate use of the APA style in references and in-text citations. In addition to the self-assessment, the learners were required to write a 400-500 word reflection on the process of delivering a presentation in front of the class, describing their feelings during the performance, and evaluating their strengths and areas for improvement. The reflections also needed to include what the students learnt from this assignment and how they were planning to improve their revised presentations.

Having studied their own performance and the teacher's feedback, four weeks later the learners were asked to prepare and record another version of the same presentation from home. Thus, they were free to practice their second presentation as many times as needed to produce a version they found satisfactory. The second recording was submitted for grading to the teacher, and the students were asked to once again reflect on any improvements they felt they had made in this new recording compared to the first one. The teacher evaluated the second version of the presentation using the same five-trait evaluation rubric as before but did not provide as detailed feedback as he had previously.

It is important to note that neither Kartchava and Nassaji (2019) nor the researcher of the current study had any control over the task design and procedure employed, which were

determined by the ESLA instructor. Using a task designed by a practicing EAP instructor with an intact class of learners alleviated concerns of ecological validity. This is furthered by the fact that the participants did not know that their recordings would be analyzed in terms of their formulaic language use and fluency development. Hence, the students' use of formulaic language occurred naturally, without a deliberate focus and expectation for its use. Thus, the impact of task repetition on improvements in the use of FS and oral fluency was evaluated for their own sake without a pedagogical intervention. This makes the current study unique because previous research on the link between formulaic language and fluency involved explicit instruction on formulaic language and intended to bring awareness to the participants about the existence of FS in the input (e.g., Boers et al., 2006; Qi, & Ding, 2011; Stengers et al., 2010; Wood, 2006; 2009; Wray & Fitzpatrick, 2010). Lewis (2000), who was one of the first advocates for teaching formulaic language, created Lexical Approach that encouraged learners to notice FS in the input during instruction (as cited in Boers et al., 2006). Lewis and others found that a focused instruction can bring about positive changes in the use of FS and oral fluency. In contrast, the present study contributes to the SLA field by attempting to measure the learners' ability to integrate FS into their language repertoire on their own through task repetition and without explicit instruction of any kind.

3.5. Data analysis

A total of ten students and 20 recordings (two recordings per student) made up the data for the current study. Two versions of the presentation performed by each student were thoroughly examined and compared to determine whether task repetition had positive effects on the development of formulaic language use and fluency. Both versions were transcribed to ease the analyses. The transcriptions did not include indications of pauses or lengthened syllables. The

analyses of formulaic language and fluency were first done by the researcher. Then, a portion of the data were analyzed by five different independent judges recruited to ascertain the researcher's objectivity. After attending the benchmarking session, the judges analyzed a representative sample of the data, with the two formulaic language judges assessing 30% of the data and the three fluency judges rating 50% of the data. The formulaic language judges analyzed a smaller sample because the nature of their analysis, which involved identifying all FS in each transcript, was more complex and time consuming than the fluency judgment of a single recording.

Considering the complicated nature of formulaic language and subjectivity of perceived fluency scoring, inter-rater reliability and intra-rater reliability were assessed in the present study. Inter-rater reliability was calculated using the simple percentage agreement (Mackey & Gass, 2005). This measurement was chosen because it is easy and appropriate for continuous data – “limited in precision only by our ability to measure them”, as mentioned by Mackey and Gass (2005, p. 243). While anything above 75% is considered “good”, values over 90% are seen as “ideal” (Mackey & Gass, 2005, p. 244) – please see the detailed description of the independent judges and the analyses they completed in sections 3.5.1. and 3.5.2 below. Intra-rater reliability, which involves the researcher reanalyzing the data at a different period of time to compare the ratings and ensure reliability of the original analysis, was measured in a similar way as the inter-rater reliability described above (Mackey & Gass, 2005).

3.5.1. Formulaic language analysis

Formulaic language is often referred to as FS which are or appear to be prefabricated and retrieved from memory as a whole (Wray, 2002). However, it can be quite challenging to identify these sequences and determine whether they are indeed formulaic or just an occasional cluster of words that happen to occur together (Wood, 2015). This is because some sequences do not follow

a general rule of fixedness and do not have an invariable form, therefore appearing to be less salient and recognizable as FS (Wray, 2002). For instance, *so far so good* is an example of a fixed FS while the FS *pull (one's) leg* can be modified depending on a sentence and intended meaning; it can also sound concrete and literal, rather than abstract, to a non-native speaker. In the spoken language, the identification of FS goes beyond the written form and considers phonological or prosodic aspects of the articulation of a FS as well as fluency criteria such as lack of pauses within a sequence (Wray, 2002).

Formulaic language constitutes a great proportion of the spoken language, which means that speakers of the target language usually share knowledge of how to appropriately use a certain FS in a particular context or situation (Altenberg, 1998). According to Altenberg (1998), FS constitute 80% of the words in London-Lund corpus of spoken English. Therefore, the frequency with which FS are used by a large group of people usually determines whether or not they are formulaic. A popular method of FS identification is the frequency-based approach, which requires analyzing and scanning a corpus (Wood, 2015). However, because the current study has a small data set, it was not possible to apply statistical identification of FS required for large data samples. Frequency, however, is not the only criterion which determines whether a sequence is formulaic or not. The FS needs to have a unitary meaning and/or function and be mentally stored in a particular way to be retrieved easily at the moment of use (Wood, 2015). For small spoken corpora, Wood (2015) recommends using comprehensive checklists to determine whether a certain sequence is formulaic or not. These checklists serve as a guideline for the researcher and facilitate the isolation and analysis of FS.

To identify FS in the participants' presentations, Wray and Namba's checklist (2013) was chosen due to its rigor, wide use among SLA scholars, and its repeated validation in research on

formulaic language with bilingual children as well as with adult native and nonnative speakers (Dickinson, 2013; McGuire & Larson-Hall, 2017). The checklist is sophisticated, yet comprehensive and consists of 11 diagnostic criteria which characterize the FS from different domains (see Appendix C). Such spectrum of different judgments assists in the analysis of the unique nature of FS. The researcher using the checklist would rate his/her perception of a word string being formulaic or not by using a Likert scale, where 1 would indicate a strong disagreement, 2 - a disagreement, 3 - neither an agreement or disagreement, 4 – an agreement, and 5 - a strong agreement. The focus on individual perception in the assessment of formulaicity is reiterated in the phrasing of the criteria, with every criterion beginning with “by my judgment”, as in “By my judgment, there is something grammatically unusual about this word string” (Wray & Namba, 2003, pp. 29-32). In the case of this study, the judgment was based on both the researcher’s intuition and knowledge of English as well as on the native speaker’s intuition in the case of the independent judges. It is also important to note that each sequence does not have to match all criteria in a given checklist to be considered formulaic (Wood, 2009). In the present study, a sequence was considered formulaic if it matched at least five criteria from the checklist. The checklist covers judgments of grammaticality of the FS, its semantic transparency, the register this word string is usually used in, and its function in discourse. There are also criteria which consider a sequence formulaic if it is marked grammatically or lexically by the speaker or applied with an action, use of punctuation, or some particular phonological pattern to give this FS special status as a unit. Additionally, this checklist considers a sequence formulaic even if it is applied inappropriately by the speaker, if it is too sophisticated linguistically for the speaker’s proficiency level, and regardless whether the speaker heard the word combination from other people before or invented it on his/her own (Wray & Namba, 2003).

The researcher transcribed and analyzed 20 presentation recordings. Every recording was analyzed by using a separate table (Table 1) that allowed for documentation of every FS found in a given student’s presentation. As mentioned earlier, the independent judges analyzed only 30% of the data which means that they used the table template to assess 6 presentations of 3 students (2 presentations per student). For every individual FS, the researcher and the judges needed to identify the number of times they felt it was used in a student’s presentation, the extent to which they agreed with every criterion in Wray and Namba’s (2003) checklist, and to provide comments to explain their assessments. Table 1 below provides an example of how the analysis of every FS in the students’ presentations looked like. For example, for the FS “in fact”, in the column “how many times it was used” the rater indicated that this FS was mentioned once in the participant’s presentation. The cell “criteria from the checklist” refers to the eleven criteria from Wray and Namba’s (2003) checklist which are listed in the table to the left of this cell. Below each criterion from the checklist, the rater assigned the value from 1 to 5 from Likert scale to show the extent of their agreement/disagreement with it. For example, the rater agreed (rating 4) with the first, the third, and the fourth criterion from the checklist; strongly agreed (rating 5) with the second, the sixth, and the eighth criterion; strongly disagreed (rating 1) with the fifth, the ninth, the tenth, and the eleventh criterion; and finally, neither agreed nor disagreed (rating 3) with the seventh criterion. In the column “comments”, they provided additional information about their judgment of this particular FS.

Table 1. *Formulaic language analysis example*

Formulaic sequence	How many times it was used	1	2	3	4	5	6	7	8	9	10	11	The criteria from the checklist	Comments

In fact	1	4	5	4	4	1	5	3	5	1	1	1	The numbers from Likert scale from 1 to 5 1-strongly disagree 2-disagree 3-neither agree nor disagree 4-agree 5-strongly agree	This phrase is used to connect two points, or to build upon a previous point. It is something that is used in academia.
---------	---	---	---	---	---	---	---	---	---	---	---	---	---	---

3.5.1.1. Formulaic language – Assessment of Inter-rater reliability. The use of FS can lead to near-nativelike turn of phrase and fluent “error-free” speech beyond L2 competence (Wood, 2015, p. 81) The phrases sound more nativelike when they are used in the same way as a native speaker would use them and are appropriate in a given context (Wood, 2015). Native speaker intuition usually assists in determining whether a particular FS is a well-known and widely used. The native speaker judgment is commonly used to identify FS in case studies with a spoken corpus (Wray & Namba, 2003). For the purpose of the present study, two native speaker judges participated in the analyses of formulaicity. The independent judges were Canadian English L1 speakers (mean age: 28) enrolled in an Applied Linguistics graduate program. At the time of the study, they had had on average 5 years of ESL teaching experience. The judges were chosen because of their L2 teaching experience, availability, and general interest in the study. Even though the judges reported having had some knowledge about formulaic language, they were trained to ensure that they understood the task and goals of the study. Moreover, one of the concerns about native speaker judgment described by Wray (2002) is that there may be a variation among judges. In order to avoid such variation, a benchmarking identification session was held, where, through

discussion, the researcher made sure that the judges were familiarized with the nuances of identifying FS. The benchmarking session lasted two hours and consisted of three steps:

- to understand the goals of the present study and to sign the confidentiality agreement (see Appendix D);
- to become familiar with the basics of formulaic language;
- to practice performing formulaic language analysis based on Wray and Namba's checklist (2003).

The judges were first presented with a general idea of the study but, in order to minimize bias in their assessments, they were not provided with the details on the participants, the number of presentations they had completed, and in what order. The judges were given time to ask any questions about the study before offering their participation. Once they agreed to participate, the judges signed the confidentiality agreement, which required them not to download or share the provided documents and to respect the confidentiality of the research participants (see Appendix D).

A general introduction to the current research on formulaic language was provided, followed by examples of FS (Step 2). To make sure that the raters understood what is meant by FS, they were asked to give their own examples of FS and explain their choices. In order to train judges in the formulaic language analysis based on the given criteria, two recordings of one student were randomly selected for the benchmarking session and examined together with the researcher (Step 3).

After the benchmarking session, the judges were each given a password protected USB key with four randomly selected recordings and transcripts of two versions of the oral presentations done by two students. The judges were asked to independently complete the analysis, using the

Wray and Namba's checklist (2003) to identify FS. As shown in Table 1, for every FS they found they filled out a Likert scale, indicating to what extent they agreed with the certain criterion from the checklist. They were also asked to provide additional comments to justify why they considered certain sequences formulaic. A sequence was judged as formulaic when it was identified as such by the researcher and at least one judge.

Taking into account the benchmarking session, the two native speaker judges assisted in the formulaic language analysis of 6 recordings (3 students). After completing the analysis, the researcher informed the judges that every student had two versions of the presentation. Interestingly, the judges confessed that they had not realized they heard the same student twice.

During the benchmarking session, the inter-rater reliability was at 100%, constituting an ideal level (Mackey & Gass, 2015). This agreement was achieved as a result of a round-table discussion among all the raters. However, the level of agreement between the researcher and judges changed to 58% when the results of the remaining analysis were compared. One of the factors which could explain the result was the lack of opportunity to have a follow-up session after the completion of the analysis (the judges were available to meet only once for the training). It can be assumed that a follow-up meeting between the judges and the researcher could have provided a chance for a discussion of the differences in the analyses. The fact that the judges were not experts on formulaicity and relied on their intuition for judgment may be another reason for the low inter-rater reliability value. A close investigation of the judges' ratings has revealed several inconsistencies. For example, the judges forgot to mention certain sequences in their analysis, even though they had identified them as such earlier; at times, they would identify one sequence as formulaic but then miss another despite the two being similar structure and function.

As for the intra-rater reliability, re-analyzing the data after four months allowed the researcher to recognize more language as formulaic. After all, formulaic language analysis is time-consuming, requires attention and high concentration, and the “absolute certainty is elusive” (Wood, 2015, p. 32). This may explain why there is no perfect consistency in the judgments collected at different times (e.g., during second attempt of the analysis the researcher found some new FS which had not noticed in the first judgment).

3.5.1.2. FS classification. The functional classification of the lexical bundles conducted by Biber et al. (2004) in university classroom teaching and textbooks was chosen for organizing all FS used by the participants in this study. The FS were classified according to their discourse functions to determine which functions were the most common and whether they contributed to the resulting fluency ratings. It is important to note that FS can be studied under different names and classified under different categories. One such category is “lexical bundles” (Biber et al., 2004), which are FS that can frequently appear in text corpora and serve particular functions in discourse (for example, *I mean you know, is based on the*, Wood, 2015). Since Biber et al. (2004) were among the first to contribute to the analysis and classification of lexical bundles, their functional taxonomy was chosen for the formulaic language analysis in this study; the taxonomy is widely used by SLA researchers. It is based on a comprehensive corpus-based analysis of lexical bundles that occur in university discourse and include four discourse functions:

1. Stance expressions: *If you want to, I was going to.*
2. Discourse organizers: *on the other hand, if you look at.*
3. Referential expressions: *there’s a lot of, in the case of.*
4. Special conversational functions: *thank you very much, what are you doing.*

Considering the fact that the participants in the current study were allowed to make their presentations less formal and academic in format, a pragmatic sub-category “filler” was added to the “special conversational function” to account for the FS that the students relied on in their discourse to fill pauses and maintain speech flow. Since lexical bundles are not idiomatic by nature, a separate category was added to the classification – figurative function – which included phrasal verbs (e.g., *blend in, rely on*) and collocations (e.g., *take for granted*). Additionally, the “Other” category was added for those FS that either did not fit the categories mentioned above or did not serve an obvious function. Hence, there was a total of six categories used to classify the FS in the participants’ presentations. These included: stance expressions; discourse organizers; referential expressions; special conversation functions; figurative function; other.

3.5.2. Fluency judgment

Fluency is usually defined as a “nativelike” ability in the L2 and as a function of “temporal variables of speech” such as: (1) rate of speech, (2) length of runs, and (3) pauses/hesitations (Wood, 2015, p. 86). McGuire and Larson-Hall (2017) have operationalized speech rate (SR) as the number of audible syllables uttered per minute and mean length of runs (MLR) as the number of syllables spoken between pauses. The terms “pauses” and “hesitations” are generally used interchangeably and “are measured by the proportion of the total speaking time spent in pauses or silence” (Wood, 2015, p. 87). For the purposes of this study, Segalowitz’s (2010, p. 165) definition of “perceived fluency” where the listener’s perceptions of cognitive fluency is based on the speaker’s utterance fluency was chosen, and holistic scores were assigned by human raters to provide a perception of fluency of a given participant. Therefore, in the present study, temporal variables of speech were not calculated, but, instead, a qualitative approach was used to measure subjective perception of speech fluency. Some studies which examined whether human holistic

fluency scores matched analytic fluency scores reported that the human judgment coincides with the results of the temporal measures of L2 fluency (Nakano, Kondo & Tsutsui, 2008; Xi & Mollaun, 2006).

A Likert scale with four different fluency criteria described by Segalowitz (2010) - speech rate, the comprehensibility of the speech, pauses and hesitations, and the general language proficiency – was developed for the purposes of this study. The fluency scale contained these criteria, in particular, because they usually help to express the impression of someone's fluency. Fluency has different dimensions – for example, a learner can speak very fast but sound indistinctively and cause a misunderstanding. In contrast, someone may speak at a normal rate but use a lot of fillers and repetitions, which could affect the whole impression about their fluency. This shows that the speed, for instance, is not the predominant indicator of someone's fluency and that normal rate of fluency does not guarantee fewer hesitations or mistakes.

The fluency judgment used here also includes a question about general proficiency. According to Lennon (1990), fluency can be construed in a “broad sense” as an equivalent to the overall proficiency, and in a “narrow sense”, as a part of language proficiency, including the smoothness and flow of speech (as cited in Bosker, Pinget, Quene, Sanders & de Jong, 2013, p. 160). It is commonly assumed that when a speaker has fewer mistakes, they may appear more fluent to the listener (Bosker et al., 2013). However, if a speaker does not make mistakes but has a lot of hesitations in their speech, it might also distract the listener, and the speaker might be perceived as less fluent in their L2. Therefore, having a fluency scale that accounts for the nuances of fluency in both broad and narrow senses was deemed important to ensure a deep and thorough analysis.

A template (Table 2) was designed to document the information required for the fluency judgment of the 20 recordings. As previously mentioned, the independent judges rated 50% of the data, which means that they used the template to assess ten presentations of five students (2 presentations per student). As Table 2 shows, the fluency scale included four different criteria, for which the researcher and independent judges provided their perceptions using the Likert scale, and a section for additional comments that allowed the raters to summarize their assessments and provide additional and unique information regarding a particular presenter’s fluency. Using example in Table 2, a rater judged this participant’s speech rate as appropriate (rating 3), agreed that his/her speech was comprehensible (rating 4), did not have a lot of pauses and hesitations (rating 4), and agreed that the language was at advanced proficiency (rating 4); the rater also provided a detailed rationale for their assessment in the “comments” section.

Table 2. *Fluency scale and illustrative example*

Statements	Likert Scale				
1. The speech rate (speed) of the speaker was appropriate and aided understanding.	1 Disagree	2 Neither agree nor disagree	3 Agree		
2. The speech of the student was smooth, comprehensible and the words were distinctive.	1 Strongly Disagree	2 Disagree	3 Neither Agree Nor Disagree	4 Agree	5 Strongly Agree
3. The student spoke without pauses, repetitions, hesitations, and fillers during the presentation.	1 Strongly Disagree	2 Disagree	3 Neither Agree Nor Disagree	4 Agree	5 Strongly Agree
4. The language of the speaker (grammar and the choice of the vocabulary) was at the advanced level of proficiency.	1 Strongly Disagree	2 Disagree	3 Neither Agree Nor Disagree	4 Agree	5 Strongly Agree

<p>5. Please, briefly explain your ratings above providing as much detail as possible.</p>	<p>The speaker's speed was quite normal. Sometimes, he was hesitant and had some repetitions but they did not distract. He did not have any fillers. He started speaking in the middle of the presentation much faster and more confidently. Overall, the student sounded very confident about his content and learnt his presentation quite well. The language was quite advanced and he used a lot of academic words. He used transition words like obviously, additionally. His presentation was logically organized and was easy to follow. He had few mistakes and errors.</p>
--	---

The ratings of each criterion for each presentation were compiled and comments summarized. The ratings of each criterion in the first version were compared with the ratings of that criterion in the second version to determine whether the participants were perceived as more fluent (e.g., whether their speed became faster or speech more distinctive in the second version based on the given ratings). The fluency ratings were compiled and compared for each participant and the group as a whole.

3.5.2.1. Fluency – Assessment of Inter-rater reliability. An assessment of learners' oral proficiency and fluency cannot be fully objective, therefore after completing the judgment on her own, the researcher held a benchmarking session with three independent judges in order to get their assessment of the data; this is a common practice in L2 fluency research. If different judges show similar appreciation of a participant's performance, this could indicate that the participant's oral proficiency is perceived more or less similarly by the different experts (Boers et al., 2006). Three independent judges (different from the ones involved in the FS analysis) were engaged to complete the fluency judgment. They were Master students (mean age: 33) with extensive ESL teaching experience. Two of them were advanced speakers of English from Mexico and Turkey and had taught English for more than 5 years; the remaining judge was a native speaker of English from Canada and had less than 5 years of teaching experience. They were chosen due to their

teaching experience, availability, and interest in the current study. Even though, as teachers, the judges were familiar with measuring L2 fluency, the researcher held a training session. The two-hour session consisted of three steps:

- to familiarize them with the present study and to sign the confidentiality agreement (see Appendix D);
- to introduce existing research on fluency measurement;
- to practice fluency judgment together with the researcher, using the checklist developed for the study (Table 2).

First, the judges were informed about the goals of the current research but did not get any specific information about the study design, the number of participants nor their oral presentation attempts. The judges were given time to ask any questions they had about the study. Once they agreed to participate, the judges signed the confidentiality agreement described above (see Appendix D). Then, the researcher familiarized the judges with the current research on fluency measurement including temporal variables of speech, and explained why the perceived fluency by human raters was chosen for this study (Step 2).

In order to practice conducting the fluency judgment (Step 3), the judges listened to two recordings done by a randomly chosen student. After listening to each recording, they assessed each criterion and wrote short comments about how fluent they had perceived the participant to be. Upon completing the assessment of each recording, the judges and the researcher had a discussion about their fluency ratings. After completing the benchmarking session, the judges were each given a password protected USB drive with the recordings of two presentations of four students randomly assigned to the judges (i.e., 8 recordings per judge).

Using Table 2, they were asked to complete the rest of the analysis independently by filling out the fluency scale and providing comments about their judgments. In total (including the benchmarking session), the judges assisted in the fluency judgment of 10 recordings (5 students). Similarly to the FS judges, several of the fluency judges did not realize that the recordings they had listened to belonged to the same 4 students, not to 8 different ones.

The inter-rater reliability at the benchmarking session was ideal being at 100%, and was achieved due to the round-table discussion of the perceived fluency judgments (Mackey & Gass, 2015). However, the inter-rater reliability agreement for the remaining data changed (42%). The first reason may lie in the fact that, similarly to the formulaic language judges, the fluency raters analyzed the portion of the data at home and there was no opportunity for a follow-up session or discussion. Second, perceived fluency is a highly subjective measure, which makes it hard to pin down the factors the raters respond to when evaluating. Analysis of the comments provided in the fluency scale, on the other hand, supplied more insight into the judges' perception of fluency, revealing similarities with the researcher's assessments most of the time. Furthermore, having English as a foreign language (EFL) teachers as fluency raters could have influenced their judgments since they could have based their assessment on their previous teaching contexts, teacher education, and their own L2 experiences. Finally, it is important to note that the fluency judges worked with the audio files only. The researcher intentionally did not provide transcripts to ensure that they based their perception of speech fluency on listening to the recording, judging the speech rate, the comprehensibility of the speech, number of pauses and the overall language proficiency without letting transcripts affect their aural perception. However, this could have influenced the resulting ratings since all judges reported not understanding some of the participants' pronunciation.

As for the intra-rater reliability, reanalyzing the data made the researcher change some of the fluency ratings. The reason for that might have been the fact that after listening to the recordings several times, the researcher started to understand the participants better, judging them as more fluent as a result.

3.5.3. Combining formulaic language analysis and fluency judgment

As a final step in the data analysis, in order to determine whether task repetition impacted the learners' FS use and fluency, the results from the FS analysis and fluency judgment for version 1 and version 2 of each presentation were compiled and compared per student. The oral presentations of those participants who improved the most were analyzed further and described in detail. The findings of the study are presented in the next chapter.

Chapter 4: Results

4.1. Overview

This study aimed to answer the following research questions:

7. What are the common formulaic sequences (FS) used by Chinese L1 EAP students and what functions did they serve in the students' presentations?
8. What are the fluency profiles of Chinese L1 EAP students in their presentations?
9. Does task repetition bring on FS and fluency gains (independently and/or together) in the presentations of Chinese L1 EAP students?

The formulaic language analysis and fluency judgment were conducted in order to evaluate the impact of task repetition on the development of formulaic language and perceived fluency gains. To answer the first research question, the participants' oral presentations were evaluated to (1) identify FS, using the Wray and Namba's checklist (2003), and (2) to classify the functions of these FS in accordance with the functional taxonomy of lexical bundles of Biber et al. (2004). The objectivity of the results was assured via an inter-rater reliability measure. To determine the participants' fluency profiles (Research Question 2), the participants' oral presentations were assessed in terms of (1) speech rate, (2) speech comprehensibility, (3) number of pauses and hesitations, and (4) general language proficiency, using a fluency scale based on Segalowitz's (2010) identification of perceived fluency. The objectivity of the fluency ratings was also determined with the help of independent raters. Finally, to determine the effect of task repetition on the participants' FS and fluency profiles (Research Question 3), the results of the formulaic language analysis and the fluency ratings were compared for all participants. In what follows, the results of the analyses undertaken are described.

4.2. Research Question 1

In order to determine the kinds of FS the Chinese L1 EAP students used and the functions those sequences served in their presentations, the checklist created by Wray and Namba (2003) was applied to identify FS (i.e., sequences which seem to be prefabricated and retrieved from memory as wholes) in both versions of each presentation. Similarly, the same checklist was used by the independent judges to verify the researcher's assessment of 30% of the data. Afterwards, the number of tokens (the total count of FS) and the variety of FS types (the qualitative characteristic indicating the extent to which the learner uses different FS types without repetition) in the two versions of each student's presentations were compared to track any changes and signs of development in the use of formulaic language (Qi & Ding, 2011). Finally, all the identified FS were assigned a function which they presumably served in the discourse. The functions were identified according to the functional taxonomy of lexical bundles developed by Biber et al. (2004). The results of the formulaic language analysis are detailed first, followed by the FS functional classification.

4.2.1. Formulaic language analysis

Identification of FS. Formulaic language is usually referred to as “formulaic sequences” (FS), which are identified, according to Wray's (2002) as: “a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated: that is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar” (p. 9). This definition was chosen in the present study to identify FS because it accounts for “individual variation” between speakers, implying that the notion of FS can be rather vague and that what might seem formulaic to one person might not necessarily be such for another (Stengers et al., 2010, p. 323). In order to account for a

rather complicated nature of formulaic language, criteria-based checklists, which rely on the decisions of judges about formulaicity, are commonly used with spoken language samples (Wood, 2015). As opposed to the frequency-based approach of identifying formulaic language in corpora, which requires a researcher to set such specifications as the “minimum length of word combinations and minimum frequency cutoffs” (Wood, 2015, p. 20) before scanning the corpus for FS, criteria-based checklists take into consideration all possible features usually associated with FS. Moreover, Wray (2002) suggests that computer corpus analysis software can suffer from such limitations as frequency cutoffs and judgments of non-formulaicity, which can be resolved by employing a criteria-based checklist. For example, in order to identify a certain sequence as formulaic, frequency cutoffs should be determined to see how many occurrences the FS has in a corpus. However, according to Wray (2002), some specific FS might have a relatively small number of occurrences in the corpus and might be used only once in certain situations, but those factors do not make them less formulaic; this implies that frequency is not the only prominent factor to be used when it comes to the identification of FS. Additionally, in the studies involving L2 speakers, a lot of FS might be seen as idiosyncratic or non-native like, i.e., not typical or standard, but under a detailed examination they are shown to have a unitary meaning or function, and thus, could also be considered formulaic (Wood, 2015).

Considering the small sample employed in the current study (n=10), human judgment was used to determine what was formulaic and what was not. To identify all the possible FS in the participants’ presentations, the Wray and Namba’s (2003) comprehensive checklist that contains 11 diagnostic criteria which describe possible features of formulaicity was used (see Appendix C for all the criteria used in the checklist). The criteria considered whether a FS was:

- grammatically irregular (i.e., *hang out*);

- semantically transparent (i.e., *meanwhile*);
- associated with a specific register (i.e., *how's it going, guys?*);
- used to perform a function in discourse or communication (i.e., *in conclusion*);
- used by the speaker when conveying the subsequent idea (i.e., *and so on*);
- phonologically coherent (i.e., *it is sooo good*);
- was marked grammatically or lexically in a way that gives it a special status as a unit (i.e., *on the other hand*);
- previously encountered (i.e., *you guys*);
- novel and used as a derivation, deliberate or otherwise (i.e., *as a last*);
- applied inappropriately (i.e., *a lot number of*);
- too sophisticated for the speaker (i.e., *come about with*);

It was decided not to have a specific requirement of applying all eleven criteria to the sequence to consider it formulaic. Instead, a sequence was marked as formulaic if it matched at least five different criteria from the checklist. A similar approach has been employed by other scholars in the field. For example, Wood (2010) used the judgment criteria only as a guide to determine formulaicity in his L2 speech corpus, choosing not to prioritize the importance of a particular criterion above the others. Moreover, certain FS were determined as formulaic even if their use was not target-like. This was the case for *traced on (early 19 centuries)*, which was considered to be a FS even when the *traced back (to the early 19th century)* form was used. Also, a sequence was determined formulaic if it sounded nonstandard but had a unitary meaning and served a particular function in the discourse, and/or had a particular way of being stored and retrieved from memory (Wood, 2015). For instance, *as a last* was considered FS in the present study even though the word *lastly* is more appropriate to express the intended meaning in

English. More examples of the FS which were applied with some errors in the present study include: *I'm gonna say about, as a last, a lot number of, how's going, thank you for my attention/thanks for it attention, traced on to, close the traffic, just to like the grandma make it, the first, out of my expectations.*

The participants also had problems balancing between certain FS which express formality as they did not always apply them in a pragmatically-appropriate manner. That is, they overused such FS as *you guys, hi guys, which may be seen as* that could be seen as too informal for an in-class presentation; at the same time, they used such highly formal FS as *Good afternoon, ladies and gentlemen, I'm very happy to be here.*

Some of the FS used by the participants could have originated from the sources they used while preparing their presentations and decided to borrow. This possibility is accounted for in Wray and Namba's (2003) checklist when they deem a sequence formulaic if it could have been encountered before, from other people and/or sources. Additionally, a lot of FS used by the participants sounded rather informal (e.g., *you guys, pretty much, kind of, how's it going*) and could have originated from other speakers' discourse, especially that of native speakers.

The results show that, as a group, the participants used 145 FS types in their oral presentations (see Appendix E for the complete list), including: *this picture shows, at first, in the same way, thanks for listening, to be honest, blend in, look like, etc.* Six of the ten participants, in particular, showed positive changes and signs of perceived improvement in the number and variety of FS they used in their second version of the presentation compared to the first. The remaining four participants, however, did not appear to demonstrate change in the use of FS. These findings are detailed further in the sections that follow.

The number, variety and repetition of FS. Two versions of each participant's presentation were compared to determine change in the use of the formulaic language. Table 3 shows the totals of (1) tokens; (2) FS types per version with repetitions of the same FS; (3) FS types transferred from Version 1 to Version 2; and (4) a total of newly occurring FS types used by the participants in Version 2. In the table, "S" stands for Student, and "V" – for Version, respectively.

Table 3. *FS group results*

	S1		S 2		S 3		S 4		S 5		S 6		S 7		S 8		S 9		S 10	
	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V	V
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
Total number of tokens without repetitions	23	27	19	25	14	15	5	4	10	12	5	10	4	5	19	22	4	11	10	15
Total number of FS types with repetitions	35	49	20	30	15	16	7	6	10	13	5	11	4	5	21	23	4	11	11	16
Total number of FS types from V1 in V2		12		13		12		2		9		3		2		11		3		6
Total number of new FS types in V2		15		12		3		2		3		7		3		10		8		9

From the numerical data, it can be seen that there is an overall upward trend in the token totals produced by the group in Version 2; the exception is Student 4 who showed a slight decrease in the total. As for the repetitions of the same FS per version, there were few, with the exception of Student 5, Student 7, and Student 9 who did not have them at all. In terms of the variety of FS in both versions, all participants used different FS. Six out of the ten participants

had a larger number of newly occurring FS in Version 2 compared to those transferred from Version 1 (e.g., Student 1, Student 2, Student 6, Student 8, Student 9, and Student 10).

Student 1, for example, employed the highest number of tokens (27) in Version 2 compared to Version 1 (23). She also had the most FS repetitions in both versions – 35 FS and 49 FS, respectively. As for the variety, she transferred 12 FS types from Version 1 to Version 2, and had 15 newly occurring FS types in Version 2, which was the highest number in the group.

Table 4 details the percentage of increase in the number of tokens per student and group, as well as the percentage of the newly occurring FS types in Version 2 in the participants' presentations.

Table 4. *The percentage of the increased total and newly occurring FS types*

Participants	Increased total in V2	Newly occurring FS types in V2
Student 1	17%	56%
Student 2	38%	52%
Student 3	7%	20%
Student 4	-20%	50%
Student 5	20%	25%
Student 6	100%	70%
Student 7	25%	60%
Student 8	16%	45%
Student 9	175%	73%
Student 10	50%	60%
Total per group	29%	49%

As Table 4 shows, overall, the participants had a positive change in the use of formulaic language in Version 2, having increased their token total by 29%, with Student 9 showing the

best result (i.e., increase of 175%). As a group, the participants used 49% of new FS in Version 2. Student 9, once again, used the highest percentage of new FS types in Version 2 (i.e., 73%).

It may also be suggested that the percentage of the increased total does not always represent quality. As shown in Table 4, Student 1, who appeared to use the highest number of tokens and newly occurring FS types in her Version 2, did not demonstrate the highest increase. This is explained by her overall performance on Version 1 which was ranked high from the start, resulting in a less impressive percentage difference. On the other hand, Student 9, whose formulaic language use was weaker in Version 1, had the highest percentage increase because of the great difference between the versions in terms of the number and types of FS used (i.e., 4 tokens in V1 and 11 tokens in V2; 8 newly occurring FS types out of 11).

4.2.2. Analysis of the FS functions

Despite the fact that the participants used more FS and more new sequences in Version 2, this did not seem to be enough to support any changes or progress in the use of formulaic language, which might have been due to the task repetition. Therefore, a more detailed analysis of the nature of those FS as well as the functions they served was required.

The current study adapted the functional taxonomy advanced by Biber et al. (2004), who grouped lexical bundles under four different functions based on their use:

- *Stance expressions* which “express attitudes or assessments of certainty that frame some other proposition”.
- *Discourse organizers* help to express “relationships between prior and coming discourse”.
- *Referential expressions* “make direct reference to physical or abstract entities”.

- *Special conversational functions* which express formality, simple inquiring, or reporting.

Since it was decided to account for all FS (not only lexical bundles) in the participants’ oral presentations, two new categories of functions and one sub-category were added to the taxonomy:

- *Filler function* was added as a subcategory to the “special conversational functions” to account for those FS that served as “time byers” and “compensatory devices for memory limitations” (Wray & Perkins, 2000, p. 16);
- *Figurative function* was added to include all FS of idiomatic nature, like phrasal verbs, idioms, some collocations, and compounds;
- *Other function* was added to group all FS whose function was either not obvious or did not fit into the existing categories.

Hence, the functional taxonomy used here contained a total of six functions, which included: (1) stance expressions, (2) discourse organizers, (3) referential expressions, (4) special conversational functions, (5) figurative function, and (6) other functions. The results of the function analysis are provided in Table 5, which gives the group total count for a particular function classification and provides examples to illustrate it; subcategories for each function are also detailed. The italicized words in the examples point to errors made by the participants while applying a particular FS.

Table 5. *FS functions*

Functions	Number per whole group	Formulaic sequences
1. Stance expressions		
Epistemic stance	6	I believe, I hope
Attitudinal/modality stance		we have to, you have better go, you have to try it, we can use two ways

2. Discourse organizers	
Topic introduction/focus	37 I'm gonna talk about (<i>say</i>), I'm going to introduce, I'm gonna be doing my presentation, we can imagine, we can see, this image shows, this picture shows, do you know, as my suggestion, I'm really interested in, just wanted to, I just want to
Topic elaboration/clarification	What does this mean, on the other hand, for example, according to, in conclusion, at first, lastly (<i>as a last</i>), and so on, even though, not only because but, in disadvantage, not only but, meanwhile, in addition, to summarize, on average, due to, because of, not to mention, in fact, making it easy to, the main reason for this is, so generally, that's not a big surprise why, besides that
3. Referential expressions	
Identification/focus	This is, that is, this kind of, in charge of (<i>in charges of</i>), so it is, the best in the world, that's all, most of
Imprecision	
Specification of attributes	26 Lots of, a little bit, a lot, there's/ are, there is a big number of (<i>a lot number of</i>), in a completely new way, in the same way, most of, at the same time, a large number of
Time/place/text reference	In the middle of, at the beginning, in past few decades, in recent decades, per month, per day, around the world, in the world
4. Special conversational functions	
Politeness	23 Good evening, hi guys, how's it going (<i>how's going</i>), thank you, guys, thank you very much, thank you for the attention (<i>for my attention/ thanks for it attention</i>), thanks, thanks for listening, I'm sorry, thank you for watching, good afternoon ladies and gentlemen, hello everybody, hi everyone, hello everyone, I'm very happy to be here
Simple inquiry	Hope you, guys, had a nice weekend
Reporting Filler	You guys, (like) you know, okay so, kind of, it's (just) like, and actually, to be honest, pretty much
5. Figurative function: phrasal verbs, idioms, collocations, compounds	38 Based on, traced back to (<i>traced on</i>), go through, give up, paid off, go down, based down, look for, step off, get to know, hang out, blend in, come from, fold down, pick up, pull out, rely on, go down, sign out, look though, come about with, stand up, to step foot on, take for granted, met difficulties, met challenges, junk food, good luck, daily lives, big issue, downtown, first choice, my personal choice, so/really yummy, high tech things, deeply interested in, long term, closed due to traffic (<i>close the traffic</i>)
6. Other	15 It's so good, personally my favourite, just like the grandma makes (make) it, out of my expectation (expectations), more and more, look like, feel like, not

their first choice, no matter, out there, on their own, still around, at all, one by one, of course

The results of the FS functional classification indicate that the figurative and discourse organizer functions were the most used, accounting for 38 and 37 FS, respectively. There were 26 FS that functioned as referential expressions and 23 FS that represented the special conversational function. The participants used 16 FS grouped under the “other” function. Finally, the stance expressions (6 FS) appeared to be the least used among this group of learners. For a detailed summary of all the functions served by the identified FS in the participants’ presentations, please refer to Appendix F.

Figure 1 provides the percentage totals for each function used by the participants. The percentages were generated by multiplying the totals in Table 5 by 100 and dividing the result by 145, which is the grand total of all FS used by this population sample.

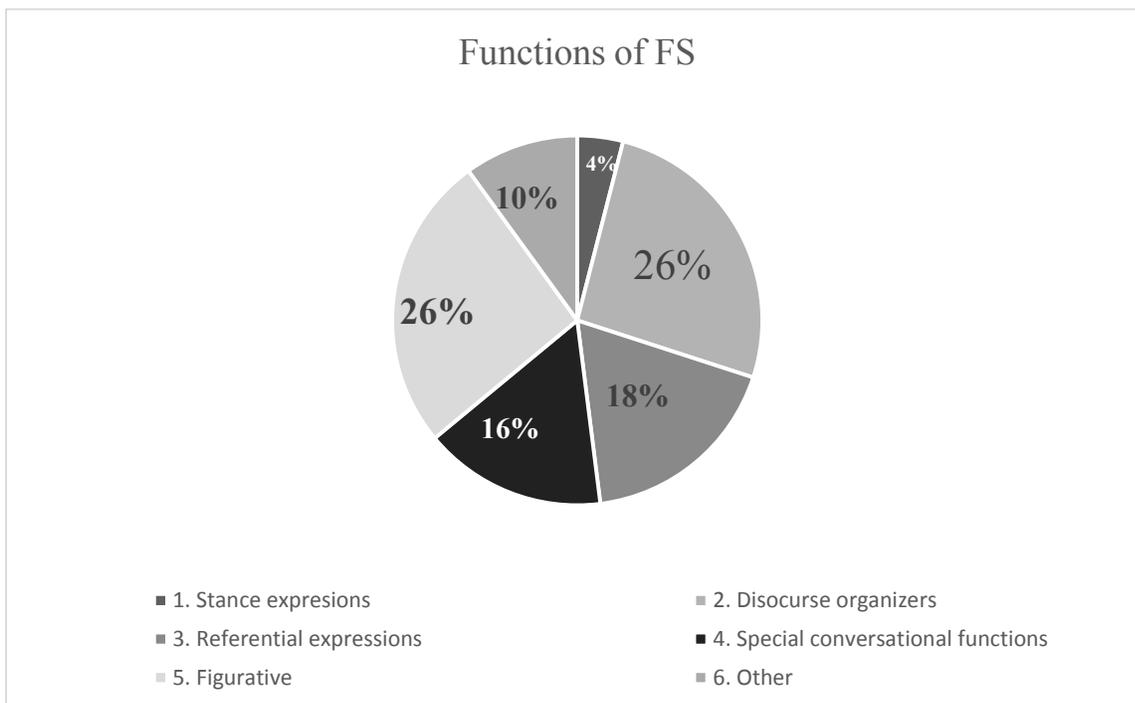


Figure 1. The percentage of the functional taxonomy of FS

Repeated FS and their functions. Some participants, as was mentioned earlier, used certain FS several times in their presentations. Looking into the nature of those FS repetitions was necessary to analyze whether they occurred randomly or if there was a reason for their use in the discourse. Table 6 illustrates the kinds of FS that were repeated by the participants and the functions these served. Only the results of those participants who repeated FS in their presentations (n=7) are displayed in the table.

Table 6. *Repeated FS and their functions in participants' presentations*

FS functions	Student 1		Student 2		Student 3		Student 4		Student 6		Student 8		Student 10	
	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2
Stance expressions														
Discourse organizers	According to (2)	For example (2)		And so on (2);										
Referential expressions			Kind of (2)	Most of (3); Kind of (2)			There are (2); A little bit (2)	There are (3)				Around the world (4)	There is/are (2)	There is/are (2)
Special conversation function	Okay so(2); it's (just) like (4); you guys (5)	You guys (5); (like) you know (14)												
Figurative function	So really yum my (3); based on (2)	So/ really yum my (2); Junk food (2)		Rely on (3)	Give up (2)	Give up (2)				Good luck (2)	Give up (2); Go down (3); Big issue (2)	Give up (2); Big issue (2)		
Other														

As Table 6 shows, the most repeated function was the figurative function, which included such FS as *give up* and *so really yummy*. While Student 3 and Student 8 appeared to particularly favor the phrasal verb *give up*, *there is/there are* was the most popular referential expression

used in the presentations of Student 4 and Student 10. The “Other” and “Stance expressions” functions were not repeated at all. Student 1, in particular, had the largest number of repetitions, with the FS serving three different functions (i.e., the discourse organizers, special conversation, and figurative functions). For example, she used FS *you guys* 5 times and *(like) you know* 14 times, which may suggest her overreliance on those FS to fill pauses, avoid hesitations, and/or keep going with the presentation.

4.2.3. Summary of the results for research question 1

The formulaic language analysis shows that, as a group, the participants used 145 FS types in their oral presentations. The results indicated that the number of FS types was bigger in Version 2 compared to the first attempt. Overall, the participants’ use of FS varied in terms of the total number and variety. All participants also used newly occurring FS types in Version 2. The FS the participants used served different discourse functions, but the figurative and discourse organizer functions dominated. The FS that the participants repeated in their presentations mainly served the figurative, referential, special conversational and discourses organizer functions.

4.3. Research Question 2

To answer the second research question (What are the fluency profiles of Chinese L1 EAP students in their presentations?), perceived fluency judgment was performed. The participants’ speech was evaluated, using the perceived fluency scale developed for the study and described in Methodology. Judgments for both versions of the presentation for each student were compiled and compared to determine whether the quality of the participants’ speech changed and if so, then how.

4.3.1. Findings from the fluency judgment

Among the multiple characterizations of fluency, Segalowitz's (2010) definition of perceived fluency was chosen for the current study. He proposes a distinction among the cognitive, utterance, and perceived fluency. While "cognitive fluency" refers to the ability of the speaker to translate thoughts to speech, "utterance fluency" corresponds to utterance features affected by cognitive fluency; "perceived fluency", in turn, is a listener's perception of a speaker's cognitive fluency based on the speaker's utterance fluency (Segalowitz, 2010). The fluency scale used to assess fluency contained the following statements and corresponding Likert-scale values (see Table 2 in Methodology for the complete tool):

1. The speech rate (speed) of the speaker was appropriate and aided understanding;
2. The speech of the student was smooth, comprehensible, and the words were distinctive;
3. The student spoke without pauses, repetitions, hesitations, and fillers during the presentation;
4. The language of the speaker (grammar and the choice of vocabulary) was at the advanced level of proficiency.

Table 7 provides a comparison of the fluency scores for both versions of each participant's presentation. The Likert-scale fluency ratings ranged from 1 to 5, where 1 indicated strong disagreement and 5 – strong agreement. It is important to note, however, that the first statement on speech rate contained only three ratings of: 1 – disagree; 2 – neither agree nor disagree; and 3 – agree. Anything above "3" was considered a high score. The rating of "3" (i.e., "neither agree nor disagree") for the rest of the statements was interpreted positively as well because it showed some change or shift in the fluency judgment even though the rater was not completely sure

whether to agree or disagree with a particular statement. “S” in the table stands for “Student” and “V” for “Version”.

Table 7. *Fluency ratings*

Fluency scale	S1		S2		S3		S4		S5		S6		S7		S8		S9		S10	
	V1	V2	V1	V2																
1. The speech rate	3	3	2	3	2	3	3	3	3	2	2	3	3	3	3	3	2	3	3	2
2. The comprehensibility and smoothness of the speech	3	4	2	4	2	4	3	4	4	4	2	2	2	3	4	5	2	3	3	4
3. The use of pauses, repetitions, and hesitations	3	3	2	3	2	4	2	4	4	4	1	4	3	3	4	5	2	4	4	4
4. Language proficiency	2	3	3	3	3	4	3	4	4	4	1	2	2	3	3	3	2	3	3	3

Overall, all participants were assessed as having higher fluency ratings for Version 2, revealing some perceived improvement compared to their scores on Version 1. Student 5 and Student 8, in particular, did not show big changes because their speech was perceived as good in Version 1 already, and therefore their fluency ratings remained either unchanged or moved slightly from “4” to “5” in Version 2.

The highest score of “5” was given to Student 8 on Version 2, which appeared to be the best result in the group. Although he initially received the lowest rating of “1”, Student 6 managed to increase his score to “4”, meaning that he had fewer hesitations and pauses.

There were 40 pairs of fluency ratings per group (10 participants were multiplied by 4 fluency statements). Each pair was analyzed and the ratings within it were compared between each other to determine whether the numbers were rising, remaining unchanged, or declining. Thus, there were found two pairs where the ratings were declining, for example, Student 5 and Student 10 in the statement about speech rate; 14 pairs did not indicate any change in scores;

and, finally, the ratings in 24 pairs were rising. Then, the percentage was calculated by dividing the number of pairs indicating rise, no change or decline (i.e., 2; 24; 14) by the total of pairs (i.e., 40) and multiplying the results by 100. Figure 2 shows the percentage of perceived improvement in the fluency ratings of the participants' presentations.

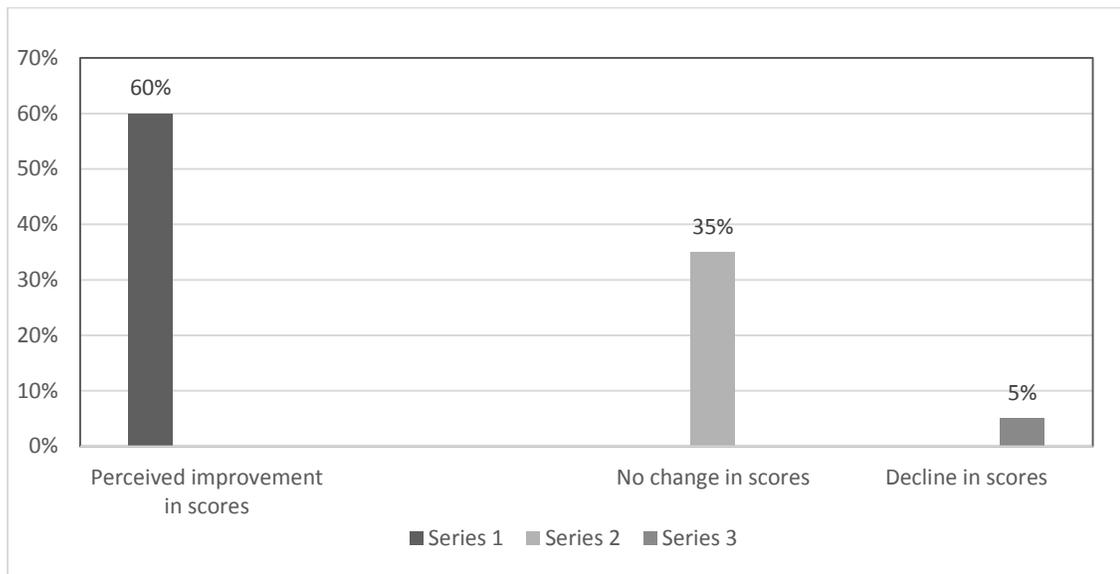


Figure 2. The percentage of the fluency scores per group

As a group, the participants' fluency scores increased by 60% in Version 2, which indicates a positive change in their perceived oral performance. Only 5% of the scores declined, and 35% of the ratings remained unchanged. The same score values for certain statements were not always interpreted to mean lack of improvement. For example, unchanged scores for speech rate were interpreted positively. The reason for this was that some participants were perceived the same in terms of their speech rate and received a positive score of "3" which indicated that their speed was normal in both versions.

The results of the fluency judgment are further detailed in the following order: (1) speech rate; (2) comprehensibility and smoothness of the speech; (3) use of pauses, repetitions and

hesitations; (4) language proficiency. Some excerpts from the “comments” section provided by the raters are also presented.

Speech rate. Speech rate, a number of syllables articulated per minute, is considered one of the best predictors of fluency (Lennon, 1990; Riggensbach, 1991; Towell et al., 1996). Overall, the ratings point to increased fluency among the participants when it comes to the speech rate exhibited in the second version of the presentations. However, the speech rate appeared to decline in the judgments of two participants (Student 5 and Student 10). The speech of Student 5, for example, seemed faster in Version 2 (the rating of “2”) while in Version 1 it was judged as normal (i.e., the rating of “3”). The speaker sounded more confident when he spoke faster, but the speed affected his speech flow making him sound monotonous. He knew the content of his presentation quite well because he did not have any fillers or pauses, but his increased speed made it difficult to follow what he was saying.

The speech rate of Student 10, on the other hand, was perceived as being slower resulting in the rating of “2” in Version 2. Even though her speed changed, her overall fluency judgment was quite positive (e.g., the rating of “4” for the comprehensibility and lack of pauses). She was careful about her pronunciation and tried to pronounce words distinctively which made her speak slower. It appears that talking slowly made it easier for her to focus on the way she pronounced words in the L2. Sometimes, it seemed like she was reading her text in Version 2, which could be another reason why she sounded slower.

Comprehensibility. The participants shared the same L1 (Mandarin) and had similar errors in articulation leading to accented pronunciation and sounding less nativelike (Derwing, Thomson & Munro, 2006). A lot of the participants were difficult to understand without the help of a transcript. With a deeper examination of their recordings and better understanding of their

accented speech through repeated listening, it became clear that the participants' pronunciation issues were the main distractor which affected the impression of their fluency. Most of the participants earned either the rating of "2" or "3" on Version 1 for comprehensibility, which was due to errors in pronunciation that made some words difficult to decipher and caused misunderstanding. Additionally, most of the participants had issues with the intonation, having no variation in tone, or with primary stress (also known as sentence stress) that they used almost on every lexical item, even on function words, which affected the smoothness of their speech (Hahn, 2004). In English, the primary stress is applied when providing new or contrastive information, and the old or given information is, usually, not stressed (Hahn, 2004). The extracts below from the "comments" section in the fluency scale reveal how accented speech and lack of variation in tone affected a rater's perception of fluency. The comments about the Version 1 of Student 2's and Student 6's presentations are provided as an example to illustrate why these participants received the rating of "2" for comprehensibility.

Student 2 (Version 1)

In terms of pronunciation, there were a lot of pronunciation errors which caused misunderstanding (e.g., "Afen countries" instead of African countries; "indictment pyramid" instead of Enlightenment period; "non tem" instead of long term). He also mispronounced the article the most of the time (de). A lot of words during a presentation were not distinctive because of the accent and mispronunciation (rating "2").

Student 6 (Version 1)

His speed was slow and his speech was not smooth. His tone was very flat without any variation in the pitch, and it was difficult to keep up with the presentation. A lot of words were not distinctive and it was hard to understand the gist of the presentation (rating “2”).

In contrast, Student 8 was rated with the highest rating of “5” on Version 2 for comprehensibility. Despite some pronunciation mistakes, his L2 suprasegmental pronunciation features such as stress, rhythm and intonation did not impede understanding and were perceived as normal by the judges, resulting in a high score (Hahn, 2004). An excerpt from a fluency judgment is provided below.

Student 8 (Version 2)

He sounded very fluent and confident. He almost did not have any hesitations or long pauses. All his pauses were natural and logical. His words were quite distinctive despite an accent (rating “5”).

Overall, Student 1, Student 2, Student 3, Student 4, and Student 10 were judged as having improved their comprehensibility and thus, sounded more comprehensive and smooth.

Pauses and hesitations. According to Riggensbach (1991), shorter and less frequent pauses make a speaker sound more fluent. The pauses which are placed not at syntactic boundaries but after every word or in unpredictable places are called juncture pauses, and they are argued to interfere with comprehension (Hawkins, 1971). Sometimes, pauses are not silent but filled with certain sounds (e.g., “uh”, “um”) and phrases (“you know”, “I mean”) and may be

referred to as hesitations and/or serve as fillers (Kowal & O’Konnell, 2004; Riggensbach, 1991). This is why in research, pauses and hesitations are usually grouped together during the analysis. The participants’ hesitations were either sounds like “uh” or fillers like “you know”, which were categorized as a FS serving the filler function.

The majority of the participants reduced the number of pauses or hesitations they made in Version 2. Student 1 and Student 7, for example, had the rating of “3” in both versions, having been perceived similarly in terms of the number of pauses and hesitations. The following extracts from the “comments” section in the fluency judgment of these two participants illustrate the complexity involved in the judgment of the pausing phenomenon.

Student 1 (Version 1)

Even though she did not have a lot of pauses in her speech, she used a lot of hesitations and fillers like “uh” instead. She used “uh/um” 38 times during her presentation. She strongly relied on those hesitations and managed to avoid making long pauses by filling them with “uhs”. The speaker also used a lot of “and” or “okay” between her sentences as fillers to make an impression of a smooth speech. She also attempted to use some transition words to organize her presentation, like first (she used firstly), before, and then. The use of fillers distracted most of the times, but despite them she still managed to sound quite confident.

Student 1 (Version 2)

As for the fillers, there were 27 “uhs” and “ums” in the presentation which filled the pauses. She also used “you know” and “guys” quite a lot trying to connect with the

audience and use those words as fillers. Surprisingly, despite a big number of hesitations, she sounded confident and rather fluent. Hesitations did not distract much and in some sentences sounded even natural.

Student 7 (Version 1)

The student spoke with a relatively normal speed. She made a lot of hesitations like “uh”. Overall, she knew her presentation very well and did not make a lot of pauses and fillers.

Student 7 (Version 2)

She sounded confident but made a lot of mistakes. Some of her errors caused misunderstanding. But, in general, her speech sounded smoothly. She did not have a lot of pauses but she used a lot of hesitations like “uhs”.

These extracts summarize the perceived pausing phenomenon exhibited by these participants. In the case of Student 1, the lack of pauses and a large number of hesitations affected the rater’s perception of the speaker’s fluency in both versions. Despite a lot of hesitations and fillers, overall, she was still perceived as confident and fluent. On the other hand, Student 7 seemed to know the content of her presentation quite well and did not have pauses but had a lot of hesitations which could be a sign of anxiety (Woodrow, 2006). It appears that a big number of hesitations can be distracting and reveal anxiety or can be redeemed by signs of speaker’s overall confidence and lack of silent pauses in the speech.

Language proficiency. The fourth statement in the fluency scale included the judgment of the speaker's language proficiency: *The language of the speaker (grammar and the choice of the vocabulary) was at the advanced level of proficiency.* As Table 7 indicates, a lot of participants received the rating of "3" ("Neither agree nor disagree that the language was advanced") for their language proficiency, indicating that the proficiency level was judged as being neither high nor low. This judgment reflects the complicated nature of the proficiency measure, which includes such varied domains as lexis and grammar. Sometimes, the judges perceived the participants as fluent because of their use of academic or formulaic language (i.e., Student 1, Student 2, Student 3, Student 8, Student 9, and Student 10).

The language proficiency of Student 2, Student 5, Student 8, and Student 10 remained unchanged which may be explained by the time interval between the presentations (i.e., 4 weeks) being not sufficient for a noticeable improvement to occur. The number of errors, however, was reduced in Version 2 which suggests that the participants managed to improve their accuracy. Overall, the majority of the participants was perceived as more proficient in Version 2, having increased their ratings by 20% (i.e., from the rating of "2" to that of "3").

4.3.2. Summary of the results for question 2

The second research question sought to discover fluency profiles of the participating participants. The fluency judgment results point to the participants being perceived as more fluent in terms of their speech rate, comprehensibility, the number of pauses and hesitations, and language proficiency. Specifically, the fluency ratings increased by 60% in Version 2. The accented speech affected understanding, and the use of stress and unnatural (for the English language) intonational patterns affected intelligibility, resulting in lower scores for speech

smoothness and comprehensibility of Version 1. Half of the participants improved in comprehensibility, receiving the score of “4” on Version 2. The majority of the participants also made fewer pauses and hesitated less when they delivered their presentations for the second time. Finally, the participants received higher scores on language proficiency, which is indicative of some improvement in the accuracy.

4.4. Research Question 3

In order to answer the third research question – *does task repetition bring on FS and fluency gains (independently and/or together) in the presentations of Chinese LI EAP students* – and to see whether task repetition had an effect on the number of FS and fluency gains in the participants’ presentations, the results from the formulaic language analysis and the findings from the fluency judgment for Version 1 and Version 2 of each presentation were compared per student. It was found that a total of six participants (Student 1, Student 2, Student 6, Student 8, Student 9, and Student 10) were judged to show improvement in both their formulaic language use and fluency development; these results are detailed next.

4.4.1. Comparison of the FS and fluency ratings

As the formulaic language analysis has shown, Student 1, Student 2, Student 6, Student 8, Student 9, and Student 10 were judged to demonstrate better results in the use of FS, varied functions of the FS, and inclusion of newly occurring FS types. Table 8 summarizes these results for each of the six participants. It provides not only the total number of tokens and the different FS types identified per student, but also details a distribution of the different FS functions per version. Additionally, the table provides the percentage of the total increase of tokens for every

student and the percentage of the newly occurring FS types. An extended table with the examples of all FS types used in both versions by the six participants is provided in Appendix G.

Table 8. *FL analysis results for the 6 improved participants*

FS functions	Student 1		Student 2		Student 6		Student 8		Student 9		Student 10	
	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2
Stance expressions	1	0	0	0	0	0	1	1	0	0	0	1
Discourse organizers	2	4	10	14	2	5	6	6	0	1	2	3
Referential expressions	5	5	5	6	1	0	2	4	1	3	3	5
Special conversational function	6	9	1	1	1	4	2	3	1	1	1	2
Figurative function	4	7	3	4	1	1	6	5	2	3	4	4
Other	5	2	0	0	0	0	2	2	0	2	0	0
Total number of tokens	23	27	19	25	5	10	19	22	4	11	10	15
The percentage of increased total		17%		38%		100%		16%		175%		50%
Total number of FS types from V1 in V2		12		13		3		11		3		6
Total number of newly occurring FS types in V2		+15		+12		+7		+10		+8		+9
The percentage of new FS types in V2 per total		56%		52%		70%		45%		73%		60%

As it can be seen from Table 8, there is an upward trend in the total number of tokens and the newly occurring FS types used by the six participants in Version 2. The number of FS types transferred from Version 1 was usually lower compared to the number of newly occurring FS types per total in Version 2. This indicates that 6 participants out of 10 had a variety of different

FS during their second attempt. Student 2 was the only exception having 13 “old” and 12 new FS out of his total of 25 FS.

The table also displays an extensive use of the different functions the FS served, with Student 2 showing the largest number of FS in each version serving the discourse organizer function. Yet, Student 9 used the FS types that served all the mentioned functions, except for the stance expressions. He increased his total by 175% having 11 tokens in Version 2. Among the 11 tokens, he had 3 FS types which he transferred from Version 1 and 8 newly occurring FS types which constituted 73% of the total. This became the highest result per group in terms of the number of newly occurring FS types in Version 2 indicating formulaic language growth.

Table 9 provides fluency scores for the six participants, revealing development in terms of (1) speech rate, (2) comprehensibility, (3) use of pauses, and (4) language proficiency.

Table 9. *FL analysis results for the 6 improved participants*

Fluency scale	S1		S2		S6		S8		S9		S 10	
	V1	V2	V1	V2								
1. The speech rate	3	3	2	3	2	3	3	3	2	3	3	2
2. The comprehensibility and smoothness of the speech	3	4	2	4	2	2	4	5	2	3	3	4
3. The use of pauses, repetitions, and hesitations	3	3	2	3	1	4	4	5	2	4	4	4
4. Language proficiency	2	3	3	3	1	2	3	3	2	3	3	3

Overall, the fluency ratings rose in Version 2, indicating a change and possibly, improvement. However, the amount of progress was unique for every student. More detailed analyses for each of the six participants are presented below.

For example, Student 8 received increased fluency ratings on all categories, revealing the best result in the group. He received the highest score of “5” for comprehensibility and lack of

pauses on Version 2. His speech rate was perceived as normal (i.e., the rating of “3”) and his language proficiency remained similar to that of Version 1 (i.e., the rating of “3”).

In what follows, a detailed analysis of each participant’s development in terms of FS use and fluency is presented.

Student 1. The fluency judgment of this student’s oral performance was rated higher on Version 2 for comprehensibility (i.e., the rating of “4”) and proficiency (the rating of “3”). She had fewer mistakes and her discourse seemed to be more organized, possibly due to the newly occurring FS types that served the discourse organizer and referential functions. Student 1 managed to have fewer pauses in Version 2 by using such FS as *you know* and *you guys*. The raters agreed that despite a large number of hesitations and fillers, she still managed to sound quite fluent, earning the rating of “3”. The excerpts from the “comments” section of the fluency scale provide more details on how the use of pauses and fillers by Student 1 was perceived by the raters on Version 1 and Version 2.

Student 1 (Version 1)

Even though she did not have a lot of pauses in her speech, she used a lot of hesitations and fillers like “uh” instead. She used “uh/um” 38 times during her presentation. She strongly relied on those hesitations and managed to avoid making long pauses by filling them with “uhs”. The speaker also used a lot of “and” or “okay” between her sentences as fillers to make an impression of a smooth speech. She also attempted to use some transition words to organize her presentation, like first (she used firstly), before, and then.

The use of fillers distracted most of the times, but despite them she still managed to sound quite confident.

Student 1 (Version 2)

As for the fillers, there were 27 “uhs” and “ums” in the presentation which filled the pauses. She also used “you know” and “guys” quite a lot trying to connect with the audience and use those words as fillers. Surprisingly, despite a big number of hesitations, she sounded confident and rather fluent. Hesitations did not distract much and in some sentences sounded even natural.

The formulaic language analysis revealed that Student 1 increased her total number of tokens in Version 2 by 17%. The results of the comparison of FS types between the two versions showed that Student 1 tried to use different FS in her Version 2, which accounted for 56% of the total number of tokens. Among the FS types that were re-used in Version 2 were: filler words and polite phrases, such as *okay so, and actually, thank you guys, you guys*; FS serving the figurative function: *so/really yummy, traced back to, junk food*; the discourse organizer *according to*; and, such referential expressions as *kind of, there is/there are* and *lots of*. She used the same FS types in both versions to express similar meanings. All six participants showed this pattern of using similar FS in both versions for expressing the same idea. The excerpts below from the presentation of Student 1 illustrates how she used the FS *more and more* in both versions to express the same idea:

Student 1 (Version 1)

Its like the healthy choice vs the junk food and the calories because **more and more** people choose to eat food like the healthy way and of course beaver tail is not their first choice of options.

Student 1 (Version 2)

So today, **more and more** people looking for healthy food. And beaver tail you know is junk food. And uhhh because they are looking for the healthy food they might not choose the beaver tails as a dessert.

In Version 2, Student 1 demonstrated increased use of FS that served particular functions. These included the use of (1) different discourse organizers to make smoother transitions between ideas: *at first*, *that's not a big surprise why*, and *besides that*; (2) referential expressions to structure discourse, *for example* and *in the same way*; and (3) FS serving the figurative function – *base down*, *look through*, *look for*, and *downtown*. Additionally, she added FS that served the special conversational functions –e.g., *how's going*, *hope you guys had a nice weekend*, and *hi guys* - to indicate formality and fillers (*like*) *you know* and *kind of*. She used the FS *you know* in the second half of the presentation 14 times relying on it as a filler. As can be seen from the excerpt below, the student used the “you know” FS to link different ideas in lieu of pauses.

Student 1 (Version 2)

You know uhmmm she likes make the bread for breakfirst (breakfast). **You know** and add the top with cinnamon sugar, butter, and gym or maybe honey. But the most popular topping is **like you know** the triditional caramel sugar. Just to like the grandma make it.

And then I want to talk about the local uhh fevers (flavors). **You know**, deres (theres) uhh different fevers (flavours) in different place.

Student 2. Student 2 was assigned higher fluency scores for the speech rate (rating “3”), comprehensibility (rating “4”), and lack of pauses and hesitations (rating “3”). His language proficiency, however, was not perceived differently and was ranked as “3”. The biggest change was in comprehensibility. Student 2 increased his comprehensibility by 40% (from the rating of “2” to “4”) as is evidenced by the following excerpts shared by one of the raters.

Student 2 (Version 1)

In terms of pronunciation, there were a lot of pronunciation errors which caused misunderstanding (e.g., Afen countries instead of African countries; indictment pyramid instead of Enlightenment period; non tem instead of long term). He also mispronounced the article the most of the time (de). A lot of words during a presentation were not distinctive because of the accent and mispronunciation.

Student 2 (Version 2)

The words were distinctive most of the time. He did start his sentences with “but”. His presentation was quite easy to follow, it was organized and logical. He used a lot of academic words and attempted to use compound sentences.

The rater's comments suggest that extensive use of academic words (FS serving as discourse organizers) might have facilitated the perception of smoother speech and influenced the impression of language proficiency and fluency.

The results from the formulaic language analysis indicate that Student 2 increased his total of tokens in Version 2 by 38% and used 52% of different FS types which did not occur in Version 1. The newly occurring FS types in Version 2 mainly served the discourse organizer function: *I'm gonna*, *we can see*, *this picture shows*, *the main reason for this is according to*, *in fact*, *even though*, and *on average*. While Version 1 did contain a few FS that served as discourse organizers, these were different from the ones used in Version 2. The re-occurring FS from Version 1 to Version 2 were: *and so on*, *in conclusion*, *do you know*, *meanwhile*, *what does this mean*, and *in disadvantage*. In Version 2, Student 2 added such referential expressions as *in recent decades*, and *in the world*; FS serving the figurative function – *deeply interested in* and *downtown*; and a greeting - *hi guys*, *how's it going*. No FS serving the “stance expression” or “other” functions were found in the speech of the student. Adding such FS as *hi guys how's it going* and *I'm gonna* at the beginning of Version 2 helped the student to organize his presentation a little better by greeting the audience and summarizing what the presentation was going to be about. The excerpts below from the introduction of the presentation show how extensive use of the different FS functions facilitated in the creation of a more organized discourse in Version 2.

Student 2 (Version 1)

Export, but, **do you know** for many afen (African) countries all of their exports are de (the) primary export are the primary commetidies (commodities). **Wa (what) das (does) this mean?** It means **most of** their export products are from local natural resources. **Making easy to** export resourses and export them to other countries without poseasy. Like the common resource good and nigera export mineral products ethopa (Ethiopia) export coffee **and so on**. **Mean while** afen (africian) countries only export the primary comodties (commodities) but also import the natural products because they don't have enough technology to produces the **high ten(tech) things**.(0:59).

Student 2 (Version 2)

Hi, how's it going guys? This is Andy, and today **I'm gonna** talk about the exports of afrin (African) countries. My major is the economecs (economics) and I am **deeply interested in** the development of afen (africian) countries economy is recent years. **According to** international trade administration a in 2014 sub-Saharan Africa is one of the fastest growing regions **in the world**. **The main reason for this is---**? **In recent decades** called the hills development their international exports for many afen (African) countries half of their peoples lives is deeply **relied on** de (the) business of export. But **do you know** for **most of** the afen (African) countries, the only export the premery (primary?) communities (commodities). **What doe(does) this mean?** It means that this country only export um natural resources. The workers can easily got these resources and the country exports the reaessources to other countries. Like the Congo export wood and oil, negria (Nigeria) export mineral products, ethopa (Ethiopia) export coffee **and so on**.

The extracts from the presentation reveal that Student 2, similarly to Student 1, used the same FS types describing the same ideas and thus transferring them from Version 1 to Version 2.

Student 6. In terms of fluency, Student 6 was rated higher on Version 2, earning the rating of “3” for the speed, the rating of “4” for the lack of pauses and hesitations, and the rating of “2” for the language proficiency. The results of the perceived fluency judgment indicate that the comprehensibility and speech smoothness was rated as a “2”, perhaps due to the monotonous intonation the student used, which made following the presentation a challenge. The language proficiency was also rated as a “2”, but the rating was still higher than for Version 1. The highest result was due to the lack of pauses, which indicated improvement by 60% (i.e., from the rating of “1” to “4” in Version 2). That rating was also the best result in the group. In Version 2, Student 6 did not pause or hesitate but continued speaking monotonously, which likely had an effect on the comprehension and overall impression of his fluency. The rater’s comments below provide evidence for such interpretation.

Student 6 (Version 1)

The speaker’s presentation was quite small in length and yet, during it he used a lot of pauses and hesitations like “uhs”. His speed was slow and his speech was not smooth. His voice was very flat and monotone and it was difficult to keep up with the presentation. A lot of words were not distinctive and it was hard to understand the gist of the presentation. There were quite a few grammar errors. Overall, the speaker created an impression of being not fluent in English or simply did not study well for his presentation.

Student 6 (Version 2)

The speaker's speed was normal but it was hard to follow, sometimes, because of a flat tone. The student did not have a lot of pauses or fillers but his lack of natural pauses made it difficult to follow his presentation, sometimes. The length of his presentation was shorter than the presentations of other participants and he seemed less prepared than the rest. It felt like the speaker was mostly reading his presentation. He also made grammatical errors.

Considering the fact that Student 6 was perceived as a student with a low level of L2 proficiency by the researcher and the judges, his growth between the versions seems quite high. He managed to improve both his formulaic language and fluency results, which could indicate some positive change in his speech due to task repetition.

In Version 2, Student 6 increased the number of tokens used by 100%. Version 1 had only 5 tokens: *not only but*, *good luck*, *thank you*, *this is*, and *I just want to*. In Version 2, on the other hand, he used 10 tokens, among which there were 7 FS types that were not used in Version 1 and constituted 70% of the total. The biggest change was in the use of FS types serving the special conversational function. He used a greeting *Hi, guys* at the beginning of his Version 2 which he did not do in Version 1, used another FS serving the pragmatic function – *thanks for listening* - at the end of his presentation, and had two fillers - *you guys* and *pretty much* - which were not used in Version 1. Among the newly occurring FS types in Version 2 there were also such discourse organizers as *I'm gonna*, *so*, *generally*, and *I'm really interested in*. Five FS

among 10 in Version 2 were used at the beginning of the presentation. The excerpts below compare the two beginnings of the student's presentation to show how an extensive use of the discourse organizer and special conversational functions in Version 2 facilitated a better structured discourse and possibly, aided fluency. Additionally, such FS like *you guys* and the contraction *I'm gonna*, due to their informal nature, created an impression of the nativelike delivery.

Student 6 (Version 1)

This is uh the first presto caass and when can the same jorney (journey) engineer danul the first four wheels cars. And my favorite car my favorite car ferri (Ferrari). Farri come in what started in nineteen forty seven in Italy. And the owner is Inno Fari and uh I like this car because is **not only** fast **but** also it is very beautiful and powerful hourse and it is very cool driving this car.

Student 6 (Version 2)

Hi guys. My name is Sherlock and today **I'm gonna** be doing my presentation on osmo history. I picked this topic because **I'm really interested in** the history of cars and **just wanted to** share it with **you guys**.

Student 8. Student 8 earned the highest score of “5” for comprehensibility and lack of pauses and hesitations on his second presentation; this was the best result in the group. Although his scores on Version 1 were equally high (i.e., the rating of “4”), even the one-point change could point to the potential benefits that task repetition can engender. The comments from the fluency judgment of this student are provided next.

Student 8 (Version 1)

The speaker spoke quite fast but still at the natural pace. He had natural pauses and his intonation was appropriate depending on a sentence either rising or falling (easy to follow). He did not have a lot of hesitations or fillers and overall, sounded pretty confident and fluent. He used quite advanced vocabulary talking about racism.

Student 8 (Version 2)

The speed of the presenter was normal and easy to follow. He sounded very fluent and confident. He almost did not have any hesitations or long pauses. All his pauses were natural and logical. His words were quite distinctive despite an accent. His English seemed upper- intermediate, he had some errors but they did not cause misunderstanding.

This student was perceived as a learner of high proficiency on both versions of the presentation; this is evidenced by the fact that his use of formulaic language and fluency development were rated high both times. However, his formulaic language analysis revealed the lowest percentage in terms of growth. This is evidenced by a 16% increase in the token total and 45% of the newly occurring FS used in Version 2. The small difference can be explained by high FS use in Version 1 and little change for Version 2. The newly occurring *I'm going to* and *we can see*, serving the discourse organizer function, seemed to help in the organization of the student's ideas in Version 2. Although, in an attempt to sound polite, he used the formal yet inappropriate FS for a presentation among peers "*good afternoon, ladies and gentlemen*", in Version 2, this was changed to the pragmatically more appropriate "*hello, everybody*". The excerpts below illustrate how the use of the different FS types, which serve the special

conversational, figurative, referential and discourse organizer functions, enabled the student to describe the same idea in a structurally more sound and, possibly fluent, way.

Student 8 (Version 1)

Good afternoon, ladies and gentlemen. Today my topic is about the racism its ah **the first** (at first) I want to introduce what is a racism. Racism is means a person or a group of person **that's** is pise or hate or repel to another person or another group of person and is always show is on words and action. And first I want to I think everybody still remember the news that has been in April 9th and its happen in the united airlines company.

Student 8 (Version 2)

Hello everybody. I think all of **you guys** know that I am an international student in Canada and I am from china and my friend told me some story that about racism that happen in Canada and United States. So today I want to talk about racism because it is still a **big issue that's around the world.** **At first** I want to say what is a racism. Racism is means a person or group of person despise, hate, and repel to another person or another group of persons show its on the words or actions. Now **I'm going to** say a news. **I believe** most of **you guys** still remember on April 9th.

The student also used the filler *you guys*, which sounds informal and more nativelike in the given context. Another interesting feature about the FS types the Student 8 used in both

versions was the phrasal verbs *give up*, *stand up*, and *sign out*, which, possibly due to their figurative function, created an impression of fluency.

Student 9. As can be seen from the extracts from the presentation, Student 9 made a lot of mistakes and mispronounced a lot of words which caused some misunderstanding. His fluency ratings, however, increased. His speech rate was perceived as more normal in Version 2 (i.e., the rating of “3”). He earned the rating of “3” for comprehensibility which may indicate that even though he pronounced the words more distinctively, he still made pronunciation errors that may have caused occasional misunderstanding. His biggest change was in the scores for the lack of pauses and hesitations, which accounted for a 40% increase (from the rating of “2” to “4”). Finally, his language proficiency was rated as “3”, indicating that he had made fewer mistakes but his overall level of proficiency was quite low. A rater’s comments below provide a comparison between the perceived judgment of Version 1 and Version 2.

Student 9 (Version 1)

The speaker’s speed was quite slow. He had a lot of pauses and hesitations like “uh”. Some words were not distinctive and it was hard to understand a sentence. The voice was monotone and the intonation was not rising or falling according to the sentence which made it hard to follow. The language was intermediate, there were a lot of mistakes. He tried to use advanced vocabulary. Overall, he sounded like he was mostly reading his presentation.

Student 9 (Version 2)

The presentation was quite long and he prepared it well. He seemed like he was reading it though. His speed was normal and he sounded confident. His speech was quite comprehensive and some pronunciation errors did not distract too much. He did not have a lot of pauses or fillers. He sounded quite fluent. His language was quite good (written text was good) but he made some mistakes. Some of his pronunciation errors caused some misunderstanding. He finished his presentation abruptly and his intonation did not always indicate the breaks between sentences. It sounded more like a flow of thoughts without a lot of natural pauses between them.

The results from the formulaic language analysis revealed that Student 9 showed a dramatic difference between the total number of tokens used in Version 1 and Version 2, constituting an increase of 175%. Not only did he increase the total number of tokens, but he also used different FS types which did not occur in his Version 1 and were at 73% of the total in Version 2. His increased percentage of the tokens total and the number of the newly occurring FS types were the highest result in the group. The FS types transferred from Version 1 to Version 2 were: the referential expression *there was*, FS serving a polite function *hi, everyone*, and the collocation *to step foot on* serving the idiomatic function. The perceived improvement on Version 2 was also due to the increase in the new functions performed by the employed FS. For example, in Version 2, Student 9 used FS types which served the discourse function (e.g., *at first*) and “other” function (e.g., *one by one* and *looks like*); these functions were not displayed by the FS used in Version 1. He also used these phrasal verbs: *pick up, fold down, and step off*, which may have helped him sound more nativelike. The two excerpts from Version 1 and Version 2 below illustrate the same portion of the presentation. It may be observed that in

Version 2 the density of FS for that paragraph is higher compared to the Version 1, which did not contain any FS.

Student 9 (Version 1)

This is picture and the first time lieutenant helmor ask the boss the intelligence of a bomb enemy that intelligence was rampic cause they enter the risk tanemant the enemy and they finally was surround by 2000 enemy soldiers who after they landed. Lieutenant Colonel Moore is the first Italian commander of service cover regiment. He let his brother face this war. In this battle the longest 3 days of there life they hold there position with imaginable courage.

Student 9 (Version 2)

The lieutenant coronel moore the first bataliane of the service cowaare cover regimen wit his brother to face this war. In a longs three days lay(they) live lay(they) help the position with unimaginable courage a military cross bender cross banded Joeseph Garowery **pick up** ground dry the battle because he watch the American soldiers **fold down one by one**.

Student 10. The findings from the fluency judgment showed that because Student 10 was perceived differently on Version 2, her ratings increased. She improved her pronunciation, and her speech was rated as more comprehensible in Version 2 (the rating of “4”), which could be due to her attempt to speak slower. However, speaking at a slower pace resulted in a decreased rating for the speech rate. She seemed to speak without pauses or fillers in both versions which

earned her the score of “4” each time. The comments below summarize her perceived fluency for the two versions.

Student 10 (Version 1)

Her speech was smooth but because of her accent some words were not comprehensive and it was hard to understand some sentences. Overall, she sounded fluent and confident. She seemed well prepared. She did not have pauses or fillers. Sometimes, she had some hesitations, but they did not distract. Her language was quite good but she made some errors.

Student 10 (Version 2)

The speaker seemed a little slow. She was careful about her pronunciation and tried to pronounce words distinctively. Maybe because she was talking slowly it was easier for her to focus on the way she pronounced words in her L2. Sometimes, it seemed like she was reading her text. Her presentation was well organized and she made logical pauses between different thoughts. It was easy to follow. She did not have long pauses and she barely made hesitations in her speech. Overall, her language seemed good, intermediate but she had some errors.

The formulaic language analysis showed that Student 10 had a 50% increase in the total number of tokens between the two versions. She used 60% of newly occurring FS types in Version 2, which included: the stance expression *I believe*, two discourse organizers - *I'm going to* and *to summarize*, and such referential expressions *lots of*, *a large number of in the middle of*, and *this is*. In Version 2, Student 10 used the “*Hello everyone*” FS, which served the special

conversational function, expressed formality, and helped to create a contextually-appropriate introduction in the presentation. In every version Student 10 used FS that served the figurative function: *get to know*, *blend in*, *downtown*, *come from*, and *hang out*. Two excerpts are provided below to exemplify how Student 10 finished her presentation in each of the versions. Version 2 displays a more structured ending due to the use of discourse organizers and stance expressions. What's more, the student made fewer errors in Version 2. For example, in Version 1 she attempted to use FS *there is* and *blend in* but made mistakes; in Version 2, she used *blend in* correctly.

Student 10 (Version 1)

Uh **on the other hand** nonono sorry, Ottawa summer also has other other interesting festivals. **There** like a la machine or Chinese fooda or Chinese fooda sorry or chinese fooda festivals and in like (international) participants can **blend** Ottawa culture easily and quickly and to summarize Ottawa summer is evident and very real various. So I hope you can enjoy this Ottawa summer and **thanks for listening**. This is the end of presentation.

Student 10 (Version 2)

Ottawa also has many other summer activities, like a Italian weekend festival, dogs festival, Chinese food festival and more. International participants can **blend in** Ottawa cultures easily by participating this activities. **To summarize**, Ottawa summer is abundant (abundant), delicious and various. **I hope** everyone will enjoy this summer. For reference I only put one reference the end of the cylinder because I have been this

festivals last year so almost information and images **come from** myself. **Thanks for listening. This is** the end of my presentation.

4.4.2. Summary of the results for question 3

Based on the analyses performed, it appears that task repetition may have augmented both the use of formulaic language and fluency of six participants out of ten. They increased the percentage of tokens and newly occurring FS types they used; they were also rated as having achieved higher fluency scores on the second iteration of their oral presentations. Yet, the nature of improvement for every student was different and rather unique. Student 6 and Student 9, for example, displayed the biggest positive change in the use of formulaic language between the versions. While all six participants used a high percentage of the newly occurring FS in Version 2, the highest result (73% increase) was reported for Student 9.

The use of the different FS types may have led to the use of varied discourse functions, which might have influenced the fluency perceptions resulting in higher scores. While Student 8 was perceived as the most fluent, Student 6 showed the biggest difference between the ratings, which constituted a 60% increase (from rating “1” to rating “4”).

Chapter 5: Discussion and conclusion

5.1. Overview

The purpose of this study was to examine the effects of task repetition on the development of formulaic language and perceived fluency in the oral presentations of Chinese L1 EAP students. The study found that overall, the participants increased the number of formulaic sequences (FS), used a variety of FS that served different functions (i.e., discourse organizers, referential and stance expressions, special conversational, figurative and “other” functions) and earned higher fluency scores after delivering the presentation the second time. The nature of the improvement, however, was different and unique for each individual. Nevertheless, six participants showed more progress in both formulaic language and perceived fluency development than the other four. It appears that task repetition may have affected the participants’ use of formulaic language and speech fluency development, thus contributing to the positive results in both areas. This chapter discusses these findings, highlights limitations, outlines pedagogical implications, and suggests future research directions.

5.2. Research question 1

In this section, the findings that answered the first research question: “*What are the common formulaic sequences (FS) used by Chinese L1 EAP students and what functions did they serve in the students’ presentations?*” are discussed. Overall, the results from the formulaic language analysis revealed that the participants used 145 FS types in their oral presentations and had an increased total of tokens and newly occurring FS types in Version 2. The FS which were identified in the presentations served six functions. To begin, the common FS used by the participants are addressed. Following this, the total number and variety of FS types in the

students' presentations are discussed in more detail. Finally, the different functions of the FS found in the participants' presentations are outlined and the purpose of repeated FS per version is discussed.

5.2.1. The common FS used by Chinese L1 EAP students

The formulaic language analysis revealed that the participants used a wide range of different FS, which is not surprising considering the “multi-faceted nature of formulaic language” and its various categories (Wray & Perkins, 2000, p. 3). The list below summarizes different categories of formulaic language and their examples found in the presentations of Chinese L1 EAP students in the current study:

- phrasal verbs: for example, *hang out*, *blend in*;
- idioms: *to take for granted*;
- collocations: *daily lives*, *big issue*;
- expressions: *of course*, *in fact*;
- lexical bundles: *I'm going to introduce*, *thank you very much*;
- compounds: *downtown*, *junk food*.

The most common FS category found in the participants' presentations was lexical bundles – “the most frequent sequences of words used in the register” (Biber et al., 2004, p. 371). Among 145 FS types found in the presentations of this group of students, ninety-three could be identified as lexical bundles. The reason for such frequent use of lexical bundles by all participants can be explained by the fact that the presentations were delivered in academic settings and there was a need for using certain lexical phrases which would carry particular functions in university discourse (Nattinger & DeCarrico, 1992). For example, the evaluation rubric employed both by the teacher and the participants for assessing performance included

“organization”, which implied structuring the presentation according to academic register norms (see Appendix B). Therefore, the participants felt the need to use lexical bundles which would help them to fulfill a number of discourse functions such as introducing the topic, elaborating on it, and concluding it, as well as expressing attitude and modality, logical relationships in the text, and addressing the audience (Biber et al., 2004). Overall, lexical bundles are abundant in many academic registers such as classroom teaching, academic prose and textbooks. This may suggest that the participants were exposed to quite a large number of lexical bundles several times through input during their study and could have transferred them to their own oral presentations.

On the other hand, the presentations delivered by the students did not serve as a typical example of an academic presentation because the topics chosen by the students were not always discipline-specific and sometimes informal (e.g., festivals, movie, Beaver tails). The informal nature of the majority of the presentations may explain an extensive use of collocations, some idioms, compounds and expressions (29 FS) and phrasal verbs (23 FS) which can be typically found in everyday discourse rather than in university registers (Biber, Johansson, Leech, Conrad, & Finegan, 1999). Some of those FS seemed more advanced than the actual L2 proficiency level of the participants and helped them sound more fluent (Pawley & Syder, 1983; Wray, 2002). Since the participants had a chance to plan and get ready for their presentations prior to delivering them, they may have borrowed some sophisticated FS from written texts. However, the participants’ knowledge of formulaic language was not tested before and after the task repetition. Therefore, it is not known exactly which FS the students knew and retrieved from memory and which ones they noticed and transferred from the written text at the moment of presenting.

5.2.2. The increased total and variety of formulaic language in the presentations

The results from the formulaic language analysis indicated that the participants in the present study used various FS and even managed to increase their number in Version 2. There is no empirical evidence in the literature which would support the effectiveness of task repetition on formulaic language development. However, the positive findings of the present study suggest that there is a possibility that task repetition may generate some variety and the increased total of FS types in L2 speech. Generally, the previous research on task repetition has demonstrated that it is useful for L2 learners and improves their language production by increasing proficiency, accuracy, fluency, and lexical complexity (Ahmadian & Tavakoli, 2011; Bygate, 1996; Gass et al., 1991). Considering positive effects of task repetition on general L2 performance found in the literature, it can be argued that it also managed to promote the changes in the participants' formulaic language use. During task repetition L2 learners build on their first attempt, and this gives them an opportunity to add lexical richness and sophistication (Gass et al., 1991). It seems that in the present study, the need and desire to improve Version 2 had also brought selectiveness and attention to the word choice, and this possibly led to a bigger variety and larger number of FS in Version 2 (Bygate, 1996; Gass et al., 1991). Moreover, addressing and developing formulaic language was not required by the task and the participants most likely were not even aware that they used more FS during their second enactment of the task. As previously mentioned, the participants' ESLA program, did not include a focused formulaic language instruction component. Thus, it may be suggested that the participants were not informed about the important role of formulaic language in oral production. These factors may suggest that the increase in total and greater variety of FS in the participants' presentations did not occur due to the pedagogical intervention or students' awareness of FS in their speech, but rather as a result of repeating the oral task (Ahmadian & Tavakoli, 2011; Bygate, 1996).

Interestingly, the analysis of the formulaic language in Version 2 revealed that students used similar FS transferred from Version 1 during their second attempt of the same task. A possible reason for transferring the same FS types from one version to another could be due to the repetition of the presentation. According to Ellis (2009b), the task repetition procedure can be organized differently: either the learners repeat the same procedure but with different content, or repeat the same task with similar content. The latter was used in the present study, as the participants delivered the same presentation twice, even though they had some freedom to lightly modify their content to improve their Version 2 after self-reflection and receiving the teacher's feedback on Version 1. In general, transferability of content could have been an advantage for the participants and allowed them to use similar language to express ideas, and sometimes even repeat utterances precisely. Thus, some of the FS could have been transferred in order to carry the same idea and serve a similar discourse function. For example, the participants used a number of lexical bundles in Version 2 which they had applied in Version 1 (i.e., *according to, in conclusion, in disadvantage, I want to say, for example, and so on, not only but, do you know, meanwhile*). In terms of functionality, such lexical bundles can be useful for organizing thoughts, and this may explain why these types of FS were used in both versions (Biber et al., 2004). Moreover, as mentioned earlier, it is likely that a portion of FS types used by the participants was taken from their prepared written text. Students prepared the presentations themselves, and they were instructed to use any available sources to prepare for the task. It is possible that they could have noticed some FS in the input and realized that they were useful for carrying a certain meaning (Boers et al., 2006). For example, according to Boers, Deconinck and Lindstromberg (2010), high-frequency FS have the best chance of being picked up by the learners autonomously. It can also be assumed that students previously heard some of those FS types and

perhaps had used them in their speech before, which could explain why they relied on them to express a particular meaning. Moreover, according to Bui and Huang (2018), familiar topics encourage more fluent L2 speech. Accordingly, the participants of the present study were allowed to deliver a presentation on the topic of their choice. This suggests that they might have had some background knowledge about their topic, and possibly knew some FS to express the ideas.

On the other hand, the high percentage of newly occurring FS types in Version 2 may suggest some perceived progress in the use of formulaic language where participants did not only repeat what they said in Version 1, but also increased the total by adding new and different FS types. In fact, six students out of the ten managed to use a wider range of FS types in Version 2, as there were more than 50% of newly occurring FS types per total number of tokens in their presentations (Student 1, Student 2, Student 6, Student 8, Student 9, and Student 10) (see Table 8 for results of percentage of newly occurring FS in participants' presentations). Content familiarity has been previously shown to affect learners' task performance and lead them to feeling more confident (Qiu & Lo, 2017). As the participants were already familiar with the presentation content, this may have allowed them to organize and express their discourse more freely. More specifically, by drawing on a broader range of the language resources, some of the presentations displayed more variety and an increased total of FS types. Additionally, newly occurring FS in Version 2 could have been generated when the learners tried to change the content of the presentation by adding new pieces of information. For example, Student 6 had 70% of the newly occurring FS types in Version 2 which is one of the highest results per group. Interestingly, this participant increased the length of his presentation and added some new content. He was the only participant out of ten who drastically modified the content of the

presentation and at the same time used newly occurring FS types to express new ideas. A change in the content of the presentation did not affect his accuracy, but he attempted to use more complex sentences.

Further discussing the increased total and variety of FS used, another factor to consider is the self-assessment and reflection that was required on Version 1 in order to improve Version 2. Due to the shift from a teacher-centered approach to learner-centered approach in teaching methodologies, self-assessment has become a popular strategy in task-based language teaching (TBLT). More specifically, self-assessment is described as a testing alternative to motivate and improve learning with the help of self-reflectiveness (Clapham, 2000; Oscarson, 2013; Tavakoli, 2010). In the present study, the participants needed to self-assess their performance twice after having delivered each version of their presentation and prior to receiving the teacher's feedback. It is important to note that the instructor's five-trait rubric, which was also used by the students, focused primarily on voice quality (volume and pace), pronunciation, intonation, word and sentence stress, eye contact and presentation style, organization, content, and the appropriate use of APA style references and in-text citations. However, the participants were free to focus on other aspects during their self-assessment which were not explicitly addressed in the rubric. Thus, there is a possibility that the self-reflection encouraged the participants to improve general L2 quality in Version 2 and attend to vocabulary and grammar, for example. Therefore, it may be argued that self-assessment could have led to producing more sophisticated language use during task repetition, and thus larger number of FS and their variety (Chen 2008).

5.2.3. FS functions in the presentations

In the present study FS types used by the participants served a variety of different functions. The results indicated that learners mainly used sequences with a figurative function

(38 FS) as well as sequences with a discourse organizer function (37 FS). FS with a figurative function were mostly phrasal verbs and some collocations (i.e., *based on, fold down, big issue, met challenges*). According to Li, Zhang, Niu, Jiang, and Srihari (2003), phrasal verbs account for one third of English verbs, which means that they are highly frequent. An examination of phrasal verbs in corpora showed that they are more abundant in spoken discourse: in conversations and informal spoken environments (Biber et al, 1999). Indeed, phrasal verbs are very common in everyday informal speech, and their synonyms borrowed from such languages as Latin, Greek, or French are usually used in more formal registers (Wood, 2015). The collocations which the students used in their presentations are also common in everyday speech, for example, *junk food, daily lives, high tech things*. Some of the collocations used had lexicogrammatical invariability (i.e., *deeply interested in* and *good luck*) while some seemed less fixed (i.e., *so/really yummy*) (Scandera, 2004). Therefore, the relatively informal nature of the presentation in the present study compared to other typical oral presentations in an academic setting might explain the frequent use of common phrasal verbs and collocations.

Discourse organizer lexical bundles usually serve two major functions: Topic introduction/focus (i.e., *I just want to*) and topic elaboration/clarification (i.e., *in fact*) (Biber et al., 2004). The reason why the discourse organizer function was used so often could be due to the fact that a big portion of the FS used in participants' presentations were lexical bundles (i.e., *I'm going to introduce, we can imagine, this image shows, in conclusion*) which are a part of university discourse and are widely used in the academic written and oral language, as observed by Biber et al. (2004). Additionally, the participants used a total of 26 referential expressions in their oral presentations (i.e., *this kind of, at the same time, in the middle of, at the beginning*). This aligns with the findings from Biber's et al. (2004) study which illustrated that referential

bundles were the most commonly used in university context, may these bundles are seen in classroom teaching, textbooks, or academic prose. An extensive use of discourse organizers and referential expressions during task repetition in the present study might suggest that the participants wanted to improve their Version 2 by structuring and organizing their speech. Similarly, Qiu and Lo (2017) illustrated in their study that the participants, who happened to be Chinese EFL students as well, wanted to make their speech more concise and logical during the second attempt of task repetition, and this resulted in a more organized discourse.

Another function which the participants used quite a few times was the special conversational function. Biber et al. (2004) found that lexical bundles serving the special conversational functions were the least frequent in the academic register, and they were only used a few times in conversations. The participants in the present study used 23 FS serving this function. This constituted 16% of the total number of FS found in the presentations. Among them were 15 polite formulas and greetings (i.e., *thank you* and *good evening*) and 8 fillers (i.e., *you guys* and *kind of*) which were added as a sub-category to this function, since fillers were not included in the functional taxonomy of Biber's et al. (2004). Even though the corpus examination showed that the special conversational functions were rare in the academic discourse, an informal nature of the presentation in the current study might explain why the students felt the need to use this function quite a few times.

Stance expressions were found by Biber et al. (2004) to be rather common in the university context, especially in classroom teaching, conversation, and textbooks. However, the participants in the present study were delivering a presentation, which is a different genre in terms of language and format compared to the types of output mentioned above. Therefore, it seems that the participants simply did not need to use the stance expression function in their

discourse as often, and this may be the reason why a small number of FS serving this function was found in the present study (i.e., *we have to, we can use two ways*).

Finally, certain functions of FS were extensively repeated throughout the presentations of some students (Student 1, Student 2, Student 3, Student 4, Student 6, and Student 10). It may be suggested that the repetition of particular functions throughout one presentation might have happened on purpose. For example, the biggest number of repetitions was found in the FS serving the figurative function. Repeating the same phrasal verbs and collocations in the same presentation may suggest that the participants wanted to add some expressiveness to their story or sound more informal (Biber et al., 1999). Another function used repeatedly was the subcategory of the special conversational function – the “filler” function. For instance, Student 1 had the largest number of the FS *you guys (like)* and *you know* in each version of her presentation, which may suggest her overreliance on those FS to fill pauses, and/or maintain the flow of her presentation. The excerpt from the presentation of Student 1 is provided below to illustrate her extensive use of repetitions and to discuss the possible reasons for having them in the discourse.

Student 1 (Version 1)

Okay so my topic is uh beaver tails ba(but) before I'm starting my topic I want to ask **you guys** have **you guys** ever heard of beaver tail or like eat it before. So what favor (flavor) you chose, okay okay okay so I like based on the four parts to introduce **you guys** to the beaver tail. The first one is what kind of beaver tail they has, like da(the) favor(flavor) and the history of beaver tail. And da(the) beaver tails competater(competitor) and da beaver tails future is based on my (50 seconds) suggestions. Okay I already showed the pictures of beaver tail theres not 10 favors(1 min) (flavors) of beer tail head so **you guys**

can order the tan(ten) favors(flavors) when you go to the store. And uh personally my favorite is number 7, its so good **you guys** you have to try it.

As can be seen from the presentation of Student 1, the FS *you guys* was repeated in almost every sentence. It is possible that speaking anxiety during performance in English (Woodrow, 2006) could have caused this overreliance on the FS *you guys*. Using this filler could have also indicated her desire to make a connection with the audience and address them as much as possible to ensure that they listen to her. Similarly, Wray and Perkins (1999) suggest that FS may serve as different devices, one of them is using formulaic language as a tool for social interaction. According to their description, it is possible that when Student 1 used fillers such as *you guys* repeatedly she felt the need to be “separated from the crowd” because of her special status as a presenter (p. 14). It could also be suggested that Student 1 did not realize how informal the sequence sounded and pragmatically speaking, applied it to a context that requires more formal speech. It may also be suggested that one of the purposes of repeated filler FS is to buy some time while recalling what goes next in the presentation. According to Wray and Perkins (1999), FS may serve as “compensatory devices for memory limitations” and “vehicles for fluency” (p.16). Similarly, in Wood’s study (2006), the participants used repetitions of certain FS “to allow for some controlled processing or conceptualization or recall of film content while continuing to speak” (p. 30). Moreover, in his study, repetition of FS helped the speakers to lengthen runs and avoid pausing. A similar pattern was observed in the current study when the participants seemed to use fillers and some discourse organizers (i.e., *according to, for example, and so on*) to avoid having too many pauses in their presentations.

5.2.4. Summary of the discussion of research question 1

Overall, the most common FS types used by the participants of this study happened to be lexical bundles, possibly due to the needs and expectations of university discourse. Another big category of formulaic language preferred by the participants was collocations and phrasal verbs which can be explained by rather informal nature of the presentations on the topics chosen by the students. As for the functions, this group of participants used the largest number of FS serving the figurative function and the discourse organizer function, respectively. An extensive use of phrasal verbs and collocations may be explained by the desire to add some expressiveness to the presentation while the reliance on the discourse organizers can suggest students' desire to structure their sentences and have a more cohesive presentation. Finally, it seems that the repetitions of the same FS throughout one presentation could have been used on purpose because they served such discourse functions as the figurative and the filler functions.

5.3. Discussion of research question 2

The current section discusses the second research question: "*What are the fluency profiles of Chinese L1 EAP students in their presentations?*" The findings from the perceived fluency judgment revealed that overall fluency ratings of the participants increased by 60% in Version 2, after the task repetition. The results are discussed in the following order: Perceived improvement in speech rate, comprehensibility, the number of pauses and hesitations, and language proficiency.

5.3.1. Speech rate

Speech rate is a temporal variable usually defined as the number of audible syllables uttered per minute (McGuire & Larson-Hall, 2017). As previously mentioned, this was measured in the current study by human raters. Previous research has shown that a holistic judgment of

speech rate tends to resemble an objective fluency measurement (Bosker et al., 2013), which is done with the help of automated speech evaluation systems (Ginther, Dimova, & Yang, 2010).

In the present study, the speech rate of the participants in their oral presentations was mainly perceived as normal and natural in Version 2. More specifically, the speed in which the utterance was produced, seemed appropriate and aided understanding. The researcher determined the perceived speech rate by comparing the two presentations of each student. In some cases, the speed had changed from slow to normal, while in others, it was normal in both versions.

Previous quantitative research has indicated that speech rate is a predictor of L2 fluency and that a faster rate of speech was found to be related to higher fluency scores and language proficiency (Kormos & Denes, 2004; Riggenbach, 1991, Towell et al., 1996). However, qualitative data suggests that the relationship between speech rate and perceived fluency appears to be more complex, as other factors may interact and influence the scores human raters assign (Préfontaine & Kormos, 2016). For example, in the present study, faster speed was somewhat distracting and affected smoothness of the speech flow, even though faster L2 speech rate is usually associated with a closer approximation to target language proficiency and may appear to be a sign of improvement (Towell et al., 1996). Similarly, Préfontaine and Kormos (2016) suggest in their qualitative investigation on the perception of fluency in second language French that when the pronunciation of the speaker is uncertain, faster speech might affect comprehension in a negative way. On the other hand, according to Derwing (1990), slow speech rate may be more preferable, especially when L2 speech is accented and thus more susceptible to misunderstanding. In the present study, the participants had some issues with pronunciation and comprehensibility, and the independent judges reported that it was easier for them when the participants spoke slightly slower, yet still naturally. Nevertheless, speaking too slowly might also be problematic because it

fails to catch the listener's attention (Préfontaine & Kormos, 2016). Overall, it seems that moderate speech rate is the most appropriate for optimal comprehensibility (Munro & Derwing, 2001). Thus, speaking with a slightly slower or normal speed helped the participants of the current study to pronounce words more distinctively and as a result receive higher scores for perceived fluency from the raters.

As evidenced by the findings of the present study, there is a positive interrelationship between speech rate and comprehensibility. For instance, the participants who spoke with a more natural speed in Version 2 also received higher scores for comprehensibility. The observed improvement in speech rate might be explained by the task repetition. The repetition of the same presentation could have affected the speed of delivery and overall confidence of the participants because they were familiar with the content. This is in line with the findings from previous research, which suggest that task repetition may lead to larger gains of speed of delivery amongst L2 learners (Ahmadian & Tavakoli, 2011; Bygate, 2001; Lambert et al., 2017; Lynch & Maclean, 2000; Wang, 2014b). Additionally, the participants were provided with the teacher's feedback after they had completed their first attempt and reflected on it. As previously mentioned, the feedback included some comments on voice quality (volume and pace), pronunciation, intonation, word and sentence stress. It can be argued that the teacher's feedback could have contributed to the improvement in the speech rate during task repetition. The feedback given to the participants on their initial performance before task repetition may have highlighted some weaknesses in the participants' spoken English. As a result, this may have brought awareness of the issues they needed to focus on (Lynch & Maclean, 2000). Knowing their problems, the participants could have paid deliberate attention to their spoken production during task repetition in order to avoid making similar errors. For instance, Sheppard's (2006)

study with Japanese participants explored the effects of task repetition accompanied by feedback. The study revealed that feedback on the initial performance positively affected fluency and accuracy of the participants. However, it should be noted that in Sheppard's (2006) study, the experimental group received a different type of feedback which was designed to draw their attention to linguistic form, and this resulted in better accuracy, lexical richness, and fluency. Thus, it can be suggested that L2 learners might improve during task repetition in the areas explicitly addressed by the teacher's feedback. For example, in the present study, the area of focus was mostly pronunciation and intonation. Thus, specifically addressing those issues led to higher scores for speech rate and comprehensibility.

Finally, the interrelationship between speech rate and comprehensibility reflected in the ratings may suggest that the majority of the participants paid deliberate attention to speech rate in order to sound more distinctive during task repetition (Derwing, 1990). Their effort might be also explained by their engagement in self-assessment of their first performance after having listened to their own recordings. A recent study by Strachan, Kennedy and Trofimovich (2018), which also involved Mandarin speakers of English, investigated the role of task repetition and self-assessment in L2 learners' awareness of their own comprehensibility in oral production. The findings of the study support the results found in the present study and suggest that the flow and normal speech rate were associated with better comprehensibility during task repetition. However, comprehensibility of their participants only changed slightly after self-assessment which could have been due to the lack of an evaluation rubric (Strachan et al., 2019). According to Casteñeda and Rodríguez-González (2011), adding an evaluation rubric to task which may or may not involve repetition, can boost effectiveness of self-reflection. As mentioned earlier, the participants of the present study had an evaluation rubric. It seems that this tool assisted them in

self-assessment of their speech since they managed to improve both speech rate and comprehensibility during task repetition. The next section provides more detail about comprehensibility of the participants.

5.3.2. Comprehensibility

The perceived fluency findings in the present study indicate that the participants had issues with comprehensibility in both versions, but seemed to slightly improve it in Version 2. Comprehensibility is a major concept in L2 pronunciation research, and it is defined as listeners' perception of how easily they comprehend L2 speech (Munro & Derwing, 1999). According to Goh (2007), pronunciation plays a crucial role in comprehensibility. Pronunciation is a complex phenomenon and might refer to both segmental features (i.e., errors that involve individual sounds) and suprasegmental features (i.e., word stress, rhythm, intonation) (Isaacs & Trofimovich, 2012). Derwing and Munro (2009) also propose the term accentedness, defined as “how different a pattern of speech sounds compared to the local variety” (p. 478). Accented speech may also affect communication because it is usually overemphasized due to its salience (Derwing & Munro, 2009). For example, the independent judges who participated in this study reported that it was challenging for them to judge fluency of the participants because of their accentedness and issues with pronunciation. The participants were all Mandarin L1 speakers, and most of them had accented speech in English which was reportedly difficult to understand without a transcript. The phonological system of Chinese is very different from that of English, and therefore, in general, L1 Chinese speakers tend to find English difficult to pronounce and may make errors in sounds, stress, and intonation (i.e., prosody) (Wang, 2014a). More specifically, many participants had issues with both segmental and suprasegmental features in the present study. For example, mispronouncing individual sounds (i.e., *favour* instead of

flavour; arregrant instead of *arrogant; Afen* instead of *African*) as well as using primary stress on every lexical item, which is not typical in the English language (Hahn, 2004), affected comprehensibility and smoothness of their speech. A survey conducted by Wang (2014a) with 100 English majors at a Chinese University revealed that pronunciation was a very common issue and accounted for 46% of the participants. Even though the participants of the current study were living in Canada, they did not engage much with non-Chinese speaking peers outside the classroom, which limited their exposure to the target language sound and practice. Will (2016) stresses that this target group of learners tend to be isolated during their study abroad and usually avoid interactions with native speakers. Yuan (2011) explains that this might be due to their insufficient English, cultural differences or anxiety.

The improvements in comprehensibility in the present study might be due to task repetition. According to Bygate (2001), L2 learners focus more on formulation and articulation during task repetition, rather than on conceptualization. This implies that they tend to focus on form rather than on ideas and content, the latter usually being the primary focus of the first attempt. Therefore, the improvement in comprehensibility might be explained by the participants' careful attention to their articulation. In addition, the teacher's feedback and self-assessment of pronunciation, intonation, word and sentence stress could have also brought participants' awareness to the issues surrounding comprehensibility. This could have motivated the participants to work on those issues during task repetition.

5.3.3. Pausing phenomenon in the participants' presentations

Pausing is a complex aspect of fluency which is associated with a temporal variable mean length of runs. This refers to the average number of syllables produced in utterances between pauses (Lennon, 1990; Riggenbach, 1991; Towell et al., 1996). As mentioned before, the number

of pauses was not calculated in the present study, but instead judged holistically. Pauses can be either silent or in a form of a hesitation. For example, hesitations can be either sounds (i.e., *uh, um*) or small phrases (i.e., *FS you know, I mean*) which are usually categorized as fillers (O'Connell & Kowal, 2004). In the present study, the participants had both silent pauses and fillers/hesitations. Previous research has shown that the increased length and number of silent pauses in L2 speech are usually more distracting to the listeners than the filled pauses, and therefore cause a breakdown in fluency (Bosker et al., 2013). In a similar way, presentations of the participants of the present study who had a large number of silent pauses in their speech were also perceived as less fluent. On the other hand, filled pauses can be a sign of anxiety (Goldman-Eisler, 1968) or a marker of certain pragmatic function in discourse (Hasselgreen, 2004). For example, Student 1 in the current study had a large number of filled pauses in both versions (i.e., *FS you know, okay so*). These filled pauses could be interpreted either as a sign of anxiety (Woodrow, 2006) or conversely as a sign of confidence and desire to attract attention of the audience (Wray & Perkins, 1999).

The findings show that the participants had a larger number of pauses in Version 1. The pauses occurred possibly because it was the first enactment of the task, and students needed more time to deliver the message and to structure discourse (Bygate, 1996). The participants might have used pauses in Version 1 as a memory device for remembering the content of their presentation as well as the order of ideas. Wray and Perkins (1999) also suggest that pauses and hesitations may serve as time buyers in oral production. Furthermore, some students seemed to have more pauses either at the beginning or the end of their presentation. The exact reason why this occurred is unclear, but it may be assumed that the participants simply knew certain parts of their presentation better than the rest. Furthermore, pause occurrence can be explained by the fact

that students could have experienced either anxiety at the beginning of the performance or fatigue at the end of it. Of course, the level of preparedness and familiarity with the content of the presentation was different for every participant. Those who performed better in class with less pauses might have rehearsed more at home and vice versa. Delivering certain parts of the presentation with more ease than the other parts could be explained not only by the level of preparedness of the participants, but also by the complexity of the language used for expressing those ideas. In general, the participants made more pauses in more complex sentences in both versions.

Syntactic location of pauses also seems to be an important factor to consider when assessing someone's fluency (Wood, 2015). Pauses, which occur anywhere in a clause or phrase, are called non-juncture pauses, and since they are not used at predictable places, they might interfere with listener's comprehension and be a sign of lower language proficiency (Cenoz, 1998). Moreover, since "the absence of hesitations within a sequence" indicates that it might be formulaic, non-juncture pauses may also suggest that L2 speaker has not automatized FS being used (Boers et al., 2006, p. 247). The participants of the present study had more non-juncture pauses in Version 1 that resulted in lower fluency ratings.

As has been previously reported in the literature, a low amount of pauses and hesitations typically indicates a higher level of L2 fluency (Riggenbach, 1991). The comparison of the results regarding the use of pauses and hesitations indicated that the majority of participants managed to decrease the number of pauses or hesitations in Version 2, which contributed to higher perceived fluency scores. The reduction of pauses in Version 2 could be due to task repetition and familiarity with the content. Bui and Huang (2018) in their study observed that "content familiarity impacts the temporal aspects of fluency, especially breakdowns in the

middle of a clause” (p. 107). However, they also concluded that content familiarity does not help with decreasing the amount of fillers or repairs in L2 speech (Bui & Huang, 2018). Another reason why the participants might have reduced the number of pauses in Version 2 could be due to their overcoming of anxiety during the second delivery of the presentation. Similar findings were found in the study conducted by Qiu and Lo (2017), who reported that participants felt more relaxed during repetition of a narrative task on familiar topics. Finally, self-assessment could have played an important role in the reduction of pauses in the speech of participants of the present study (Chen, 2008). It can be argued that self-reflection brought awareness of pausing phenomenon and possibly, stimulated to decrease the number of pauses and hesitations in Version 2 in order to increase speech rate and improve oral performance, in general.

5.3.4. Language proficiency

In the current study, the fluency ratings assigned to the participants in relation to their language proficiency (non-temporal aspects of speech such as grammar and vocabulary) increased on Version 2. Even though the changes in the language proficiency turned out to be less noticeable when compared to the temporal variables discussed above, the increased scores still indicated some perceived progress in accuracy and lexical complexity as a part of language proficiency development due to task repetition (Gass et al., 1991). Complexity and accuracy are related to attention to the structure of L2 language and reflect learners’ attempts at producing more advanced language (Robinson, 2001; Skehan, 1998).

As pointed out by Lennon (1990), fluency in the so-called broad sense can mean general language proficiency implying that a fluent speaker has a high command of the foreign language. Consequently, language proficiency can usually be mistaken for fluency. Considering this idea, language proficiency was included in the fluency scale in order to give raters an opportunity to

provide a more comprehensive perception of the participants' fluency possible. Thus, they attended not only to the temporal variables of speech during their judgment, but also focused on L2 resources (syntax and lexis) by observing grammatical errors and the richness of lexical items in the presentations. One of the major drawbacks to using automatic scoring systems for measuring oral fluency, is that language proficiency construct is usually underrepresented in them (Chapelle & Chung, 2010). This limitation of the objective fluency measurement seems to highlight the benefits of the holistic judgment employed in the current study which involves the assessment of language proficiency, and thus considers multidimensional nature of L2 fluency, expressed by Segalowitz (2010).

The most obvious progress seen in language proficiency was in the reduced amount of grammatical errors in Version 2, which may suggest a positive effect of task repetition on accuracy. These results appear to align with the findings from previous research, which suggests that task repetition may lead to higher accuracy in L2 speech (Ahmadian & Tavakoli, 2011; Bygate, 1996; Lynch & Maclean, 2000; 2001). As previously mentioned, L2 learners usually tend to focus on form more during task repetition; this explains why their accuracy improves (Bygate, 1996; van de Guhte et al., 2016). The changes in overall proficiency also occur because L2 learners need less time for L2 speech processing during task repetition (Wang, 2014b). Another reason why language proficiency slightly improved in Version 2 might be explained by the students' involvement in self-assessment, which is considered a valuable tool for subjective L2 proficiency evaluation (Suzuki, 2015). Again, the evaluation rubric used for self-assessment and for the teacher's feedback did not include any comments on language resources. Still, it may be assumed that listening to the recordings could have raised the participants' linguistic self-

awareness and encouraged improvement of certain parts of the Version 2 presentations. This may have also contributed to some positive results in accuracy.

Overall, the reasons for overall small number of noticeable changes in the participants' interlanguage – defined as developing second language knowledge (Selinker, 1972) – might be explained by the complex phenomenon of language proficiency. Improving language proficiency after one task repetition seems unrealistic, and therefore was not the intent of this study.

Additionally, even though the teacher commented on some issues with grammar or vocabulary use, addressing and correcting all errors was not their aim. Instead, the teacher's feedback was oriented more towards general issues with pronunciation, presentation skills, and academic language because this group of learners needed more practice in delivering presentations and improving their oral production in English, especially their comprehensibility. This could be another reason why the participants did not have a very noticeable progress in the language proficiency development during task repetition – their main focus was more on L2 oral skills rather than on forms they used.

5.3.5. Summary of the discussion of research question 2

Overall, the analysis of the fluency profiles of Chinese L1 EAP students in their oral presentations revealed that as a group, they had some common issues with accented speech, low comprehensibility, and a big number of pauses and hesitations. Task repetition seemed to have a positive effect on speech rate, comprehensibility, the number of pauses, and language proficiency resulting in higher perceived fluency scores. Furthermore, it appears that the self-assessment and teacher's feedback employing the same evaluation rubric also brought awareness to the issues with the pronunciation, slow speed, or the large number of pauses, and motivated students to improve them in their Version 2. There were also found interrelationship between

normal speech rate and comprehensibility, as well as between speech rate and pausing phenomenon. Finally, the changes in language proficiency were less noticeable compared to the positive change found in the use of temporal variables. However, since task repetition seemed to stimulate the participants to focus more on form in Version 2, they had some perceived progress in increased lexical complexity and accuracy.

5.4. Discussion of research question 3

The present section discusses the results obtained for the third research question “*Does task repetition bring on FS and fluency gains (independently and/or together) in the presentations of Chinese L1 EAP students?*” Six participants out of ten showed more noticeable changes in both the use of formulaic language and perceived fluency development by having a high percentage of tokens and newly occurring FS types, as well as higher fluency scores on the second iteration of their oral presentations. Possible reasons explaining the changes in the performance of the six students are offered, with the task repetition being the leading factor for the observed progress. Then, the findings of a positive effect of task repetition on formulaic language of the six students are compared with the existing research on formulaic language instruction. The impact of task repetition on speech fluency is also elaborated upon. Finally, the role of formulaic language in oral fluency of the six students is discussed.

5.4.1. The role of task repetition in formulaic language and fluency development

The results of the present study indicated that task repetition seemed to promote an increased number and some variety of FS types, while also fostering fluency gains amongst all participants’ presentations. However, more noticeable progress in both areas during task repetition was observed in the speech of six students in particular. This is in contrast to the other

four participants who had some perceived improvement in fluency, but did not show signs of development in formulaic language performance between the two versions. It should be noted that the formulaic language of the four participants did not worsen, but rather there was no difference in terms of variety, and also no noticeable increase in the number of FS. Based on these findings, it can be argued that task repetition may promote both formulaic language use and a more fluent speech. Yet, L2 learners sometimes see fluency gains independently from formulaic language development while repeating the oral task (Bygate, 1996; Bygate & Samuda, 2005; Lynch & Maclean, 2000, 2001). The case of the six students showing more improvement than the rest seems to be representative of an L2 classroom, with some learners responding better to the task conditions than others. Considering the fact that 10 participants out of 26 were assigned to this study randomly, the result of 60% of the students showing noticeable progress in both formulaic language use and fluency development during task repetition seems to be promising.

It should be noted that the perceived changes in the oral performance of the six participants were determined by their initial performance before task repetition. Version 1 was viewed as a starting point for developing formulaic language and speech fluency on Version 2. Interestingly, the six students who showed improvement during task repetition fell into two different categories: those who already had a rather high command of English and whose oral performance was quite good on Version 1, and those whose proficiency seemed much lower and who used the smallest number of FS per group in Version 1. For example, Student 1, Student 2, Student 8, and Student 10 who showed good results in Version 1, doubled their FS number, increased the variety of the sequences, and were assigned higher fluency scores in Version 2. This may suggest that task repetition helped them further improve their performance. On the

other hand, the oral performance of Student 6 and Student 9 seemed quite poor on Version 1, but they managed to show positive changes in both FS use and fluency during the repeated task. In fact, Student 6 and Student 9 showed the highest percentage in the increase of newly occurring FS types and fluency development during task repetition amongst the whole group. It may be argued that after having assessed their first presentation and having received the teacher's feedback, Student 6 and Student 9 became aware of their weak performance and made an effort to improve it in Version 2. There is a possibility that they were also interested in receiving a better grade. For instance, Student 6 increased the length of his presentation by adding new content that probably also contributed to his perceived progress. Overall, the following findings suggest that the extent to which formulaic language use and fluency might improve under the conditions of task repetition seems to be rather unique and highly individual. Based on the results of the current study, it appears that both strong and weak students may consider task repetition as an opportunity to improve their L2 oral performance (Bui & Huang, 2018).

The reason why task repetition affected the oral performance of the six students out of ten is probably because they responded to other factors which could have also contributed to the improvement in their oral performance. As mentioned before, among them could be learners' reflection on their first attempt (Chen, 2008; Strachan et al., 2019). For example, self-assessment could have helped some students not only with their fluency development, but also could have brought linguistic self-awareness (Young & West, 2018) and possibly, stimulated lexical richness and complexity leading to some development in formulaic language. Then again, the six students may have felt more than others the need to organize their discourse by using certain FS during task repetition (Qiu & Lo, 2017). Even though the motivation of the participants was not measured, it may be argued that task repetition could have brought on task motivation in the six

students, thus contributing to their task performance (Dörnyei, in press). This could be due to the task engagement of the six students, who could have found task participation enjoyable, or perhaps they were interested in reaching an individual goal upon completion (Henry, Dörnyei, & Davidenko, 2015). The emotional factor in task engagement may also be explained by the fact that the participants reported having no previous experience recording themselves in English and thus, they found the task repetition very motivating and useful for evaluating their own progress.

Thus, it is unknown which factors besides task repetition turned out to be determining for the perceived progress in both areas on Version 2. However, it can be assumed that since the six participants had some changes in both formulaic language use and fluency during task repetition, they could have been responsive to the majority of the factors. As well, they have possibly found task repetition engaging, which increased their task motivation and led to improvements in their oral production (Henry et al., 2015).

5.4.1.1. Task repetition and formulaic language. As previously mentioned, the participants did not know that their use of formulaic language would be analyzed, and using formulaic language was not an explicit goal within the task. Thus, it may be assumed that the students were not even aware that the increased total number of tokens and variety of FS happened at all. Moreover, since the ESLA program does not address formulaic language explicitly, it appears that the majority of participants were not aware of formulaic language and its important role in oral production (Boers et al., 2006; Stengers et al., 2011; Wood, 2009). Nonetheless, even though the learning experience of the participants may suggest that there was no awareness raising instruction on formulaic language and thus, did not really have resources to improve, six of the participants still used formulaic language in their oral presentations extensively, and even increased its number and variety during task repetition. It can be argued

that the participants used formulaic language in their presentations regardless of instruction because of the abundance and frequency of FS in the English language. For instance, previous corpus-based research has found that formulaic language constitutes 80 percent of English spoken language (Altenberg, 1998). Therefore, participants may have encountered some FS which they used in their presentations through repeated exposure to input and acquired these sequences without necessarily realizing their formulaic nature. Thus, the six participants modified and improved their formulaic language performance without realizing it under the effect of task repetition and with no pedagogical intervention. Considering the importance of practice and automatization of FS reported in the literature (Gatbonton & Segalowitz, 2005; Logan, 1988; Wray & Fitzpatrick, 2008; Wood, 2015, p. 152), it seems that task repetition procedure has a potential for promoting acquisition and use of formulaic language.

5.4.1.2. Task repetition and perceived fluency. In contrast to the lack of empirical evidence in the literature on the role of task repetition in formulaic language, the connection between task repetition and L2 fluency development has been evidenced by previous research (Ahmadian & Tavakoli, 2011; Bygate, 1996; Gass et al., 1991; Lynch & Maclean, 2001). The increased fluency scores of the participants of the current study for speech rate, comprehensibility, the number of pauses and hesitations, and language proficiency during task repetition seem to support its positive effect on L2 speech shown in the literature. Overall, all participants were perceived to be more fluent on Version 2. However, the six students in particular had more difference in the fluency scores between two versions when compared to the other participants. As previously mentioned, other factors could have also contributed to the improvement caused by task repetition. Most importantly, having self-reflection and the teacher's feedback afterwards made a difference, as it helped raise awareness of the issues which

students had (Castañeda & Rodríguez-González, 2011; Sheppard, 2006). This gave the participants resources for improving their Version 2. Furthermore, since students had an opportunity to record Version 2 from home, they may have used this as a chance to review and evaluate their speech as many times as they needed in order to be satisfied with their attempt, thus increasing self-awareness and determining what areas they needed to focus on (Young & West, 2018). This may lead to the assumption that the six participants whose performance seemed more improved on Version 2 could have engaged in task repetition procedure multiple times having a chance for more speaking practice as a result.

5.4.2. The role of formulaic language in speech fluency

The findings of the present study showed that task repetition, along with other factors such as self-assessment, for example, brought on gains in both formulaic language performance and perceived fluency in the oral presentations of the six students out of ten. However, it can be also argued that the interrelationship between formulaic language use and speech fluency themselves could be another reason why the six participants had some improvements in these two areas simultaneously during task repetition.

It has been fairly well-established in the literature that formulaic language is an important part of spoken language and may facilitate fluent speech of L2 learners (Wood, 2015). Mastery of FS seems to assist learners to come across as more native-like, and thus, more fluent (Pawley & Syder, 1983; Towell et al., 1996; Wood, 2006, 2009, 2010). For example, an interesting study conducted by Wood (2010), which investigated the role of formulaic language in L2 fluency, demonstrated that the use of FS may have facilitated the reduction of pauses and increased length of runs, thus having a positive effect on fluency of the participants. Similarly, in the present study, the use of some FS (i.e., fillers) helped some participants to reduce pauses and receive

higher fluency ratings as a result. Furthermore, the six participants used many discourse organizers and referential expressions in Version 2 which helped them to structure their discourse and also made them sound more fluent, proficient and confident during task repetition (Qiu & Lo, 2017). Another interesting pattern and direct link between formulaic language use and perceived fluency was found in the performance of Student 6 and Student 9 who both, as previously mentioned, showed the largest increased percentage for their fluency scores, tokens and newly occurring FS types during task repetition. The analysis of their first attempt revealed that they had a rather small number of FS types and their fluency was perceived as very low by the researcher and the raters as well (e.g., Student 6 used 5 FS and Student 9 had 4 FS). In contrast, when these participants increased their variety and total number of FS on Version 2, they also received much higher fluency scores. This may suggest that possibly their use of formulaic language during task repetition may have affected the raters' perception of their speech fluency somehow.

Finally, it is important to note that in the current study, whenever the judges or the researcher assigned higher fluency scores for speech rate, comprehensibility, and the use of pauses they also tended to highlight the participant's use of lexical bundles and other useful expressions in their comments about language proficiency. This observation seems to support the idea that the use of formulaic language may contribute to higher fluency (Pawley & Syder, 1983; Wood, 2006, 2009). The six students who managed to improve in both formulaic language and fluency development were the ones who received those comments from the raters.

5.4.3. Summary of the discussion of research question 3

The discussion of the research question 3 demonstrated that task repetition may indeed bring on FS and fluency gains both independently and together in the presentations of Chinese

L1 EAP students. Four participants out of ten did not show noticeable difference in the use of formulaic language but had higher fluency scores during task repetition. In contrast, the six students out of ten managed to show positive results in both areas by increasing variety and total number of FS and receiving higher fluency gains during task repetition. This may suggest that they could have responded better to other diverse factors besides task repetition aiming to improve their performance, such as the teacher's feedback on pronunciation and presentations skills and self-assessment of their initial performance. Moreover, the six participants could have been more motivated than the rest to advance their Version 2 and found task very engaging. The comparison of the task repetition procedure and the previous research on formulaic language instruction suggests that task repetition has potential benefits for acquisition, practice and retrieval of FS and may promote some formulaic language use in L2 speech even without any pedagogical intervention. The findings of the present study, which revealed that almost all participants received higher fluency scores on the second iteration of their oral presentations, support a positive effect of task repetition on L2 speech found in the literature. Finally, the positive interrelationship between high fluency scores and the increased variety and total number of FS in the performance of the six participants on Version 2 demonstrates how an extensive use of FS and their functions may also contribute to the change in fluency perception.

5.5. Pedagogical implications

Several pedagogical implications might be drawn from the results of the present study for curriculum designers, teachers and learners of English as a second/foreign language (ESL/EFL). To begin, the findings of this study suggest a positive effect of task repetition on fluency and the number and variety of FS in L2 speech. Ferris and Tagg (1996a) argue that in EAP classes,

teachers tend to focus more on writing and reading skills, and thus task repetition might be suggested for EAP instructors as a way to focus on speaking and to give students an opportunity to practice delivering oral presentations. On the other hand, ESL programs are usually oriented towards the development of the communicative competence and already have a strong focus on speaking (Savignon, 2002). Nevertheless, ESL practitioners might benefit from including task repetition in the curriculum as well because it will be a useful practice for students which would foster their communicative competence and an opportunity to improve their accuracy by focusing on form (Bygate, 2001). In contrast, unlike EAP or ESL instructors, some EFL teachers are known for their resistance to TBLT because they are concerned about its capability of focusing on specific grammar forms (Sato, 2010), even though supporters of this approach view focus on form during meaningful communication as a way of practicing and learning grammar (Ellis, 2009a). Implementing task repetition in those curricula could be viewed as an opportunity not only to introduce a new method of teaching a foreign language into the classroom, but also a chance to improve students' accuracy, complexity, fluency, and possibly, formulaic language (Ahmadian & Tavakoli, 2011; Bygate, 1996; Lynch & Maclean, 2001).

The findings of this research could also be meaningful for formulaic language pedagogy. Existing approaches to teaching formulaic language involve a focused instruction which raises awareness of FS in input (Lewis, 1997; Willis, 1990). However, the findings of the current study demonstrate that even without deliberate attention to the phenomenon of formulaic language, L2 learners might still have FS in their speech and even improve their number and variety due to task repetition. Indeed, the pedagogical methods related to formulaic language seem rather old-fashioned following a present-practice-produce approach where the teacher structures the lesson by introducing the target language explicitly and engaging students in a series of different control

activities designed for its practice and use (Wood, 2015). On the other hand, task repetition, as a part of a TBLT methodology, can be a different and new way of including formulaic language in classroom practice. Interestingly, it appears that raising awareness of FS in input does not always lead to automatization (Wood, 2015), therefore some researchers have focused on including that element in the formulaic language instruction (Gatbonton & Segalowitz, 2005; Wray & Fitzpatrick, 2010). Consequently, task repetition may be used as a technique which would encourage memory and automatization of FS. However, it is important to note that the chances for more noticeable progress in the use of formulaic language would probably be higher during task repetition if L2 learners were aware of FS in their speech before repeating the task.

Overall, it seems that the strength of the present study is that it has been carried out in the real EAP classroom and neither Kartchava and Nassaji (2019) nor the researcher of the present study had any control over the task repetition procedure and design which were employed by the instructor. For that reason, it can be argued that this research is ecologically valid and its findings have important implications for language pedagogy.

5.6. Limitations and future research suggestions

Despite the positive findings of the present study, some limitations need to be acknowledged. First, the identification of formulaic language in transcripts or spoken texts is challenging and “regardless of the measures... absolute certainty is elusive” (Wood, 2015, p. 32). Having native speakers of English as independent raters, who based their judgment on intuition and on Wray and Namba’s (2003) comprehensive checklist, seemed appropriate for a small sample of the present study (Wray, 2002). However, the comparison of the findings from the researcher’s analysis and the judges’ analysis, revealed that there was some variation between

judges in terms of the number of FS which was identified. Even though a discussion at the benchmarking session was included to address what Wray (2002) identifies as the problematic potential for variation among judges, there were still some differences in the findings after the judges individually marked sequences as formulaic. The variation might have been resolved if there had been a chance to have a follow-up discussion with the judges. It is also possible that factors such as fatigue and inconsistencies between individual's assessments of what constitutes formulaic language played a role in the final results. For example, many of the FS identified by the researcher were simply missed by the judges, and vice versa.

Second, it should be noted that despite the fact that there was a large number of lexical bundles in the participants' presentations (serving the discourse organizer function, referential expression function and stance expression function), finding them was not an explicit goal of the current study. Lexical bundles represent a unique category of formulaic language: they are not necessarily retrieved as wholes, unlike other FS; they are usually longer than other FS and contain three or more words; and they are identified by means of corpus analysis software programs, using a specific frequency cutoff (Wood, 2015). Moreover, according to Schmitt, Grandage and Adolphs (2004), not all lexical bundles are stored in mind as FS. It was decided in the present research, for the aforementioned reasons, to avoid focusing on the identification of all possible lexical bundles in the presentations. Therefore, there is a chance that some lexical bundles in the participants' presentations were either missed by the independent judges and the researcher or were not identified as FS because they did not seem to be salient FS. As a result, it could be argued that the actual number of lexical bundles in the presentations was different from that which was reported.

Another challenge faced by the current study is the lack of a catalogue of categories which describes FS use and functions in L2 spoken data in a specific genre (Wood, 2015). The functional taxonomy of Biber et al. (2004) is the only established, standard type classification for categorizing lexical bundles and was, therefore, chosen in the current research. However, lexical bundles and their categories are not necessarily relevant to discussion of formulaic language as a whole. Therefore, it seemed important to complement the taxonomy of Biber et al. (2004) with further functional categories to take into consideration other types of FS encountered in the participants' presentation. Several new functions were created, namely the figurative function, the other function and, finally, the filler function, which was included under the category of the special conversational functions.

Another limitation of the current study concerns perceived fluency judgment. There is no tool in the literature which can serve as a guide to measure perceived L2 fluency. The Likert-type fluency scale employed by the researcher and the independent raters was created for the purposes of this study and took into consideration Segalowitz's (2010) fluency criteria: speech rate, the comprehensibility of the speech, pauses and hesitations, and general language proficiency. Even though the independent raters participated in the present study to ensure data triangulation, there was some variation found between their fluency scores and the researcher's ratings. For that reason, it could be argued that the use of an analytic measurement system for assessing fluency might have strengthened the findings in the present study.

Finally, the conditions of the participants' presentations were different: the first presentation was delivered in front of the class, while the second one was recorded from home. Even though it is a common classroom procedure, which adds to the ecological validity of this study, different conditions of the presentation delivery might have played some role in the

findings. Future research could consider this limitation and make the two conditions equal. Moreover, future research might choose to include a larger sample and compare the effect of task repetition on formulaic language and perceived fluency between the control and experimental groups. For example, in the present study the participants did not receive an explicit instruction on formulaic language but in a future study an experimental group could be educated on the important role of FS in a fluent speech and be encouraged to notice them in the input. Another issue that could be addressed in the future is adding objective fluency measurement to perceived fluency judgment in order to reach more objective results. Ultimately, there is a call for more research which would investigate the role of formulaic language in speech fluency and for more studies which would attempt to categorize the functions of formulaic language in discourse.

References

- Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, fluency, and complexity of EFL learners' oral production. *Language Teaching Research*, 15(1), 35–59.
- Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford, UK: Oxford University Press.
- Barnard, R., & Scampton, D. (2008). Teaching grammar: A survey of EAP teachers in New Zealand. *New Zealand Studies in Applied Linguistics*, 14(2), 59–82.
- Berman, R. (2002, May). Toward a common concept of English for academic purposes. In L. Cheng (Chair), *Concepts and contexts in EAP: Preparing NNS of English for success in the academic community*. Symposium conducted at the Canadian Association of Applied Linguists, Congress of the Social Sciences and Humanities, Toronto, Canada.
- Berman, R., & Cheng, L. (2001). English academic language skills: Perceived difficulties by undergraduate and graduate students, and their academic achievement. *Canadian Journal of Applied Linguistics*, 4(1–2), 25–40.
- Biber, D., Conrad, S., & Cortes, V. (2004). If you look at...: Lexical bundles in university teaching and textbooks. *Applied Linguistics*, 25(3), 371–405.
- Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Harlow, England: Longman.
- Bod, R. (2000). The storage vs. computation of three-word sentences. Paper presented at AMLaP2000, University of Leiden, Leiden, the Netherlands.
- Boers, F. (2014). A reappraisal of the 4/3/2 activity. *RELC Journal*, 45(3), 221–235.

- Boers, F., Deconinck, J., & Lindstromberg, S. (2010). Choosing motivated chunks for teaching. In S. De Knop, F. Boers, & T. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 239–256). Berlin, Germany: Mouton de Gruyter.
- Boers, F., Eyckmans, J., & Stengers, H. (2007). Presenting figurative idioms with a touch of etymology: More than mere mnemonics? *Language Teaching Research*, *11*(1), 43–62.
- Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research*, *10*(3), 245–261.
- Bosker, H. R., Pinget, A., Quene, H., Sanders, T., & de Jong, N. H. (2013). What makes speech sound fluent? The contributions of pauses, speed and repairs. *Language Testing*, *30*(2), 159–175.
- Bui, G., Ahmadian, M. J., & Hunter, A. (2019). Spacing effects on repeated L2 task performance. *System*, *81*(1), 1–13.
- Bui, G., & Huang, Z. (2018). L2 fluency as influenced by content familiarity and planning: Performance, measurement, and pedagogy. *Language Teaching Research*, *22*(1), 94–114.
- Bygate, M. (1996). Effects of task repetition: Appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and change in language teaching* (pp. 136–146). Oxford: Macmillan Heinemann.
- Bygate, M. (2001). Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching and testing* (pp. 23–48). Harlow: Longman.
- Bygate, M. (2016). Sources, developments and directions of task-based language teaching. *The Language Learning Journal*, *44*(4), 381–400.

- Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task repetition. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 37–74). Amsterdam: John Benjamins.
- CAEL. (n.d.). Retrieved from <https://www.cael.ca/>
- Castañeda, M., & Rodríguez-González, E. (2011). L2 speaking self-ability perceptions through multiple video speech drafts. *Hispania*, 94(3), 483–501.
- Cenoz, J. (1998). *Pauses and communication strategies in second language speech*. ERIC Document ED 426630. Rockville, MD: Educational Resources Information Center.
- Chambers, F. (1997). What do we mean by fluency? *System*, 25(4), 535–544.
- Chapelle, C., & Chung, Y. (2010). The promise of NLP and speech processing technologies in language assessment. *Language Testing*, 27(3), 301–315.
- Chen, Y. M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language Teaching Research*, 12(2), 235–262
- Cheng, L., & Fox, J. (2008). Towards a better understanding of academic acculturation: Second language students in Canadian universities. *The Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes*, 65(2), 307–333.
- Clapham, C. (2000). Assessment and testing. *Annual Review of Applied Linguistics*, 20, 147-161.
- Clark, H. H., & Fox Tree, J. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84, 73–111.
- Coulmas, F. (1979). On the sociolinguistic relevance of routine formulae. *Journal of Pragmatics*, 3(3-4), 239–266.
- Credit ESL. (n.d.). Retrieved from <https://carleton.ca/slals/esl/credit/>
- de Bot K (1992). A bilingual production model: Levelt's speaking model adapted. *Applied Linguistics*, 13(1), 1–24.

- Dechert, H. W. (1980). Pauses and intonation as indicators of verbal planning in second-language speech productions: Two examples from a case study. In H.W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 271–285). The Hague: Mouton.
- Dechert, H. W., & Raupach, M. (1987). *Psycholinguistic models of production*. Tübingen, Germany: Gunter Narr Verlag.
- De Jong, N., & Perfetti, C. A. (2011). Fluency training in the ESL classroom: An experimental study of fluency development and proceduralization. *Language Learning*, 61(2), 533–568
- Derwing, T. M. (1990). Speech rate is no simple matter: Rate adjustment and NS-NNS communicative success. *Studies in Second Language Acquisition*, 12(3), 303–313.
- Derwing, T. M., & Munro, M. J. (2009). Putting accent in its place: Rethinking obstacles to communication. *Language Teaching*, 42(4), 476–490.
- Derwing, T., Rossiter, M., Munro, M., & Thomson, R. (2004). Second language fluency: Judgments on different tasks. *Language Learning*, 54(4), 655–679.
- Derwing, T. M., Thomson, R. I., & Munro, M. J. (2006). English pronunciation and fluency development in mandarin and Slavic speakers. *System: An International Journal of Educational Technology and Applied Linguistics*, 34(2), 183–193.
- Deschamps, A. (1980). The syntactical distribution of pauses in English spoken as a second language by French students. In H.W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 255–262). The Hague: Mouton.
- Dickinson, P. (2013). 'B/w U & me': The functions of formulaic language in interactional discourse on twitter. *The Linguistics Journal*, 7(1), 7–38.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.

- Dörnyei, Z. (in press). Task motivation: What makes an L2 task engaging? In Z. E. Wen & M. J. Ahmadian (Eds.), *Researching L2 task performance and pedagogy: In honour of Peter Skehan* (pp. 53–66). Amsterdam: John Benjamins.
- Ejzenberg, R., 2000. The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In: Riggenbach, H. (Ed.), *Perspectives on fluency* (pp. 287–314). Michigan: University of Michigan Press.
- Ellis, N. C., Simpson-Vlach, R., & Maynard, C. (2008). Formulaic language in native and second language speakers: Psycholinguistics, corpus linguistics, and TESOL. *TESOL Quarterly*, 42(3), 375–396.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Ellis, R. (2009a). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics*, 19(3), 221–246.
- Ellis, R. (2009b). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 30(4), 474–509.
- Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text*, 20(1), 29–62.
- ETS TOEFL. (n.d.). Retrieved from <https://www.ets.org/toefl/contact/canada>
- Evans, S., & Green, C. (2007). Why EAP is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6(1), 3–17.
- Eyckmans J., Boers, F., & Stengers, H. (2007). Identifying chunks: Who can see the wood for the tress? *Language Forum*, 33(2), 85–100.
- Ferris, D., & Tagg, T. (1996a). Academic oral communication needs of EAP learners: What subject-matter instructors actually require. *TESOL Quarterly*, 30(1), 31–58.

- Ferris, D., & Tagg, T. (1996b). Academic listening/speaking tasks for ESL students: Problems, suggestions, and implications. *TESOL Quarterly*, 30(2), 297–320.
- Fillmore, C. (1979). On fluency. In C. Fillmore, D. Kempler, and W. Wang (eds.), *Individual differences in language ability and language behavior* (pp. 85–101). New York: Academic Press.
- Foster, P. (2001). Rules and routines: A consideration of their role in task-based language production of native and non-native speakers. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 75 – 93). Harlow, UK: Longman.
- Fox Tree, J. (2001). Listeners' use of u and uh in speech comprehension. *Memory and Cognition*, 29(2), 320–326.
- Fox, J., Cheng, L., & Zumbo, B. D. (2014). Do they make a difference? The impact of English language programs on second language students in Canadian universities. *TESOL Quarterly*, 48(1), 57–85.
- Freed, B. F. (1995). What makes us think that students who study abroad become fluent? In B.F. Freed (Ed.), *Second language acquisition in a study abroad context* (pp. 123–148). Philadelphia: John Benjamins.
- Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, 20(3), 321–340.
- Fulcher, G. (1996). Does thick description lead to smart tests? A data-based approach to rating scale construction. *Language Testing*, 13(2), 208–238.
- Gass, S., Mackey, A., Alvarez-Torres, M.J., & Fernandez-Garcia, M. (1999). The effects of task repetition on linguistic output. *Language Learning*, 49(4), 549–581.

- Gibbs, R., & Gonzales, G. (1985). Syntactic frozenness in processing and remembering idioms. *Cognition*, 20(3), 243–259.
- Ginther, A., Dimova, S., & Yang, R. (2010). Conceptual and empirical relationships between temporal measures of fluency and oral English proficiency with implications for automated scoring. *Language Testing*, 27(3), 379–399.
- Goh, C. C. M. (2007). *Teaching speaking in the language classroom*. Singapore: SEAMEO Regional Language Centre.
- Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. New York: Academic Press.
- Granger, S. (1998). Prefabricated patterns in advanced EFL writing: Collocations and formulae. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis and applications* (pp. 145–160). Oxford: Clarendon Press.
- Grosjean, F. (1980). Temporal variables within and between languages. In H. W. Dechert & M. Raupach, (Eds.), *Towards a cross-linguistic assessment of speech production* (pp. 39–53). Frankfurt: Lang.
- Hahn, L. D. (2004). Primary stress and intelligibility: Research to motivate the teaching of suprasegmentals. *TESOL Quarterly*, 38(2), 201–222.
- Hasselgreen, A. (2004). *Testing the spoken English of young Norwegians: A study of test validity and the role of 'smallwords' in contributing to pupils' fluency*. Cambridge, UK: Cambridge University Press.
- Hawkins, P. R. (1971). The syntactic location of hesitation pauses. *Language and Speech*, 14, 277–288.

- Henry, A., Dornyei, Z., & Davydenko, S. (2015). The anatomy of directed motivational currents: Exploring intense and enduring periods of L2 motivation. *The Modern Language Journal*, 99(2), 329–345.
- Hunter, A. M. (2017). *Fluency development in the ESL classroom: The impact of immediate task repetition and procedural repetition on learners' oral fluency* (Unpublished doctoral dissertation). St. Mary's University/University of Surrey, Guildford, England.
- IELTS Official Test Centre. (n.d.). Retrieved from <https://ieltscanada.ca>
- IELTS Scoring Guide. (n.d.). Retrieved from <https://www.ieltsessentials.com/global/results/assessmentcriteria>
- Isaacs, T., & Trofimovich, P. (2012). Deconstructing comprehensibility: Identifying the linguistic influences on listeners' L2 comprehensibility ratings. *Studies in Second Language Acquisition*, 34(3), 475–505.
- Jan, C-W., Li, B-J., & Lin, C-C. (2014). The Use of the Mother Tongue in Chinese EFL Classrooms. *Journal of China University of Science and Technology*, 58, 161–181.
- Johnstone, R. (1988). *Communicative Interaction: A Guide for Language Teachers*. London: CELT.
- Jones, M. & S. Haywood (2004). Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 269–291). Amsterdam: John Benjamins.
- Kartchava, E., & Nassaji, H. (2019). The role of task repetition and learner self-assessment in technology-mediated task performance. *International Journal of Applied Linguistics*. Advanced online publication. <https://doi.org/10.1075/itl.19013.kar>

- Kim, S. (2006). Academic oral communication needs of East Asian international graduate students in non-science and non-engineering fields. *English for Specific Purposes*, 25(4), 479-489.
- Koponen, M., & Riggenbach, H. (2000). Overview: Varying perspectives on fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 5–24). Ann Arbor, MI: The University of Michigan Press.
- Kormos, J. (2006). *Speech production and second language acquisition*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32(2), 145–164.
- Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language speech processing. *Studies in Second Language Acquisition*, 39(1), 167–196.
- Laufer, B. & Waldman, T. (2011). Verb-noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning*, 61, 647–672.
- Lennon, P. (1984). Retelling a story in English. In H. W. Dechert, D. Möhle & M. Raupach (Eds.), *Second language productions* (pp. 50–68). Tübingen: Gunter Narr Verlag.
- Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning* 40(3), 387–412.
- Levelt, W. (1989). *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Lewis, M. (1993). *The lexical approach*. Hove, UK: LTP.
- Lewis, M. (1997). Pedagogical implications of the lexical approach. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 255–270). Cambridge: Cambridge University Press.

- Li, W., Zhang, X., Niu, C., Jiang, Y., & Srihari, R. (2003). An expert lexicon approach to identifying English phrasal verbs. In *Proceedings of the 41st annual meeting of the ACL* (pp. 513–520). Sapporo: Japan.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, 95(4), 492–527.
- Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4(3), 221–250.
- Lynch, T., & Maclean, J. (2001). A case of exercising: Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 141–162). Harlow: Longman.
- Mackey, A., & Gass, S. M. (2015). *Second language research: Methodology and design*. London: Routledge.
- McGuire, M., & Larson-Hall, J. (2017). Teaching formulaic sequences in the classroom: Effects on spoken fluency. *TESL Canada Journal*, 34(3), 1–25.
- Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111–129.
- Moon, R. (1998). Frequencies and forms of phrasal lexemes in English. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications* (pp. 79–100). Oxford: Clarendon.
- Munro, M. J., & Derwing, T. M. (1999). Foreign accent, comprehensibility, and intelligibility in the speech of second language learners. *Language Learning*, 45(1), 285–310.
- Munro, M. J., & Derwing, T. M. (2001). Modeling perceptions of the accentedness and comprehensibility of L2 speech in the role of speaking rate. *Studies in Second Language Acquisition*, 23(4), 451–468.

- Myers, L. (1996). The contribution of genre theory to theme-based EAP: Navigating foreign fiords. *TESL Canada Journal*, 13(2), 33–45.
- Nakano, M., Kondo, Y., & Tsutsui, E. (2008). Fundamental research on automatic speech evaluation of L2 English. Presented at the Distance Learning and the Internet Conference, Tokyo, Japan.
- Nattinger, J. R., & DeCarrico, J. S. (1992). *Lexical phrases and language teaching*. Oxford, UK: Oxford University Press.
- O'Connell, D.C., & Kowal, S. (2004). The history of research on the filled pause as evidence of *The Written Language Bias in Linguistics* (Linell, 1982). *Journal of Psycholinguistic Research*, 33(6), 459–474.
- Oscarson, M. (2013). Self-assessment in the classroom. In A. Kunnan (Ed.), *The companion to language assessment vol. II: Approaches and development* (pp. 712–729). New York: Wiley-Blackwell.
- Pawley, A. & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and communication* (pp. 191–226). New York, NY: Longman.
- Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-drawing techniques. *Language Learning Journal*, 40(1), 65–79.
- Préfontaine, Y., & Kormos, J. (2016). A qualitative analysis of perceptions of fluency in second language French. *IRAL, International Review of Applied Linguistics in Language Teaching*, 54(2), 151–169.
- Qi, Y., & Ding, Y. (2011). Use of formulaic sequences in monologues of Chinese EFL learners. *System*, 39(2), 164–174.

- Qiu, X., & Lo, Y. Y. (2017). Content familiarity, task repetition and Chinese EFL learners' engagement in second language use. *Language Teaching Research*, 21(6), 681–698.
- Raupach, M. (1980). Temporal variables in first and second language speech production. In H. W. Dechert, & M. Raupach (Eds.), *Temporal variables in speech: Studies in honour of Frieda Goldman-Eisler* (pp. 271–285). The Hague: Mouton.
- Riggenbach, H. (1991). Towards an understanding of fluency: A microanalysis of nonnative speaker conversation. *Discourse Processes*, 14(4), 423–441.
- Robinson, P. (2001). Task complexity, task difficulty and task production: Exploring interactions in a componential framework. *Applied Linguistics*, 22(1), 27–57.
- Sato, R. (2010). Reconsidering the effectiveness and suitability of PPP and TBLT in the Japanese EFL classroom. *JALT Journal*, 32(2), 189–200.
- Sajavaara, K. (1987). Second language speech production: Factors affecting fluency. In H. W. Dechert & M. Raupach (Eds.), *Psycholinguistic models of production* (pp. 45–65). Norwood, NJ: Ablex.
- Savignon, S. (2002). *Interpreting communicative language teaching*. New Haven, CT: Yale University Press.
- Schmidt, R.W. (1983). Interaction, acculturation, and the acquisition of communicative competence: A case study of an adult. In N. Wolfson & E. Judd (Eds.), *Sociolinguistics and language acquisition* (pp. 137–174). Rowley, MA: Newbury House.
- Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-derived recurrent clusters psycholinguistically valid? In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 127–151). Philadelphia, PA: John Benjamins.

- Schmitt, N., & Underwood, G. (2004). Exploring the processing of formulaic sequences through a self-paced reading task. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing, and use* (pp. 173–190). Philadelphia, PA: John Benjamins.
- Segalowitz, N. (2010). *The cognitive bases of second language fluency*. New York: Routledge.
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10(2), 209–231.
- Sheppard, C. (2006). *The effect of instruction directed at the gaps second language learners noticed in their oral production* (Unpublished doctoral dissertation). University of Auckland, New Zealand.
- Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding more fuel to the fire: An eye-tracking study of idiom processing by native and non-native speakers. *Second Language Research*, 27(2), 1–22.
- Skehan, P. (1998). *A cognitive approach to language learning*. Oxford: Oxford University Press.
- Skehan, P. (2003). Task-based instruction. *Language Teaching*, 36(1), 1–14.
- Stengers, H., Boers, F., Housen, A., & Eyckmans, J. (2010). Does “chunking” foster chunk-uptake? In S. De Knop, F. Boers & A. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 99–117). Berlin, Germany: Mouton de Gruyter.
- Strachan, L., Kennedy, S., & Trofimovich, P. (2019). Second language speakers’ awareness of their own comprehensibility. *Journal of Second Language Pronunciation*.
- Suzuki, Y. (2015). Self-assessment of Japanese as a second language: The role of experiences in the naturalistic acquisition. *Language Testing*, 32(1), 63–81.
- Tavakoli, M. (2010). Investigating the relationship between self-assessment and teacher-assessment in an academic context. *Asian EFL Journal*, 12(1), 234–260.
- Thomson, H. (2017). Building speaking fluency with multiword expressions. *TESL Canada Journal*, 34(3), 26–53.

- Towell, R., Hawkins, R., Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics* (17), 84–119.
- Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences: acquisition, processing, and use* (pp. 153–172). Philadelphia, PA: John Benjamins.
- van de Guchte, M., Braaksma, M., Rijlaarsdam, G., & Bimmel, P. (2016). Focus on form through task repetition in TBLT. *Language Teaching Research*, 20(3), 300–320.
- Wajnryb, R. (1990). *Grammar dictation*. Oxford: Oxford University Press.
- Wang, Z. (2014a). Developing Accuracy and Fluency in Spoken English of Chinese EFL Learners. *English Language Teaching*, 7(2), 110–118.
- Wang Z. (2014b). On-line time pressure manipulations: L2 speaking performance under five types of planning and repetition conditions. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 27–62). Amsterdam: John Benjamins.
- Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning*, 63(1), 91–120.
- Will, N. L. (2016). From isolation to inclusion: Learning of the experiences of Chinese international students in U.S. *Journal of International Students*, 6(4), 1069–1075.
- Willis, D. (1990). *The lexical syllabus: A new approach to language teaching*. London: Harper Collins.
- Willis, J. (1996). *A framework for task-based learning*. Harlow: Longman
- Wood, D. (1998). *Making the grade: An interactive course in English for academic purposes*. Toronto: Prentice Hall Allyn and Bacon.

- Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *The Canadian Modern Language Review*, 63(1), 13–33.
- Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *Canadian Journal of Applied Linguistics/Revue Canadienne De Linguistique Appliquee*, 12(1), 39–57.
- Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence, and classroom applications*. London/New York: Continuum.
- Wood, D. (2015). *Fundamentals of formulaic language: An introduction*. London: Bloomsbury.
- Wood D., & Namba, K. (2013). Focused instruction of formulaic language: Use and awareness in a Japanese university class. In *The Asian Conference on Language Learning Official Conference Proceedings 2013*, Osaka, Japan.
- Woodrow, L. (2006). Anxiety and speaking English as a second language. *RELC Journal*, 37(3), 308–328.
- Wray, A. (2000). Formulaic sequences in second language teaching: principle and practice. *Applied Linguistics*, 21(4), 463–489.
- Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.
- Wray, A. & Fitzpatrick, T. (2008). Why can't you just leave it alone? Deviations from memorized language as a gauge of nativelike competence. In F. Meunier & S. Granger (Eds.), *Phraseology in language learning and teaching* (pp. 123–148). Amsterdam: John Benjamins.
- Wray, A., & Fitzpatrick, T. (2010). Pushing learners to the extreme: The artificial use of prefabricated material in conversation. *Innovation in Language Learning and Teaching*, 4(1), 37–51.

- Wray, A., & Namba, K. (2003). Use of formulaic language by a Japanese-English bilingual child: A practical approach to data analysis. *Japanese Journal for Multilingualism and Multiculturalism*, 9(1), 24–51.
- Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication*, 20(1), 1–28.
- Xi X., & Mollaun P. (2006). *Investigating the utility of analytic scoring for the TOEFL academic speaking test (TAST)*. Princeton, NJ: Educational Testing Service.
- Young, E. H., & West, R. E. (2018). Speaking practice outside the classroom: A literature review of asynchronous multimedia-based oral communication in language learning. *EuroCALL Review*, 26(1), 59–78.
- Yu, X. (2009). A formal criterion for identifying lexical phrases: Implication from a classroom experiment. *System: An International Journal of Educational Technology and Applied Linguistics*, 37(4), 689–699.
- Yuan, W. (2011). Academic and cultural experiences of Chinese students at an American university: A qualitative study. *Intercultural Communication Studies*, 20(1), 141–157.
- Yuan, F. & Ellis, R. (2003). The effect of pre-task planning and online planning on fluency, complexity, and accuracy in L2 oral production. *Applied Linguistics*, 24, 1–27.

Appendix A

Secondary data clearance



Office of Research Ethics
5110 Human Computer Interaction Bldg | 1125 Colonel By Drive
Ottawa, Ontario K1S 5B6
613-520-2600 Ext: 4085
ethics@carleton.ca

CERTIFICATION OF INSTITUTIONAL ETHICS CLEARANCE

This document is an administrative clearance corresponding to approved protocol **#106793** from the Carleton University Research Ethics Board. The Carleton University Research Ethics Board-A (CUREB-A) has granted administrative clearance for the research project described below and research may now proceed. CUREB-A is constituted and operates in compliance with the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans* (TCPS2).

Ethics Protocol Clearance ID: Project # 109310

Research Team: Yana Lysiak (Primary Investigator)
Dr. Eva Kartchava (Research Supervisor)
David Wood (Research Supervisor)

Project Title: The impact of task repetition on the development of formulaic language in L2 speech fluency among Chinese EAP students

Funding Source (If applicable):

Effective: **August 03, 2018**

Expires: **August 31, 2019.**

Restrictions:

This certification is subject to the following conditions:

1. Clearance is granted only for the research and purposes described in the application.
2. Any modification to the approved research must be submitted to CUREB-A via a Change to Protocol Form. All changes must be cleared prior to the continuance of the research.
3. An Annual Status Report for the renewal of ethics clearance must be submitted and cleared by the renewal date listed above. Failure to submit the Annual Status Report will result in the closure of the file. If funding is associated, funds will be frozen.
4. A closure request must be sent to CUREB-A when the research is complete or terminated.
5. Should any participant suffer adversely from their participation in the project you are required to report the matter to CUREB-A.

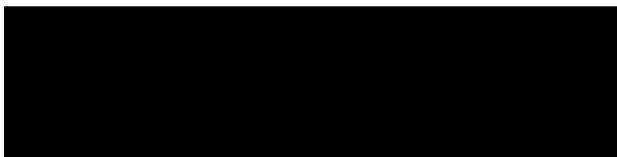
Failure to conduct the research in accordance with the principles of the *Tri-Council Policy Statement: Ethical Conduct for Research Involving Humans 2nd edition* and the *Carleton University Policies and Procedures for the Ethical Conduct of Research* may result in the suspension or termination of the research project.

Upon reasonable request, it is the policy of CUREB, for cleared protocols, to release the name of the PI, the title of the project, and the date of clearance and any renewal(s).

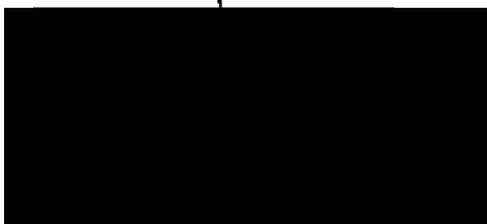
Please contact the Research Compliance Coordinators, at ethics@carleton.ca, if you have any questions.

CLEARED BY:

Date: August 03, 2018



Bernadette Campbell, PhD, Chair, CUREB-A



Andy Adler, PhD, Vice-Chair, CUREB-A

Appendix B

Evaluation rubric from the task repetition in class

Oral Presentation of Independent Learning Project

Name: _____ Total Score: / 3.00 or %

Presentation Title: _____ Time: _____

Voice Quality: spoke clearly, at a good volume and pace	1	2	3	4	5	6	
Pronunciation: Clear pronunciation. Did specific errors occur in pronunciation, intonations, word or sentence stress?	1	2	3	4	5	6	
Eye Contact and Presentation Style: looked at the class and did not read their notes. Used effective body and hand gestures	1	2	3	4	5	6	
Organization: clear introduction, body, and conclusion. It followed the time limit.	1	2	3	4	5	6	
Content: Did the presentation display knowledge and research on the topic	1	2	3	4	5	6	
APA: in-text citations and references.	1	2	3	4	5	6	

General Comments:

Appendix C

Wray and Namba's checklist

1. By my judgment, there is something grammatically unusual about this word string.
2. By my judgment, part or all of the word string lacks semantic transparency.
3. By my judgment, this word string is associated with a specific situation and/or register.
4. By my judgment, the word string as a whole performs a function in communication or discourse other than, or in addition to, conveying the meaning of the words themselves.
5. By my judgment, this precise formulation is the one most commonly used by this speaker/writer when conveying this idea.
6. By my judgment, the speaker/writer has accompanied this word string with an action, use of punctuation, or phonological pattern that gives it special status as a unit, and/or is repeating something s/he has just heard or read.
7. By my judgment, the speaker/writer, or someone else, has marked this word string grammatically or lexically in a way that gives it special status as a unit.
8. By my judgment, based on direct evidence or my intuition, there is a greater than-chance-level probability that the speaker/writer will have encountered this precise formulation before, from other people.
9. By my judgment, although this word string is novel, it is clear derivation, deliberate or otherwise, of something that can be demonstrated to be formulaic in its own right.
10. By my judgment, this word string is formulaic, but it has been unintentionally applied inappropriately.

11. By my judgment, this word string contains linguistic material that is too sophisticated, or not sophisticated enough, to match the speaker's general grammatical and lexical competence (Wary & Namba, 2003, pp. 29-32).

Appendix D

Confidentiality agreement

CONFIDENTIALITY ACKNOWLEDGEMENT & UNDERTAKING

Study PI: Yana Lysiak, MA candidate ALDS, School of Linguistics and Language Studies.

Project Title: The impact of task repetition on the development of formulaic language in L2 speech fluency among Chinese EAP students.

I understand that the integrity of the research, the principal investigator and the University depends on my acting in an impartial, ethical, confidential, and respectful manner. I hereby agree to:

1. Respect and maintain the confidentiality of all research participants. I will not reveal the name(s), or any information about such persons at any time;
2. Keep secure and confidential all documents and information, whether in the form of papers, books, files, documents, electronic communications, or in any other form or format, (including oral communication) which comes into my possession or knowledge in my capacity as a research assistant and relates to the project;
3. Not to download documents to any DropBox or Cloud System and only to download documents to a computer hard drive if absolutely necessary and with appropriate security protections;
4. Not divulge, disclose, or communicate, in any form or manner, directly or indirectly, confidential information to any person, firm, corporation, or other entity, other than a person otherwise appropriately designated to receive such confidential information;
5. I will not retain any documents and information related to my role as a research assistant. Ensure that all materials in my possession are returned to the principal investigator, as applicable, upon the completion of my involvement in the research study.
6. To erase or destroy all research information in any form or format regarding the research study that is not returnable to the principal investigator (e.g. information stored on my computer hard drive or in emails) after consulting with the research team.

I have read, understand, and agree to abide by the terms of this acknowledgement and undertaking.

DATED this ____ day of _____, 20__

Signature

Print Name

Appendix E

The full list of FS per group

1. I believe
2. I hope
3. we have to
4. you have better go
5. you have to try it
6. we can use two ways
7. I'm gonna talk about (*say*)
8. I'm going to introduce
9. I'm gonna be doing my presentation
10. we can imagine
11. we can see
12. this image shows
13. this picture shows
14. do you know
15. as my suggestion
16. I'm really interested in
17. just wanted to
18. I just want to
19. what does this mean
20. on the other hand
21. for example
22. according to
23. in conclusion
24. at first
25. lastly (*as a last*)
26. and so on
27. even though
28. not only because but
29. in disadvantage
30. not only but
31. meanwhile
32. in addition
33. to summarize
34. on average
35. due to
36. because of
37. not to mention
38. in fact
39. making it easy to
40. the main reason for this is

41. so generally
42. that's not a big surprise why
43. besides that
44. this is
45. that is
46. this kind of
47. in charge of (*in charges of*)
48. so it is
49. the best in the world
50. that's all
51. most of
52. lots of
53. a little bit
54. a lot
55. there's/ are
56. there is a big number of (*a lot number of*)
57. in a completely new way
58. in the same way
59. most of
60. at the same time
61. a large number of
62. in the world
63. in the middle of
64. at the beginning
65. in past few decades
66. in recent decades
67. per month
68. per day
69. around the world
70. good evening
71. hi guys how's it going (*how's going*)
72. thank you, guys
73. thank you very much
74. thank you for the attention (*for my attention/ thanks for it attention*)
75. thanks
76. thanks for listening
77. I'm sorry
78. thank you for watching
79. good afternoon, ladies and gentlemen
80. hello everybody
81. hi everyone
82. hello everyone
83. I'm very happy to be here

84. hope you, guys, had a nice weekend
85. you guys
86. (like) you know
87. okay so
88. kind of
89. it's (just) like
90. and actually
91. to be honest
92. pretty much
93. based on
94. traced back to (*traced on*)
95. go through
96. give up
97. paid off
98. go down
99. based down
100. look for
101. step off
102. get to know
103. hang out
104. blend in
105. come from
106. fold down
107. pick up
108. pull out
109. rely on
110. go down
111. sign out
112. look though
113. come about with
114. stand up
115. to step foot on
116. take for granted
117. met difficulties
118. met challenges
119. junk food
120. good luck
121. daily lives
122. big issue
123. downtown
124. first choice
125. my personal choice
126. so/really yummy

127. high tech things
128. deeply interested in
129. long term
130. closed due to traffic (*close the traffic*)
131. It's so good
132. personally my favourite
133. just like the grandma makes (*make*) it
134. out of my expectation (*expectations*)
135. more and more
136. look like
137. feel like
138. not their first choice
139. no matter
140. out there
141. on their own
142. still around
143. at all
144. one by one
145. of course

Appendix F

The results of the formulaic language analysis

FS functions	Student 1		Student 2		Student 3		Student 4		Student 5		Student 6		Student 7		Student 8		Student 9		Student 10	
	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2	V1	V2
Stance expressions	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	1
Discourse organizers	2	4	10	14	3	3	0	1	1	2	2	5	1	2	5	3	0	1	2	3
Referential expressions	5	5	5	6	2	3	2	1	5	6	1	0	0	0	2	4	1	3	3	5
Special conversation function	6	8	1	1	3	3	2	2	1	1	1	4	3	2	2	3	1	1	1	2
Figurative function	4	7	3	4	5	5	0	0	2	2	1	1	0	1	6	5	2	3	4	4
Other	5	2	0	0	1	1	1	0	1	2	0	0	0	0	2	2	0	3	0	0
Total number of tokens without repetition	23	26	18	25	14	15	5	4	10	11	5	10	4	5	19	22	4	11	10	15
Total number of tokens with repetition	35	49	19	30	15	16	7	6	10	11	5	11	4	5	21	23	4	11	11	16
Total number of FS types from V1 in V2		12		13		12		2		8		3		2		11		3		6
Total number of new FS types in V2		+15		+12		+3		+2		+3		+7		+3		+10		+8		+9

Appendix G

FS used by participants who showed some progress

FS functions	S1		S2		S6		S8		S9		S10	
	V 1	V2	V1	V2	V1	V2	V 1	V2	V1	V2	V1	V2
1. Stance expressions	Have to try it						I believe	<i>I believe</i>				I hope
2. Discourse organizers	<i>According to;</i> As my suggestion	<i>According to;</i> At first; That's not a big surprise why; Besides that	<i>And so on;</i> <i>In conclusion;</i> Because of; <i>Do you know;</i> <i>What does this mean;</i> We can imagine; Making it easy to; Meanwhile; ; This image shows; <i>In disadvantage</i>	<i>And so on;</i> <i>In conclusion;</i> ; I'm gonna; <i>Do you know;</i> <i>What does this mean;</i> <i>Meanwhile;</i> ; We can see; This picture shows; The main reason for this is; According to; In fact; Even though; On average; <i>In disadvantage</i>	<i>Not only but;</i> <i>I just want to</i>	<i>Not only but;</i> I'm gonna; So generally; I'm really interested in; <i>I just wanted to</i>	<i>At first (the first);</i> Not only because; Because of; Even though; <i>Lastly "as a last";</i> I (just) want to say	<i>At first;</i> <i>Lastly "as a last";</i> <i>Not only because;</i> I'm going to; We can see; <i>I want to say</i>		At first;	<i>For example;</i> On the other hand	I'm going to; <i>For example;</i> To summarize;

3. Referential expressions	<i>Kind of; There are/there is; Lots of; The best in the world; In the world;</i>	<i>Kind of; There are/there is; Lots of; For example; In the same way;</i>	<i>Most of; Kind of; Per month; Per day; In past few decades;</i>	Most of; Kind of; Per month; Per day; In recent decades; In the world;	This is		<i>That is; There is/there are;</i>	<i>That is; Most of; There is/there are; Around the world;</i>	<i>There are</i>	<i>There was; In the world; In charges of;</i>	Lot of; <i>There are/there is; A lot;</i>	Lots of; There are; A large number of; In the middle of; This is
4. Special conversation function	<i>You guys; It's (just) like; To be honest; Okay so; And actually; Thank you, guys</i>	<i>You guys; Hope you guys had (have) a nice weekend; Hi guys; Thank you, guys; Like you know; How's it going (how's going); Okay so; And actually; Kind of</i>	Thank you	Hi, guys, how's it going;	Thank you	Hi guys; You guys; Pretty much; Thanks for listening	Good afternoon, ladies and gentlemen; Thank you so much;	Hello everybody ; You guys; Thanks;	<i>Hi, everyone</i>	<i>Hi, everyone</i>	<i>Thanks for listening;</i>	Hello everyone; Thanks for listening
5. Figurative function	<i>Based on; Traced back to (traced on); So/really yummy; Junk food;</i>	Base down; So yummy/really yummy; Traced back to (trace back to); Junk food; Look through;	<i>Rely on; High tech things; Long term;</i>	<i>Rely on; Deeply interested in; High tech things; downtown</i>	<i>Good luck</i>	<i>Good luck;</i>	<i>Give up; Go down; Big issue; Sign out; Daily lives; Stand up;</i>	<i>Give up; Big issue; Sign out; Pull out; Stand up;</i>	<i>Based on; To step foot on</i>	Pick up; Fold down; Step off; To step foot on	<i>Get to know; Downtown; Hang out; Blend in</i>	<i>Get to know; Downtown; Blend in; Come from;</i>

		Look for; Downtown										
6. Other	Personally my favourite; It's so good; <i>More and more</i> ; Of course; Not their first choice	<i>More and more</i> ; Just to like the grandma makes (make) it					Still around; Feel like;	At all; Look like;		One by one; Looks like;		
Total number of tokens	23	27	19	25	5	10	19	22	4	11	10	15
Total number of FS types from V1 in V2		12		13		3		11		3		6
Total number of new FS types in V2		+15		+12		+7		+10		+8		+9