School of Education

The Effects of Task-Based Instruction on Vietnamese EFL Learners' Speech Production Capacity

Thi Huyen Thanh Do

0000-0002-0268-6128

This thesis is presented for the Degree of

Doctor of Philosophy

of

Curtin University

April 2021

Declaration

To the best of my knowledge and belief this thesis contains no material previously published by any other person except where due acknowledgment has been made.

This thesis contains no material which has been accepted for the award of any other degree or diploma in any university.

Human Ethics The research presented and reported in this thesis was conducted in accordance with the National Health and Medical Research Council National Statement on Ethical Conduct in Human Research (2007) – updated March 2014. The proposed research study received human research ethics approval from the Curtin University Human Research Ethics Committee (EC00262), Approval Number #HRE2017-0249

Signature: Thi Huyen Thanh Do

Date: 29 April 2021

Abstract

This study investigates the effects of different instructional options on Vietnamese EFL students' speech production capacity to address a practical issue regarding English Language Teaching (ELT) in Vietnam, namely students' inability to communicate fluently when using their second language (L2) resources in real-time. The theoretical motivation for the study is that while some previous research on input-based tasks (e.g., Shintani, 2011, 2016) has argued that input-based instruction could result in comparable development of both comprehension and production skills in comparison with output-based instruction for beginning-level young learners, other studies (e.g., Shintani et al., 2013; Tanaka, 1999, 2001) have suggested that combining input-based and output-based tasks during instruction can maximize L2 learning. Thus, the main aim of the study is to determine whether input-based task instruction is enough or whether combined input-output task-based instruction is more effective in terms of developing beginning-level learners' ability to perform production tasks fluently. If the former is the case for beginning-level Vietnamese EFL learners, it would be of considerable practical importance for instructional design and curriculum renewal in Asia where the instructional contexts are better suited to input-based task instruction than output-based task instruction in terms of class sizes, teachers' proficiency levels, learners' background and preferences, and administrative constraints on teaching and testing. To achieve this aim, the study compares the pre-test and post-test speech performances of learners who received (1) input-based task instruction only, (2) combined input-output task-based instruction, and (3) non-task-based narrow reading and listening instruction. This is done to examine the effects of each instructional option on beginning-level university EFL learners' speech production capacity in the performance of a repeated oral monologic task, and whether or not any observed effects would transfer to the performance of a parallel task.

The study employed a mixed within- and between-subjects design with the participation of 102 Vietnamese university EFL learners who were randomly assigned to two task-based instructional groups (Input group and Combined group) and a non-task-based reading and listening group (Comparison group). The study

was conducted in three phases over a five-week period. In Phase 1, all the participants undertook the pre-test. In Phase 2, the treatment phase, all three groups received the same amount of instruction time (three 90-minute lessons on a weekly interval), but each group received a different treatment. The Input group lessons involved in input-based tasks only in which the participants were not required to produce output. The Combined group lessons involved both input-based and output-based tasks, which required the learners to produce output. Meanwhile, the Comparison group lessons involved narrow reading and listening of an audiobook with no specific instruction related to the target tasks. After treatment, the third phase was conducted in Week 5 with two post-tests, the Same-Task-Post-Test and the Parallel-Task-Post-Test. Participants' oral performances on the pre-test and post-tests were audio-recorded, transcribed, and then coded for four dependent variables: speech rate, between-clause pausing, within-clause pausing, and self-repairs. These were taken as measures of overall fluency, conceptualization, formulation, and monitoring, respectively. After all the data were checked and corrected for normality and homogeneity of variance to ensure that no assumptions of MANOVA were violated, statistical analyses were conducted to answer the research questions.

The results reveal that for beginning-level university-age EFL learners the combined input-output task-based instruction served as the best option to improve their speech production capacity during the performance of a repeated task. The gains in speech production capacity for the combined input-output group were explained by way of automatization of L2 knowledge and skills through output-based practice. Increased automaticity might have eased attention demands for the conceptualization and formulation stages of L2 speech production, thus facilitating parallel processing of these stages and resulting in corresponding less between-clause and within-clause pausing, and a faster speech rate. The study also indicates that there was meaningful transfer of practice effects from both forms of task-based instruction to the performance of a parallel task, but only in terms of between-clause pausing (conceptualization). This gain was attributed to the increased familiarity with the task structure and task demands that lead to a

reduction in conceptual demands and a corresponding lower frequency of between-clause pauses. Overall, the results of the study indicate that the combination of both input-based and output-based tasks in the syllabus was more effective in improving beginning-level university-age EFL learners' speech production capacity than the use of input-based instruction regardless of whether the input was provided through tasks or reading and listening practice.

These findings bear theoretical and pedagogical implications regarding the role of task-based instruction in the development of L2 speech production capacity, the use of pause location and self-repairs as speech production measures, and the importance of parallel task repetition on subsequent task performances.

Acknowledgements

This study would not have been accomplished without the support of a number of people and organizations. First of all, I would like to express my deepest gratitude to my supervisors, Associate Professor Craig Lambert and Professor Rhonda Oliver, for their professional guidance, continuous support, and valuable feedback throughout the time of my PhD journey. They have motivated, challenged, and encouraged me at every phase of the study. It is my great honour to have the opportunity to work with those wonderful academics, who have made the completion of my thesis possible.

I greatly appreciate the Vietnamese Government, Vietnamese Ministry of Education and Training, and Curtin University for the financial support for my doctoral study. Without their sponsorship, my dream of doing a PhD course in Australia would not have been realized. I am so honoured to have received the CIPRS-MOET scholarship for this study.

I would sincerely thank the Board of Management of Hanoi University of Industry (HaUI), and all the involved staff of the Faculty of Foreign Languages, HaUI for enabling and assisting me to undertake data for this study. I am also indebted to my colleagues and students at the Faculty of Foreign Languages, HaUI for their enthusiastic cooperation and excellent support during my field trips. They all have played an important part in this research study.

Many thanks also go to my fellow friends at Curtin Humanities Graduate Research Hub and my Vietnamese friends in Perth who have been sharing all the ups and downs along my PhD journey, making it a memorable and enjoyable life experience.

Last but not least, this thesis is dedicated to my late father, who is no longer able to see its completion. My deepest thanks go to my beloved family, my mother, my parents-in-law, my sister and brother, and especially my husband and my two sons for being by my side, loving and caring for me unconditionally during this long journey.

Table of Contents

List of Figures

List of Tables

List of Abbreviations

ACT	Adaptive Control of Thought
BP_freq	Frequency of Between-clause Pauses
CbG	Combined Group
CBI	Comprehension-Based Instruction
CLT	Communicative Language Teaching
CpG	Comparison Group
DFA	Discriminant Function Analysis
EFL	English as a Foreign Language
ELT	English Language Teaching
ESL	English as a Second Language
GLM	General Linear Model
InG	Input Group
ISLA	Instructed Second Language Acquisition
L1	First Language
L2	Second Language
MANOVA	Multivariate Analysis of Variance
NFLP2020	National Foreign Language Project 2008-2020
PBI	Production-Based Instruction
PTR	Parallel Task Repetition
Rp_freq	Frequency of Repairs
SLA	Second Language Acquisition
SR	Speech Rate
STR	Same Task Repetition
TBLT	Task-Based Language Teaching
TSLT	Task-Supported Language Teaching
WP_freq	Frequency of Within-clause Pauses

Chapter 1. Introduction

In this chapter, the background of this study is described including information about the status of teaching English in Vietnam and the motivation for the research. A brief overview of the theories related to Task-Based Language Teaching (TBLT), second language (L2) speech production, and fluency is also provided. Finally, the study design, its significance, and the thesis outline is presented.

1.1 Background and Personal Motivation

With English becoming a global language and an indispensable tool in the process of integration and internationalization, the ability to speak English fluently has become the main goal of many L2 learners. However, for many learners, especially those in a country like Vietnam where English is spoken as a foreign language, this goal can be difficult to achieve. Yet, for the last two decades, English has been identified by the Vietnamese government as the main foreign language to be taught across all education levels in Vietnam (Hoang, 2013). In fact, at the tertiary level, it is expected that most university graduates need to be sufficiently proficient in English to use it confidently in their daily communications and to communicate effectively in English in global, multi-cultural, and multi-lingual environments (MOET, 2003, 2008).

To date, however, English education in Vietnam has not been successful in developing learners' ability to reach these goals. Teaching practices tend to focus on linguistic elements such as grammatical structures, vocabulary and syntactic rules, rather than developing students' communicative competence. As a result, many university graduates are not fluent in speaking English, nor can they communicate in English in everyday situations despite several years' study (Hoang, 2008; Le, 2011). It is common to hear someone saying that they can understand what they read and hear, but that they cannot speak aloud what they want to express. They may know the grammatical rules and vocabulary very well, but have difficulty transforming what is known about the language into comprehensible utterances in real-time communication. According to researchers (e.g., Hoang, 2013; Hoang, 2014; Le & Barnard, 2009; Mai & Iwashita, 2012), students' low level of English proficiency in Vietnam is due to: (1) ineffective teaching and learning approaches, (2) teachers' limited language proficiency, (3) lack of English language environment, (4)

classroom constraints and (5) teaching and testing mismatch. It will be helpful to address each of these factors impacting ultimate level attainment in English separately as background to the present study.

1.1.1 Ineffective Teaching and Learning Approaches

In line with Confucian ideology, Vietnamese students, like many other Asian learners, are often viewed as passive learners, believing that book knowledge and what they receive from their teachers is all they need to learn a given subject (Phuong-Mai et al., 2005; Subramaniam, 2008; Thi Tuyet, 2013; Wong, 2004). As such, the belief is that the teachers' job is to lecture and the students' job is to rote memorize this information, rather than to construct new knowledge or participate in critical discussions. On the one hand, these beliefs encourage students to study hard to gain theoretical knowledge, but on the other hand, they can hinder them from becoming independent and critical-thinking communicators (Le, 2011). As Nguyen (2002) indicates, Vietnamese students "are very traditional in their learning styles: they are quiet and attentive, good at memorizing and following directions, reluctant to participate" (p. 4). Exacerbating this situation is the fact that Vietnamese education is heavily influenced by exam-oriented practices with language teachers often adopting examination-focused learning activities for communicative purposes. As a consequence, Vietnamese students are often more motivated to pass examinations than to improve their communicative skills (Bock, 2000).

With regard to English language teaching approaches, the prevailing method for many decades in Vietnam has been the Grammar-Translation (or at least methods that align with this), which focuses on knowledge about the English language (linguistics forms and structures), rather than communicative competence (Hoang, 2008; Le, 2007). In light of this, the common practice is for teachers to present grammatical structures deductively and students to memorize the forms and rules and apply them into follow-up exercises. As such, memorization of grammatical rules, mechanical drills, and repetitions are considered as the preferred learning strategies to help students achieve accuracy, fluency, and self-confidence (Bernat, 2004; Le, 2012; Oanh & Hien, 2006). This culture of teaching and learning can explain why many Vietnamese students have a good command of meta-linguistic knowledge and can gain high scores in grammar-oriented examinations, but still have difficulty making themselves understood in oral communication.

To counter this issue, in the early 2000s, a Communicative Language Teaching (CLT) approach was introduced into English classrooms together with a new curriculum that promoted learner-centred, communicative, task-based language teaching (MOET, 2006). In the Vietnamese context, CLT is used to "(a) make communicative competence the goal of language teaching and (b) develop procedures for the teaching of the four language skills that acknowledge the interdependence of language and communication" (Richards & Rodgers, 2001, p. 155). In 2008, the government of Vietnam approved the policy entitled *National Foreign Language Project 2008-2020* (NFLP2020) which has a strong commitment to improving the foreign language competence of Vietnamese school-aged students. The ultimate goal of the project is that 'most young Vietnamese graduates of professional secondary schools, colleges and universities will have a good command of foreign language which enables them to independently and confidently communicate, study and work in a multilingual and multicultural environment of integration' (MOET, 2008, p. 1). In order to achieve the stated goal, a lot of effort has been given to developing curriculum and pedagogy, and to providing teacher training to support the implementation of these. CLT, which emphasizes learners' opportunities to use English for communicative purposes, is now considered the backbone to achieving the aim of the project (Le & Yeo, 2016; Le, 2015; Nguyen, 2011) and has been well supported by the teachers' positive attitudes and their efforts to adapt their teaching practices. However, the application of a Western teaching model into a Confucian-oriented educational system has also encountered a lot of challenges and seemed to be ineffective due to many other constraints (Hoang, 2013; Mai & Iwashita, 2012), which are discussed below.

1.1.2 Teachers Limited Language Proficiency

One of the major challenges facing Vietnamese English language education is the lack of teachers who are sufficiently competent to ensure the teaching success (Hoang, 2008). While teachers' language proficiency is a key to the successful enactment of the NFLP2020 project, there remains a shortage of qualified teachers at all educational levels (Bui & Nguyen, 2016; Le, 2015). Researchers have also pointed out the fact that English teachers' language proficiency is lower than the minimum level required (Le, 2015; Nguyen & Mai, 2015). The teachers' limited language proficiency and their inability to use English

effectively for classroom instruction affect the quality of teaching and learning English in Vietnam (Manh et al., 2017).

1.1.3 Lack of Favourable Language Environment

Because English is only used as a foreign language (EFL) in Vietnam, Vietnamese students' English learning is often restricted to the classroom context. Hoang (2013) claims that 'there is a serious lack of a natural environment for students to practise communicating in English' (p. 76). Although in their English classes students may have opportunities to practise speaking English with other students during communicative tasks and activities, outside their classrooms, there is almost no opportunity for communication in English, especially with competent or native speakers. Consequently, students' communicative competence cannot be enhanced.

1.1.4 Classroom Constraints

Within the classrooms, other constraints also exist, including large class sizes and poor classroom facilities (Hien & Loan, 2019; Le, 2011). Large class sizes of 30 to 65 mixed-ability students cause a lot of challenges to teachers in terms of classroom management, and can create difficulties in their attempts to implement communicative activities. The lack of classroom equipment such as computers, projectors, speakers, etc. also inhibits language teaching and learning processes in Vietnam. Consequently, teachers end up choosing easier teaching techniques such as teacher-fronted lecturing rather than CLT and so students do not have opportunities to improve their communicative competence through interaction.

1.1.5 Mismatch between Teaching and Testing

While the communicative approach is recommended for teaching, recent studies show that most exams are still grammar-based and test linguistic competence or lexico-grammatical knowledge in written form rather than communicative competence (Hoang, 2010; Mai & Iwashita, 2012). Language tests do not require learners to reflect their communicative competence appropriately (Nguyen & Le, 2013). Furthermore, due to the washback effects of examinations, students and teachers are preoccupied with linguistic knowledge instead of developing communicative skills.

Although considerable effort has been invested by all the parties involved, there are still many unresolved issues and the quality of English Language Teaching (ELT) in Vietnam is considered far from satisfactory: Vietnamese students' command of English remains rather limited, and most of the graduates still cannot use it effectively as a means of communication, even after a long period of learning English (Pham, 2004). Therefore, the present study attempts to address the problems of improving students' ability to use English for real-time communication, especially their oral fluency by investigating the effects of different task-based instructional options on Vietnamese EFL students' speech production capacity.

1.2 Theoretical Overview

TBLT, which emerged from and constitutes a strong version of Communicative Language Teaching (R. Ellis, 2003), has become a dominant approach to second and foreign language teaching all over the world. TBLT is currently adopted as an alternative to traditional pedagogies in many Asian Pacific countries with the aim to promote students' communicative competence through tasks (Adams & Newton, 2009). TBLT has been introduced and has a particular currency in Vietnam as it informs the current foreign language curriculum reforms and recommendations (Newton & Bui, 2020). In TBLT, tasks are the principal component in syllabus design, classroom instruction, and assessment.

Although different researchers hold different views about TBLT (Bygate et al., 2001; Ellis, 2015; Long, 1985a, 2016; Nunan, 1989; Prabhu, 1987; Skehan, 1998b; Willis, 1996), they all support the idea that engaging in tasks can promote learners' communicative ability, their fluency and their acquisition of linguistic structures.

1.2.1 Tasks and Second Language Acquisition

There have been a number of studies showing evidence that tasks can be effective tools for learning (Lee, 2000; Oliver, 2002; Pica et al., 1993; Richards & Rodgers, 2001). According to Richards and Rodgers (2001), "tasks are believed to foster processes of negotiation, modification, rephrasing, and experimentation that are at the heart of second language learning" (p. 228).

One rationale for the use of tasks is the opportunities for interaction that they provide, and how this relates both to learners' processing of input and their output

performance. For example, Oliver (2002) described how even for children tasks promote the type of learner interaction where they can receive meaningful input, use the target language and in doing so receive feedback on their production, which in turn, might lead them to modify their output in their L2 – all features argued by Long (1996) to be facilitative of language learning. Ellis (2008) stated "tasks, or rather the performance of them, create opportunities for learning" (p. 819). Within TBLT, tasks are deemed not only as necessary but also as sufficient for learners to learn a language (R. Ellis, 2003). That is, within TBLT tasks are an indispensable part of the approach.

A second rationale for the use of tasks is in the management of learners' attentional resources during task performance. Skehan (2009) argues that factors in the design and implementation of tasks are capable of increasing or decreasing the demands that tasks make on attention in relation to different stages of L2 speech processing (Kormos, 2006). Tasks, by definition, focus learners on meaning, but different approaches to task instruction may reduce the attention that learners need to devote to ***what*** they need to say in performing tasks, thus freeing up limited attentional resources to ***how*** they can say it efficiently. In other words, reducing conceptualization and encoding demands appears to facilitate the development of their speech processing ability with concomitant gains in fluency.

The present study is informed by a cognitivist approach and investigates beginning-level university EFL learners' capacity to complete monologic narrative tasks. As such, the data for the study are not highly interactive, but instead explore the effects of different approaches to task-based instruction on the development of learners' fluency in L2 speech production processes.

1.2.2 Input-Based and Output-Based Tasks

Tasks can be input-based or output-based as long as they satisfy the four criteria for a 'task' (Ellis, 2009). A simple way to distinguish between the two is as follows: In an input-based task, learners are presented with oral or written language input (through listening or reading) which they must comprehend in order to be able to achieve the outcome of the task - they are not, however, required to produce the language. In contrast, output-based tasks require learners to produce language, either in oral or written form.

With regard to the roles of input and output in second language acquisition, numerous studies have compared the effects of input and output on the acquisition of particular grammatical features (see Erlam, 2003) and vocabulary (Ellis, 1995; Ellis & He, 1999; Ellis & Heimbach, 1997; Ellis et al., 1999; Loschky, 1994; Shintani, 2011, 2012a, 2012b). Although the results have been inconclusive, research generally suggests that input-based tasks are more suitable for beginning level learners and for acquiring 'new' grammatical structures, while output-based tasks may be more beneficial in helping more advanced learners achieve greater control over 'partially acquired' structures (Ellis, 2015, 2017).

While some previous research (e.g., Shintani, 2011, 2012a, 2016) has provided evidence that input-based tasks can be as effective as production-based activities in EFL vocabulary learning for young beginners, resulting in better performance in comprehension tests and the same levels of achievement in production tests, other studies (e.g., Shintani, Li, & Ellis, 2013; Tanaka, 1999, 2001) have suggested combining input-based and output-based instruction for optimal L2 learning. Therefore, in the current research, the aim is to explore how input-based task instruction versus combined input-output task-based instruction supports Vietnamese EFL learners' fluency development and fluency transfer in relation to stages of speech production in a theoretical model (Levelt, 1989, 1995, 1999) which has been adapted to L2 speech production by De Bot (1992) and Kormos (2006), and operationalized by Lambert et al. (2017).

1.2.3 L2 Speech Production

The modular model of speech production proposed by Levelt (1989, 1995, 1999) is the most widely used theoretical framework in second language production research (e.g., De Bot, 1992; Kormos, 2006). According to the model, there are four stages of speech production, namely the *conceptualization, formulation, articulation,* and *monitoring* stages. In the conceptualization stage, the speaker prepares a preverbal plan by selecting the information to be conveyed in the message. In the formulation stage, the preverbal plan is converted into linguistic form through lexical and grammatical encoding. In the last stage of speech production, articulation, the phonetic plan is buffered and sounds and syllables of actual speech are produced to convey the message. Monitoring can occur at any of these stages to ensure the appropriateness and accuracy of the overt speech.

Although the stages are expected to be in hierarchical order, the modular model assumes that 'processing is incremental' (Kormos, 2006): as soon as the formulator receives the preverbal plan, the conceptualizer starts with the next chunk despite the fact that the previous chunk is still being processed. When different processing components operate simultaneously, parallel processing is considered to take place, provided that processing mechanisms are sufficiently automatic (Kormos, 2006; Levelt, 1989, 1995, 1999).

According to Lambert et al. (2017), the formulation stage may be largely automatic for proficient speakers and allow for parallel processing with other modules. However, lower proficiency speakers, especially second language learners, may encounter different degrees of breakdown in parallel processing due to their difficulty with lexical retrieval and grammatical encoding (De Bot, 1992; Kormos, 2006). This is because they can only produce fluent speech if the speech production mechanisms function efficiently, including "the automaticity of encoding processes, the conceptual demands of planning the content of the message, and the skillful handling of performance breakdowns" (Lambert et al., 2017, p. 170). The issues related to L2 speech production will be addressed in Section 4.2.

1.2.4 Fluency

Fluency is one of the three aspects of production, namely fluency, accuracy and complexity (see Section 2.3.2 for more information on the relationship of these factors). Fluency can be approached from both a broad and a narrow sense. In the broad sense, fluency can be seen as overall oral proficiency, whereas fluency in the narrow definition refers to smoothness and ease of oral linguistic delivery (Lennon, 1990). According to Skehan (1996), fluency "concerns the learner's capacity to produce language in real time without undue pausing or hesitation" (p. 22).

Fluency is often considered to be a multidimensional construct. Segalowitz (2010) distinguished three notions of fluency: cognitive fluency (the speaker's abilities to efficiently plan and execute his speech), utterance fluency (the fluency that can be measured objectively in a sample of speech), and perceived fluency (the impression that listeners have of the fluency of a sample of speech). Utterance fluency can be conceptualized as (i) speed fluency, (ii) repair fluency, and (iii) break-down fluency, and can be objectively measured (Lambert & Kormos, 2014; Skehan, 2003; Tavakoli & Skehan, 2005). Speed fluency is what

relates to the rate of delivery, and is usually measured by calculating pruned speech rate (the number of pruned syllables per second) (Ellis & Barkhuizen, 2005), articulation rate (the number of syllables per second excluding pausing time), or mean length of run (the mean number of syllables between two silent pauses). Break-down fluency relates to pausing behaviours and can be measured by pause frequency, pause length, and pause location (De Jong, 2018; De Jong & Bosker, 2013). Repair fluency is usually operationalized as the frequency with which the speakers make repetitions and corrections and can be measured by the ratio of self-repairs (De Jong et al., 2013; Lambert et al., 2017). These aspects of fluency and their relationship will be discussed in more detail in Section 5.1 and 5.2.

As Lambert et al. (2017) point out, it is difficult to relate these fluency measures to a specific speech production mechanism. However, it can be argued that speech rate represents the whole speech production process (overall fluency); between-clause pauses relate to breakdowns in the conceptualization stage; whereas within-clause pauses can signal breakdowns during the formulation stage (Götz, 2013; Towell et al., 1996). Finally, overt self-repairs might represent the availability of attentional resources for monitoring (De Jong, 2016b; Kahng, 2018; Kormos, 1999, 2006; Lambert et al., 2017; Saito et al., 2018). Details of utterance fluency measures in relation to cognitive processes underlying L2 speech production will be discussed in Section 5.4.

With regard to fluency development, a number of studies have examined the effects of task repetition on learners' planning opportunities (e.g., Wang, 2014) and on immediate gains in L2 fluency (Lambert et al., 2017). Both of these studies investigate learners' fluency in terms of speech rate, frequency of clause-final and mid-clause filled pauses, and overt self-repairs and reformulations. They suggest that task repetition has beneficial effects on fluency with evidence of a large increase in speech rate and a decrease in clause-final pauses, mid-clause pauses, and self-repairs. Furthermore, Lambert et al. (2017) were able to relate the findings to specific stages of L2 speech production (conceptualization, formulation, and monitoring) and used this to determine that task repetition has positive, but differential effects on three L2 speech production stages in terms of the number of repetitions and at different levels of proficiency. However, they only examined the effects of massed task repetition on immediate gains in L2 fluency and examined only filled pauses

during the same task repetition. The role of task repetition will be further addressed in Section 5.3.2.

The present study, using an experimental pre-test/post-test design, investigates the impact of input-based task instruction versus combined input-output task-based instruction on learners' fluency development when using a repeated task and then using a similar task to see if the gains transfer to the performance of a parallel task. Thus, in the present study, the measures of fluency are adapted from Lambert et al. (2017) in relation to hypothesized stages of L2 speech processing (Kormos, 2006; Levelt, 1989, 1995, 1999): (1) speech rate (a composite measure of L2 speech production), (2) between-clause pausing (a measure of conceptualization), (3) within-clause pausing (a measure of formulation or encoding), and (4) self-repairs (a measure of monitoring).

1.3 Significance of the Study

Theoretically, the study is intended to contribute to our understanding of the development of L2 learners' speech production capacity through task-based language teaching. There have been few studies investigating the impact of task-based instruction on learners' speech production capacity. Whilst Shintani (2011) has studied the impact of comprehension-based versus production-based instruction on vocabulary and grammar acquisition for young beginner learners, her study did not address learners' speech processing. The present study does so, exploring the impact of input-based task instruction versus combined input-output task-based instruction on Vietnamese university learners' speech production capacity. The study also examines whether the effects documented by Shintani (2011) for true beginners are the same with EFL false-beginner university learners (i.e., input-based task instruction is equal to output-based instruction) or it is better to combine the two kinds of instruction as suggested by Tanaka (1999, 2001) and Shintani et al. (2013). It also responds to a call by Bygate (2016) for studies that enable TBLT as a 'researched pedagogy' to "report the range of teaching learning procedures available, the outcomes and the relative efficiency of the approach" (p. 397). The study is expected to add to the existing body of empirical research on TBLT in Vietnam, most of which has focussed on the constraints in the implementation of TBLT, especially regarding teachers' beliefs and perceptions (e.g., Barnard & Viet, 2010; Cao, 2018; Chu & Oliver, 2017).

Pedagogically, this study may provide directions for teachers about how to select and sequence tasks to suit learners' level of proficiency, particularly to develop their speech processing capacity. It is hoped that the findings from this research can provide grounds for potential improvement of task-based English teaching and learning in Vietnam in general and the enhancement of students' communicative competence in particular.

1.4 Thesis Outline

This thesis consists of nine chapters. Following the Introduction chapter, Chapter 2 provides a theoretical background to the study. Chapter 3 reviews the literature related to Input-based and output-based task instruction as independent variables. Chapter 4 discusses in detail aspects of second language speech production. Chapter 5 is devoted to second language fluency development and measures of fluency as dependent variables of the study. Chapter 6 presents the research questions and methods used for data collection and data analysis. Chapter 7 and Chapter 8 report on the results of the study. Chapter 9 discusses the findings of the study, contributions of the research, both theoretically and pedagogically, as well as limitations of the study and recommendations for further research. Chapter 10 concludes the thesis with a summary of key findings and implications.

Chapter 2. Theoretical background

There has been considerable and dynamic development in the field of second language acquisition (SLA) since its beginning in the 1960s. In this time there has been an extensive number of studies and the construction of various theories to explain how a second language (L2) is learned and the factors that influence this process. This chapter reviews the theoretical positions that are considered influential in SLA and discusses how these different views may contribute to language pedagogy more generally, and in particular to the focus of the current research. Theoretical issues regarding TBLT, L2 speech production and L2 speech fluency will be addressed in Chapters 3, 4 and 5.

2.1 What is Second Language Acquisition?

Second language acquisition is a complex process influenced by social factors, learners' language processing mechanisms (the main focus of the current study), and individual learner factors (Ellis, 2008). Whilst in early times 'acquisition' as a subconscious process was distinguished from 'learning' – which is considered a conscious process – many today no longer make this distinction (Long, 1996). However, Ellis (2015) does suggest that "Acquisition refers to the incidental process where learners 'pick-up' a language without making any conscious effort to master it; whereas learning involves intentional effort to study and learn a language" (pp. 6-7). Therefore, although the distinction today is blurred, in this thesis Ellis' categorization will be followed: namely, that acquisition occurs naturally through daily communication while learning takes place consciously mostly through formal instruction. Furthermore, the current study is underpinned by the belief that different approaches to formal instruction will have different effects on learners' language development and particularly on their fluency.

With respect to instruction, this will also vary according to language context. Ellis (2015), for instance, makes a distinction between 'second' and 'foreign' language acquisition on the basis of the contextual differences and opportunities where the language is used. Second language learning takes place in a context where the target language is used as a means of community communication outside the classrooms while foreign language learning refers to the language learning that is done mainly through formal instruction in the classroom and in settings where learners have limited use of the target language in daily

situations. This means that the term 'L2 acquisition' can be used to cover both these types of language learning and the manner in which it is done – including 'naturalistic acquisition' and 'instructed leaning' (Ellis, 2015). In this way, SLA can be described as the acquisition of any language other than a learner's mother tongue in either foreign or second language contexts. In the current study, the term SLA is used to reflect this and specifically refers to the context of studying English as a foreign language in Vietnam.

2.2 Input, Interaction and Output in SLA

There are many internal and external factors that impact upon second language acquisition. Among these, input, interaction, and output are deemed facilitative, if not necessary, in the acquisition of a second language. These are described next with an example of a related theory provided after each.

2.2.1 Input and the Input Hypothesis

Over time various definitions of input have been proffered beginning with Corder (1967), one of the pioneer SLA researchers, who defined 'input' as what is available to the learner, but distinguished it from 'intake' which he described as that internalized by the learner. Nunan (2004) views 'input' as "the spoken, written and visual data that learners work with in the course of completing a task" (p. 47). VanPatten and Williams (2015) define input as "language the learner hears (or reads) and attends to for its meaning" (p. 9) and add the caveat that language that is not responded to with respect to its meaning is not considered input. According to Ellis (2015), "the term 'input' refers to the samples of the oral or written language a learner is exposed to. This constitutes the 'data' that learners have to work with to construct their interlanguage" (p. 12). Input can be either in form of non-interactive reading or listening texts or in the form of 'models' that learners receive during interaction (Ellis, 2008). A key construct in the current study is input (and hence its discussion in this chapter), further, as input can be generated through interaction, this key theoretical position is also described later in this chapter.

Input has been considered critically important in "understanding the what and why of second language acquisition" (Swain, 1985). The importance of input in second language acquisition is encapsulated in numerous theories of SLA. However, differences exist between these theories according to the extent of its role. Ellis (2008) and VanPatten and

Williams (2015) suggest that although different theories attribute different roles to input in the language learning process, they all acknowledge the need for language input. According to Krashen (1982, 1985), input is essential and sufficient for L2 acquisition. Similarly, Gass and Mackey (2015) posit that the input which learners are exposed to is the "sine qua non of acquisition" (p. 181). VanPatten and Williams (2015) also argue that "acquisition will not happen for learners of a second language unless they are exposed to input" (p. 9). However, VanPatten (1996) while accepting the crucial role of input, maintains that simple exposure to input is not adequate to bring about L2 acquisition. As such, he claims that input processing is needed to help learners establish form-meaning connections from the input, and proposes Processing Instruction as an input-based grammar instruction approach in SLA (VanPatten, 1996, 2004). Ellis (2008) provides an account of the role of input in SLA from 1) behaviourist, 2) mentalist, 3) internationalist, and 4) the sociocultural viewpoints:

1) Behaviourists give little countenance to the learner's internal processing and instead view input as composed of the speaker's stimuli and feedback. In this way, L2 acquisition is dependent on external factors and can be shaped by manipulating the input to make both the stimuli and feedback appropriate.

2) From the mentalist point of view, in contrast, input is essential, but insufficient for L2 acquisition to occur by itself. That is, language input does not provide sufficient information for the learner to arrive at the rules of the L2, and instead they suggest that input only serves as a catalyst to activate internal processing mechanisms.

3) Interactionist theories emphasize the importance of input especially that provided through verbal interaction because it contains the data the learners need for acquisition.

4) Finally, sociocultural theories explain the relationship between input and acquisition differently. In sociocultural SLA, L2 acquisition is seen as "an inherently social practice that takes place within interaction" (p.206). That is, acquisition occurs socially *in* rather than *through* interaction and development takes place later as the learner internalizes new language features and is able to produce the form independently.

These contrasting views support the claim by Doughty and Long (2003) that the difference between SLA theories is due to the conceptualization of how input is processed by language learners. As indicated above, it is the processing of input that is important in the present study.

Gass (1997) also acknowledges the key role of input. However, she differs from Ellis (2008) in that she suggests that what remains contested is the type and amount of input that is necessary for second language development to occur. Initially, Gass (1997) described the role of input for learning from four different perspectives: 1) the input-interaction perspective, 2) the input hypothesis perspective, 3) the universal grammar approach, and 4) based on the information processing model. These are described as follows:

1) In the input-interaction position, language input provided to the learner needs to be manipulated through interaction for language acquisition to occur. This is because it is through interaction and negotiation of meaning that learners focus on particular parts of the language, which will later be integrated into their interlanguage.
2) In contrast, in the Input Hypothesis perspective (Krashen, 1985), input is seen to play the primary role in L2 acquisition. However, comprehensible input must be at an $i + 1$ level for language development to occur.
3) Universal grammar views language as a set of abstract principles, so there must be something in addition to language input to enable easy and speedy language learning. This means input alone is not sufficient. What may reduce the task of learning is the innate capacity that helps language learners acquire the second language.
4) The information processing model assumes that human mind has a limited processing capacity and, therefore, in order for learning to occur, the learner must first notice what needs to be learned when receiving the input. It is the gap between the learner's interlanguage system and the target language that needs attention. In this model, input provides information necessary for language construction.

As can be seen in this brief overview, although different theories consider input from different perspectives and attach different levels of importance to it and its role in second language acquisition, they all share the belief that input is an essential component that

constitutes data for the learner's internal processing and acquisition (VanPatten & Williams, 2015). It should be noted, however, some like Long (1985a), and Gass et al. (2013) argue that whilst comprehensible input is necessary, it is not sufficient for the acquisition process. The importance of input is further discussed below beginning with one of the earlier influential theoretical positions in SLA, namely the **Input Hypothesis** (Krashen, 1982, 1985) or **Comprehension Hypothesis** (Krashen, 2003).

Krashen (1985) claimed that comprehensible input is necessary and efficient as the driving force for L2 acquisition and that "speaking is a result of acquisition and not its cause" (p. 2). On this basis, the Input Hypothesis proposes that humans acquire a language by simply receiving comprehensible input and understanding the messages, and especially "if input is understood, and there is enough of it, the necessary grammar is automatically provided" (Krashen, 1985, p. 2). As such, language need not be taught because acquisition will occur naturally so long as the learner receives adequate comprehensible input. It is noted that not all input is processed for acquisition, but rather the learners do so when the language input is 'a little beyond' their current level of language competence. It is also argued that input modifications that occur through contextual and extra-linguistic clues can make input comprehensible for learners, and simplified input (e.g., foreigner or teacher talk) or modified input can be beneficial for acquisition. Thus this hypothesis is underpinned by the following principles: "comprehensible input is the essential ingredient for second-language acquisition", it needs to be at one stage above the learner's current level of competence (i.e., $i + 1$); it must be comprehended; and there is sufficient amount of it (Krashen, 1985, pp. 2-4).

Subsequently, the Input Hypothesis is frequently referred to as the Comprehension Hypothesis (Krashen, 2003, 2008). The Comprehension Hypothesis states that we not only acquire language but also develop literacy (e.g., reading and writing ability, vocabulary knowledge) when we comprehend what we hear and read (Krashen, 2008). In other words, it is contended that language acquisition is a subconscious process, and input must be comprehended for language acquisition to occur. Furthermore, the Comprehension Hypothesis suggests that narrow input, that is "a great deal of input in a narrow range of subjects" (Krashen, 2014) is more beneficial for second language acquisition. Narrow input can be provided through narrow reading (Krashen, 1981; Schmitt & Carter, 2000) or narrow

listening (Dupuy, 1999; Krashen, 1996), which means learners focus on one theme at a time, or works of the same author consistently over a period of time rather than a variety of texts of different topics. It is argued that rich and repeated exposure to common vocabulary and discourse within a topic or by the same author makes input more comprehensible (Krashen, 2014). Moreover, massive input around one subject matter allows the build-up of background knowledge, and increases the acquirer's familiarity with context, which are also believed to be facilitators of comprehension, and thus language acquisition (Krashen, 1985). While there has been supporting evidence for narrow reading (e.g., Cho & Krashen, 1994, 1995; Kang, 2015) as well as for narrow listening alone (e.g., Caspino, 2005; Dupuy, 1999; Krashen, 1996), combining narrow reading and listening, i.e., providing narrow input in both aural and written forms simultaneously, is also found to be facilitative of language development (e.g., Chang, 2009; Kimura & Ssali, 2009).

Over the long period of time since its inception, Krashen's Input Hypothesis has spawned a considerable number of studies on different types of language input and their influence on SLA. Ellis (2008) summarises such studies in relation to two claims of the Input Hypothesis: (1) whether modified input promotes comprehension, and (2) whether comprehension leads to acquisition. Research has shown mixed results in response to the first claim. While some studies report positive effects (e.g., Long & Ross, 1993), others show no difference or negative effects of modifying input (e.g., Ellis et al., 1994; Griffiths, 1990; Leow, 1993; Loschky, 1994; Parker & Chaudron, 1987). With regard to the second claim, the studies that have investigated the relationship between comprehension and acquisition also show inconclusive results (Ellis, 1995; Leow, 1993; Mason & Krashen, 1997; Rodrigo et al., 2004). These mixed results may be explained, at least in part, by the different research designs used in these studies. Even so, the findings generally indicate that acquisition is likely to occur if input is comprehended by the learner, but comprehension does not necessarily guarantee acquisition (Ellis, 2008). Gass et al. (2013) also agree that "Although understanding alone does not guarantee that learning will occur, it does set the scene for learning" (p. 343).

Despite the contribution of the Input Hypothesis to the development of the SLA field, there remain several questions about it, and a number of authors have challenged key aspects. First, questions have been raised about the rather vague term 'current level' (*i*) and

'the next level of competence' (*i* + 1) for a learner. For instance, Gass (1997) claims that Krashen "gave us little indication of how we can know whether specific input is indeed at the *i* +1 level or at the *i* + 23 level" (p. 100). Earlier McLaughlin (1987a) argued that it is difficult to define the concept of a learner's level and this limits its application into the classroom, even when individual differences of competence are taken into consideration. Furthermore, if levels can be determined, there are also difficulties in providing learners with appropriate input at *i* + 1 level to promote their language acquisition. Other researchers have argued that language input is not limited to comprehensible input and not necessarily within the learner's *i* + 1 level of competence. Gass and Selinker (2000), for example, point out that language learners are exposed to large amounts of language input and may well notice input that is beyond the *i* + 1 range because of frequency, prior knowledge, and attention. Gass et al. (2013) also questions the issue of quantity as it is not clear what Krashen means by sufficient quantity of input: "How do we know whether the quantity is sufficient or not? One token, two tokens, 777 tokens?" (p. 132). It has been argued by other critics that second language acquisition is achieved not only by means of comprehensible input, but also through other kinds of input such as comprehended and incomprehensible input. In fact, White (1987) claims that it is incomprehensible input, not comprehensible input that is necessary for L2 development. This is because when learners receive some incomprehensible input or when comprehension difficulties occur, they notice the inadequacy of their own interlanguage systems and make modifications to their language. In this way, the noticing of their own inadequacy serves as a kind of negative feedback, another feature deemed important in L2 acquisition process (see Section 2.2.4 for a description of this). Gass et al. (1998) also share this view by arguing that incomprehensible input helps learners recognise the gap between their interlanguage grammar and that of the target language, and the need for linguistic modification. The incomprehensibility of some language structures can draw learners' attention to specific features that need to be acquired. Developing this further, Ellis and Shintani (2014) point out that "the main problem with the Input Hypothesis rests in the claim that acquisition will automatically take place if learners comprehend input" (p. 177), arguing that learners can use contextual clues in a top-down approach to process the input for comprehension, and so without attention to the input. They go on to argue that when learners approach input in this way, no learning can occur. Extending these arguments, the current study investigates

the effects of combined input-based and output-based instruction versus input-based instruction, relative to non-task-based narrow input instruction on the development of cognitive processes underlying L2 learners' speech production to determine whether input alone is sufficient or the combination of both input and output opportunities will be better for L2 speech production.

2.2.2 Interaction and the Interaction Hypothesis

Interaction is defined as "the oral exchanges a learner participates in – with native speakers or with other learners – which provide both 'input' and opportunities for 'output' (i.e. use of the L2 in production" (Ellis, 2015, p. 12). Although the current study does not focus on interaction per se, because it provides these opportunities within both the input and output conditions (key design constructs of this study), interaction is described in detail next.

Long (1981) proposed the **Interaction Hypothesis**, defining interaction in relation to input: "Input refers to the linguistic forms used; by interaction is meant the functions served by those forms, such as expansion, repetition, and clarification" (p. 259). In this early version, Long (1981, 1983) concurred with Krashen's (1985) claim that comprehensible input served as data for second language acquisition and emphasized the importance of interaction in making input comprehensible. He argued that interactional modifications occurring in the negotiation of meaning were more important because they provided learners with information about those linguistic structures they found problematic.

A number of authors indicate that along with input, interaction is a fundamental component in second language acquisition. Specifically, those subscribing to an interactionist perspective emphasize the role of interaction in providing access to those features facilitative of language learning, namely opportunities to receive input in the form of models (exemplars of linguistic features), produce output, and get feedback on their production (Ellis, 2008; Long, 1983, 1996; Mackey, 1999; Oliver, 1998, 2002; Pica, 1996). Through the process of interaction, and specifically when communication breakdown is overcome through a process of negotiation for meaning (Long, 1981, 1983; Oliver, 1998), 'facts' of the L2 are made salient (Ellis, 2008). Because of this, Gass et al. (2013) view conversational interaction, coupled with the input, as the driving force of language

development and state that "conversational interaction in an L2 forms the basis for the development of language, rather than being only a forum for practice of specific language features." (p. 378). This view is shared by many (e.g., Gass, 1997; Gass et al., 1998; Oliver, 1998, 2002; Pica, 1996), including Mackey et al. (2012) who state that: "the interactional 'work' that occurs when a learner and his/her interlocutor (whether a native speaker or more proficient learner) encounter some kind of communication breakdown is beneficial for L2 development (p.349)". In general, researchers agree that interaction, especially negotiated interaction, plays an important role as a facilitator in learners' L2 acquisition because it makes input comprehensible, facilitates communication, and increases the saliency of L2 forms.

Long's proposal has triggered numerous studies and empirical research investigating the impact of interactional modifications on comprehension by making input comprehensible. There is evidence from a number of studies that negotiated interaction facilitates comprehension (e.g., Ellis et al., 1994; Gass & Varonis, 1994; Loschky, 1994; Pica et al., 1987). These studies typically compare the effects of different types of input: baseline (unmodified) input, pre-modified/ modified input, and interactionally modified input. The results show that interactionally modified input results in better comprehension than pre-modified or unmodified input. Although there has been a concern that these positive results may be due to more 'time on task' for the interactionally modified input treatment, overall it is evident that interaction leads to a greater level of comprehensible input.

The early version of the Interaction Hypothesis, however, has not been without its critics. Their concerns relate to how comprehension contributes to L2 acquisition. These concerns were addressed in Long's (1996) updated Interaction Hypothesis. Ellis (2008) indicates that while the early version accounts for the importance of comprehensible input provided through negotiation, the later version explains how interaction facilitates acquisition by involving the learner's internal learning mechanism, which is considered an advance on the early version.

Specifically, in his update version of the Interaction Hypothesis, Long (1996) takes account of all constructs said to be facilitative of SLA: comprehensible input, modified output, interaction, and noticing, which he describes as interlinked and facilitative of language acquisition. Furthermore, these constructs are captured within the parameters of

data for acquisition, which he describes as including both positive and negative evidence (negative feedback including such things as negotiation for meaning, recasts, metalinguistic comment, and explicit correction). As Oliver (1995) outlines, positive evidence provides learners with models of correct target forms, whereas negative evidence has a role to play by showing learners what is ungrammatical and unacceptable in the target language. For instance, through interaction, learners may notice non-salient forms, direct their attention to the gap in their own interlanguage system, and modify their output, all of which enable learning to take place. However, Long (1996) is very cautious - claiming that negative feedback facilitates, but not guarantees acquisition.

The updated Interaction Hypothesis has inspired a large number of studies investigating whether and how interaction facilitates acquisition (see Mackey, 2007). Mackey and Goo (2007) carried out a meta-analysis that examined the effects of interaction on language learning. They indicated that negotiated interaction had positive effects on vocabulary and grammar acquisition, but that the effects were stronger on the former than the latter. They also noted that negotiation benefits both the speaker and the listener involved in the interaction. It was revealed that such interactions in the laboratory and foreign language contexts result in larger effects than those conducted in the classroom and second language contexts. Interestingly, however, Ellis et al. (2006) and Mackey (1999) found even stronger effects in the delayed post-test. This may be explained by “the fact that sometimes information is not immediately integrated into the learner’s interlanguage, but is put into storage and only fully activated later” (Ellis, 2015, p. 157). Whilst many of these studies were undertaken orally, Iwasaki and Oliver (2003) and Ziegler (2016) reported that negotiated interaction also occurs in computer-based interaction, with online and face-to-face negotiation showing similar effects on acquisition. Overall, interaction has been consistently observed to serve as a facilitator of language acquisition.

The updated version of the Interaction Hypothesis, however, also has been subject to a number of criticisms. First, it is argued that the Interaction Hypothesis may not be able to explain what can and cannot be learned through negative evidence, nor how all linguistic aspects are acquired by learners (Ellis, 2008; Gass, 2003). Second, different types of negative feedback maybe not suitable for all levels of proficiency, for example, beginner learners may

not negotiate effectively due to a lack of language resources (Pica, 1996). Oliver (2000) also describes how the age of the learners affects the utility of the feedback.

Despite these caveats, the Interaction Hypothesis has been the focus of numerous studies and remains a key theoretical perspective in the field of SLA. As Ellis (2008) indicates: "no theory of L2 acquisition is complete without an account of the role played by interaction and the IH continues to be one of the most convincing statements of this role to date" (p. 260). Although the current study does not investigate interaction per se, learners participating in the output-based tasks have opportunities for interaction; and, therefore, it is important to be cognizant of the role it plays in SLA.

2.2.3 Output and the Comprehensible Output Hypothesis

Output or comprehensible output is the third component deemed essential for successful second language learning. Output refers to the language a learner produces, either by speaking or writing (Ellis, 2015). According to VanPatten (2003), output is the language the learner produces with a communicative intent. Swain (1985) uses the term 'comprehensible output' to refer to the need when a learner is "pushed toward the delivery of a message that is not only conveyed, but that is conveyed precisely, coherently, and appropriately" (p. 249). She claims that comprehensible input, though crucial, is not sufficient for acquisition because learners can use context to comprehend input without using the syntax, and suggests that what is lacking is output.

Swain (1985, 1995, 2005) proposed the **Comprehensible Output Hypothesis** as an addendum to the Input Hypothesis. From her observation from French immersion programs in Canada, Swain was able to conclude that 'comprehensible input' alone was not sufficient for L2 acquisition. She argued that the immersion students, in spite of seven years of receiving comprehensible input, still did not fully acquire the target language. This was not because of insufficient comprehensible input, but rather because of limited opportunities for comprehensible output. Swain (1985) claimed that opportunities to produce output, especially 'pushed output', facilitates acquisition and that comprehensible output "is a necessary mechanism of acquisition independent of the role of comprehensible input" (p. 252). The point is that in comprehending a message learners can use top-down strategies with little syntactic processing, but production requires learners to process syntactically

(i.e., bottom-up processing) and pay attention to form if they are pushed to produce concise, coherent, and socio-linguistically appropriate messages. Swain (2005) put it simply: "the output hypothesis claims that the act of producing language (speaking or writing) constitutes, under certain circumstances, part of the process of second language learning" (p. 471).

Swain (1995, 2005) suggested three main functions of output:

1) the *noticing/triggering function*: when learners produce the language, they may realise a mismatch between what they want to say and what they are able to say, and so they recognise their linguistic problems;
2) the *hypothesis testing function*: when learners produce output, they can test "a hypothesis about comprehensibility or linguistic well-formedness" and receive feedback from an interlocutor;
3) the *metalinguistic/reflective function*: learners consciously reflect on their learning, and thereby "control and internalize linguistic knowledge"

The role of output in L2 acquisition has been and continues to be controversial. While Krashen (1982) acknowledges only one indirect role for output, namely generating comprehensible input, Swain (1985, 1995, 2005) argues that output is a 'cognitive tool' in second language learning that has several roles independent of comprehensible input; and that the role of output changes from being the 'product' to the 'process' of learning. She points out that the minimum role of output is "to provide opportunities for contextualized meaningful use, to test out hypotheses about the target language, and to move the learner from a purely semantic analysis of the language to syntactic analysis of it" (Swain, 1985, p. 252).

Similarly, Gass and Selinker (2008) consider four ways that output can contribute to the learning process: (1) getting feedback; (2) testing hypotheses; (3) improving automaticity; and (4) moving meaning-based processing to grammar-based processing. These are briefly described, in turn, next:

1) *Getting feedback* provides learners with important information about their performance. Feedback to learners' about their production can be given explicitly (e.g., metalinguistic explanation) or implicitly (e.g., recasts), through

input-providing or output-prompting interactions (Ellis, 2008). It is noted that different forms of corrective feedback may have different impacts on learning (Mackey & Goo, 2007).

2) Output can serve as a way of *testing a hypothesis*, especially when it occurs in a negotiation sequence. It is through negotiation and feedback that learners have opportunities to think about the language they produce, test their initial hypotheses, make modifications, and arrive at a correct form. Pica et al. (1989) suggest that learners "test hypotheses about the second language, experiment with new structures and forms, and expand and exploit their interlanguage resources in creative ways" (p. 64)

3) Output has a part to play in the development of *fluency and automaticity* (i.e., fast, unconscious, and effortless processing). In language performance, automaticity means that one is able to take control over linguistic knowledge, bringing together a number of skills from perceptual, cognitive, and social domains. Gass and Selinker (2008), for instance, claim that "the consistent and successful mapping (practice) of grammar to output results in automatic processing" (p. 345)

4) Output "may force the learner to move from semantic processing to *syntactic processing*" (Swain, 1985, p. 249). This means processing language at the semantic level is not sufficient; learners need to move to processing at the syntactic level, which is essential for accurate production of language.

There is considerable evidence that provides support for the role of output in L2 acquisition. According to Ellis (2008), evidence can be indirect or direct. Indirect evidence is found in the studies that examine whether or not, and how much modified output is produced. The studies of this type (e.g., Mackey et al., 2003; Swain & Lapkin, 1995; Van den Branden, 1997) have shown that learners, both adults and children, produce and modify their output in response to corrective feedback such as clarification requests. Direct evidence comes from the studies that investigate the effects of pushed output in the acquisition of linguistic forms. However, studies so far have reported mixed results. While some like Ellis and He (1999); De la Fuente (2002); Izumi (2002); and Nobuyoshi and Ellis (1993) suggest that pushed output has effects on learners acquisition, others (e.g.,

McDonough & Mackey, 2006; Smith, 2005) find no relationship between output that includes different types of negotiation and acquisition. Shehadeh (2002) points out that despite a large amount of research into the Comprehensible Output Hypothesis, "no definitive conclusions can be made, because the question of whether and how learners' output, or output modification, helps with L2 learning is still largely unanswered" (p. 601). On this basis, other theorists argue that output has no role in acquisition. For example, Krashen (1998) argues that:

1) output, especially comprehensible output, is too scarce for real contribution to linguistic competence,
2) acquisition can occur without language production, and
3) there is no direct evidence of output leading to language acquisition.

Despite such claims, generally output is deemed to facilitate language acquisition in that output provides learners opportunities for using language, getting feedback, testing hypotheses, and moving to syntactic processing. As Ellis (2015) points out, it is now agreed that both comprehensible input and comprehensible output are necessary for learners to acquire high levels of linguistic competence in a second language. In the present study, effects of input-based instruction using either tasks or narrow reading and listening are compared with instruction combining both input-based and output-based tasks on the development of learners' L2 speech production.

2.2.4 Noticing and the Noticing Hypothesis

As can be seen, a fundamental contribution of interaction and output is that they can serve to promote noticing – and in particular noticing the gap between the learner's own output and the target language. The **Noticing Hypothesis** was developed by Schmidt (1990, 1994, 2001) and highlighted the important role of 'noticing' and 'noticing-the-gap' in L2 acquisition. He described 'noticing' as conscious attention to the formal features in the input and 'noticing-the-gap' as identifying the difference between the input and the output that the learner is able to produce. While in Schmidt's earlier work (1990, 1994), noticing was discussed in terms of 'consciousness', in his 2001 publication, 'attention' and its role as a necessary factor for all aspects of L2 learning was addressed:

> The allocation of attention is the pivotal point at which learner-internal factors (including aptitude, motivation, current L2 knowledge, and processing ability) and learner-external factors (including the complexity and distributional characteristics of input, discoursal and interactional context, instructional treatment, and task characteristics) come together. What then happens within attentional space largely determines the course of language development, including the growth of knowledge (the establishment of new representations), fluency (access to that knowledge), and variation. (Schmidt, 2001, p. 11)

It can be seen that Schmidt's earlier work (1990) represents the strong form of the Noticing Hypothesis, arguing that "subliminal language learning is impossible, and that noticing is the necessary and sufficient condition for converting input to intake" (p. 129). However, Schmidt's (2001) later work reflects a weaker version, acknowledging that some (though not much) learning might be possible without conscious attention: "people learn about the things they attend to and do not learn much about the things they do not attend to" (p. 30). He also claims that attention is "the necessary and sufficient condition for long-term memory storage to occur" (Schmidt, 2001, p. 20) and it must be specifically focused, directed to relevant aspects of language, and be for a particular learning purpose. It is also noted that noticing needs not to be intentional, it can occur incidentally while learners focus on comprehending input.

Schmidt's view about the role of conscious noticing and attention in L2 acquisition has been the focus of considerable debate. On the one hand, there are those who reject the importance of consciousness in processing input, claiming that acquisition is a subconscious process (e.g. Krashen, 1981, 1985). On the other hand, considerable support is provided by a number of other researchers (e.g. Ellis, 2008; Lynch, 2001; Skehan, 1995). In fact, both Swain (1995, 2005) and Long (1996) take account of conscious noticing as a facilitative connection between input and acquisition. Swain (1995), for instance, claims the noticing function of output: "in producing the target language (vocally or subvocally), learners may notice a gap between what they want to say and what they can say, leading them to recognize what they do not know, or know only partially, about the target language (p. 125)". That is, by noticing their linguistic problems or the gap in their interlanguage system, learners are then able to identify what they need to learn. This can facilitate the activation of the previous knowledge or the generation of new knowledge. Long (1996) emphasizes the importance of learners' attention to linguistic form claiming that negative feedback

(such as through negotiation of meaning and recasts) benefits acquisition "because it connects input, internal learner capacities, particularly selective attention, and output in productive ways" (p. 451-452).

The role of noticing has also been supported by Ellis and Shintani (2014) who suggest that "those linguistic forms in the input that are noticed (i.e. consciously attended to) are more likely to be acquired than those that are not noticed. Acquisition is also enhanced when learners notice-the-gap (i.e. pay attention to the difference between the input and their own output)" (p. 187).

Learners' attention to linguistic form, that is their noticing, can be induced by a number of means, including negative feedback, and saliency raising measures such as interactional modifications. It can also be promoted by way of input modifications, typically by enriching the input with a high frequency of the target forms, enhancing the input (phonologically or typographically), or elaborating texts to make linguistic features in the input salient to the learners (Ellis & Shintani, 2014; Sharwood Smith, 1993; VanPatten, 1996). A number of studies have looked at how these input modification techniques benefit noticing and acquisition. According to Ellis and Shintani (2014), although these studies vary in their methodology and report mixed findings, they reach some common conclusions, namely:

> Modified input, whether enriched, enhanced or elaborated, can help learners to notice and acquire both unknown words and grammatical features and to use partially acquired features more accurately.
>
> In general, modified input is more effective if learners' attention is explicitly drawn to the target items. Enriched input alone is less effective as it results in only low-level awareness of the items.
>
> Noticing appears to be related to learning; learners learn when 'detection' of the target items has taken place. (p. 186)

In general, learners are more likely to pay attention to form when linguistic features of input are made salient, and consequently learning is more likely to occur if leaners notice the gap in their linguistic knowledge. Therefore, in the current study, input is enriched with high-frequency provision of the target form (e.g., verbs in past simple) to induce learners' attention.

Generally (but not always), research on those features described above – input, interaction, output, and noticing - are informed by cognitive approaches to SLA. The next section discusses these in more detail.

2.3 Cognitive Approaches to SLA

Cognitive SLA is "a branch of SLA that examines the mental processes involved in the acquisition of a second language" (Ellis, 2015, p. 171). Within cognitive SLA, there are two popular approaches: emergentist or constructionist approaches and processing approaches (Mitchell & Myles, 2004).

Emergentist or **constructionist approaches** to second language learning draw on a usage-based view of language development. That is, language development is driven by communication, and language learners do not need to have an innate, language-specific acquisition device. As N. Ellis (2003) points out, constructivists argue that "the complexity of language emerges from associative learning processes being exposed to a massive and complex environment" (p. 84). In their views, learning occurs on the basis of associative processes and involves the analysis of patterns in the language input. Language development is seen as the result of rule-like regularities that emerge from the billions of associations that are made during language use (N. Ellis, 2003; Mitchell & Myles, 2004). That is, second language acquisition occurs through language usage, by extracting regularities from language input, rather than the construction of abstract rules. These models have much less currency for the current study and, therefore, a further detailed description is not provided.

Processing approaches investigate how L2 learners process linguistic information, and how their L2 processing ability develops over time. Processing approaches to SLA are commonly concerned with the operation of processing mechanisms and capacities of the brain when dealing with second language learning. According to Mitchell and Myles (2004), the two most popular approaches are Information-processing models and Processability Theory. While information-processing models (Anderson, 1980, 1983, 1985; McLaughlin, 1987b, 1990; Skehan, 1998a) examine how different memory stores deal with second language information, and how the automatization and restructuring of this information take place, Processability Theory (Pienemann, 1998, 2005) investigates specifically the

factors that affect the way learners process the second language input. Processability Theory is not used to inform the current study because it is more concerned with the sequence of acquisition of grammar structures and the hierarchical order of processability, and does not address how L2 forms are integrated into the learner's interlanguage system, whereas the present study focuses on how information in the input is processed and automatized through different stages of speech production during learners' performance of monologic tasks. Instead, information-processing models are used as a basis in the current study and, therefore, are described next, followed by the Limited Attentional Capacity Model (Skehan, 1998a).

2.3.1 Information-Processing Models

One of the earliest and most influential information-processing models in SLA is Anderson's **Adaptive Control of Thought (ACT) Model** (Anderson, 1980, 1983, 1985). In terms of knowledge representation, this model distinguishes declarative knowledge from procedural knowledge. Declarative knowledge is *knowing about* something, and requires attention of the speaker, whereas procedural knowledge is *knowing how* to do something, which does not require attention. There are three memory stores in the model: two long-term memory stores (a declarative memory and a procedural memory) and one limited-capacity memory store (a working memory). Anderson posits that both declarative knowledge and procedural knowledge are processed via working memory. For rapid performance (e.g., speech production) to take place, it is necessary for declarative knowledge to be converted into procedural knowledge because procedural knowledge is attention-free and can be processed in larger units without taking up much capacity in the working memory. Anderson claims that learning involves the conversion of declarative knowledge into procedural knowledge and practice has a role to play in this process. According to Anderson (1980, 1983, 1985), the transforming process consists of three stages: (1) the cognitive stage, (2) the associative stage, and (3) the autonomous stage. In the cognitive stage, a description of the procedure is learned. For example, the learner would learn that 'walked' consists of 'walk' and '-ed', but might not be able to produce 'walked' correctly in conversation. The associative stage involves working out the way to perform the skill. For example, the learner may learn from 'walked' and 'cleaned' that they both consist of '-ed' when used in past tense, and realize that adding '-ed' to the verb is the

procedure for generating past tense verbs. Anderson (1983) noted that it is in the associative stage that errors are likely to occur. For instance, the learner may produce erroneous past tense verbs (e.g., goed) by overgeneralizing the adding '-ed' rule. In the autonomous stage, the skill gets automatized and does not require any declarative knowledge. For example, the learner's action of adding '-ed' to past tense verbs become increasingly automatic. With regard to foreign language learning, Anderson (1980) sees the difference between foreign language learners with native speakers in terms of how these kinds of knowledge develop. The process of learning a foreign language is described as below:

> When we learn a foreign language in a classroom situation, we are aware of the rules of the language, especially just after a lesson that spells them out. One might argue that our knowledge of the language at that time is declarative. We speak the learned language by using general rule-following procedures applied to the rules we have learned, rather than speaking directly, as we do in our native language. Not surprisingly, applying this knowledge is a much slower and painful process than applying the procedurally encoded knowledge of our own language. Eventually, if we are lucky, we can come to know a foreign language as well as we know our native language. At that point, we often forget the rules of the foreign language. It is as if the class-taught declarative knowledge had been transformed into a procedural form. (Anderson, 1980, p. 224)

It can be inferred that when declarative knowledge has been transformed into procedural knowledge, it is accessed more quickly and efficiently. As a result, the learner's speech becomes more fluent. As Towell et al. (1996) claim, Anderson's ACT model provides an account of how declarative linguistic knowledge may be turned into procedural linguistic knowledge which is crucial for speech production and fluency development. However, it may be not possible for L2 learners to achieve the same level of procedural knowledge as the native speakers. While first language (L1) learners mostly reach the autonomous stage, L2 learners normally reach the associative stage. Anderson (1980, 1983, 1985) also acknowledges that although many L2 learners have a fairly high degree of procedural knowledge and can use L2 rules without conscious awareness, they still do not reach full autonomy.

Another information-processing model of SLA is McLaughlin's (1987b, 1990) **Information Processing Model**. This model is not dissimilar to Anderson's ACT model, but with different terminology. McLaughlin draws on the idea that learners have limited

information-processing capacity and calls on the need for *automatization* and *restructuring*. Automatization refers to learners' control over linguistic knowledge while restructuring means the changes to internalized representations in terms of organization when new information is integrated into the learner's existing interlanguage system as a result of learning. Like Anderson, McLaughlin's (1987b, 1990) model views second language learning as the acquisition of a complex cognitive skill. This process of language learning is summarized as follows:

> To learn a second language is to learn a skill, because various aspects of the task must be practised and integrated into fluent performance. This requires the automatization of component subskills. Learning is a cognitive process, because it is thought to involve internal representations that regulate and guide performance … As performance improves, there is constant restructuring as learners simplify, unify, and gain increasing control over their internal representations (Karmiloff-Smith 1986). These two notions – automatization and restructuring – are central to cognitive theory. (McLaughlin, 1987b, pp. 133-134)

From an information processing perspective, learning a language is like learning other skills. Initially, learners need to resort to controlled processing, which requires attentional control and is limited by processing capacity of the short-term memory. To maximize the processing capacity, learners routinize their skills through practice. Practice results in changes in interlanguage by learners creating more chunks of information that are available for automatic processing, which requires minimal attentional control and therefore reduces the load on the learner's processing capacity. Language performance requires the learner to bring together a number of subskills in perceptual, cognitive, and social domains. The more these skills become automatized, the more attentional resources are freed up for processing new information and learning more complex language. Then, "automatic processes can work in parallel, activating clusters of complex cognitive skills simultaneously" (Mitchell & Myles, 2004, p. 101). Ultimately the move from controlled to automatic processing via practice is deemed to be learning. The continuing movement from controlled to automatic processing leads to a constant restructuring of the L2 learner's linguistic system.

As can be seen from these two information-processing models, despite different terminology, both models take into account two kinds of knowledge: declarative knowledge and procedural knowledge. In addition, these kinds of knowledge are connected in that

declarative knowledge which requires controlled processing can be converted through practice into procedural knowledge which then allows for automatic processing. Thus, these models explain the step-by-step nature of learning: when knowledge becomes proceduralized/autonomous, it is accessed automatically, quickly, and efficiently without exhausting the working/short-term memory with limited processing capacity. Then new declarative knowledge can be attended to and eventually become proceduralized via practice.

2.3.2 The Limited Attentional Capacity Model

Skehan's (1998a) cognitive approach to language learning also draws on the assumption of limited processing capacity, but in a different way. He claims that L2 learners (like native speakers) possess a *dual-mode* system, consisting of a rule-based system and an exemplar-based (memory-based) system. The rule-based system includes generative rules which require more processing, and are used to formulate complex and well-formed sentences. The exemplar-based system consists of discrete lexical items and formulaic chunks of language, which are required for fast, fluent language performance, especially in real-time communication. According to Skehan (1998a), it is necessary that L2 learners develop and maintain the capacity for dual-mode processing during performances if they progress close to the target norms of competence. Skehan (1998a) posits that the two systems co-exist and that "language users can move between these systems, and do so quite naturally" (p. 54). In this way, Skehan (1998a), in his **Limited Attentional Capacity Model** or **Trade-off Hypothesis**, advocates a single-resource model of attention and proposes a dual-mode processing model which is composed of an exemplar-based system (prioritizing fluency) and a rule-based system (prioritizing accuracy and complexity). This model claims that due to learners' limited attentional capacity, "learners cannot attend to everything equally" (Skehan & Foster, 1999, p. 96) and may encounter difficulty focusing on both meaning and form. Thus, while producing language, learners need to prioritize either meaning by accessing their exemplar-based system or form by accessing their rule-based system. Consequently, learners cannot focus on all three dimensions of language performance (fluency, complexity, and accuracy), and trade-off effects between fluency (focus on meaning) and complexity and accuracy (focus on form) may occur during the course of speech production. For example, in performing cognitive complex tasks, learners

may achieve greater fluency accompanied by greater accuracy or greater complexity, but not both. Then, task design and implementation variables need to be taken into consideration so as to have balanced L2 instruction that alternatively directs learners' attention to different aspects of task performance.

2.4 Summary

This chapter provides an overview of those theoretical perspectives that inform the current study and in particular the contribution of input, interaction, output, and noticing. Their connection to various theories, especially key hypotheses, was also reviewed. The next chapter discusses a particular pedagogy called Task-Based Language Teaching, and specifically how through the use of tasks, opportunities are provided for these theoretical constructs to occur.

Chapter 3. TBLT and SLA

This chapter first reviews what is meant by 'task' and describes different types of tasks. Next, task-based language teaching (TBLT) is compared with task-supported language teaching (TSLT) in terms of how tasks are used in each, how learners' attention is directed to form within tasks, and the theoretical underpinnings of both, especially in terms of how L2 is learned. The design of a task-based lesson, together with the teacher and learners' roles is discussed. Finally, the distinction between input-based and output-based task instruction is outlined, followed by a description of the relevant empirical studies.

3.1 Definitions of Tasks

To understand what is meant by task-based language teaching, it is first necessary to define what 'task' means.

In the last three decades, tasks have emerged as a crucial construct in SLA research and second language pedagogy. They have been defined differently by various researchers, sometimes more broadly, other times in a narrower way, and then with respect to how they contribute to learning and pedagogy. Nunan (2006) distinguished real-world/target tasks (those in the real world outside the classroom) from pedagogical tasks (those used in the classroom).

In a broad sense, Long (1985b, p. 89) defined a task as "a piece of work undertaken for oneself or for others, freely or for some reward" with examples of tasks including painting a fence, dressing a child, filling out a form, etc. Therefore, according to this definition, tasks can be any real-world activity performed with or without the use of language.

In a narrower sense, tasks are those activities that require the use of language and involve different processes in task performance, normally in the classroom (Bygate et al., 2001; Nunan, 2006; Richards et al., 1985; Skehan, 1996). According to Richards et al. (1985), a task is:

> an activity or action which is carried out as the result of processing or understanding the language (i.e., as a response). For example, drawing a map while listening to a tape, and listening to an instruction and performing a command, may be referred to as tasks. Tasks may or may not involve the production of language. A

> task usually requires the teacher to specify what will be regarded as successful completion of the task (p.289).

This definition of tasks takes a pedagogical perspective and focuses on meaning rather than linguistic structure. In a similar way, Nunan (2006) described a task as a classroom activity that focuses on meaning rather than form and requires learners to comprehend, manipulate, produce or interact in the target language. He also indicated that a task needs to be an independent 'communicative act' and have 'a sense of completeness'.

Prabhu (1987) offered a more pedagogically oriented definition of a task describing it as "an activity which required learners to arrive at an outcome from given information through some process of thought, and which allowed teachers to control and regulate that process" (p. 24). In this way, particularly the inclusion of the phrase 'some process of thought', his definition is oriented towards cognitive processes.

Developing this further, Skehan (1996) defined tasks by including reference to meaning, to an implied pedagogical perspective, and a real-world connection. Specifically, he stated that a task is "an activity in which: meaning is primary; there is some sort of relationship to the real-world; task completion has some priority; and the assessment of task performance is in terms of task outcome" (p. 38).

Bygate et al. (2001) also defined 'task' referencing these aspects, reflecting language use, with emphasis on meaning and achievement. They claim that this 'basic, all-purpose definition' of task can be modified depending upon the purposes for which tasks are used (e.g., testing or learning).

Although R. Ellis (2003) agreed that the definitions of tasks could be variable to some degree, he argued for a need for an inclusive definition which can identify "the essential commonalities in tasks, irrespective of their actual use". R. Ellis (2003, 2009) then proposed such a definition with four criteria, which was further elaborated in Ellis and Shintani (2014) as follows:

1. The primary focus should be on 'meaning' (i.e., learners should be mainly concerned with encoding and decoding messages not with focusing on linguistic form).
2. There should be some kind of 'gap' (i.e., a need to convey information, to express an opinion or to infer meaning).

3. Learners should largely rely on their own resources (linguistic and non-linguistic) in order to complete the activity. That is, learners are not 'taught' the language they will need to perform the task, although they may be able to 'borrow' from the input the task provides to help them perform it.
4. There is a clearly defined outcome other than the use of language (i.e., the language serves as a means for achieving the outcome, not as an end in its own right). Thus, when performing a task, learners are not primarily concerned with using language correctly, but rather with achieving the goal stipulated by the task (pp.135-136).

In brief, for a language-teaching activity to be a task, it must meet these four essential characteristics: (1) a focus on meaning, (2) some kind of gap, (3) learners using their own linguistic resources, and (4) having a non-linguistic communicative outcome.

These criteria of 'tasks' mean it is possible to distinguish them from 'situational grammar exercises'. Though the purpose of both 'tasks' and 'exercises' is to learn a language, the former are meaning-focused activities in which participants act as language users, whereas the latter focus only on form (i.e., the structure of the language) with the participants functioning mainly as learners. In the former, learners are often not told what the linguistic feature is and thus engage in meaning and language use. However, in the latter, learners are informed quite explicitly of the focused language item and are more likely to focus on how to use it correctly (Ellis & Shintani, 2014). This distinction between tasks and exercises is also crucial for pedagogy – teachers will claim they are using tasks, when in fact they are having their learners engage with language exercises (Oliver & Bogachenko, 2018). To take account of all these aspects, in the current research, R. Ellis's (2003, 2009) definition of tasks, as described above, especially within a TBLT approach, is adopted.

3.2 Types of Tasks

Tasks may be of different types, and researchers have spent considerable effort trying to identify which most effectively promote language learning. However, as with the definition of tasks, the ways the different task types can be classified also vary. For example, Willis (1996) divided tasks, according to the kinds of operation the learners need to perform: (1) listing; (2) ordering and sorting; (3) comparing; (4) problem-solving; (5) sharing personal experiences and (6) creatives tasks with a detailed description of operation for each type.

By contrast, much earlier, Prabhu (1987, pp. 46-47) differentiated tasks based on the kind of cognitive activity involved. He grouped tasks into three main categories namely information-gap, reasoning-gap, and opinion-gap. Information-gap activity involves “a transfer of given information from one person to another – or from one form to another, or from one place to another – generally calling for the decoding or encoding of information from or into language”. Reasoning-gap activity involves “deriving some new information from given information through processes of inference, deduction, practical reasoning, or a perception of relationships or patterns”. Opinion-gap activity involves “identifying and articulating a personal preference, feeling, or attitude in response to a given situation”.

Pica et al. (1993) categorized tasks as either one-way or two-way exchanges, and based on this proposed five types of pedagogic tasks: (a) jigsaw; (b) information gap activities; (c) problem solving; (d) decision making; and (e) opinion gap. They also suggested that the most effective tasks are those in which:

1. Each interactant holds a different portion of information which must be exchanged and manipulated in order to reach the task outcome.
2. Both interactants are required to request and supply this information to each other.
3. Interactants have the same or convergent goals.
4. Only one acceptable outcome is possible from their attempts to meet the goal. (p.17)

On this basis, they asserted that jigsaw and information gap tasks are more effective in promoting language learning than other types of tasks

More recently, R. Ellis (2003, 2009) classified tasks as unfocused or focused tasks. Unfocused tasks are those that are not designed to elicit a specific form, but rather provide learners with opportunities to use different forms, while focused tasks are designed to induce learners to use a particular linguistic feature. However, the target linguistic feature of a focused task is not explicitly taught, such as in the way exercises treat particular forms, and instead the task continues to meet the four necessary criteria (as described above).

Tasks have also been classified as input-based and output-based (see Section 1.2.2). Input-based tasks often involve receptive skills and are in the form of reading and listening activities which do not require learners to produce language. In contrast, output-based tasks are those that engage learners in language production (R. Ellis, 2003).

In the current research, both input-based and output-based tasks are used because the aim is to compare the effects of input-based tasks versus a combination of input-based and output-based tasks on learners' fluency development. These tasks must of course satisfy all four criteria of a task as stated above.

3.3 Task-Based versus Task-Supported Language Teaching

Task-based language teaching has its origins in Communicative Language Teaching (CLT), which emerged in the 1970s as an alternative to traditional approaches to language teaching (Ellis, 2017). Within CLT, functional language use was considered the centre of language teaching. Richards and Rodgers (2001) defined CLT as "a broad approach to teaching that resulted from a focus on communication as the organizing principle for teaching rather than a focus on mastery of the grammatical system of the language" (p. 36). CLT, as an approach rather than a method, considers developing learners' communicative competence as the ultimate goal of language teaching through a variety of classroom tasks and activities. This learner-centred approach focuses primarily on learners' use of language, fluency development, and their ability to communicate in different settings (R. Ellis, 2003; Richards & Rodgers, 2001; Van den Branden et al., 2009). Later, CLT developed into a weak and a strong version (Howatt, 1984, as cited by R. Ellis, 2003, p. 28). The weak version focused on 'learning-to-communicate' and considered communicative tasks as means of practicing the language items at the production stage. However, the strong version emphasized 'learning-through-communication' claiming that learners had opportunities to discover how language is used in communication (R. Ellis, 2003). In many ways, the difference between the weak and strong form of CLT parallels the distinction between TSLT and TBLT. According to Ellis (2017), both approaches employ tasks, but the ways they use tasks, direct learners' attention to form, and their theoretical underpinnings of how L2 is learned differ considerably.

In TSLT, tasks are incorporated into traditional structural approaches to teaching as a necessary component, but they are not an essential part of the language program. In this way, teachers use tasks as a supplement to other methods of instruction. For example, in traditional structural methodology, which often involves a three-stage presentation-practice-production procedure, tasks can be used in the third stage of the process. In the first stage, a language item is presented to the learners through examples and/or

explanation. It is in this stage that the learners' attention is focused on specific forms. In the second stage, the target item is practised by means of different exercises. Tasks are then often used at the free production stage to provide communicative practice for the target language items (R. Ellis, 2003).

In this way, TSLT views language as a consisting of discrete linguistic units (e.g., lexical items or grammatical rules), which need to be acquired sequentially by the learners before they are able to use them for communication (R. Ellis, 2003; Ellis, 2017; Van den Branden et al., 2009). The syllabus is specified in structural terms with a list of linguistics features or notions/functions to be taught. Learners are also made aware of what target features to be learnt so that the can use them correctly. Learning in TSLT approach is, therefore, deemed intentional.

As such TSLT is underpinned by skill-learning theory (Anderson, 2000; DeKeyser, 1998), which claims that declarative knowledge of a skill or a linguistic form can be transformed into procedural knowledge through practice. As R. Ellis (2003) points out, "communicative practice serves as a device for proceduralizing knowledge of linguistic structures that have been first presented declaratively" (p. 146). From this point of view, tasks have an important role to play in TSLT by creating opportunities for learners to use the target forms communicatively, enabling them to proceduralize the declarative knowledge of a language. Tasks which are used for the purpose of practising pre-determined linguistic features must be, by this intent, focused tasks.

On the other hand, TBLT views tasks as the primary units of teaching in their own right, which constitute a both necessary and sufficient basis for language learning (R. Ellis, 2003; Richards & Rodgers, 2001). Proponents of TBLT (e.g., Long, 1985b, 1996; Prabhu, 1987) argue that effective teaching should be entirely based on tasks and task performance by engaging learners in real language use. The utility of such an approach is supported by Ellis (2008) who states "tasks, or rather the performance of them, create opportunities for learning" (p. 819).

TBLT, as defined by Ellis and Shintani (2014), is an approach that "aims to develop learners' communicative competence by engaging them in meaning-focused communication through the performance of tasks" (p.135). TBLT calls for holistic learner-driven learning, and a focus on both meaning and form (Ellis, 2009; Skehan, 2003; Van den Branden et al.,

2009). In TBLT, there is no explicit presentation of the target items, and when learners perform the tasks, their attention is induced rather than explicitly directed to linguistic forms. Furthermore, in TBLT attention to form is secondary to the primary focus on meaning. Research (Ellis, 2015; Skehan, 2003) suggests that the way tasks are designed and implemented can steer learners' attention to form in a way that is likely to promote acquisition. As Ellis and Shintani (2014) indicate, "a key principle of TBLT is that even though learners are primarily concerned with constructing and comprehending messages, they also need to attend to form for learning to take place" (p. 135).

In TBLT, therefore, learning is claimed to take place incidentally. Learners acquire the target features while focusing their attention primarily on meaning. Language is treated as "a 'tool' for achieving a communicative outcome rather than as an 'object' to be studied, analysed and displayed" (Ellis & Shintani, 2014, p. 136). The underpinning theories of language learning utilising a TBLT approach include the Interactionist perspective and usage-based language learning. According to the Interactionist proponents (e.g., Gass & Mackey, 2015; Long, 1996), interaction facilitates learners' acquisition by making input understandable, providing feedback, and pushing learners to produce comprehensible and, based on the feedback they receive, modified output. Usage-based theorists (Tomasello, 2003) claim that language structure emerges from language use. Learning starts with holistic ready-made chunks of language and then learners combine these to make more complex constructions. Because of this, teaching linguistic features to learners before their performance of a task is not supported. Instead, both linguistic and communicative competence is deemed to be developed by learners' engaging in meaning-focused communication, such achieved through the performance of tasks. As Ellis (2017) indicates, the aim of TBLT is to promote learners' language acquisition by engaging learners in natural language use. In this way, TBLT encompasses 'learning through doing' by providing learners with opportunities for meaning-focused communication using their existing linguistic and non-linguistic resources. As a consequence, task-based interactions enable the learning of new linguistic features which are acquired incidentally. In brief, tasks have an important role in promoting learns' incidental acquisition of language items; and developing learners' communicative and linguistic competence. To achieve this, learners need to access their

own resources to achieve the outcomes of the tasks in which focus on meaning is primary and the gap in information or opinions facilitates communication.

3.4 Design of a Task-Based Lesson

Various frameworks for designing a task-based lesson have been proposed by different researchers (e.g., R. Ellis, 2003; Prabhu, 1987; Skehan, 1996; Willis, 1996), but generally they all have three phases: pre-task, during/main-task, and post-task. According to Ellis (2006), of these phases only the during/main-task phase is compulsory. The pre-task and post-task phases are not obligatory, but they have a role to play in ensuring the effects of task performance in language acquisition. The three phases are described in detail next:

The pre-task phase can be used as a preparation for the learners' performance of the main task. Ellis (2006, p. 21) suggests four options for the pre-task stage: (1) performing a task similar task, (2) providing a model of how to perform the task, (3) engaging learners in non-task preparation activities, or (4) strategic planning of the main task performance. As Skehan (1996) notes, as attentional capacity is limited, pre-task activities should ease the cognitive load, releasing attentional capacity for the learners to focus on the language to be used.

The during/main-task phase involves what happens around the task itself. The task-as-workplan is now turned into "a task-in-action and a task-in-interaction" (Van den Branden, 2016). According to Ellis (2006), there are different instructional options with regard to teachers' plans for how the task is to be done (task-performance options), and teachers' and students' online decisions about how to perform the task (process options). Task-performance options include whether: (1) to require students to perform the task under time limits, (2) to let students access the input while performing a task, and (3) to introduce some surprising element into the task. Process options, on the other hand, involve online decisions about how the discourse arising from the task performance is enacted. These options cannot be planned, but must be undertaken while the task is performed.

The post-task phase involves activities following a task performance. In Ellis' (2006) view, this phase can be used to have learners (1) repeat the task performance, (2) reflect on how they performed the task, and (3) pay attention to form, especially the problematic

forms. Examples of post-task activities suggested by Ellis (2006) include, but are not limited to: (1) reviewing of learners' errors, (2) consciousness-raising tasks, (3) production practice activities, (4) noticing activities.

These varieties of options are summarised by Ellis and Shintani (2014) as shown in Table 3.1. It should be noted that this task-based framework, as described above, is suggestive, and there is not a pre-determined or a fixed structure for a task-based lesson (see Lambert, 2020 for discussion and examples). Instead, there is room for teachers' creativity and decision-making about what to add to the basic format of a lesson, which minimally consists of the during/main-task phase. Including either the pre-task or post-task phase or both to the lesson is optional. "Once the basic structure of the lesson has been decided, the specific option(s) to be included in each phase of the lesson can be considered" (Ellis, 2006, p. 262).

Table 3.1

Implementation Options in the Different Phases of a Task-based Lesson

Phase	*Options*	*Description*
Pre-task phase	1. Modelling performance of the task	Students listen or watch the task being performed by 'experts'.
	2. Pre-teaching language	The teacher presents language that will be useful for performing the task.
	3. Schema-developing	The teacher elicits and extends students' knowledge of the topic of the task.
	4. Strategic planning	The students are given time to prepare to perform the task before they actually perform it.
Main-task phase	1. Time pressure	Students are given only a limited amount of time to perform the task.
	2. Contextual support	Students are allowed to access the input data when they perform the task.
	3. Explicit instruction	The teacher takes time out from the performance of the task to explicitly teach a linguistic feature that is useful for performing the task.
	4. Surprise element	Additional information relevant to the task is provided after the students have started to perform the task.
Post-task phase	1. Repeat performance	Students are asked to repeat the task.
	2. Report	Students are asked to report the outcome of the task to the whole class.
	3. Language work	Students complete language exercises related to linguistic problems that they experienced when performing the task.

Note. From *Exploring Language Pedagogy Through Second Language Acquisition Research* (p. 142), by Ellis and N. Shintani, 2014, New York: Routledge. Copy right 2014 Rod Ellis and Natsuko Shintani.

3.5 Roles of the Teacher and Learners in TBLT

The teacher and learners play a crucial role in ensuring the successful implementation of any teaching methodology including TBLT. Their roles in TBLT classrooms have been highlighted in the literature of TBLT (e.g., R. Ellis, 2003; 2017; Van den Branden, 2009, 2016).

Researchers have suggested various roles of the teacher in the implementation of TBLT. Willis (1996), for example, suggests a variety of roles for a teacher including being a monitor, adviser, chairperson, language guide, and facilitator - which highlights the considerable demands placed on a teacher implementing TBLT. In response to Swan's (2005) criticism that "the thrust of TBLT is to cast the teacher in the role of manager and facilitator of communicative activity rather than an important source of new language" (p. 391), Ellis (2017) argues that in TBLT the teacher's role must be much more than merely managing and facilitating tasks. The teacher is required to "perform a variety of roles including those of manager and facilitator of communication but also the traditional roles of corrector and provider of new language" (Ellis, 2017, p. 114).

The two roles that Van Avermalet et al. (2006) believe the teacher should take are those of a motivator and a supporter. Specifically, they argue that in order for the learner's actual learning to take place, the teacher needs to maintain two core actions throughout three stages of a lesson (i.e., the planning stage, the performance stage, and the post-task assessment stage). These actions are (1) motivating learners to actively engage in task completion, and (2) interactionally supporting learners' performance in a way that promotes their negotiation of meaning, input comprehension, output production, and focus on form, all of which are deemed central in the acquisition of a second language (Van Avermalet et al., 2006, p. 175).

Similarly, Samuda (2009) also argues that the role of the teacher in a task-based lesson is to find "ways of working with tasks to guide learners toward the types of language processing believed to support L2 development" (p. 120). She emphasizes that while "a role for the task is to create opportunities for the formulation and negotiation of meaning", the teacher's role is "to 'lead from behind' (Gibbons, 1998) the task to complement and support what the task has set in motion" (p. 380). Specifically, in the complementary relationship with 'task', the teacher can help by guiding learners' attention to form-meaning mappings. This, according to Samuda (2009), can be achieved by the teacher's use of language from the input to: (1) create alignment with learner groups; (2) scaffold an implicit language focus; (3) introduce an explicit focus on new form-meaning relationship; and (4) frame negative feedback.

Recently, Van den Branden (2016) has approached the role of the teacher as an active contributor to the development of TBLT from three perspectives: (a) as a mediator of the students' language learning; (b) as a change agent in the innovation of second language education; and (c) as a researcher. As a mediator of language learning, the teacher's decisions and actions at different phases of a task-based lesson can promote students' learning and enhance the effectiveness of TBLT. Van den Branden (2009) points out that in the during-task phase, the teacher has at least three important roles to play:

> The teacher remains a crucial interactional partner in task-based language classrooms, by taking the role of motivator (i.e., launching the students into action by constructing joint projects), organizer (making sure that students know what they are expected to do and organizing temporal and spatial aspects of task performance), and, last but not least, conversational partner and supporter, as the more proficient, knowledgeable interlocutor who can feed the language-learning needs of different students in a wide variety of ways. (p. 284)

As a change agent, the teacher is considered a key figure in the implementation of TBLT because what the teacher actually does when working with tasks is crucial to ensure the success of a lesson or a program. Finally, the teacher as researcher needs to be "actively involved in classroom-based research on task-based language learning and teaching" (Van den Branden, 2016, p. 167) so that they can contribute to the development of TBLT as a researched pedagogy.

Together with the teacher, learners are considered indispensable actors on the stage of TBLT. As a learner-centred approach to instruction, TBLT requires the learners to play an independent and active role of a language user/ communicator in contrast to being a passive recipient of content provided by a teacher in traditional classrooms. Nunan (2004) emphasizes 'active learning' as one of the seven principles for TBLT, pointing out that "learners learn best through doing – through actively constructing their own knowledge rather than having it transmitted to them by the teacher" (p. 36). It is noted that it is the learners, not the teacher, who does the work. He also states that it is possible for learners to apply their own learning strategies, plan and monitor their own learning; and that learning outcomes will be impacted by learners' perceptions about what they should contribute to task performance. Ellis and Shintani (2014) state that "for tasks to work as they are intended, students need to function primarily as 'communicators' rather than as 'learners'" (p.143). They also suggest that learners may need to "move backwards and forwards

between these two roles" (p. 143), especially when they work collaboratively in groups to achieve the task outcome.

In general, teacher roles and learner roles are seen as two sides of a coin. Assigning learners a more active role in TBLT requires the teacher to take a different role rather than simply being a knowledge provider as happens in traditional classrooms. The role of teachers in TBLT is varied, but remains crucial regardless of whether they are a facilitator, mediator, motivator, organizer, partner, or supporter. Furthermore, according to Ellis and Shintani (2014), the teacher's role will change depending on whether the task is carried out individually, in pairs, or in groups. In many ways, the teacher can be considered the learners' most special interlocutor, especially when functioning as a task participant.

3.6 Input-Based versus Output-Based Task Instruction

The primary concerns of Instructed Second Language Acquisition (ISLA) researchers are "what is the best way to learn and/or teach an additional language?" (Loewen & Sato, 2017, p. 3) and how instruction can best promote language learning (Ellis, 2005b). However, it remains controversial because as Ellis (2005b) points out:

> There is no agreement as to whether instruction should be based on a traditional focus-on-forms approach, involving the systematic teaching of grammatical features in accordance with a structural syllabus, or a focus-on-form approach, involving attention to linguistic features in the context of communicative activities derived from a task-based syllabus or some kind of combination of the two. (p. 210)

Although there are different theoretical views about how much influence instruction can have on second language learning, most ISLA researchers agree that instruction does have an impact on L2 acquisition. Questions, however, remain as to whether it should be input-based or output-based, and form-focused or meaning-focused.

3.6.1 Distinguishing Input-Based and Output-Based Instruction

It is now generally accepted that in order for successful language learning to occur, learners need access to extensive L2 input as well as opportunities for producing output. On this basis, it has been proposed that task-based instruction should involve tasks that enhance both skill development and incidental language learning (Ellis & Shintani, 2014). However, debate remains as to whether these should be input-based tasks (comprehension-

based instruction - CBI) or output-based tasks (production-based instruction - PBI). The main distinction between CBI and PBI is whether or not it requires language production (Ellis & Shintani, 2014; Shintani et al., 2013; Tanaka, 1999, 2001). As discussed in Chapter 2, it is without a doubt that input plays a crucial role in language learning as the "sine qua non of acquisition" (Gass & Mackey, 2015, p. 181). What remains more controversial is the role that output plays, with some researchers supporting a strong, independent, and direct role (DeKeyser, 1997; Swain, 1985, 1995, 2005) while others advocating a weaker, supporting, and indirect role (Krashen, 1982, 2003; VanPatten & Williams, 2015).

Input-based instruction has been operationalized as the one in which learners are provided with input, and are not prohibited from producing output (Shintani, 2011; Shintani & Ellis, 2010; Shintani et al., 2013). Learners need to comprehend the given input in order to achieve the task outcome. Though learners are not required to produce output, they are free to engage in both social and private speech when working with the input. It is assumed that learners will pick up new language forms from the input providing that they can comprehend it and notice the inherent specific linguistic features. One type of input-based approach is VanPatten's (1996) **Input Processing Instruction**. Underpinning this approach is the belief that instruction alters the way learners perceive and process input (processing principles) resulting in input being more likely to become intake, thus facilitating the development of the learners' interlanguage systems. According to this model of input processing, when learners attend to a grammatical structure in the input and process the message, they are likely to establish a form-meaning connection, which is essential for language acquisition to occur. Therefore, processing instruction normally consists of three components: (1) explicit explanation of the target form, (2) explicit strategy training to help learners overcome possible processing problems, and (3) structured input activities to push the form-meaning mapping process and the conversion of input into intake.

In contrast, production-based activities both present learners with input and require them to produce output (speaking or writing). In this way, output-based interactions are often two-way between the teacher and learners or between learners as peer interactions. Both CBI and PBI can involve focused instruction (tasks designed to provide contexts for the use of a certain language item) or unfocused instruction (there is no pre-selection of the target language features). Unfocused input-based task instruction often provides learners

with general samples of the language such as through the use of listening or reading tasks, whereas unfocused output-based task instruction employs production tasks to elicit learners' free production of the language using their own resources.

3.6.2 Studies on Input-Based and Output-Based Instruction

There have been a lot of studies investigating the effects of input-based instruction versus output-based instruction on second language learning, mostly with regard to grammar and vocabulary acquisition. However, the results have been mixed, with evidence for the efficacy of both input-based instruction and output-based instruction, as well as for a combination of both types.

Some studies have found better results for output-based treatment groups over input-based groups (e.g., Erlam, 2003; Rassaei, 2012). Erlam (2003) investigated the effects of structured-input instruction and output-based instruction on the acquisition of a French grammatical item. The results revealed that both instructional groups outperformed the control group, and the output group performed better than the structured-input group in both comprehension and production tests. Similarly, Rassaei (2012) compared the effects of two types of input-based instruction (input textual enhancement and input enrichment) with output-based instruction (output production) on the development of L2 knowledge for EFL learners. 129 Iranian learners enrolled in five intact EFL classrooms were designated as four experimental groups (Textual enhancement, Input enrichment, Meaningful output, and Explicit instruction) and one control group in the study. The results suggested that while both input-based and output-based instruction can result in the development of L2 knowledge, output-based instruction was more effective than input-based instruction. The study also showed that only the input-based instruction which involved the use of textual enhancement gave rise to L2 development, indicating that exposure to input alone was not sufficient for L2 acquisition, and some kinds of intervention (e.g., textual enhancement in this case) was needed to facilitate the development of L2 knowledge.

On the other hand, other studies have shown that input-based instruction had similar or better results over output-based instruction for the development of both grammar and vocabulary knowledge (e.g., Shintani, 2011; Shintani & Ellis, 2010). Shintani and Ellis (2010) investigated the effects of input-based instruction and output-based

instruction on the incidental acquisition of English plural *-s* for young Japanese learners. The results showed that although both types of instruction lead to incidental acquisition of the grammar item (English plural *-s)*, with significantly better results than the control group on both the comprehension and production tests, there was evidence for the advantage for input-based instruction over output-based instruction in promoting incidental acquisition for young beginner learners. This advantage for the input-based instruction group was explained by the teacher's feedback which enhanced the form-function mappings of the item as learners attempted to comprehend the input. In the same vein, Shintani (2011) compared effects of input-based and output-based instruction on the acquisition of vocabulary for young beginning-level Japanese EFL learners. The results showed that the input-based instruction group performed better on comprehension tasks, and surprisingly, as well on the production tests as the production-based instruction group although they had fewer opportunities for production. Shintani explained that it may be because the input-based instruction provided opportunities for richer interaction (e.g., learners' degree of discourse control; negotiations of meaning) that promoted learning than occurred in the production-based group.

Tanaka (1999, 2001) investigated whether the combination of comprehension and production practice would have better effects on learners' receptive and productive knowledge. The results of his studies showed that by combining comprehension and production practice, the effect was greater than for providing comprehension or production practice alone in both comprehension and production tests. Moreover, the effect was durable over time for both forms of practice. It appears that when combined, "comprehension and production practice complemented each other to promote learning of the structure" (Tanaka, 2001, p. 23). From these results, he proposed that combining comprehension and production practice may enhance learning rather than when implemented separately because the combination can increase learners' comprehension and production abilities not only in the short-term, but also over time. On this basis, it has been recommended that the two forms of practice with their own unique role, should be incorporated in order to promote learners' success in second language learning. More recently, Smith (2015) conducted a study with teenager Japanese EFL learners to examine (1) whether combining comprehension practice and production practice lead to gains in

grammar knowledge, and (2) which combining order (i.e., alternating comprehension and production practices versus delaying production practice) would be more effective for the development of both receptive and productive knowledge of the English simple past *–ed*. The results revealed that both combined groups had significant improvement over time and the effects for both combining options were comparable. The results of Smith's (2015) study were consistent with Tanaka's studies (1999, 2001), confirming the benefits of combining input-based and output-based practice for L2 grammar instruction, regardless of the order in which they are combined.

The efficacy of both input-based and output-based instruction was generally supported by Shintani et al. (2013) when they conducted a meta-analysis of comparative studies that investigated the effects of the two types of instruction on L2 learning. The results of 35 studies included showed that (1) both CBI and PBI had positive effects on both receptive and productive knowledge; (2) for receptive knowledge, CBI was more effective than PBI in the immediate post-tests but not in the delayed tests; (3) for productive knowledge, CBI and PBI were equally effective in short-term but PBI showed better effects in the long-term. Overall, such outcomes did not lend support to skill-learning theory (DeKeyser, 2015), which claimed that CBI benefited learners' receptive knowledge development; and PBI was more effective for developing learners' productive knowledge. Nor did the results of this meta-analysis suggest the superiority of one type of instruction over the other on the acquisition of both receptive and productive knowledge. However, Shintani et al. (2013) did suggest that CBI was beneficial for the acquisition of new target features, especially for beginner-level learners, while PBI was more advantageous for more advanced learners in developing partially acquired structures. They also proposed that "grammar instruction will be most effective if it involves a combination of comprehension-based and production-based activities" (Shintani et al., 2013, p. 323); and called for further empirical research to be done in relation to that proposal.

Given that input-based instruction and output-based instruction have distinct advantages in L2 acquisition, it is important to determine how to combine the benefits of both types as complementary, rather than contradicting instructional methods for learners' optimal learning. The present study seeks to determine whether Shintani's (2011) claim that input-based task instruction is equal to output-based instruction even on production tasks

or whether Tanaka's (1999, 2001) and Shintani et al.'s (2013) suggestion of combining two types of instruction has greater advantages, especially for low-level university EFL learners.

3.7 Summary

This chapter has outlined the key constructs in TBLT. Various definitions and types of tasks have been discussed. In addition, TBLT was distinguished from TSLT based on: the way tasks are used, the focus on form, and, the theories of learning. Last but not least, input-based instruction versus output-based instruction with their unique benefits were compared. On this basis, it is suggested that a combination of both instructional types may be most beneficial for low-proficiency EFL learners' language development.

Chapter 4. Second Language Speech Production

This chapter discusses theories of speech production and L2 speech fluency. First, the modular model of L1 speech production is reviewed as a basis for understanding the adapted version of this model for L2 speech production. Next, details of each stage of L2 speech production are presented in comparison with L1 speech production. Finally, the main differences between L1 and L2 speech production are summarized at the end of the chapter.

4.1 L1 Speech Production

There is no doubt that when learning a second language, one of the most important goals is being able to speak the language fluently. Understanding how speech is produced is the key to helping learners improve this dimension of their language ability. Such knowledge provides a better understanding of how to develop and measure learners' access to language resources during L2 task performance.

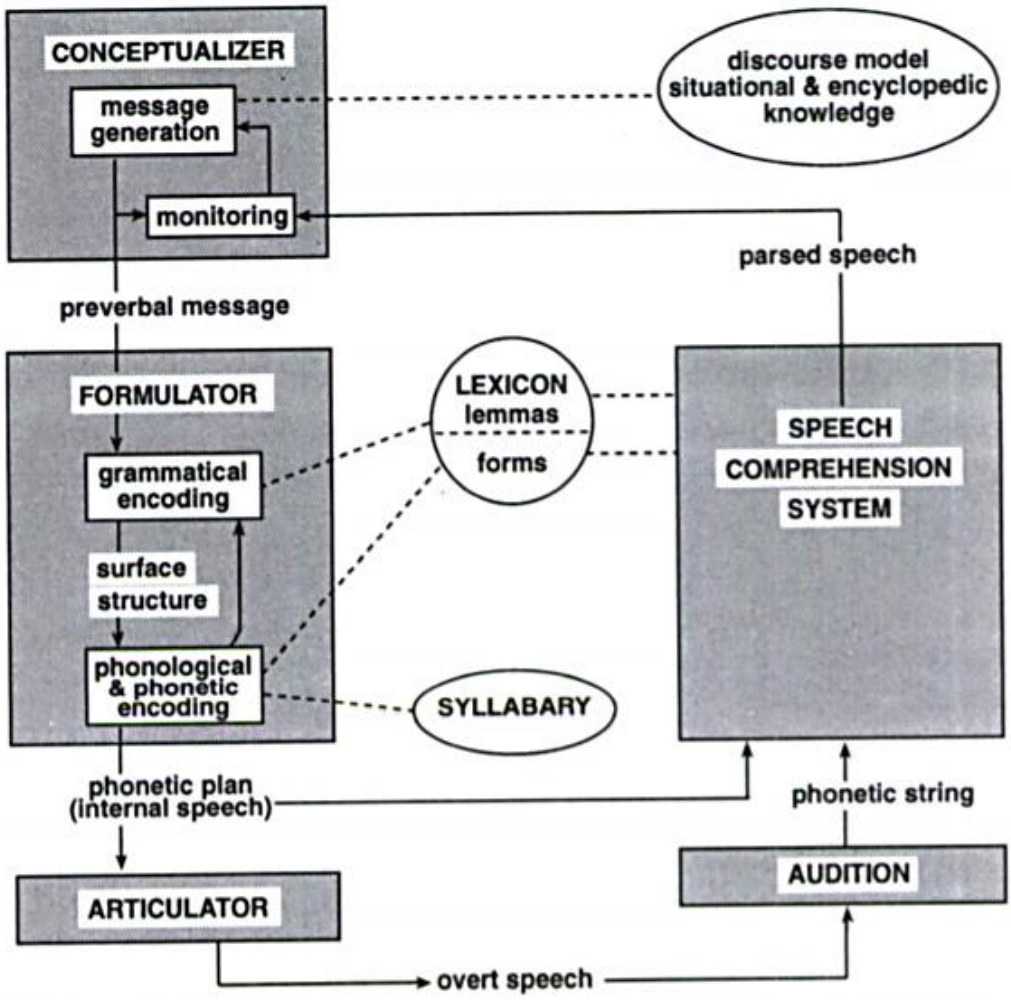
A highly influential and well-established model of speech production is that of Levelt (1989, 1995, 1999), which is labelled "the blue print for the native speaker" (Levelt, 1989, p. 9). Levelt's model posits that speech production is modular and includes four independent and sequential stages, namely *conceptualization*, *formulation, articulation,* and *monitoring* with various sub-processes involved in each.

Figure 4.1 (see below) presents the various processing components involved in the production of speech with the boxes representing processing modules, while the circles and ellipses represent knowledge stores. The primary modules within this model are the Conceptualizer, the Formulator, the Articulator, and the Speech-Comprehension System. According to (Levelt, 1989, 1995, 1999), the *Conceptualizer* has two functions, namely generating preverbal messages, and monitoring the whole of speech production. Next, the *Formulator* is responsible for formulating the language representation of the message (i.e., encoding it grammatically, and phonologically) to produce a phonetic plan (internal speech). The *Articulator* is then in charge of executing the phonetic plan and converting it into overt speech. Finally, the *Speech-Comprehension System* allows the monitoring of speech production to take place by making both internal and overt speech available to the conceptual system. Each of these components receives some kind of input and produces a

certain kind of output, which then serves as input for the next component. Furthermore, the information flows unidirectionally between the components (e.g., forward from the Conceptualizer to the Formulator, but not backward from the Formulator to the Conceptualizer). In brief, producing speech involves the speaker's conceptualizing the message, encoding the message, and finally articulating it. Monitoring can occur during each of these stages. These production stages are accomplished by the speaker accessing various knowledge sources. According to Levelt's (1995) model, there exist three knowledge stores: the store of World Knowledge, the Lexicon, and the Syllabary. The first store contains the speaker's external and internal knowledge of the world (also called *encyclopedic knowledge*), which is accessed for conceptualization. The second store, the Lexicon, is a repository of lexical entries, each of which is composed of two parts, the lemma and the lexeme. In the lemma, a lexical entry's meaning and syntax are represented as declarative knowledge, whereas the lexeme consists of procedural knowledge of a lexical entry's morphology and phonology. The entry's meaning and syntax are essential for grammatical encoding, whereas its morphology and phonology are used for phonological encoding. Finally, the Syllabary contains gestural scores that are used to produce the syllables of the actual speech.

Figure 4.1

A Diagram of the Processing Components Involved in the Generation of Speech



Note. From "The Ability to Speak: From Intentions to Spoken Words," by W. J. M. Levelt, 1995, *European Review*, *3*(1), p. 14. Copyright 1995 by John Wiley & Sons, Ltd.

4.1.1 L1 Conceptualization

In the *conceptualization* stage, a preverbal message is generated when speakers select information from their world knowledge and use this to convey in their message. The generation of preverbal messages from intention includes two sub-stages: macro-planning and micro-planning (see Figure 4.1).

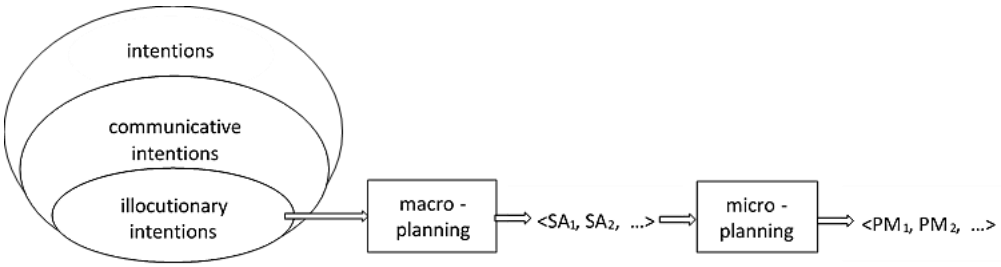
Macro-planning involves selecting information instrumental to realizing the communicative goal, and dealing with the linearization problem (i.e., ordering information for expression). This process starts with choosing a communicative intention or goal, and

elaborating the intention into a sequence of illocutionary intentions (sub-goals), and for each of these, the speaker has to select and order information in the way it will be expressed. The output of the macro-planning phase is a sequence of *speech act intentions* or *speech acts* (i.e., an assertion, command, or question). For example, if the intention is to give a person the direction to a museum, the speaker needs to break it down into sub-goals such as "first direct her to the city center, then inform her about the precise location of the museum" (Levelt, 1989, p. 109). Then to realize these sub-goals, the speaker has to plan and order the contents of speech acts (e.g., "to go to the city center first take the freeway, then turn right at the second traffic light") (Levelt, 1989, p. 109).

Following the macro-planning stage of conceptualization, the *micro-planning* stage then involves further shaping the speech acts into the format of a preverbal message. The content of each speech act is assigned a specific informational perspective and structure (e.g., mood, tense-aspect, and argument structure) that will help the addressee recognize the communicative intention. The output of micro-planning phase is a preverbal message for each corresponding speech act (see Figure 4.2). The preverbal message is also the output of the conceptualization stage.

Figure 4.2

From Intention to Preverbal Message



SA = Speech act, PM = Preverbal message

Note. From *Speaking: From Intention to Articulation* (p. 110), by W. J. M. Levelt, 1989, Cambridge, MA: MIT Press. Copyright 1989 Massachusetts Institute of Technology.

In brief, the conceptualization stage involves macro-planning (breaking a communicative intention down to one or more individual speech acts and ordering them for expression) and micro-planning (shaping each speech act into a preverbal message). Levelt (1989) describes macro-planning and micro-planning as two incremental stages in the

process of message encoding. In other words, it is not necessary for all macroplanning to be completed before microplanning can begin, and the two processes can alternate with one another or occur simultaneously during the conceptualization phase of pre-verbal planning. It should also be noted that the preverbal message contains conceptual information (e.g., semantics, style, register) that is not yet linguistic, and will constitute the input for the next processing component in Levelt's model, that is, for the Formulator, to work on.

4.1.2 L1 Formulation

The process of formulating can start as soon as the preverbal message has been prepared for expression. In the *formulation* stage, the preverbal message (a conceptual structure or 'preverbal plan') is converted into a phonetic plan (a linguistic structure) through two processes: grammatical and phonological encoding.

The first phase of the formulation stage, *grammatical encoding*, involves accessing a lexical item's lemma information (i.e., a lexical entry's meaning and syntax) stored in the lexicon (see Figure 4.1), and relevant syntactic building procedures to produce a surface structure or "an ordered string of lemmas grouped in phrases and subphrases of various kinds" (Levelt, 1989, p. 11). The first step of encoding a message linguistically is to "retrieve appropriate words for its lexical concepts" (Levelt, 1995, p. 17). A lemma will be activated when its meaning matches fragments of the preverbal message. Lemma information includes the meaning or concepts being conveyed (e.g., "*give* involves some actor X causing some possession Y to go from actor X to recipient Z") (Levelt, 1989, p. 11). It also contains syntactic features such as the lemma's syntactic category together with its grammatical functions and information needed for syntactic encoding (e.g., *give* is categorized as a verb that can take a subject, a direct object, and an indirect object) (Levelt, 1989, p. 11). When a lemma is activated based on its meaning, the syntax is made available, which in turn triggers relevant syntactic building procedures. For example, when the lemma *'give'* is activated, it means that there is a subject, a direct object, and an indirect object; and verb-phrase-building procedures will be activated to build the verb phrase '*gave Mary the book'* (Levelt, 1989, p. 11). When all the lemmas have been accessed and all the relevant syntactic building procedures have been activated, a surface structure is formed (e.g., *John gave Mary the book)*. In this way, a surface structure is the output of grammatical encoding, and the input for the second phase, *phonological encoding*.

Phonological encoding results in an articulatory phonetic plan for the whole utterance, which involves syntactic activation of lexical forms that are used to express them. In addition to the lemma information (i.e., meaning and syntax), a lexical item also contains lexeme information about its morphology and phonology. For example, "*dangerous* consists of a root (*danger*) and a suffix (*ous*),… it contains three syllables of which the first one has the accent, and… its first segment is /d/" (Levelt, 1989, p. 12). The output of phonological encoding is a phonetic plan, which is also called *internal speech* because it is not yet overt speech but "an internal representation of how the planned utterance should be articulated—a program for articulation" (Levelt, 1989, p. 12). For example, when encoding phonologically the utterance '*John gave Mary the book',* the speaker may be aware that the syllable /buk/ will be stressed. The end product of the Formulator module is the phonetic plan which functions as input for the Articulator, the next module in Levelt's model of L1 speech production.

4.1.3 L1 Articulation

In the third stage of speech production, *articulation*, the phonetic plan is realized as the sounds and syllables of speech are produced using the requisite motor skills. The outcome of the articulation stage is then the overt speech that the learner produces during an L2 task performance.

4.1.4 L1 Monitoring

Finally, *monitoring*, which is done by the Conceptualizer, involves checking the accuracy and appropriateness of the output of any of the three modules in the model. While the other three stages (conceptualization, formulation, and articulation) operate independently in a sequential input-output manner (i.e., the output of one stage functions as input for the next stage), monitoring can occur at any point within these stages. This takes place by means of the Speech-Comprehension System, which allows access to both internal and overt speech as well as lemma and form information in the lexicon to parse the speech and evaluate it. When the monitoring system inside the Conceptualizer detects any meaning or form deviations between the original intention and the parsed speech, it may interrupt the speech stream, reformulate the preverbal message and send a repair message to the Formulator. This is one source of a speaker's false starts, hesitations, and self-repairs.

Self-monitoring and self-repairs can occur even before the utterance is fully articulated if the speaker detects trouble with internal speech.

The whole process of L1 speech production in Levelt's (1995) model (see Figure 4.1) can thus be summarized as going through three main stages: generating a preverbal message (conceptualization), creating a phonetic plan through grammatical and phonological encoding (formulation), and producing overt speech (articulation) with self-monitoring taking place at any of these stages. These stages of speech production involve the work of three modules: the Conceptualizer, the Formulator, and the Articulator. These modules act as autonomous specialist processing components that do not require sharing of processing functions to carry out their respective roles. That is, each module's operation is independent of the other modules with its own characteristic input, and there is no direct exchange of information between them. Feedback between modules, however, is provided through monitoring.

Of importance to the present study are the notions of *serial processing* and *parallel processing*. Serial processing means that "each fragment of information will have to be processed in stages, going from the conceiving of messages to articulation", whereas parallel processing means "all processing components can work in parallel, albeit on different fragments" (Levelt, 1989, p. 28). As such, different parts of an utterance can be at different processing stages at the same time. In other words, different stages of speech production can take place simultaneously provided that the processing in one of the modules is sufficiently automatized for the other to proceed uninhibited. For example, when the Conceptualizer has passed its output to the next component (the Formulator), it can start with another piece of input at the same times as the previous output is being simultaneously encoded by the Formulator. However, this can only occur if the activities of the Formulator are sufficiently automated to allow the speaker's attention to move forward with conceptualization, rather than being diverted into lexical retrieval and grammatical encoding. For L1 speakers, the activities of the Formulator are largely automatized. While message conceptualization and monitoring require the L1 speaker's attention and memory resources, grammatical and phonological encoding and articulation are usually automatic processes that require little attention and can take place in parallel (Levelt, 1989; Poulisse,

1997). These features of parallel processing and automaticity together allow for speedy real-time L1 language production.

4.2 L2 Speech Production

De Bot (1992), Kormos (2006), and Segalowitz (2010) have adapted Levelt's (1989, 1995, 1999) modular model of L1 speech production to the conditions governing L2 speech production.

Firstly, De Bot (1992) adapted Levelt's model to account for bilingual speech production by introducing additional requirements beyond those of a monolingual processing model. With respect to the Conceptualizer, De Bot (1992, p. 21) argued that the first of the two production phases of conceptualization (macro-planning) is not language-specific, whereas the second phase (microplanning) is language-specific. In other words, macroplanning may involve activating relevant concepts which are normally identical or shared between languages, whereas microplanning is likely to take place in the intended language only. For De Bot, information about the language in which an utterance is produced is, therefore, specified in the preverbal message which is the outcome of the Conceptualizer and the input to the Formulator.

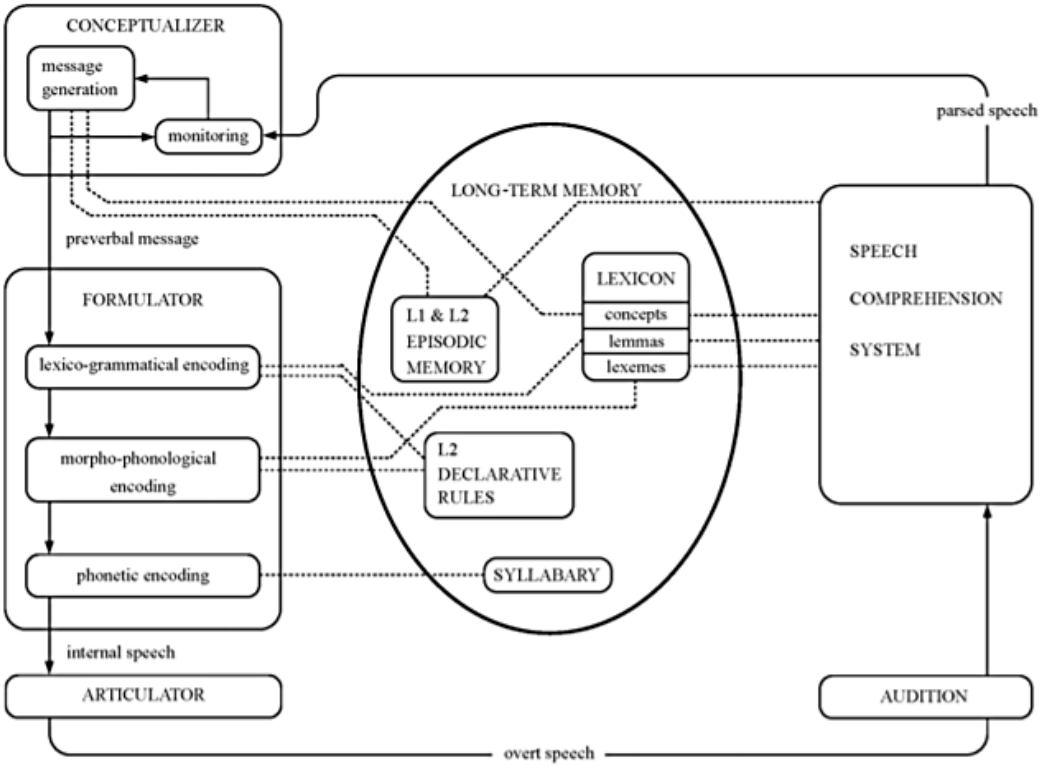
In terms of the Formulator module, De Bot (1992) argued that the way L2 formulator operates in L2 production is almost the same as the L1 formulator does in L1 production. In both cases, the preverbal plan, which is the output of the Conceptualizer, is converted into a phonetic plan, which is the input of the Articulator. However, De Bot also indicated that each language possesses its own distinctive micro-planning and formulator, but these formulators draw on a single lexical store (the lexicon) where both L1 and L2 lexical items are stored together. De Bot thus proposed the idea of parallel phonetic plans that form the output of separate formulators in the L1 and the L2. The respective formulators send their own phonetic plan to the Articulator. For DeBot, the Articulator is not language-specific and has "an extensive set of sounds and pitch patterns from both languages" (De Bot, 1992, p. 17) that are used to produce overt speech.

More recently, Kormos (2006) provides a comprehensive model of L2 speech production (see Figure 4.3) in relationship to Levelt's (1989, 1999) model of L1 speech production. Like De Bot (1992), Kormos (2006) acknowledges that production mechanisms

in both the L1 and the L2 are similar, consisting of separate specialist processing modules (i.e., the Conceptualizer, the Formulator, and the Articulator). Furthermore, she acknowledges that L2 speech production in two modules can take place simultaneously provided that production in one of the modules is sufficiently automatic to allow the production in the other to proceed unimpeded. However, Kormos (2006) proposes some modifications in knowledge stores to account for L2 speech production. In particular, L2 speech production draws on one large memory store, namely the Long-Term Memory, instead of three knowledge stores (the store of World Knowledge, the Lexicon, and the Syllabary) as presented in Levelt's (1995, 1999) L1 speech production model (see Figure 4.1). This long-term memory store is composed of several sub-stores: an episodic memory, a semantic memory (the lexicon), a syllabary, and a declarative knowledge memory (Kormos, 2006, p. 167) (see Figure 4.3). The episodic memory contains temporally organized experiences in one's life. In contrast, the semantic memory consists of concepts (both linguistic and non-linguistic), lemmas (syntactic information), and lexemes (morpho-phonological information) in hierarchical order (Kormos, 2006, p. 167). This means that "in speech production activation flows from the conceptual to the lemma and finally to the lexeme level, whereas in speech comprehension activation flows in the opposite direction" (Kormos, 2006, p. 168). Next, the syllabary contains the automatized gestural scores which are used to produce syllables for internal speech. It is noted that all these three knowledge stores (the episodic memory, the semantic memory, and the syllabary) are shared between the L1 and the L2. In other words, there is a single episodic memory, a single semantic memory, and a shared syllabary for both L1 and L2. Finally, Kormos (2006) argues for the existence of a fourth and important knowledge store in L2 production, a declarative knowledge memory, which contains information about syntactic and phonological rules in L2. This memory of declarative knowledge is L2-specific because, as Kormos points out, for L2 speakers, many syntactic and phonological rules in L2 are not yet automatized and are stored as declarative knowledge, while in L1 production these rules are almost automatic. Kormos also notes that, except for the addition of the new declarative knowledge store and the reorganization of the four knowledge sub-stores into one large memory store (the Long-Term Memory), there is no significant difference between the bilingual speech production model and the one constructed for monolingual speakers.

Figure 4.3

The Model of Bilingual Speech Production

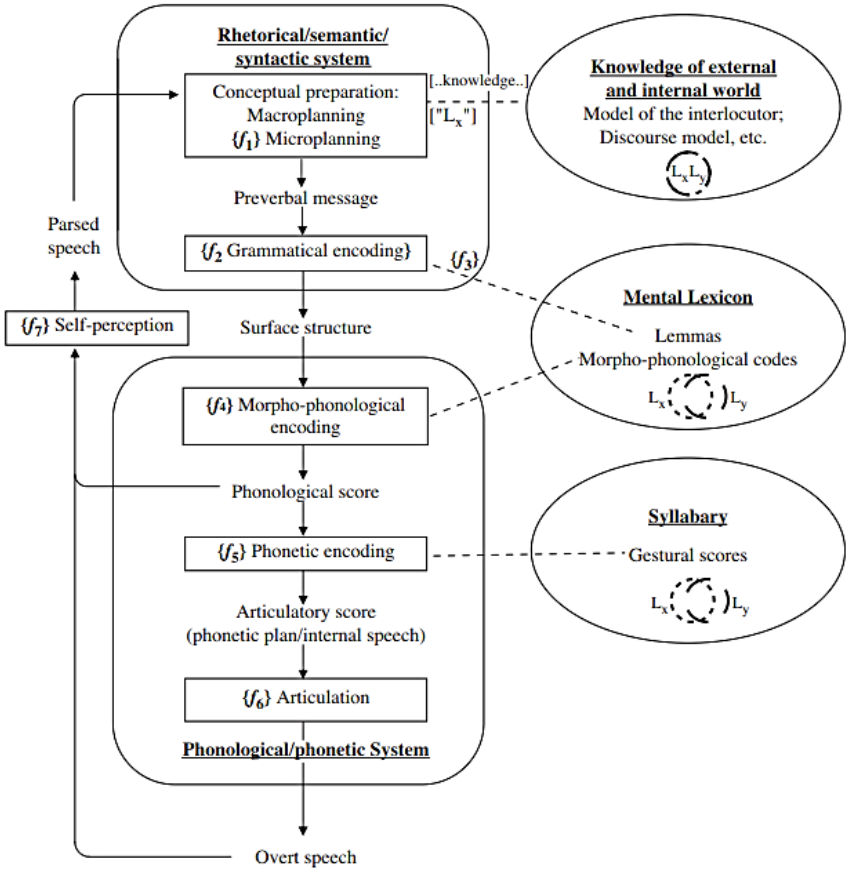


Note. From *Speech Production and Second Language Acquisition* (p.168), by J. Kormos, 2006, Mahwah, NJ: Lawrence Erlbaum Associates. Copyright 2006 by Lawrence Erlbaum Associates, Inc.

Following De Bot (1992) and Kormos (2006), and in an attempt to provide a systematic understanding of L2 cognitive fluency, Segalowitz (2010) proposes an updated and integrated model of the L2 speaker by adapting Levelt's (1999) L1 speech production model and incorporating De Bot's (1992) amendments regarding L2 speech production. Specifically, in this model Segalowitz identifies seven potentially critical points where underlying processing issues could lead to dysfluencies in L2 speech. These points are called 'fluency vulnerability points' and are marked with *f* symbols (see Figure 4.4). The seven fluency vulnerability points include: (*f*1) microplanning, (*f*2) grammatical encoding, (*f*3) lemma retrieval, (*f*4) morpho-phonological encoding, (*f*5) phonetic encoding, (*f*6) articulation, and (*f*7) self-perception (for details, see Segalowitz, 2010, p. 9). Encountering difficulties at any of these points may lead L2 speakers to interrupt their speech fluidity.

Figure 4.4

Model of the L2 Speaker



Note. From *Cognitive Bases of Second Language Fluency* (p. 9), by N. Segalowitz, 2010, New York: Routledge. Copyright 2010 by Taylor & Francis.

In brief, researchers agree that L2 production is similar to L1 production in that it also has four important processing components: (a) conceptualization, (b) formulation, (c) articulation, and (d) monitoring. However, it is also agreed that L2 production is distinct from L1 production in certain aspects due to the influence of the L1 on the L2, as well as the limited attentional capacity that speakers bring to the task of L2 production. The three most important differences between L1 and L2 speech production outlined by De Bot (1992) and Kormos (2006) are: (1) the influence of the L1 on L2 processing, (2) the incomplete nature of L2 knowledge, and (3) learners' access to L2 resources and corresponding reduced rate of speech. Differences between L1 and L2 speech production are also due to the vulnerability points in L2 speech processing (Segalowitz, 2010) which L1 speakers do not encounter because they can resort to automatic and parallel processing. These differences explain why

L2 speakers need to resort to the use of serial processing, L1 transfer and the use of communication strategies during real-time L2 speech production. Details of each stage of L2 speech production will be discussed in Sections 4.2.1-4.2.4 below.

4.2.1 L2 Conceptualization

Conceptualization in an L2, like in L1, involves the planning of what a speaker wants to say to realize a communicative intention. In this stage, the speaker makes decisions about the content of the preverbal message and its organization by selecting information from his or her world knowledge and organizing it into an initial pre-verbal structure before deciding on the language that will be used to express it. In line with Levelt's model (1989, 1995, 1999), conceptualization in an L2 can be broken down into two sub-stages, namely macro-planning and micro-planning (De Bot, 1992; Segalowitz, 2010). Macro-planning involves deciding what to say based on the speaker's knowledge of the world, ordering the information, making decisions about register choice (i.e., formal or informal speech), and specifying which language to use for (parts of) an utterance. Micro-planning finalizes the pre-verbal message by adding a specific informational perspective and structure (e.g., mood, tense-aspect, and argument structure) to each of the speech acts. The question that Kormos (2006) raises about this stage is "whether speakers formulate parallel speech plans—a plan for L1 and another one for L2—or a single speech plan in which each concept is labelled with a language tag" (p. xx). The idea of parallel preverbal plans in L1 and L2 was proposed by De Bot (1992) but was rejected a year later by De Bot and Schreuder (1993). Since then, most researchers (e.g., Kormos, 2006; Poulisse, 1997; Poulisse & Bongaerts, 1994) argue for a single preverbal plan that specifies both conceptual information and the language to be used to express the message (language tag). The choice of language for communication is indicated by adding language cues to the activated concepts. As such, each concept may have a language cue, and a preverbal message that is composed of a string of concepts may have a variety of language cues (Kormos, 2006). Kormos (2006) illustrates this by giving an example of encoding the sentence "*The policeman fined the motorist"*, pointing out that "a German-English bilingual speaker might add a language cue + English to the concepts of POLICEMAN and MOTORIST, whereas the concept of FINE might receive a tag + German" (Kormos, 2006, p. 170).

In brief, with the exception of the language specification (i.e., indication of the language in which [parts of] an utterance will be produced) in the preverbal message, conceptualization in the L2 remains the same as in L1 speech production. The conceptual information together with the language tag specified in the preverbal message will drive relevant grammatical and phonological encoding in the next stage, Formulation.

4.2.2 L2 Formulation

In this stage, the preverbal message (i.e., the outcome of the Conceptualizer) activates items in the mental lexicon corresponding to different chunks of the intended message, and this preverbal plan is then formulated into a phonetic plan (internal speech) through lexico-grammatical, morpho-phonological and phonetic encoding (see Figure 4.3).

In lexico-grammatical encoding, the conceptual specifications in the preverbal plan activate corresponding lemmas in the lexicon (i.e., lexical encoding), then this lexical-syntactic information is used to build up the surface structure (i.e., syntactic encoding). With respect to lexical encoding, the questions that Kormos (2006) raises are (1) "whether the conceptual specifications contained by the preverbal plan activate only L2 items in the lexicon, or whether L1 and L2 words both receive activation", and if so, (2) "whether the fact that both L1 and L2 words are activated also means that these words are both candidates for lexical encoding" (Kormos, 2006, p. xxi). At this point, it is necessary to distinguish *activation* from *selection* and note that activation does not entail selection. Kormos (2006) cites early studies (e.g., Costa, 2005) suggesting that only words of the selected language are activated. However, she provides evidence from later research (e.g., Colomé, 2001; Hermans et al., 1998; Poulisse, 1997, 1999; Poulisse & Bongaerts, 1994) to support the position that (1) not only L2 but L1 lemmas are also activated to some extent, and (2) both L1 and L2 activated lemmas compete for selection, but the lemmas whose features best match the conceptual specifications and the language cue will be selected. Kormos (2006) then concludes that "in this bilingual speech production model, lexical encoding means the matching of the conceptual specifications and the language cue with the appropriate lexical entry in the mental lexicon" (p. 170). The assumption is that only the lemmas in the selected language are selected. However, the speakers may replace the L2 specification with an L1 item possibly due to lack of an appropriate L2 lexical item or erroneous activation (La Heij, 2005; Poulisse, 1999; Poulisse & Bongaerts, 1994). When the concepts erroneously activate

both L1 and L2 lemmas, which may entail further activation of phonological forms of both L1 and L2 words, L1 words are selected for further processing because they are generally more frequently used than L2 words.

Regarding syntactic encoding, Kormos (2006, p. 171), like De Bot (1992), argues that L2 production is not significantly different from L1 production in that it is "lexically driven" and comprises several sequential sub-phases. As in L1 production, the first phase of L2 syntactic encoding involves the activation of syntactic features associated with a lexical item (e.g., the syntactic category, grammatical information) (see section 4.1.2 for examples). The second phase involves employing syntactic encoding procedures to build up phrases and clauses from activated words, then arranging these phrases into an appropriate sequence for an utterance. While the first phase draws on the speakers' declarative knowledge, the second phase involves applying their procedural knowledge. At this stage, like L1 speakers, proficient L2 speakers normally have automatic access to procedural knowledge of syntactic and morphological rules. However, for lower proficiency level learners, automaticity is not always possible. When a form is not fully procedualized, learners may resort to the use of communication and transfer strategies which require additional attention to the formulation module of speech production and prevent the learners from parallel processing in other modules. As a result of the breakdown that occurs in parallel processing while L2 speakers search for language to express their ideas, they are forced to serially process their speech one stage at a time from conceptualization to formulation. This of course results in a breakdown in fluency and a decrease in speech rate so that L2 learners cannot meet the demands of real-time speech production on tasks (Skehan, 2009) as learners first think of what they want to say and then take additional time to think about how to say it.

In the next phase, the morpho-phonological form of the lexical items is activated and syllabified in its syntactic context, and the parameters for loudness, pitch, and duration are set. The most important issue that needs addressing at this stage is "whether the phonological form of non-selected but nonetheless activated words also receives activation, that is, whether activation can cascade from the lemma to lexeme (phonological word form) level" (Kormos, 2006, p. xxiii). With regard to this question, Kormos (2006) argues that the phonological form of words in the non-selected language is also activated, and both L1 and L2 lexemes compete for selection. She also argues that cascading of activation occurs

between the lemma and the lexeme level during L2 production. In brief, Kormos proposes that the processing mechanisms of L2 phonological encoding work in a similar way as in L1 speech production, but that key adjustments need to be made to account for the fact that the L2 speaker already speaks an L1 and that aspects of the L1 will impact the processing of their L2 through cascading activation from lemma to lexeme.

Following phonological encoding, the phonemes of words are activated in a serial manner. According to Kormos (2006), representations of these L1 and L2 phonemes are stored together in a single network in the syllabary. She maintains that memory representations for identical phonemes in L1 and L2 are shared, whereas "phonemes that are different in L1 and L2 are stored as separate representations" (p. 173). Kormos suggests that L2 learners may often use L1 phonemes in place of similar L2 phonemes, or they may apply phonological processes associate with the pronunciation of the L1 in encoding to L2 phonology. This results in the accent in which they speak the L2.

In the formulation stage of speech production, it becomes obvious that the difference between monolingual and bilingual speech processing involves the influence of L1 on the L2. This difference seems to be unavoidable as most of the knowledge stores (e.g., conceptual memory, the lexicon, and the syllabary) are shared in L1 and L2, and thus both L1 and L2 linguistic items compete for selection (Kormos, 2006; La Heij, 2005; Poulisse, 1999; Poulisse & Bongaerts, 1994). This competition can result in unintentional switching between L2 to L1 forms. However, drawing on L1 and L2 forms can also be intentional when the speaker decides to use an L1 item for an L2 specification in the preverbal message due to a lack of an appropriate L2 lexical item or because that L1 item better matches the conceptual specifications. Another result of the competition between L1 and L2 forms in L2 speech production is the transfer of L1 production rules when a production process that is appropriate to the L1, but not to the L2, is employed. However, applying the procedural knowledge of L1 rules when encoding an L2 item can lead to both positive and negative semantic, syntactic or phonological transfer.

4.2.3 L2 Articulation

Articulation is the third processing module in both L1 and L2 speech production. At the articulation stage, articulatory gestures, or the physical motor skills involved in

producing syllables, are retrieved and activated. This results in the overt speech of l2 learners. In L1 speech production, syllables are assumed to be the basic units of articulatory execution, and the phonetic plan is composed of numerous syllable programs. For bilingual speakers, however, whether these syllable programs are automatized or not is likely to depend on the speaker's L1 as well as their level of proficiency in their L2. In line with De Bot's (1992) view, Kormos (2006) hypothesizes that lower proficiency L2 speakers are largely dependent on L1 syllable programs, whereas higher proficiency L2 speakers can engage separate sets of motor skill programs for their L1 and their L2.

4.2.4 L2 Monitoring

The fourth and final process in L2 speech production is monitoring, which deals with learners' attention to their own output to check the accuracy and appropriateness of the output from each of the three primary speech production modules (conceptualization, formulation, and articulation). According to Kormos (2006), monitoring in L2 speech production involves the same mechanisms as speech comprehension (see Figure 4.3) and proceeds in a similar way to that in L1 (see Figure 4.1). Specifically, there are three monitor loops that are in charge of checking the outcome of production processes. The first loop evaluates the output conceptually with the original intentions of the preverbal plan. The second loop attends to the accurate formulation of the message in linguistic form by engaging internal speech and focusing attention on the phonetic plan that was the outcome of the formulator before it was articulated. The last loop involves attention to overt speech produced to communicate the message by means of the acoustic-phonetic processor to check for pronunciation or other articulation problems. Upon detecting any trouble in the output with regard to any of these loops, "the monitor issues an alarm signal, which, in turn, triggers the production mechanism for a second time starting from the phase of conceptualization" (Kormos, 2006, p. 173). The speaker then stops the speech stream, and either repairs the utterance or reformulates it entirely depending on the nature of the problems, the speakers' L2 resources, and the time constraints imposed by the context. Dörnyei and Kormos (1998) propose that L2 learners can either modify messages in these various ways or abandon the intended message completely, and resort to resource deficit strategies such as paraphrase or circumlocution. In all of these cases, breakdowns in speech processing may occur while speakers reconceptualize, reformulate, and rearticulate their

messages, resulting in pauses and repetitions to buy themselves time to meet these additional processing demands.

Despite similar processing mechanisms, monitoring in the L1 differs from monitoring in the L2 in several ways. The causes of these differences are due to attentional demands (Kormos, 2006). While formulation and articulation in L1 production are largely automatic (attention-free) and can operate in parallel with conceptualization and monitoring, L2 speech processing requires attention and serial processing at almost all levels (e.g., conceptual, lexical, syntactic, phonological and articulatory), although articulation is likely to be largely automatic for all except the lowest proficiency L2 learners. With limited attentional resources, L2 speakers, therefore, have less attention to spare for monitoring than L1 speakers, and they must decide what to prioritize in order to reduce the cognitive burden associated with monitoring in an L2 (Kormos, 2011). As all aspects of speech production require attention at some point for the learners' balanced, ongoing development, instruction must be designed to support learners by ensuring that attention is alternated between relevant aspects of their developing L2 speech processing capacity (Skehan, 1996, 2009).

4.3 Summary

This chapter has reviewed theories of speech production. Models of L1 and L2 speech production have been compared with respect to processing mechanisms underlying each stage of conceptualization, formulation, articulation, and monitoring. While it is agreed that most aspects of L2 production can be explained by the model of L1 production proposed by Levelt (1989, 1995, 1999) (see Section 4.1), there are several differences between L1 and L2 speech production in relation to the influence of the L1 on the L2, the speaker's incomplete linguistic knowledge of the L2, the degree of automaticity and the limited attentional capacity of L2 speakers. These issues explain why L2 speech contains more hesitations and is generally not as smooth and rapid as L1 speech. These characteristics of speech are referred to as speech fluency, which will be discussed in the next chapter.

Chapter 5. Second Language Speech Fluency

This chapter discusses the theories of L2 speech fluency. Definitions of fluency are discussed in terms of cognitive fluency, utterance fluency, and perceived fluency (Segalowitz, 2010). Next, the relationship between utterance fluency and cognitive fluency, and the development of L2 fluency are examined. Finally, measures of utterance fluency in terms of speed fluency, breakdown fluency, and repair fluency are outlined in relation to the different stages of L2 speech production, followed by the rationale of the measures to be used in the present study.

5.1 Definitions of L2 Fluency

Speaking a language fluently is one of the ultimate goals when learning an additional language as an adult, but it is also considered the most difficult skill for many learners to master, particularly in foreign language contexts like Vietnam where learners have little, if any, opportunity to use the language outside of the classroom. It is important to define what it means to be fluent in an L2 and how fluency is measured.

Researchers have different views about what fluency means. Fluency has been defined in broad and narrow senses.

In the broad sense, fluency is often used as a synonym for overall oral proficiency, or even for general linguistic proficiency (Chambers, 1997; Lennon, 1990). For example, when people say “She speaks English fluently” in a conversation, they typically mean that the person has a good command of English and can use the language efficiently. Meanwhile, fluency in the narrow sense is “the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention under the temporal constraints of online processing” (Lennon, 2000, p. 26). Similarly, Skehan (1996) also claims that fluency “concerns the learner’s capacity to produce language in real time without undue pausing or hesitation” (p. 22). In line with this point of view, Segalowitz (2013) suggests that “the qualities that make speech fluent include fast speech, and the relative absence of undue hesitations, pausing, repetition, and repairs” (p. 240). In general, fluency in this narrower sense refers to the capacity to communicate the intended message rapidly, effortlessly, and smoothly without disfluencies such as pauses, hesitations, repetitions, and repairs. In second language research, this narrower notion of fluency is recognized as a key construct of L2 oral

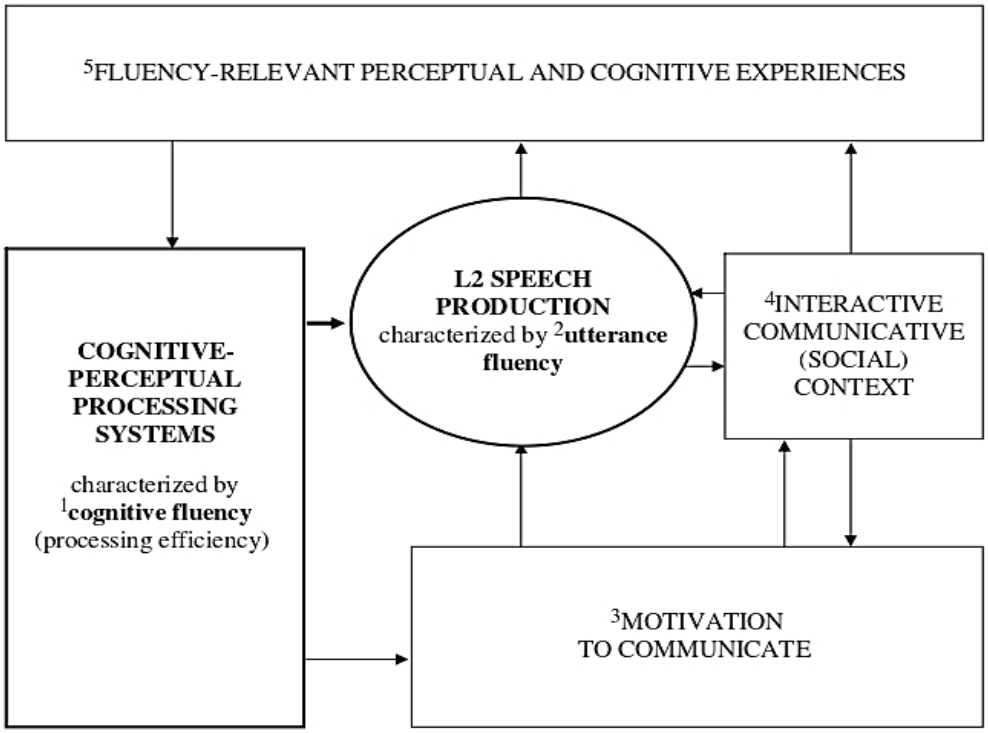
proficiency and a means of examining the L2 speech production process as discussed in Chapter 4. The term *fluency* will be used in the narrow sense throughout the remainder of this thesis.

According to Segalowitz (2010), in order to fully understand the nature of L2 fluency, a cognitive science approach is needed. He does this using a dynamic systems framework (Triadic Framework) for L2 fluency to show the relevant factors that jointly determine fluency (see Figure 5.1). In this framework, speech fluency is examined from a cognitive and social perspective. The four major sources of influence that contribute to L2 speech production (characterized by utterance fluency) include: (1) *the cognitive-perceptual processing systems* (characterized by cognitive fluency), (2) the *motivation to communicate*, (3) *the interactive communicative or social context*, and (4) *the perceptual and cognitive experiences*. These components are related and interact with one another to create a multidimensional dynamical system that is "continuously changing and adapting over time" (Segalowitz, 2010, p. 164). Specifically, the efficiency of the cognitive-perceptual processing systems can influence motivation to communicate. It seems that the more difficult a communication task is, the less willing the L2 speaker is to engage in it. The way that the speaker engages in the communicative act might then affect the nature of the social context and vice versa. The social context influences L2 speech production by setting the sociolinguistic demands associated with communication. In addition, the social context is related to the cognitive-perceptual systems of the speaker in that a social context can generate experiences relevant to L2 fluency development for the L2 speaker; these experiences in turn can shape the speaker's cognitive fluency. Finally, cognitive processing efficiency determines how fluent the utterances are, and it is the relationship between cognitive fluency (the highlighted rectangle) and utterance fluency (the highlighted oval) that characterizes L2 fluency. Overall, Segalowitz's (2010) view of fluency can be described as follows:

> Fluency in a second language is an observable characteristic of realtime speech behavior. This behavior reflects the execution of the neurological and muscular mechanisms that a speaker has developed over an extended period of time through socially contextualized, communicative activities. The operation of these mechanisms reflects the cognitive and emotional states of the speaker at the time of speaking. (pp. 6-7)

Figure 5.1

Framework for L2 Fluency



The act of producing L2 speech (central oval) is subject to at least four broadly defined influences that interact with one another, creating a system that is continuously changing and adapting over time, and that has the features of a dynamical system. L2 fluency is characterized by the relationship between cognitive and utterance fluency (the highlighted rectangle and oval). System components include: the cognitive-perceptual systems underlying speech production, motivation to communicate, the interactive communicative or social context created by the act of communicating, and the perceptual and cognitive experiences resulting from communicating and from the social context. These experiences affect utterance fluency by shaping cognitive fluency.

Notes

[1] Cognitive fluency features include processing speed, stability, and flexibility in the planning, assembly, and execution of utterances in terms of lexical access, and the use of linguistic resources (linguistic affordances) to express construals, handle sociolinguistic functions, and pursue psychosocial goals.

[2] Utterance fluency features include speech rate, hesitation, and pausing phenomena, etc.

[3] Motivation includes willingness to communicate, beliefs about communication, language and identity, and the concept of the L2 self. Motivation influences speech production and the selection of social contexts in which to speak.

[4] The social context influences speech production by setting the cognitive task demands associated with communication and is the source for learning about linguistic affordances.

[5] Experiences include frequency of exposure, opportunities for repetition practice, etc.

Note. From *Cognitive Bases of Second Language Fluency*. (p. 164), by N. Segalowitz, 2010, New York: Routledge. Copyright 2010 Taylor & Francis.

Segalowitz (2010) further points out that *L2 fluency* is a multidimensional construct, and that it is important to make a distinction among three aspects of L2 fluency: cognitive fluency, utterance fluency, and perceived fluency.

Cognitive fluency is defined as "the speaker's ability to efficiently mobilize and integrate the underlying cognitive processes responsible for producing utterances" (Segalowitz, 2010, p. 48). Producing an utterance involves fast and efficient coordination of many cognitive processes such as conceptual planning, lexical and grammar encoding, articulation and monitoring as discussed in Section 4.2. It is generally assumed that L2 fluency development depends on the extent to which these cognitive processes are automatic (Hilton, 2008; Kormos, 2006; Segalowitz, 2010). Fluent L2 speakers are those who are capable of efficiently converting their intended messages into comprehensible sounds within the time constraints of real-time oral communication (De Jong, 2016b). Lower-proficiency L2 speakers may have difficulty producing fluent speech, because the production of L2 speech for them requires controlled and serial processing of content and language. In contrast, proficient L2 speakers, like L1 speakers, are typically able to speak smoothly and rapidly due to the fact that their access to L2 resources during encoding is sufficiently automatic to allow them to process, in a parallel manner, both what they want to say and how they want to say it. Furthermore, Segalowitz (2010) proposes that three most important L2-specific elements of cognitive fluency are processing speed, processing stability, and processing flexibility in terms of lexical access and the use of linguistic resources. The importance of these features is described as follows:

> Processing speed is a necessary but not sufficient feature for defining cognitive fluency. Processing stability may be more important than processing speed as a feature of cognitive fluency. Flexibility of attention-based processing may be a more important feature of cognitive fluency than is processing speed, especially in grammatical processing. (Segalowitz, 2010, p. 106)

Utterance fluency refers to speech fluidity as a property of the actual speech. This fluidity is reflected in the observable features of speech such as speech rate, pauses, hesitations, and repairs. Utterance fluency can be objectively measured with respect to three aspects: speed fluency, break-down fluency, and repair fluency (De Jong et al., 2013; Lambert & Kormos, 2014; Skehan, 2003; Tavakoli & Skehan, 2005). Speed fluency is defined as the rate of oral delivery (e.g., speech rate, articulation rate). Break-down fluency refers to

the number, location, and duration of pauses and hesitations. Repair fluency equates with the number of corrections (e.g., replacements and reformulations) that the speakers make (De Jong et al., 2013). With regard to speech production, speed fluency is related to automaticity in accessing L2 knowledge and provides a composite measure for all stages of speech production discussed in Section 4.2 (conceptualization, formulation, articulation, and monitoring). In contrast, breakdown fluency provides a finer-grained picture of speech processing in that it reflects the automaticity of either conceptual content or its formulation depending on where in an utterance the breakdown occurs. The role of location of breakdowns in the measurement of specific stages of speech processing will be discussed in more detail in Section 5.4.2. Finally, repair fluency reflects the monitoring process in L2 speech production (De Jong, 2016b; Dörnyei & Kormos, 1998; Götz, 2013; Kormos, 1999, 2006; Lambert et al., 2017; Levelt, 1989; Segalowitz, 2010; Skehan, 2003, 2009; Tavakoli & Skehan, 2005) and attention to the reformulations of initially encoded and articulated concepts (Lambert et al., 2020). As such, utterance fluency is seen as the manifestation of underlying cognitive fluency.

The third aspect of fluency, *perceived fluency*, refers to the subjective impressions that L2 speech produces in listeners (Lennon, 1990; Segalowitz, 2010). Research has demonstrated that utterance fluency is often correlated with perceived fluency, particularly measures of speed fluency such as speech rate and phonation time ratio (Bosker et al., 2013; Derwing et al., 2004; Kormos & Dénes, 2004; Préfontaine, 2013; Rossiter, 2009). Thus, perceived fluency is a subjective correlate of utterance fluency (Segalowitz, 2016).

In sum, L2 fluency can be defined as a skilled performance, which reflects the efficiency of operation of the L2 speaker's underlying cognitive processes responsible for producing utterances (cognitive fluency). This performance can be objectively measured by oral features of utterances such as speech rate, pauses, hesitations, and self-repairs (utterance fluency), or subjectively rated by the listener's inferences about the ease and flow of the speech (perceived fluency). These three aspects of fluency (cognitive, utterance, and perceived fluency) are related to each other, but according to Segalowitz (2010), in a much narrower sense, "L2 fluency is characterized by the relationship between cognitive and utterance fluency" (p. 164). Furthermore, if we want to do something about developing L2 fluency, it is important to go beyond a description of the overt speech and investigate the

underlying mechanism of L2 utterance fluency and what makes L2 speakers fluent. Therefore, the present study is mainly concerned about the link between utterance fluency and cognitive fluency. As such, objective measures of utterance fluency in relation to cognitive fluency and speech production stages are the focus of the study.

5.2 Relating Utterance Fluency to Cognitive Fluency and Speech Production

Relating utterance fluency to cognitive fluency is of great importance in determining which aspects of utterance fluency are reliable indicators of cognitive fluency. As Segalowitz (2010) points out, cognitive fluency refers to efficient (i.e., fast and accurate) operation and the degree of automaticity of cognitive processes underlying speech production such as conceptual planning, lexical retrieval, grammatical encoding, and monitoring. When these processes are not fully automatized, as in the case of low-proficiency L2 learners, speech production is inefficient as attention is required at all stages of speech production. The processes of producing L2 speech will then occur in a serial manner, one after the other, which results in less fluent and more effortful speech with more disfluencies (Kormos, 2006; Segalowitz, 2010). According to Segalowitz's (2010) cognitive science approach to fluency, "the speaker's ability to mobilize and temporally integrate, in a nearly simultaneous way, the underlying processes of planning and assembling an utterance (the cognitive fluency) results in that utterance having the particular properties it does have (the utterance fluency)" (p. 169). In other words, L2 utterance fluency is dependent upon L2 cognitive fluency.

The relationship between L2 utterance fluency and cognitive fluency is normally investigated in two ways (De Jong et al., 2013). The first way is to investigate fluency gains within speakers over time (e.g., Riazantseva, 2001; Saito et al., 2018; Towell et al., 1996). If learners improve in a specific measure of utterance fluency, it can be assumed that this development is related to development in L2 cognitive fluency. The second way is to measure L2 utterance fluency and cognitive fluency separately and then relate measures of utterance fluency to measures of cognitive fluency. Utterance fluency is usually measured by calculating observable aspects of speech such as speech rate, pauses, and repairs. Cognitive fluency is often measured indirectly by means of stimulated recall (e.g., Kahng, 2014), or more directly by measuring linguistic knowledge (e.g., vocabulary, grammar, and pronunciation) and processing skills (e.g., speed of syntactic encoding, lexical access, and

articulation) which are supposed to underlie cognitive fluency (e.g., De Jong et al., 2013; Kahng, 2020; Segalowitz & Freed, 2004).

The studies that investigate fluency gains show that some measures of utterance fluency are quite good indicators of L2 cognitive fluency while others are not. For instance, Towell et al. (1996) investigated the development of fluency in 12 advanced students of French studying abroad over a period of three years. The participants individually watched a film and then re-told the story of the film in a recording booth. The same task was performed in French (L2) in the second year (Time 1) and the third year of the course after a year of residence abroad (Time 2), and then in English (L1) in the fourth year of the course (Time 3) for the purpose of comparing L1 and L2 performance. The temporal measures used in their study were speaking rate, articulation rate, phonation/time ratio, mean length of pause, and mean length of run. The results showed fluency gains in mean length of run, speaking rate, and articulation rate, but not in phonation/time ratio and mean length of silent pause. They concluded that fluency development was not the result of the increase in speed rate or the decrease in the amount of pausing, but due to the increase in mean length of run, “the most important of the temporal variables contributing to this development” (Towell et al., 1996, p. 84). Based on Anderson’s (1983) idea of 'proceduralization of knowledge' and Levelt’s (1989) model of speech production, the authors explained that increase in mean length of run was related to the change in the speed with which linguistic knowledge (e.g., knowledge of syntax and lexical phrases) can be accessed for online speech production. In other words, gains in fluency were attributable mainly to the increased degree of proceduralization of L2 linguistic knowledge, or L2 cognitive fluency as indicated by Segalowitz (2010). Based on further qualitative analysis of the scripts, the authors suggested that the formulator was where that proceduralization took place. More recently, Saito et al. (2018) conducted a cross-sectional study on the development of different utterance fluency aspects based on Kormos’s (2006) idea of the link between different utterance fluency measures (e.g., frequency of final-clause pauses, mid-clause pauses, and repairs) and cognitive processes underlying speech production (e.g., conceptualization, formulation, monitoring). The results suggested that improvement in utterance fluency in terms of final-clause pauses, mid-clause pauses, and articulation rate reflected development in conceptualization, formulation, and overall speech production respectively.

To date, there have been several studies investigating cognitive fluency with separate measures (i.e., measuring processes underlying speech production) and relating it to utterance fluency. Segalowitz and Freed (2004) examined gains in utterance fluency aspects and related the results to measures of L2 cognitive fluency. Utterance fluency measures included speech rate, mean length of run without silent pauses, mean length of run without filled pauses, and longest fluent run (no silent or filled pauses). Cognitive fluency was measured in terms of speed of lexical access (measured by reaction time), efficiency of lexical access (measured by the coefficient of variation of reaction time), and speed and efficiency of attention control. The results showed that the mean length of run without fillers was significantly correlated to lexical access speed and lexical access efficiency. They also found a significant but negative relation between gains in speech rate and efficiency of attention control. From these results, they claimed that gains in utterance fluency may reflect underlying L2-specific cognitive processing abilities. De Jong et al. (2013) also employed separate measures of utterance fluency, and L2 linguistic knowledge and skills supposedly underlying cognitive fluency to identify which L2 utterance fluency measures are indicators of L2 cognitive fluency. In their study, utterance fluency was measured in terms of breakdown fluency (number of silent pauses, mean duration of silent pauses, and number of filled pauses), speed fluency (mean duration of syllables, i.e., speaking time divided by the total number of syllables), and repair fluency (number of corrections and repetitions). Cognitive fluency was investigated in terms of L2 linguistic knowledge (e.g., vocabulary, grammar, and pronunciation) and processing skills (e.g., speed of sentence building, lexical retrieval, and articulation) with the authors' acknowledging that cognitive fluency includes non-linguistic factors such as working memory capacity. The results showed that most measures of utterance fluency were related to at least one of the measures of cognitive fluency (either linguistic knowledge or processing skills). While mean duration of syllables was most strongly related to linguistic knowledge and skills (50% of the explained variance), mean length of silent pauses was the least related (5% of the explained variance). It was then assumed that mean syllable duration was a strong indicator of L2 cognitive fluency whereas mean duration of silent pause was not. With evidence from previous studies, De Jong et al. (2013) came to the conclusion that the fluency gains over time were related to the development of L2 cognitive fluency, rather than personal speaking style or general cognitive abilities. Later, Kahng (2014) examined cognitive fluency

qualitatively by using stimulated recall and compared results with quantitative evidence from measures of utterance fluency. The results showed that speed fluency was strongly correlated with L2 oral proficiency and cognitive fluency. Furthermore, silent pause rate within clauses was found to have one of the strongest correlations with L2 speaking scores. This finding was corroborated by the results from stimulated recall showing that lower proficiency learners reported more issues regarding L2 declarative knowledge on grammar and vocabulary (formulation) than the higher proficiency learners. Kahng's (2014) findings were also consistent with the claim in previous studies that pauses within clauses were related to processing difficulties in speech production. More recently, Kahng (2020) investigated the relation of L2 utterance fluency with L2 cognitive fluency and L1 utterance fluency. This study replicated De Jong et al.'s (2013) study with similar measures of L2 utterance and cognitive fluency, but controlled for L1 base measures for cognitive fluency and investigated pauses in terms of mid-clause and end-clause pauses and their relationships with cognitive fluency separately. The results showed that most L2 utterance fluency measures were predicted by "the combination of L2-specific cognitive fluency measures and the equivalent L1 utterance fluency measure" (p. 476), but the number of mid-clause silent pauses and mean syllable duration are largely predicted by L2 cognitive measures. Specifically, the number of mid-clause silent pauses and the mean syllable duration were significantly correlated with speed of L2 syntactic encoding and lexical retrieval while the correlation with L1 fluency measures was weak. In contrast, the mean length of silent pauses and the number of clause-final filled pauses did not show significant correlations with any L2 cognitive measures. With regard to measures of L2 cognitive processing, speed of syntactic encoding and lexical retrieval exhibited association with L2 utterance fluency measures, whereas pronunciation duration did not. These findings are in line with De Jong and Bosker (2013), who suggest that L2 utterance fluency measures, especially mid-clause silent pauses and mean syllable duration, are more strongly correlated with the formulation (syntactic encoding and lexical retrieval) than with the articulation stage of speech production.

In brief, together the above empirical studies suggest that aspects of L2 utterance fluency are closely related to L2 cognitive fluency, especially at the formulation stage of speech production. Measures of L2 utterance fluency, especially speech rate, mean syllable

duration, mean length of run, and number of mid-clause silent pauses are strong indicators of L2-specific knowledge and cognitive skills which are necessary for lexical, syntactic, and phonological encoding at the formulation stage. Gains in these utterance fluency measures are attributable to the development of cognitive processes such as lexical retrieval and syntactic encoding. In other words, oral features of utterances reflect the operation of underlying cognitive processes and vice versa, cognitive fluency can be inferred from the analysis of utterance fluency measures (Segalowitz, 2010). Based on these findings, the present study investigates the relationship between L2 utterance fluency and cognitive fluency by measuring fluency gains over time, and then relates the progress in specific measures of utterance fluency to the development in corresponding cognitive processes underlying stages of L2 speech production.

5.3 L2 Fluency Development

While it is important to understand what L2 fluency entails and the relationship between aspects of fluency, key concerns for many L2 learners and teachers include what factors impact L2 fluency development, and how to improve L2 fluency.

5.3.1 Factors Affecting L2 Fluency Development

As fluency is a multidimensional and dynamic construct, it is influenced by many factors. The influence can be from internal sources, such as the speaker's L1 speaking style and proficiency level (De Jong, 2016b; Derwing et al., 2009; Riazantseva, 2001; Saito et al., 2018), automaticity of knowledge and skills, and vocabulary knowledge (De Jong & Perfetti, 2011; De Jong et al., 2013; Hilton, 2008; Towell et al., 1996) or from external sources such as task structure and demands (Skehan & Foster, 1999; Tavakoli & Skehan, 2005), planning opportunities, and task/content familiarity (Bui & Huang, 2016; Bui, 2014; Ellis, 2005a). Among these factors, L2 fluency development is mostly influenced by (1) L1 individual speaking style and preference (e.g., a speaker's speech rate, pausing behaviours) (De Jong et al., 2013; Derwing et al., 2009; Duran-Karaoz & Tavakoli, 2020; Kahng, 2014, 2020; Towell et al., 1996), and (2) the extent to which the cognitive processes underlying speech production are automatic (Hilton, 2008; Hunter, 2017; Kormos, 2006; Segalowitz, 2010). While a person's fluency might be a trait (i.e., a relatively permanent and specific to an individual) that is hard to change and L1 fluency behaviour can transfer to L2 performance (Derwing et

al., 2009), increased L2 proficiency can help learners overcome these transfer effects (Riazantseva, 2001). Moreover, the automaticity of linguistic knowledge and skills can also be improved through practice and proficiency development. As mentioned in Chapter 4 (Section 4.2) and earlier in this chapter (Section 5.1), proficient L2 learners, like L1 speakers, can speak fluently because the largely automatized and effortless access to linguistic resources allows them to conceptualize a message and encode it in linguistic forms at the same time (parallel processing). For lower proficiency level learners, however, the fact that these processes are not fully automatized, requiring more effort and attentional control causes them to serially process their speech one stage at a time (i.e., formulation occurring separately after conceptualization). The difficulties associated with linguistic encoding during real-time communication naturally lead to a slower and more disrupted L2 speech with a range of disfluencies like silent pauses, fillers (e.g., *uh* and *um*), repetitions, false starts, and self-repairs (De Jong et al., 2013; Derwing et al., 2009; Towell et al., 1996). Therefore, in order to improve L2 learners' speech fluency, it is important to facilitate the conversion of declarative knowledge into procedural knowledge as a key to enable parallel processing of different cognitive processes underlying L2 speech production because when knowledge becomes proceduralized, it is accessed automatically, effortlessly and efficiently (Anderson, 1980, 1983; McLaughlin, 1987b, 1990). It is suggested that practice has a role to play in this process in that practice can lead to changes in interlanguage by creating more chunks of information that are available for automatic processing.

5.3.2 Task Repetition and L2 Fluency Development

Among different forms of practice, task repetition is one task implementation variable that has been investigated in a large number of studies. According to Bygate and Samuda (2005), task repetition means 'repetitions of the same or slightly altered tasks - whether whole tasks, or parts of a task' (p. 43). Task repetition can be seen as a form of pre-task planning in which the speaker not only has the opportunity to plan the message in terms of conceptualization and formulation but also an opportunity to actually perform the task, going through all the stages of speech production from conceptualization, formulation, articulation to monitoring the output. Thus, the first performance of the task can be regarded as a preparation for subsequent performances (Ellis, 2005a) because "part of the work of conceptualisation, formulation and articulation carried out on the first occasion is

kept in the learners' memory store and can be reused on the second occasion" (Bygate, 2001, p. 29).

Previous research has consistently shown that task repetition is widely used in L2 classrooms and has beneficial effects for improving L2 performance in general and L2 fluency in particular (Ahmadian, 2011, 2012; Ahmadian et al., 2017; Ahmadian & Tavakoli, 2011; Bygate, 1996, 2001; Bygate & Samuda, 2005; De Jong & Perfetti, 2011; Fukuta, 2016; Hunter, 2017; Lambert et al., 2020; Lambert et al., 2017; Lynch & Maclean, 2000). The impact of task repetition on L2 production has recently been investigated with reference to three major stages of Levelt's (1989, 1993) model of speech production (i.e., *conceptualization, formulation*, and *articulation*, see Section 4.1). First, having a prior performance means a lot of relevant work on conceptualization, formulation, and articulation has been done before the speakers perform the task for the second time (Bygate & Samuda, 2005). Task repetition, thus, eases online processing demands because L2 speakers can build upon the content conceptualized on the first performance, the language used to express that content, as well as the task structure and relevant procedures of the task being performed. Second, repetition of tasks can provide learners with the opportunity to familiarise themselves with task demands, strengthen form-meaning connections, proceduralize relevant linguistic knowledge on the first performance and make it available for automatic processing during subsequent performances. It can be said that repeated practice facilitates automatized access to linguistic knowledge and frees up more attentional resources.

It is important to distinguish two types of task repetition: same task repetition (STR) and parallel task repetition (PTR) or task-type practice in Bygate's (2001) term. STR refers to the repetition of both procedures and content of the same task (Ahmadian, 2011; Ahmadian & Tavakoli, 2011; Bygate, 1996, 2001; De Jong & Perfetti, 2011; Lambert et al., 2020; Lambert et al., 2017), whereas PTR involves the repetition of the same procedures of a given task type but with different content (Bygate, 2001; De Jong & Perfetti, 2011; Lambert et al., 2020). While research has consistently indicated that fluency is robustly affected by STR (Ahmadian, 2011; Ahmadian & Tavakoli, 2011; Bygate, 2001; De Jong & Perfetti, 2011; Lambert et al., 2017; Lynch & Maclean, 2000; Wang, 2014), the effects of PTR on learners'

interlanguage development are mixed (Bygate, 2001; De Jong & Perfetti, 2011; Hunter, 2017; Lambert et al., 2020).

Bygate (2001) conducted a study with 48 overseas students to investigate the effects of PTR (i.e., repeating either narrative tasks or interview tasks) on the structure and control of oral language during the performances of a repeated task, and a parallel task. The students were separated into three groups of 16, namely a narrative group, an interview group, and a control group. During treatment, the two experimental groups performed six parallel tasks of one task type (either narrative tasks or interview tasks) on a fortnightly interval. The results revealed that PTR had a positive impact on the second performance of a repeated task, but did not benefit the performance of new versions of the same task type. In other words, PTR had significant effects on the performance of a repeated task, but there was no transfer of practice effects. Bygate explained the results by way of the significant interaction between repetition and task-type practice: "performance on a repeated task is primed by some kind of build-up from the experience of attending to tasks of the same type" (Bygate, 2001, p. 267). On the one hand, previous experience of a task 10 weeks earlier (STR) can be sufficient to affect subsequent performances of the same task by releasing spare capacity on the part of the speaker. On the other hand, PTR can have an impact on subsequent performance in a similar way at least when the task is repeated because of "a residual gain" from task-type exposure.

Similarly, De Jong and Perfetti (2011) investigated the role of speech repetition in fluency development for EFL students. The participants were divided into two groups to receive three fluency training sessions in either STR or PTR using two versions of the 4-3-2 task sequence. During a treatment session, while the STR group talked about one topic three time within time limit of 4, 3, and 2 minutes respectively, the PTR group performed the same task but talked about three different topics. The results showed that fluency improved for both groups during training sessions but only the STR group maintained the observed gains on the post-test. While short-term effects for both types of repetition were attributed for lexical and structure priming and possible effects of time pressure, De Jong and Perfetti (2011) explained that the transfer effects occurred for the STR group because of "the proceduralization of linguistic knowledge due to repeated use" (p. 560). Meanwhile,

the lack of long-term transferable effects for the PTR condition was possibly due to the impact of varied topics.

Recently, Lambert et al. (2020) investigated the relative impact of four task preparation options on L2 learners' speech production. The study involved the participation of 104 Japanese EFL learners, who were required to complete an oral opinion task after 10 minutes of preparation. During this preparation time, the STR group repeated the same task as the assessment task there times, the PTR group repeated three versions of the same task type with different content, the first language (L1) planning group planned content for the task using their L1 (Japanese), and the second language (L2) planning group planned the content for the task using their L2 (English). The effects of these task preparatory options were compared with regard to speech rate (overall fluency), between-clause pausing (conceptualization), within-clause pausing (formulation), and repairs (monitoring). Results showed that in terms of speech rate, the STR and PTR groups were highly comparable and outperformed both the L1 and L2 planning groups. There was also a transfer of observed fluency gains across performances for both the STR and PTR groups. These results suggest that different forms of task preparation might be sequenced to support L2 learners' speech production by alternately easing conceptualization, formulation, and monitoring demands during task performance.

5.4 Measuring Utterance Fluency

Several studies have attempted to establish objective measures of fluency (Bosker et al., 2013; De Jong, 2016b, 2018; De Jong & Bosker, 2013; De Jong et al., 2013; Derwing et al., 2009; Derwing et al., 2004; Götz, 2013; Hilton, 2008; Kahng, 2014; Kormos & Dénes, 2004; Préfontaine, 2013; Riggenbach, 1991; Rossiter, 2009; Towell et al., 1996). Although the variables used in the measurement of fluency differ from study to study, De Jong (2018) presents eleven objective measures of utterance fluency that are most frequently used in studies on fluency (see Table 5.1).

Table 5.1

Frequently Used Measures of Utterance Fluency

Measure	Formula
Speech rate	Number of syllables/total time
Prune speech rate	(Number of syllables – number of disfluent syllables)/total time
Articulation rate	Number of syllables/speaking time[a]
Pace	Number of stressed syllables/total time
Mean length of utterance	Total speaking time/number of utterances[b] *or* Number of syllables/number of utterances[b]
Number of silent pauses (per minute)	Number of silent pauses/total time or speaking time[a]
Mean duration of silent pauses	Pausing time/number of silent pauses
Phonation time ratio	Speaking time/total time
Number of filled pauses (per minute)	Number of filled pauses/total time or speaking time[a]
Number of repetitions (per minute)	Number of repetitions /total time or speaking time[a]
Number of repairs (per minute)	Number of repairs and restarts/total time or speaking time[a]

[a]Speaking time is equal total time minus silent pausing time
[b]Number of utterances is equal to the number of silent pauses *plus* 1

Note. From "Fluency in Second Language Testing: Insights from Different Disciplines," by N. H. De Jong, 2018, *Language Assessment Quarterly*, 15 (3), pp. 237-254. Copyright 2018 by Nivja H. De Jong

According to Skehan (2003) and Tavakoli and Skehan (2005), utterance fluency should be measured for three sub-dimensions: speed fluency, breakdown fluency, and repair fluency. More specifically, De Jong (2016b) and Lambert et al. (2017) argue that utterance fluency can be measured by speed fluency, location of breakdowns, and repairs in utterances. Towell et al. (1996) suggest that temporal variables of fluency are related to the stages of speech production. The following section addresses details of utterance fluency measures and their relationship with stages of L2 speech production.

5.4.1 Measures of Speed Fluency

The most frequently used measures of speed fluency are temporal measures such as speech rate and articulation rate.

Speech rate

Speaking rate (SR) is the most frequently used general measure of L2 fluency and can be unpruned or pruned speech rate. Unpruned speech rate refers to the mean number of unpruned syllables (i.e., including all dysfluencies such as false starts, repetitions, corrections, and non-lexical filled pauses) per second or minute (including silent pause time) (Chambers, 1997; Kormos & Dénes, 2004; Tavakoli, 2016). Pruned speech rate refers to the number of pruned syllables (i.e., the total number of syllables excluding all dysfluencies such as false starts, repetitions, corrections, and non-lexical filled pauses) per second or minute (including silent pause time). Pruned speech rate is typically argued to better reflect underlying cognitive fluency overall than unpruned speech rate (Derwing, 2017; Ellis & Barkhuizen, 2005; Rossiter, 2009; Skehan, 2014).

Articulation rate

Articulation rate represents the mean number of syllables per second or minute (excluding pause time). It is calculated by dividing the total number of syllables articulated (including filled pauses and dysfluencies) by the amount of speaking time (i.e., total sample time minus pause time) (De Jong, 2018; Kormos & Dénes, 2004; Towell et al., 1996).

Among the speed fluency measures, speech rate and pruned speech rate are reported to be strongly associated with both L2 proficiency and perceived fluency (Bosker et al., 2013; Derwing et al., 2004; Kahng, 2014, 2018; Kormos & Dénes, 2004; Towell et al., 1996), whereas articulation rate is not. As regards L2 speech production, speech rate provides a composite measure of the all stages of L2 speech production (De Jong, 2016a; Götz, 2013; Towell et al., 1996), and articulation rate can be related to automatization within the articulator (Kahng, 2014; Skehan, 2009; Towell et al., 1996). As De Jong (2016a) suggests, “if a researcher needs one single measure that encompasses all aspects of fluency at the same time, the measure of pruned speech rate is the best choice” (p. 214). Following this suggestion, in the present study, pruned speech rate is used as a composite measure of speech fluency that encompasses the efficient working of all stages of L2 speech production as the aim is to investigate the development of learners’ cognitive fluency as measured by utterance fluency under different task-based instruction conditions.

5.4.2 Measures of Breakdown Fluency

Breakdown fluency indicates disruptions in the flow of speech and is typically characterized by frequency, length, and location of pauses and hesitations (De Jong et al., 2013; Götz, 2013). The commonly used measures of breakdown fluency are frequency and location of filled/unfilled pauses.

Frequency of pauses

Frequency of pauses is calculated by dividing the total number of filled and unfilled pauses by the total time or total number of syllables (De Jong, 2018).

Pauses are often classified into filled and unfilled pauses. Filled pauses are typically non-lexical fillers such as *um, er, uh,* whereas unfilled pauses are periods of non-articulation (silent time) in the speech. However, not all silent pauses are indicative of disfluency because both L1 and L2 speakers may need to take short pauses for articulatory functions such as breathing (Rose, 2017). It is, therefore, necessary to have a lower-bound threshold for the definition of a silent pause. De Jong and Bosker (2013) recommend a threshold for defining silent pauses as a measure of L2 fluency with a lower cut-off threshold of 250-300 milliseconds as having the highest correlation with variation in L2 proficiency. They conclude that in L2 research, a cut-off point of 250 milliseconds is a good choice as shorter silent pauses represent weaker correlations with L2 proficiency. Thus, frequency of unfilled pauses can be calculated as the total number of silent pauses of minimum 250 milliseconds divided by the total number of syllables in a speech sample. In contrast, the frequency of filled pauses can be measured in terms of the total number of filled pauses divided by the total number of syllables in a speech sample. In the present study, frequency of pauses is calculated by dividing the total number of both filled and unfilled pauses by the total number of pruned syllables.

Location of pauses

Pauses can also be classified into within-clause (mid-clause) pauses and between-clause (final-clause) pauses with respect to whether the pauses occur in the middle of a clause or at a boundary between two clauses. As Davies (2003) suggests, it is the distribution of pauses, rather than the number of pauses that may be indicative of L2 speakers' proficiency. Between-clause pauses can reflect a variety of underlying processes

including time required to conceptualize the next clause, differences in speaking style, rhetorical effects, and articulation phenomena such as the need to take a breath. Skehan, Foster, and Shum (2016) suggested that dysfluencies (including pauses) between clauses were associated with Conceptualizer-linked discourse-level issues and are evidence of macro-planning. Interestingly, research has shown little difference for between-clause pausing between high and low proficiency speakers (De Jong, 2016b; Kahng, 2020) as the need for conceptualization and the individual differences involved tend to be the same for all speakers regardless of proficiency in the language that they are speaking. However, within-clause pauses have been found to be sensitive to L2 proficiency with lower-proficiency speakers pausing more within clauses than higher-proficiency speakers, particularly when speech breaks down before challenging lexical items (De Jong, 2016b; Kahng, 2014, 2020). For this reason, it is typically argued that when individual differences in speaking style are controlled, pauses at different locations (i.e., between clauses vs. within clauses) are associated with different cognitive processes underlying different stages of L2 speech production. While between-clause pauses provide a measure of conceptualization, within-clause pauses provide a measure of encoding problems (formulation) (see Section 4.2). Frequency of between-clause pauses can be calculated as the number of both filled and unfilled pauses between clauses divided by the total number of pruned syllables in a speech sample. Frequency of within-clause pauses, on the other hand, can be calculated by dividing the number of both filled and unfilled pauses within clauses by the total number of pruned syllables.

As the aim of the present study is to investigate how L2 speech fluency is influenced at different stages of speech production, the frequency of between-clause pauses and within-clause pauses are used as measures of conceptualization and formulation (encoding) respectively. Pause location was determined based on Analysis of Speech Unit (AS-unit) and clause boundaries (Foster et al., 2000). “An AS-unit is a single speaker’s utterance consisting of an independent clause, or sub-clausal unit, together with any subordinate clause(s) associated with either” (Foster et al., 2000, p. 365). An independent clause has minimum a clause that includes a finite verb, whereas an independent sub-clausal unit consists of at least one phrase which can be extended to a full clause or a minor utterance (e.g., *Oh poor woman, Thank you very much*) that is common in speech; and a subordinate clause consists

of minimum a finite or non-finite verb and at least another clause element (e.g., subject, object). Based on clause boundaries, filled and unfilled pauses were categorized as either between-clause or within-clause pauses.

5.4.3 Measures of Repair Fluency

Repair fluency refers to "occasions where the speaker attempts to make changes to what is being said" (Skehan, 2014, p. 20). Some researchers (e.g., De Jong et al., 2013; Skehan, 2003; Tavakoli & Skehan, 2005) suggest that repair fluency should be measured in terms of hesitations, repetitions, corrections, and reformulations. Others (e.g., De Jong, 2018; Lambert et al., 2017; Lennon, 1990), however, argue that repetitions, like filled pauses and hesitations, are more related to planning and encoding issues whereas self-repairs (e.g., reformulations and replacements) represent monitoring process during the speech production. Kormos (1999) also suggests that the frequency of self-repairs represents attention available for the monitoring process in L2 speech production. Recently, Lambert et al. (2020) argue that although both within-clause pauses and self-repairs can be indicative of learners' attention to form, it is possible to make distinct inferences from them. While within-clause pauses indicate the amount of attention that learners need for "initial encoding of their ideas" (lexical and syntax encoding), self-repairs represent the degree of attention they paid to the "correction of initially encoded language" (monitoring) (Lambert et al., 2020, p. 8). As such, in the present study, only actual repairs to initially encoded discourse are counted as repairs, as the goal was to measure learners' monitoring of initially encoded language. Self-repairs include reformulations and replacements. Reformulations are operationalized as modifications of syntax, morphology, or word order whereas replacements refer to the change of lexical items by either adding or deleting something (Skehan, 2014). Following Lambert et al. (2017), repair fluency is measured by dividing the number of overt self-repairs or self-corrections by the total number of pruned syllables in the speech sample.

5.4.4 Measures of Fluency Used in the Present Study

Recent research also suggests that measures of fluency need to be chosen carefully because one measure may overlap with others (Kormos, 2006; Skehan, 2014; Tavakoli, 2016; Tavakoli & Skehan, 2005). With respect to the relation between utterance fluency

measures and cognitive processes of speech production, as discussed in Section 5.2, research suggests that speech rate relates to all stages of L2 speech production, the number of between-clause pauses reflects breakdowns in the conceptualization stage, the number of within-clause pauses is linked to breakdowns in the formulation stage, and the number of repairs (i.e., reformulations and replacements) represents monitoring. Measuring fluency with regard to speech rate, between-clause pauses, within-clause pauses, and repairs can provide insights into how cognitive processes underlying L2 speech production develop. Based on these suggestions and in order to fulfil the aim of relating utterance fluency to cognitive processes of L2 speech production (De Jong et al., 2013; Kahng, 2014, 2020; Kormos, 2006; Levelt, 1989, 1995, 1999; Saito et al., 2018; Segalowitz, 2010), the current study adapts the measures of fluency used by Lambert et al. (2017). These four measures include: (1) speech rate (a composite measure of overall L2 speech production), (2) frequency of between-clause pauses (a measure of conceptualization), (3) frequency of within-clause pauses (a measure of formulation or encoding), and (4) frequency of self-repairs (a measure of monitoring) (Table 5.2). The count of frequencies of disfluencies (pauses and self-repairs) is to correct for the differences in length of speech samples (De Jong, 2016a).

Table 5.2

Fluency Measures Used in the Study

L2 Speech Processing Stage	Measures	Abbreviation	Formula
Overall speech production	Speech rate	SR	Number of pruned syllables/ total sample time (seconds)
Conceptualization	Frequency of between-clause pauses	BP_freq	Number of filled and unfilled pauses between clauses/ number of pruned syllables
Formulation	Frequency of within-clause pauses	WP_freq	Number of filled and unfilled pauses within clauses/ number of pruned syllables
Monitoring	Frequency of self-repairs	Rp_freq	Number of reformulations and replacements/ number of pruned syllables

In the present study, the aim is to investigate the effects of different approaches to task-based instruction on the development of learners' fluency in L2 speech production processes. As such, L2 performance was measured with respect to the four utterance fluency variables as presented in Table 5.2 before and after the treatment was given to the learners (i.e., pre-test vs. post-test). It was hypothesized that the difference between the post-test and pre-test in terms of utterance fluency measures would reflect the development of different cognitive processes underlying L2 speech production. Specifically, after the treatment, if the learners had higher speech rates, lower frequencies of within-clause pauses and self-repairs (the frequency of between-clause pauses may not be different), they would have improved their L2 speech processing capacity by moving from serial processing to parallel processing. Providing the learners with different forms of instruction (i.e., input vs. combined input-output) could have different impact on the degree of the move from serial processing to parallel processing. The learners' performances on a parallel task in the post-test were also measured to examine if the gains in fluency transfer to a parallel-task performance.

5.5 Summary

This chapter has reviewed the theories of L2 speech fluency. Various definitions of L2 fluency are presented in terms of cognitive, utterance, and perceived fluency. This is followed by the discussion on the relationship between utterance and cognitive fluency. Finally, objective measures of utterance fluency are outlined according to three dimensions of speed, breakdown, and repair fluency, on the basis of which, measures for the current study are selected to reflect different cognitive processes of L2 speech production.

Chapter 6. Methods

This chapter presents the research questions, followed by the methods used in this study including information about the research design, participants, materials, and procedures of data collection. The last section outlines how the data analysis was done, including transcription and coding, interrater reliability, data screening, and statistical analysis. The aim throughout is replicability of the study, either in whole or in part.

6.1 Research Questions

This study aimed to examine between-group and within-group effects of input-based task instruction versus combined input-output task-based instruction on EFL learners' speech production capacity in monologic task performances. The purpose of the study was to determine (1) the effects of input-based task instruction versus combined input-output task-based instruction on EFL learners' speech production capacity in the performance of a repeated oral monologic task (Research Question 1), and (2) whether or not any observed effects would transfer to the performance of a parallel task (Research Question 2). Furthermore, the study examined which measure of speech production capacity best predicted group membership (Research Question 3). The answers to Research Questions 1-3 will be addressed in Chapter 7. Finally, significant differences in speech production within each group over time (within-subject effects) were investigated in Research Questions 4-6. The answers to these questions will be addressed in Chapter 8.

The specific questions are as follows:

1. What are the differential effects of **input-based task instruction** versus **combined input-output task-based instruction** in comparison with narrow input instruction on **learners' speech production capacity** when repeating the **same narrative task**?
2. What are the differential effects of **input-based task instruction** versus **combined input-output task-based instruction** in comparison with narrow input instruction on **learners' speech production capacity** when performing **a parallel narrative task**?
3. Which measures of speech production capacity are the best predictors of group membership for the participants?

4. How does learners' speech production capacity change over time in the performance of a repeated task and a parallel task when they were involved in **narrow input instruction**?
5. How does learners' speech production capacity change over time in the performance of a repeated task and a parallel task when they are involved in **input-based task instruction**?
6. How does learners' speech production capacity change over time in the performance of a repeated task and a parallel task when they are involved in **combined input-output task-based instruction**?

6.2 Research Design

The research design was informed by a pilot study, which was carried out with an aim to assess the feasibility of the research design. The pilot study was also done to trial the instructional and testing materials and to determine anticipated effect sizes for the preliminary power analysis. Specifically, the pilot study employed a pre-test/post-test design and involved the participation of six English major freshmen at a university in Vietnam over five days. Results revealed that the research design was feasible, but that the researcher needed to recruit more participants than planned because of a potentially high dropout rate at the university where the study was conducted. In terms of materials, it was found out that the testing materials and the instructional materials were suitable for the learners in terms of difficulty as all the participants were able to complete the tasks within the time given. However, it was noted that the first task (picture sequencing) needed to include some ready-numbered pictures so that the participants could keep track and complete the task more easily. Based on the pilot data, Cohen's *d* was calculated to get the anticipated effect sizes. However, as the sample size of the pilot study was too small, power analysis was based on the number of participants required to reliably test the meaningful effect sizes in the planned post-hoc pairwise comparisons. According to Cohen (1988, p. 36), 34 participants in each group was deemed sufficient to provide power for the null hypothesis to be reliably tested for medium to large effect sizes (between d = 0.6 and d = 1.0).

This study employed an experimental, mixed within- and between-subjects design including three phases (pre-test, treatment, and post-test) with three participating groups:

input-based task instruction group (*Input Group* for short, or InG), combined input-output task-based instruction group (*Combined group* for short, or CbG), and a non-task-based narrow input group (*Comparison Group* for short, or CpG). The first independent variable in the study was the between-subjects variable *Group* at three levels namely (1) Input Group, (2) Combined group, and (3) Comparison Group. The second independent variable was the within-subject variable *Time* at three levels: (1) Pre-Test, (2) Same-Task Post-Test, and (3) Parallel-Task Post-Test. The dependent variables were four fluency measures related to hypothesized stages of L2 speech processing (Kormos, 2006; Levelt, 1989, 1995, 1999): (1) *speech rate* (a composite measure of L2 speech production), (2) *between-clause pausing* (a measure of conceptualization), (3) *within-clause pausing* (a measure of formulation or encoding), and (4) *self-repairs* (a measure of monitoring or attention to form) (Kahng, 2014; Kormos, 1999; Lambert et al., 2020; Lambert et al., 2017; Towell et al., 1996).

The study was conducted in three phases over a five-week period. In Week 1, all the participants undertook the pre-test (Oral Narrative Task 1). Next, the treatment took place over a period of three weeks (Weeks 2, 3, 4). The instruction time was carefully controlled across the three groups to consist of a total of 4.5 hours. Each class met for three 90-minute periods at weekly intervals during which time each group was given a different type of instruction. Specifically, the Input group performed only input-based tasks, the Combined group engaged in both input-based and output-based tasks, and the Comparison group received incidental input through narrow reading and listening without any specific task-based instruction. After treatment, the third phase was done in Week 5 with all participants undertaking two post-tests, a same-task-post-test and a parallel-task-post-test (Oral Narrative Task 1 and 8).

The overview of the research design can be seen in Table 6.1.

Table 6.1

Research Design

Phase	Week	Input Group	Combined Group	Comparison Group
Pre-Test	Week 1		Pre-Test (Narrative 1)	
Treatment	Week 2	Input lesson 1 (90') (Narrative 2, 3)	Input-output lesson 1 (90') (Narrative 2, 3)	Narrow reading and listening lesson 1 (90') (audiobook chapters)
	Week 3	Input lesson 2 (90') (Narrative 4, 5)	Input-output lesson 2 (90') (Narrative 4, 5)	Narrow reading and listening lesson 2 (90') (audiobook chapters)
	Week 4	Input lesson 3 (90') (Narrative 6, 7)	Input-output lesson 3 (90') (Narrative 6, 7)	Narrow reading and listening lesson 3 (90') (audiobook chapters)
Post-Test	Week 5	Same-Task Post-Test (Narrative 1) Parallel-Task Post-Test (Narrative 8)		

6.3 Participants

The study involved the participation of 102 EFL university-age learners (*n*=34 for each condition). The size of the groups was based on the observed differences and the anticipated effect sizes found in the pilot study. 34 participants for each group provided power for the null hypothesis testing with medium to large effect sizes between d = 0.6 and 1.0 in the post-hoc pairwise comparisons (Cohen, 1988, p. 36).

The participants were drawn from the English major freshmen population at a university in Hanoi, Vietnam. Taking part in the study was completely voluntary and participants signed a consent form at the beginning of the study. Information regarding their age, gender, and English learning experience was obtained using a background questionnaire. The participants were randomly allocated to three groups of 34. Descriptive statistics for the three groups are presented in Table 6.2.

Table 6.2

Descriptive Statistics for Background Information of the Participants

	Group	N	Range	Minimum	Maximum	Mean	SD
Age	Input Group	34	2	18	20	18.44	0.56
	Combined Group	34	2	18	20	18.79	0.59
	Comparison Group	34	1	18	19	18.44	0.50
	Total	102	2	18	20	18.56	0.57
Years of Leaning English	Input Group	34	4	8	12	9.82	1.24
	Combined Group	34	6	7	13	9.65	1.25
	Comparison Group	34	7	4	11	9.68	1.34
	Total	102	9	4	13	9.72	1.27

Of the 102 participants, only 15 (14.7%) were male while 87 (85.3%) were female which is representative of the enrolments in this discipline area. The participants' ages ranged between 18 and 20 years (M=18.56, SD=0.57). Most of the participants had approximately ten years of English learning experience at their primary and secondary schools (M=9.72, SD=1.27) and had an elementary proficiency level (Basic user, A2 CEFR) as indicated by their entrance results. According to the CEFR global scale, learners at A2-level can do the following:

> can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need. (Council of Europe, 2001, p. 5)

To establish the comparability of the three groups in terms of their speaking capacity, the speaking scores of the previous semester were compared (see Table 6.3). These scores were based on two progress tests (weighted 20% each) and one final test (weighted 60%). Progress Test 1 and the final test shared the same format including three parts: Part 1 required students to orally answer questions about themselves, their family, friends, work and interests; Part 2 required students to make a conversation with another student based on a situation card; and Part 3 asked students to speak about a given topic and answer one or two follow-up questions. Progress Test 2 required students to make an

oral presentation on one of the topics in the syllabus. Although these tests did not directly verify learners' abilities to relate narrative stories orally, it was felt that this combined range of tests would provide a valid measure of learners' abilities to complete the task used in the study as well as provide a valid indication of their initial comparability in terms of oral skills. They provided a valid measure of learners. A general linear model (GLM) analysis using SPSS25 for Windows was conducted on the speaking scores after all scores were confirmed to be normally distributed. The results show that there was no statistically significant difference in speaking scores among the three groups, Pillai's Trace F (4, 198) = .348, p = .845, η_p^2 = .007. This means the three groups of participants were comparable with regard to their speaking ability based on their previous semester's scores.

Table 6.3

Descriptive Statistics for Speaking Scores of the Participants

Group	N	Range	Minimum	Maximum	Mean	SD
Input Group	34	3.5	5.15	8.65	7.40	0.88
Combined Group	34	3.95	5.2	9.15	7.27	0.98
Comparison Group	34	3.4	5.2	8.6	7.18	0.75
Total	102	4	5.15	9.15	7.28	0.87

Furthermore, in order to ensure the comparability of the groups in terms of their fluency before treatment sessions, a one-way Multivariate Analysis of Variance (MANOVA) analysis was conducted on the pre-test data. The results using Pillai's trace reveal no significant main effect between groups for fluency, F (8, 194) = .501, p = .855, η_p^2 = .020. This indicates the comparability of the groups in terms of their levels of fluency before the treatment sessions.

Levene's Test of Equality of Error Variances shows that the variances of each variable were equal across the groups, speech rate, F (2, 99) = 2.42, p = .94; frequency of between-clause pauses, F (2, 99) = .25, p = .77; frequency of within-clause pauses, F (2, 99) = 2.85, p = .62; and frequency of repairs, F (2, 99) = 1.67, p = .193). This means that the assumption of homogeneity of variances was met for all four dependent variables.

The univariate ANOVA results confirm that there were no significant differences among the three groups on any of the four variables (speech rate, F (2, 99) = .075, p = .928, η_p^2 = .002; frequency of between-clause pauses, F (2, 99) = .149, p = .861, η_p^2 = .003;

frequency of within-clause pauses, $F(2, 99) = .571$, $p = .567$, $\eta_p^2 = .011$; frequency of repairs, $F(2, 99) = .152$, $p = .859$, $\eta_p^2 = .003$) (see Table 6.4). Therefore, it can be concluded that the three groups were comparable on all fluency measures at the beginning of the study.

Table 6.4

Between-Groups Differences in the Pre-Test

Dependent variables	df	Mean square	F	Sig.*	Partial eta squared
SR (sr)	2	0.003	0.075	0.928	0.002
BP_freq (lg10)	2	0.005	0.149	0.861	0.003
WP_freq (sr)	2	0.019	0.571	0.567	0.011
Rp_freq (sr)	2	0.001	0.152	0.859	0.003

Note. SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs, sr = square root transformation, lg10 = log10 transformation.

*The mean difference is significant at the .05 level.

6.4 Materials

6.4.1 Pre-Test and Post-test Materials

An oral narrative task was used for each of the testing sessions to elicit meaning-focused extended discourse. The task included a request to tell a story based on 15 pictures adapted from Lambert and Robinson (2014) with a one-sentence opening prompt (e.g., Please start with the following sentence: *One day, Mr. I <u>was walking along</u> and he <u>passed</u> a fresh pie in someone's window*). The stories are about the main character, Mr. I, who intends to take someone else's belongings, and it ends badly in each version.

Two versions of this task were created for the pre-test and the post-tests. In the Pre-Test, Narrative 1 was used (see Appendix D). The post-test included the task used in the Pre-Test (Narrative 1) plus a new parallel task (Narrative 8) (see Appendix E). Narrative 1 and Narrative 8 were parallel in that they both involved the same main character (Mr. I) with a similar array of characters, objects of focus and objects within the setting, and they both contained the same number of elements (15 pictures representing observable events) which were divided into three sub-narratives to ease the performance demands.

6.4.2 Treatment Materials

An additional six versions of the narrative task based on the same main character, Mr. I, were created for training learners in the two experimental groups. Each treatment session involved the narration of two parallel stories with 15 pictures for each. Specifically, the materials for the Input group included input-based tasks only, whereas the Combined group materials were composed of both input-based and output-based tasks. The Comparison group received non-task-based instruction that involved narrow listening and reading of an audiobook consisting of 30 chapters.

6.4.2.1 Materials for the Input Group

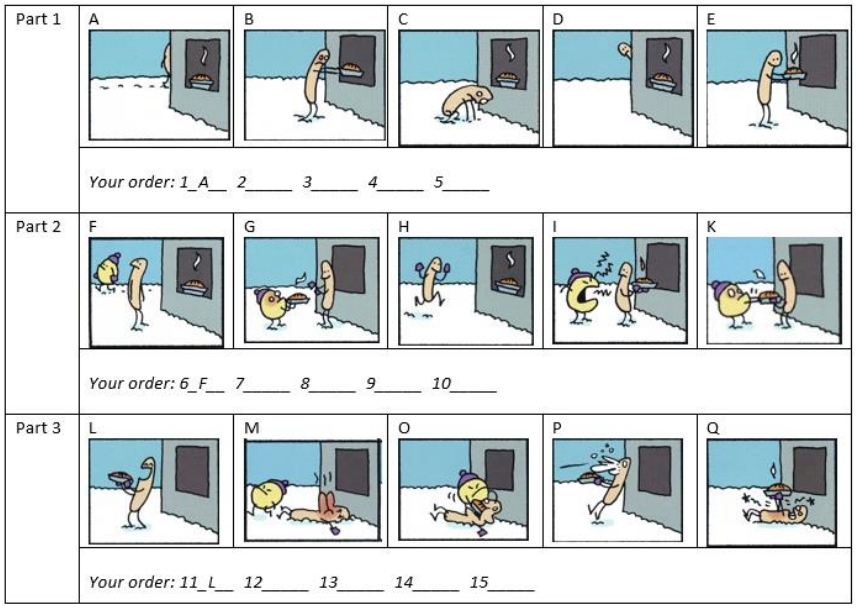
The Input group used materials based on listening and reading tasks (see Appendix F). None of the tasks required students to produce language in either the oral or the written mode. The treatment materials for the Input group were sequenced so that the tasks for the first story were completed before the students proceeded to the tasks for the second story.

The task sequence related to each story all began with a brief introduction about the story provided by the researcher. The first task (Task 1) was a listening and sequencing pictures task (see Figure 6.1). The 15 pictures, adopted from Lambert and Robinson (2014), were divided into three sub-narratives of five pictures each to ease the task demands.

Figure 6.1

Example of Input Task 1

Pre-task (5 minutes)
Listen to the teacher give a brief introduction about the story and listen to the recording of the story once to get some ideas what the story is about.
Task 1.1 (15 minutes)
- Look at the pictures for one minute and try to figure out what is in each picture.
- Listen to the recording of Story 1B again and sequence the pictures in each part of the story according to what you hear.
- Write a letter of the corresponding picture next to each number to show your answer. Number 1, 6 and 11 are already done for you.



Task 2 was a combined reading and listening task. The task included 15 items that directed the participants' attention to the use of past tense verbs and linking words that were relevant to developing the ability to complete an oral narrative task. An example of this is '*When he was far away from the house, he <u>put/ was putting</u> the pie on the ground, sat down, and began to eat it*.'

Finally, Task 3 was a reading task with five multiple-choice questions following the script of the story (see Figure 6.2). These questions were structured in such a way as to focus the learners' attention on the main ideas of the story and how these can be best expressed.

Figure 6.2

Example of Input Task 3 Questions

Questions

1. How does the speaker describe picture 1?

A. One snowy winter's day, Mr. I was walking along and he walked past a pie in the window of a house.
B. On a snowy winter's day, Mr. I was walking along and he walk past a pie in the window of a house.
C. One snowy winter's day, Mr. I walking along and he walked past a pie in the window of a house.

2. Which sentence describes the picture below?

A. There was no one there because he picked up the pie with both hands.
B. There was no one there so he picked up the pie with both hands.
C. There was no one there but he picked up the pie with both hands.

6.4.2.2 Materials for the Combined group

The materials for the Combined group included two input-based tasks and one output-based task. Task 1 and Task 2 were exactly the same tasks as Task 1 and 3 for the Input group. Task 3 for the Combined group was an output-based story-telling task (see Figure 6.3). There were two separate worksheets for each pair of the students in the Combined group (see Appendix G). Details of how these materials were implemented are outlined in Section 6.5 below.

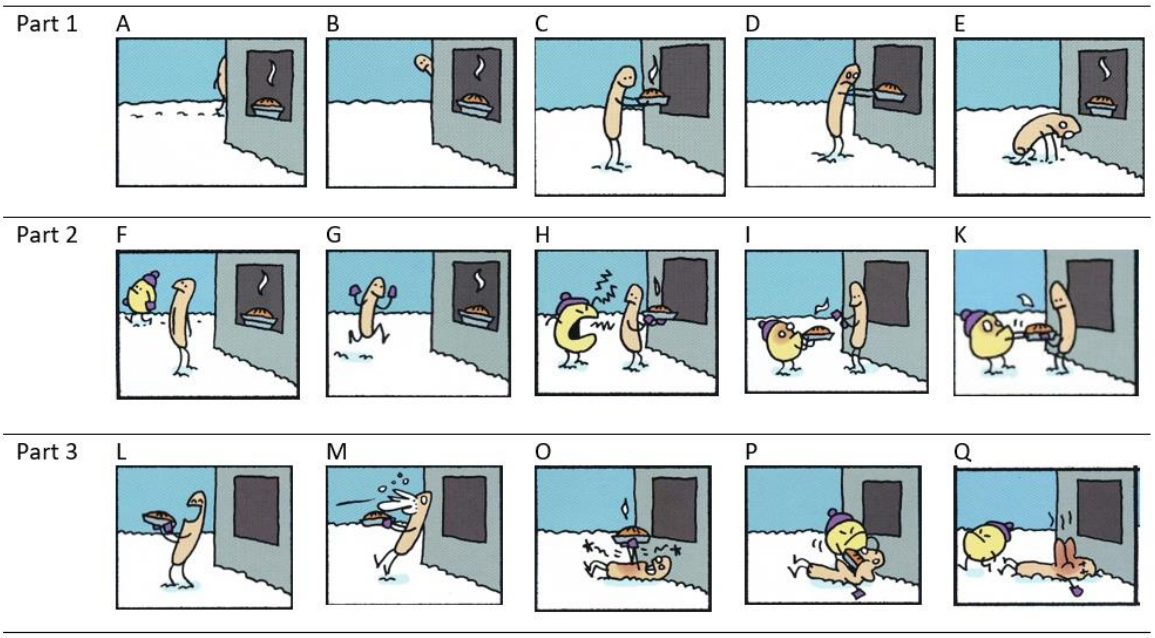
Figure 6.3

Example of Output Task

Task 3.1. (10 minutes)

Look at the pictures and retell the story about Mr. I to your partner using your own English. Please start with the following sentence: ***One snowy winter's day, Mr. I was walking along and he passed the pie in the window of a house.***

You have 3 minutes to prepare. Do not use your notes or the written version of the story when you do this.



Task 3.2. (10 minutes)

Listen to your partner retell Story 1B and check if s/he understood it correctly.

6.4.2.3 Materials for the Comparison Group

The Comparison group used an audiobook, namely *'Charlie and the Chocolate Factory*'. The entire story had 30 chapters with accompanying audios. The chapters were allocated in such a way that they were finished being read to the learners in the three class sessions. This novel included many instances of regular and irregular past tense verbs, which were considered incidental input. The 30-chapter audiobook treatment was considered as an equivalent amount of instructional contact time (not amount of instruction) to those in the input and combined input-output groups. It was also comparable in that it potentially exposed learners to narrative discourse in a similar way to what the learners in the other two conditions were using.

6.5 Procedures

6.5.1 Participant Recruitment and Group Establishment

Participants were recruited from eight English major classes. The researcher visited every class to discuss the research project and recruit participants, explaining the information in the participant information sheet (see Appendix A). Next, those who were willing to participate were asked to sign the consent form (see Appendix B) and complete a background questionnaire (see Appendix C). This was used to gain personal information about the participants' ages, gender, and English learning experience. The researcher then emailed a pre-test timetable to 102 students who confirmed their participation in the study. These participants were then randomly allocated into three groups of 34, namely the Input group, the Combined group, and the Comparison group. Each group was then divided into three sub-groups of nine to thirteen participants for ease of implementing the treatments.

6.5.2 Testing and Treatment

The study took place over five weeks and consisted of three phases: the pre-test (Week 1), the experimental treatment (Week 2-3-4), and the post-test (Week 5). Data were collected from testing sessions in meeting rooms, and treatment sessions were conducted in classrooms on the university campus. All the testing and treatment sessions were run by the researcher who was also an English teacher in the program.

1) The pre-tests were conducted individually in the first week prior to the treatment (Week 1). The test began with oral instructions, after which students had the opportunity to ask questions to ensure they knew what to do. After two minutes of pre-task planning time with a one-sentence opening prompt, the students then narrated a 15-frame picture story in English using their own linguistic resources. Their performances were audio-recorded using Sony ICD-UX560 Digital Voice Recorders, and subsequently transcribed for analysis.

2) In the experimental phase, each sub-group participated in three 90-minute lessons following one of the three treatment options. The total instructional contact time for all three groups was 4.5 hours during Weeks 2, 3, and 4. As discussed in Section 6.4.2 above, the lessons for the Input group involved input-based listening and reading tasks only, and the students were not obliged to produce output. The Combined group received similar

input-based tasks followed by output-based tasks where the students were required to produce output. Meanwhile, the Comparison group was exposed to narrow input through non-task-based listening and reading of an audiobook.

In the Input condition, the teacher first gave a brief introduction about the theme of the story and had the students listen to the story for the first time to get some ideas regarding what the story was about. The students could ask any questions for clarification. Next, they listened to a recording in order to sequence the scrambled pictures by writing the letter for each picture in the numerical sequence provided (Task 1). Then, in Task 2, the students read the story and guessed which language was most likely to be used in the story. They indicated their choice by putting a tick *(√)* above one of the words or phrases given to fill in gaps in the story. Next, they listened to the story again, circled the words they heard for each numerical gap, and checked if their guesses were right. Finally, in Task 3, the students were required to read the story and answer the five multiple-choice questions that followed. The materials were delivered task-by-task to ensure that the students completed the sequencing and listening tasks before they were exposed to the written story. The tasks were timed carefully to ensure all tasks were completed according to the allocated minutes as planned.

The Output condition participants received the same instruction time allowance for the two input-based tasks (Task 1 and 2), and in addition, they performed an output-based task (Task 3). The Pre-Task, Task 1, and Task 2 were completed in the exact way as the Pre-Task, Task 1, and Task 3 of the Input group. Specifically, first, the students received an introduction about the theme of the story. This was followed by the first listening of the story so that the students have some ideas of the story plot (Pre-Task). Then they listened to the story again and completed the picture-sequencing task (Task 1). Next, they read the story and answered the same set of multiple-choice questions as the Input group (Task 2). Finally, the students worked in pairs, and following the correct order of the pictures, they narrated their story orally so that their partner could take notes, and then retell that story (Task 3). Two stories were done in parallel and the students had opportunities to tell both stories using their own linguistic resources, one with the pictures given in the correct order and one from their notes. In this way, Task 3 was different from the Input condition. Details of the treatment activities for each experimental condition are shown in Table 6.5.

In the Comparison condition, the students listened to the English audiobook *'Charlie and the Chocolate Factory*' chapter-by-chapter and read along with the printed novel during each of the three sessions without any focus on the target tasks.

3) Finally, all post-tests were administered in Week 5, after the treatment was completed. The testing procedures were the same as in the Pre-Test. However, in the post-test phase, the students had to complete two narrative tasks, one task that was the same as in the Pre-Test (Same-Task Post-Test) and a new parallel task (Parallel-Task Post-Test). Half of the students did Same-Task Post-Test first and then Parallel-Task Post-Test, and the other half did Parallel-Task Post-Test first, followed by the Same-Task Post-Test to distribute practice effect equally across the two tests. The students' speech was audio-recorded for subsequent transcription and analysis.

Table 6.5

Treatment Session Activities

Comparison Group Lesson (90’)	Input Group Lesson (90’)	Combined Group Lesson (90’)	
		Student A	Student B
Listening to an audiobook and reading along	*Story 1* - Pre-task 1: Introduction (5’) - Task 1.1: Listen to the audio and sequence pictures (15’) - Task 1.2: Read the story and choose one of the underlined words to complete the story, then listen to check the answers (15’) - Task 1.3: Read the story to answer questions (10’) ---------- *Story 2* - Pre-task 2: Introduction (5’) - Task 2.1: Listen to the audio and sequence pictures (15’) - Task 2.2: Read the story and choose one of the underlined words to complete the story, then listen to check the answers (15’) - Task 2.3: Read the story to answer questions (10’)	- Pre-task 1.1: Introduction (Story 1) (5’) - Task 1.1: Listen to the audio and sequence pictures (Story 1) (15’) - Pre-task 1.2: Introduction (Story 2) (5’) - Task 1.2: Listen to the audio and take notes (Story 2) (15’) ---------- - Task 2: Read the story to answer questions (Story 1) (10’) ---------- - Task 3: Retell Stories (Output) (40’) *Story 1* + 3.1 Retell Story 1 (10’) + 3.2 Listen to Student B retelling Story 1 and check (10’) *Story 2* + 3.3 Listen to Student B retelling Story 2 and take notes (10’) + 3.4 Retell Story 2 (10’)	- Pre-task 1.1: Introduction (Story 1) (5’) - Task 1.1: Listen to the audio and take notes (Story 1) (15’) - Pre-task 1.2: Introduction (Story 2) (5’) - Task 1.2 Listen to the audio and sequence pictures (Story 2) (15’) ---------- - Task 2: Read the story to answer questions (Story 2) (10’) ---------- - Task 3: Retell Stories (Output) (40’) *Story 1* + 3.1 Listen to Student A retelling Story 1 and take notes (10’) + 3.2 Retell Story 1 (10’) *Story 2* + 3.3 Retell Story 2 (10’) + 3.4 Listen to Student A retelling Story 2 and check (10’)

6.6 Data Analysis

6.6.1 Transcribing and Coding

All participants' task performances (*n* = 306) from the pre-tests and post-tests were transcribed in normal orthography by a research assistant who was an EFL teacher in Vietnam. These transcripts were then verified for accuracy by the researcher using the PRAAT 6.053 software function "To TextGrid (silences)" (Boersma & Weenink, 2019). Unpruned and pruned versions of transcripts were then prepared. Unpruned transcripts included all the uttered syllables, and disfluencies such as false starts, hesitations, repetitions, filled pauses (e.g. 'uh', um), and self-repairs (reformulations and replacements) (Ellis & Barkhuizen, 2005) (see Appendix K). In contrast, the pruned transcripts involved the removal of all these disfluencies.

Semi-automatic analysis (i.e., combining automatic and manual coding) was carried out using the PRAAT 6.053 software function "To TextGrid (silences)" (Boersma & Weenink, 2019). Specifically, the total time of each speech sample and the length of silent pauses of 0.25 seconds or more were identified automatically, whereas the location of filled and unfilled pauses, the total number of within-clause and between-clause pauses, and the number of self-repairs were coded manually by the researcher. It was also necessary for the length of silent pauses to be verified manually for accuracy. All pause boundaries (the beginnings and ends of each run of speech and pause) were checked and adjusted as necessary by repeated listening and visual inspection of the zoomed spectrogram and waveforms in the PRAAT software. As the screen view zoomed in, the spectrogram with the pitch and intensity display allowed for easy identification of the beginning and end of each pause (see Figure 6.4). The length of each pause was then added to the corresponding location in the transcript, which was opened in another window at the same time. Pauses shorter than 0.25 seconds were excluded because they were deemed to be micropauses (Riggenbach, 1991), and there was not a strong relationship between these pauses and variation in L2 speech production ability (De Jong & Bosker, 2013). Next, each transcript was broken down into AS-units (Foster et al., 2000). Based on the clause boundaries, it was possible to identify which pauses (both filled and unfilled pauses) were in the middle (within-clause pauses) or at the end of the clause (between-clause pauses). All false starts,

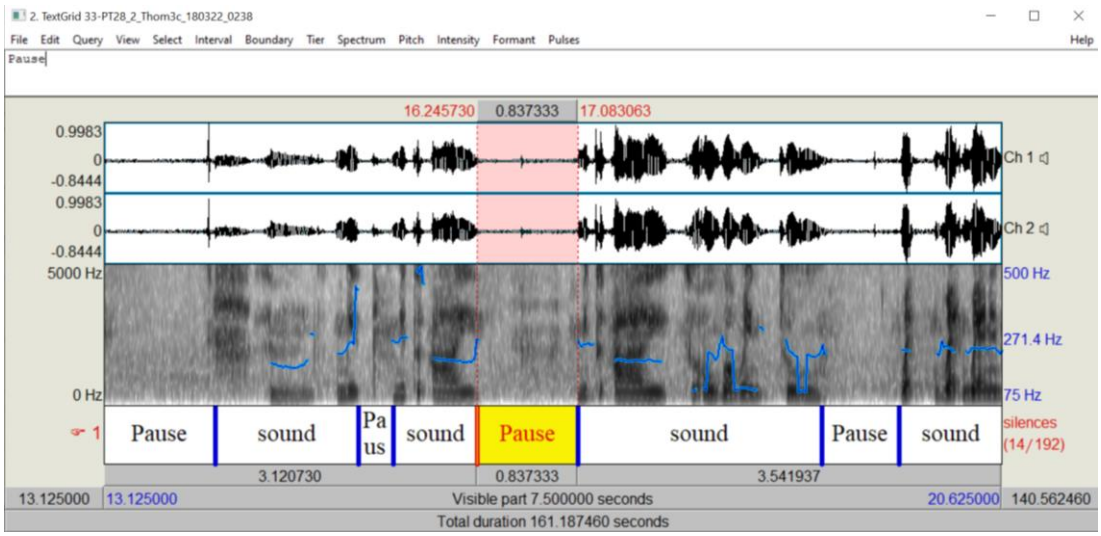
repetitions, hesitations, reformulations, and replacements were also identified and marked in the appropriate places in the transcripts. In the coded transcripts, silent pauses were marked in parentheses according to duration (in seconds), dysfluencies such as false starts, hesitations, and repetitions were put in brackets <…>, and self-repairs such as reformulations and replacements were put inside brackets […]. AS-unit boundaries were marked by an upright slash (|), clause boundaries were marked by a double colon (::); filled and unfilled pauses within clauses were marked blue whereas filled and unfilled pauses between clauses were marked red for easy coding (see Appendix H). The unpruned coded version of transcripts then included all information regarding the length and location of pauses, self-repairs and dysfluencies (see Appendix I). Finally, pruned transcripts were produced by removing all dysfluencies including filled pauses, false starts, repetitions, hesitations, reformulations, and replacements. The following is an example of an unpruned transcript and a pruned version of the same transcript.

Unpruned utterances: *He decided to uh <steal> steal the pie (1.56) |When he touched [a] (0.33) the pie (0.38):: a (0.62) pink guy [go through] (0.53) ah went through (1.61)|*

Pruned utterances: *He decided to steal the pie. When he touched the pie, a pink guy went through*.

Figure 6.4

PRAAT Sound File and Textgrid



6.6.2 Calculating Variables

As outlined in Section 5.4.4 (see Table 5.2), the four dependent variables of the study were (1) speech rate (as a measure of overall L2 speech production capacity), (2) the frequency of between-clause pauses (as a measure of the conceptualization stage of L2 speech production), (3) the frequency of within-clause pauses (etc. formulation), and (4) frequency of self-repairs (etc. monitoring). In order to compute these four measures, a range of fluency sub-measures were coded as detailed below. Among these sub-measures, the total sample time was automatically extracted using PRAAT, the number of pruned syllables were based on pruned transcripts, whereas the other three measures (number of filled and unfilled pauses between clauses, number of filled and unfilled pauses within clauses, and number of self-repairs) were coded using unpruned transcripts.

6.6.2.1 Speech Rate

Speech rate was calculated as the number of pruned syllables divided by the total sample time (in seconds).

The number of pruned syllables was counted using an online syllable counter (syllablecounter.net) (Lambert et al., 2020). After pruned transcripts were prepared (excluding all of the disfluencies such as filled pauses, false starts, repetitions, hesitations, reformulations and replacements), they are entered into the online syllable counter (syllablecounter.net) to determine the final syllable count. The total number of pruned syllables was counted from the first syllable after the prompt sentence to the last syllable of the last meaningful utterance of the pruned transcript. This number was used to calculate all four variables of the present study. Specifically, pruned speech rate was calculated by dividing the number of pruned syllables by the number of seconds of the sample. The number of pruned syllables was also used as the divisor for all other three measures.

The total sample time (in seconds) of the speech sample was drawn from the PRAAT TextGrid file. It was the number of seconds between the first syllable after the prompt sentence and the last syllable of the last meaningful utterance in the transcript. Unfinished utterances at the end of the transcript that did not make sense or contribute to the content were excluded from the analysis (e.g. '*So he*'). Similarly, ending utterances like '*All done*', '*Thank you'*, or '*I'm finished*' were also discounted.

6.6.2.2 Frequency of Between-Clause Pauses

Frequency of between-clause pauses was calculated by dividing the number of filled and unfilled pauses between clauses by the number of pruned syllables.

The number of filled and unfilled pauses (minimum 0.25 seconds) between clauses was manually counted separately by the researcher using the unpruned transcripts. These numbers were added up to make the total number of between-clause pauses, and then they were entered into the spreadsheets and SPSS for later calculation.

6.6.2.3 Frequency of Within-Clause Pauses

Frequency of within-clause pauses was calculated by dividing the number of filled and unfilled pauses within clauses by the number of pruned syllables.

The number of filled pauses and the number of unfilled pauses (minimum 0.25 seconds) within clauses were counted separately after the unpruned version of transcripts with all information regarding pause length and location was prepared. The values were tabulated into the spreadsheet and added to calculate the total number of within-clause pauses.

6.6.2.4 Frequency of Self-Repairs

Frequency of self-repairs was calculated by dividing the number of self-repairs (reformulations and replacements) by the number of pruned syllables.

The number of self-repairs (e.g., *a pink man <is> was coming*) and reformulations (e.g., *Mr. I <took> uh decided to take a piece of cake*) was also manually coded based on the information in the unpruned transcripts. The data were entered into the spreadsheet for calculation.

In brief, the steps for coding the data are summarized in Table 6.6 below:

Table 6.6

Steps of Coding Data

Steps	Data Coding Activities	Outcome
1	Transcribe	Raw orthographic transcripts
2	Verify transcripts	Raw orthographic transcripts
3	Identify total sample time using PRAAT	Number of seconds of each speech sample
4	Identify silent pauses of 0.25 seconds and add to the raw transcripts using PRAAT	Unpruned transcripts with silent pause information (length and position)
5	Break unpruned transcripts into AS-units	Unpruned transcripts in AS-units
6	Identify and mark filled and unfilled pauses within and between clauses	Within clause pauses marked blue Between clause pauses marked red
7	Identify and mark self-repairs (reformulations and replacements) and other dysfluencies	Repairs put in brackets [...] Other dysfluencies put in brackets <...>
8	Count number of filled and unfilled pauses within and between clauses	Number of filled and unfilled pauses within clauses Number of filled and unfilled pauses between clauses
9	Count number of self-repairs (reformulations and replacements)	Number of repairs (reformulations and replacements)
10	Remove all dysfluencies including filled pauses, false starts, repetitions, hesitations, repairs	A pruned version of transcripts
11	Run syllable counts for pruned transcripts	Number of pruned syllables
12	Tally all values into spreadsheets for subsequent calculation of variables and entry into SPSS	Data spreadsheet

6.6.3 Interrater Reliability

To establish interrater reliability, one group's performances (34 transcripts) were coded independently by the researcher and a research assistant after careful discussion and practice on three variables that required manual coding: (1) number of unfilled and filled pauses within clauses, (2) number of unfilled and filled pauses between clauses, and (3) number of self-repairs (replacements and reformulations). Pearson Product Moment Correlation Coefficients showed high interrater reliability for all manually coded measures: number of unfilled and filled pauses within clauses (r = .996), number of unfilled and filled pauses between clauses (r = .992), and number of self-repairs (r = .959). After interrater reliability was established, the researcher coded the rest of the transcripts.

6.6.4 Preliminary Data Screening

Score distributions were first screened for normality and homogeneity of variance using SPSS25. The normality of the score distributions was determined by looking at the z-scores of skew and kurtosis and inspection of histograms and Q-Q plots. Score distributions for each variable were considered to be normally distributed if *z*-scores for skew and kurtosis did not exceed ± 1.96 (George & Mallery, 2010; Gravetter & Wallnau, 2014). Preliminary data screening shows that most distributions were positively skewed. However, it was possible to correct the skewness in the distributions with either square root or log10 transformation to have normally distributed scores for parametric statistical testing. Based on the shape of the raw score distributions (Tabachnick & Fidell, 2005, pp. 86-88), speech rate, frequency of within-clause pauses, and frequency of self-repairs for each group were corrected with square root transformation, whereas frequency of between-clause pauses was corrected with log10 transformation. After transformation, all score distributions were approximately normally distributed with the transformed z-scores for skew and kurtosis within ±1.96.

Box's M Test of Equality of Covariance Matrices was used to test homogeneity of covariance. The results show that a Box's M Test of Equality of Covariance Matrices was not significant, F (156, 26115) = .948, p = .667, indicating that the observed covariance matrices of the dependent variables were equal across groups.

The assumption of homogeneity of variance was assessed by examining whether the *p* value in the Levene's Test of Equality of Error Variances had exceeded the alpha .05. The results of Levene's test show non-significant values for most of the distributions (p >.05), except for the measure of self-repair fluency in Parallel-Task Post-Test, F (2, 99) = 3.41, p = .037 (see Table 6.7). This indicates that variances were homogeneous for most of levels of the dependent variables. However, when group sizes are equal, MANOVA is fairly robust against violations of this assumption (Field, 2009). Therefore, in the present study, MANOVAs using SPSS25 for Windows were conducted to seek answers to the research questions (noting that all three groups had the same number of participants, n = 34).

Table 6.7

Levene's Test of Equality of Error Variances (Based on Mean)

Test	Variable	Levene statistic	df1	df2	Sig.
Pre-Test	SR (sr)	2.424	2	99	0.094
	BP_freq (lg10)	2.857	2	99	0.062
	WP_freq (sr)	0.250	2	99	0.779
	Rp_freq (sr)	1.671	2	99	0.193
Same-Task Post-Test	SR (sr)	0.440	2	99	0.645
	BP_freq (lg10)	2.147	2	99	0.122
	WP_freq (sr)	1.043	2	99	0.356
	Rp_freq (sr)	1.381	2	99	0.256
Parallel-Task Post-Test	SR (sr)	0.268	2	99	0.765
	BP_freq (lg10)	0.388	2	99	0.680
	WP_freq (sr)	0.281	2	99	0.756
	Rp_freq (sr)	3.412	2	99	**0.037**

Note: SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs, sr = square root transformation, lg10 = log10 transformation.

6.6.5 Statistical Analyses

With regard to Research Questions 1 and 2, a one-way MANOVA was performed on the post-test scores, followed by univariate tests and post-hoc pairwise comparisons to examine the differences between the three groups at the beginning and at the two post-tests. In order to further examine the relationship between dependant variables and identify which of the four speech production variables at which test point best differentiated groups (Research Question 3), a follow-up Discriminant Function Analysis (DFA) was conducted. As suggested by Field (2009), following up a MANOVA with both univariate tests and discriminant function analysis provides a fuller picture of the data. Eight score distributions of the four fluency measures (speech rate, frequency of between-clause pauses, frequency of within-clause pauses, and frequency of self-repairs) at two test points (Same-Task Post-Test and Parallel-Task Post-Test) were used for the MANOVA and DFA.

To investigate Research Questions 4, 5, and 6, a repeated-measures MANOVA was performed on split-file data to determine within-subject effects over the three test points for each group. The independent variable was the within-subject variable *Time* at three levels (Pre-Test, Same-Task Post-Test, and Parallel-Task Post-Test). The four dependent variables were speech rate, between-clause pausing, within-clause pausing and self-repairs. This test created separate results for each group so that specific differences across three tests for each group on each measure were located.

Because the assumption of homogeneity of variance was violated for the measure of self-repairs in Parallel-Task Post-Test, Pillai's trace was used as it is the most conservative test for avoiding Type 1 error in the GLM and also quite robust to violations of the assumptions in case of equal sample sizes (Field, 2009, pp. 604-605). For pairwise comparisons on frequency of self-repairs in Parallel-Task Post-Test, Games-Howell tests were used because it was the most conservative test in the GLM for avoiding Type II error when equality of variance is violated. Any effects on the measure of self-repairs were interpreted with caution. Fischer's Least Significant Difference (LSD) tests were chosen for pairwise comparisons on all other measures. LSD tests do not make adjustments for comparisons as the tests are used following MANOVA, which can be argued to protect against inflating Type 1 error rate as a result of post-hoc analyses (Cramer & Bock, 1966; Lambert et al., 2020). Finally, effect sizes for three-group comparisons were interpreted using partial eta squared (η_p^2) following Cohen (1988) who suggested that η_p^2 values of .01, .09, and .25 represent small, medium, and large effect sizes, respectively. For all post-hoc pairwise comparisons between groups, Cohen's *d* was calculated to indicate the size of an experimental effect because it is very important for meta-analysis on fluency in the future. According to Cohen (1992), the effect sizes *d* = 0.2, 0.5, and 0.8 represent small, medium, and large effect sizes respectively**.** Therefore, if the means of two groups do not differ by 0.2 standard deviations or more, the difference is trivial, even if it is statistically significant. In this study, partial eta squared (η_p^2) was used for multi-group comparisons and Cohen's *d* was used for pairwise comparisons because it was difficult to obtain the value of *d* for comparisons of more than two groups. While partial eta-squared indicates the % of the variance in the dependent variable attributable to a particular independent variable, Cohen's *d* indicates the size of the difference between two means in standard deviation units.

6.7 Summary

In this chapter, details of the research methods have been presented. The research design, participants, materials with detailed procedures, and methods of data analysis including transcribing, coding, and interrater reliability were also explained. Finally, preliminary data screening results were reported and statistical procedures used in the study were outlined. The next two chapters will present the quantitative results of all the statistical tests that were used to answer the research questions.

Chapter 7. Results for Between-Groups Differences

This chapter presents the results for differences between groups after treatment. The chapter first reports the descriptive statistics for the dependent variables across groups and tests. Next, Research Questions 1 and 2 are addressed with the presentation of the relative effects of input-based task instruction and combined input-output task-based instruction in comparison with non-task-based narrow input instruction, and the differences between the two task-based instructional approaches in developing learners' speech production capacity. The last part of this chapter reports the results of the discriminant function analysis with regard to Research Question 3. Hence the three research questions that this chapter addresses are:

RQ1. What are the differential effects of **input-based task instruction** versus **combined input-output task-based instruction** in comparison with narrow input instruction on **learners' speech production capacity** when repeating the **same narrative task**?

RQ2. What are the differential effects of **input-based task instruction** versus **combined input-output task-based instruction** in comparison with narrow input instruction on **learners' speech production capacity** when performing **a parallel narrative task**?

RQ3. Which measures of speech production capacity are the best predictors of group membership for the participants?

7.1 Descriptive Statistics

As discussed in Chapter 5 (Table 5.2) and Chapter 6, the four dependent measures used in the study are speech rate (SR), frequency of between-clause pauses (BP_freq), frequency of within-clause pauses (WP_freq), and frequency of self-repairs (Rp_freq). Each measure was taken three times (i.e., once before the treatment and twice after), so that there are 12 score distributions included in the statistical analysis. Table 7.1 shows the descriptive statistics for all measures in each condition at the Pre-Test and both Post-Tests based on untransformed means (see Appendix M for full descriptive statistics of transformed means).

Table 7.1

*Descriptive Statistics for Speech Production Measures across Groups and Tests**

Measures	Group	N	Pre-Test		Same-Task Post-Test		Parallel-Task Post-Test	
			Mean	SD	Mean	SD	Mean	SD
SR	Input	34	1.14	0.41	1.21	0.30	1.16	0.33
	Combined	34	1.14	0.53	1.29	0.32	1.16	0.33
	Comparison	34	1.09	0.34	1.08	0.33	1.09	0.34
BP_freq	Input	34	0.28	0.13	0.26	0.11	0.23	0.09
	Combined	34	0.30	0.13	0.26	0.09	0.23	0.07
	Comparison	34	0.29	0.13	0.30	0.15	0.27	0.10
WP_freq	Input	34	0.42	0.22	0.40	0.18	0.42	0.19
	Combined	34	0.51	0.31	0.41	0.19	0.46	0.21
	Comparison	34	0.45	0.22	0.49	0.25	0.49	0.23
Rp_freq	Input	34	0.01	0.01	0.01	0.01	0.01	0.01
	Combined	34	0.02	0.02	0.01	0.01	0.01	0.01
	Comparison	34	0.02	0.02	0.03	0.02	0.02	0.01

Note. SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs

*Values represent untransformed means.

As can be seen from Table 7.1, both Input group and Combined group produced higher speech rates on the Same-Task Post-Test than on the Pre-Test. However, in the performance of a parallel task, the mean speech rates of both groups decreased to almost equal the Pre-Test means. The Comparison group displayed a stable speech rate of around 1.09 pruned syllables per second across three performances. With regard to between-clause pausing, both experimental groups paused less between clauses on the Same-Task Post-Test than on the Pre-Test, and the frequency went down even further in the Parallel-Task Post-Test. Meanwhile, the scores for the Comparison group only went up and then down again to a small degree. Regarding within-clause pausing, Input group and Combined group showed a similar fluctuating trend with a decrease in scores between Pre-test and the Same-Task Post-Test, and an increase in scores between the Same-Task Post-Test and the Parallel-Task Post-Test, whereas scores of the Comparison group in both post-tests were higher than on the Pre-Test. In terms of self-repairs, the Input group had stable self-repair scores across

three tests, the Combined group had less self-repairs in both post-tests than on the Pre-Test, while the Comparison group scores fluctuated slightly over the three tests. In brief, therefore, it seems that both experimental groups displayed a somewhat opposite trend to the Comparison group. Details of the differences between groups are presented in the following section.

7.2 Results for Between-Groups Differences after Treatment

As presented in Section 6.3 (Chapter 6), the three groups were comparable before treatment sessions with respect to the fluency measures as determined by the pre-test task performance. The test scores for speech rate, frequency of between-clause pauses, frequency of within-clause pauses, and frequency of self-repairs in the two post-tests (i.e., the Same-Task Post-Test and the Parallel-Task Post-Test) were compared to determine whether there were significant differences in speech production capacity between groups at each post-test point after treatment. A one-way MANOVA was then performed with *Group* as a between-subjects factor at three levels (Input group, Combined group, and Comparison group) and eight dependent variables (speech rate, frequency of between-clause pauses, frequency of within-clause pauses, and frequency of self-repairs at two post-tests).

One-way MANOVA results, using Pillai's trace, show a significant main effect for group, F (16, 186) = 2.494, p = .002, η_p^2 = .177. This means that there were significant differences in speech production capacity between the groups after treatment on a linear combination of all eight dependent variables (4 on the Same-Task Post-Test; 4 on the Parallel-Task Post-Test). In the following sections, the results of follow-up univariate tests and post-hoc pairwise comparisons using LSD tests, together with Cohen's d values, provide more information on the source of this effect.

7.2.1 Between-Groups Differences on the Same-Task Post-Test

In order to answer Research Question 1, the univariate results for the four dependent variables on the Same-Task Post-Test were examined. The results reveal that the three groups were significantly different with regard to speech rate with a small effect size (F *(2,99)* = 3.788, p = .026, η_p^2 = .071), and frequency of self-repairs with a medium effect size (F *(2,99)* = 9.270, p = .000, η_p^2 = .158). The differences between groups did not reach a significance level for the frequency of between-clause pauses, and the frequency of within-

clause pauses, and effect sizes were very small (p = .355, η_p^2 = .021 and p = .169, η_p^2 =.035, respectively). Table 7.2 summarizes the univariate results for the Same-Task Post-Test.

Table 7.2

Between-Groups Differences in Speech Production Measures on the Same-Task Post-Test

Variable	df	Mean square	F	Sig.*	Partial eta squared
SR (sr)	2	0.083	3.788	**0.026***	0.071
BP_freq (lg10)	2	0.033	1.046	0.355	0.021
WP_freq (sr)	2	0.041	1.813	0.169	0.035
Rp_freq (sr)	2	0.034	9.270	**0.000***	0.158

Note. SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs, sr = square root transformation, lg10 = log10 transformation.

*The mean difference is significant at the .05 level.

Follow-up post-hoc pairwise comparisons using Fischer's LSD tests together with Cohen's *d* values revealed more detailed information about the differences between the three groups on each of the four measures.

7.2.1.1 Between-Groups Differences in Speech Rate on the Same-Task Post-Test

Overall, after treatment both experimental groups had higher speech rates than the Comparison group in the performance of the Same-Task Post-Test. However, only the difference between the Comparison group and the Combined group where the students had the opportunity to produce the new language that they heard in performing the task reached statistical significance. Specifically, follow-up post-hoc pairwise comparisons using Fischer's LSD tests (Table 7.3) show that the mean scores for speech rate were significantly different between the Comparison group and the Combined group with a medium effect size (p = .008, d = .65). In contrast, the differences between the Comparison group and the Input group were not significantly different, with a small effect size (p = .092, d = .41). There was no significant difference between the Input group and the Combined group (p = .31), and the effect size was small (d = .26).

Table 7.3

Groups Comparisons in Speech Rate on the Same-Task Post-Test

Dependent variable	Group comparison	Mean difference	Std. error	Sig.*	Cohen's *d***	95% CI for difference	
						Lower bound	Upper bound
Speech rate[sr]	CpG - InG	-0.061	0.036	0.092	0.41	-0.108	0.035
	CpG - CbG	-0.097	0.036	0.008*	0.65	-0.169	-0.027
	InG - CbG	-0.037	0.036	0.310	0.26	-0.122	0.049

Note. InG = Input group, CbG = Combined group, CpG = Comparison group.

[sr] Values represent square root transformation.

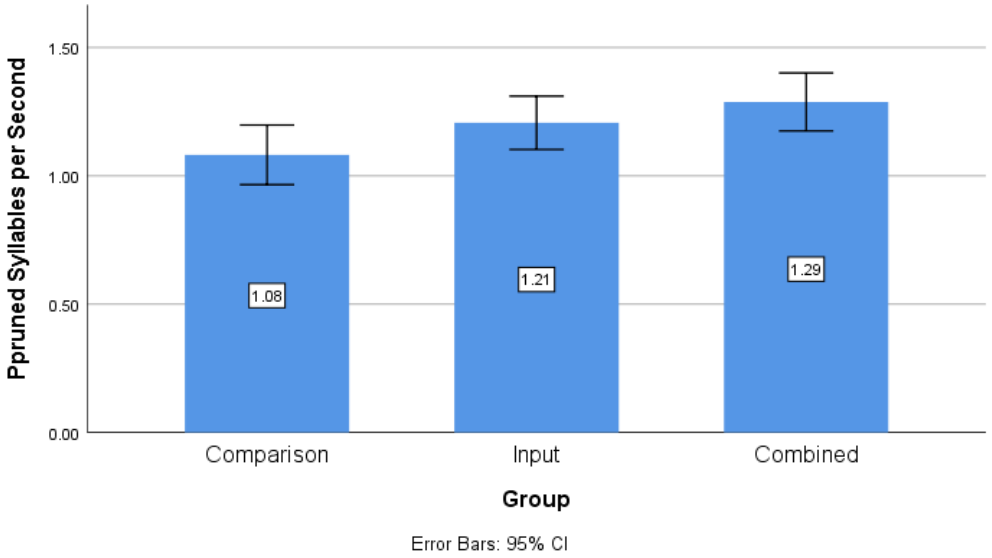
*The mean difference is significant at the .05 level.

** Cohen (1988) recommends *d*=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

The differences in speech rate of the three groups on the same-task post-test are illustrated in Figure 7.1 below. The figure displays the untransformed means and the 95% confidence interval of those means, which shows that after treatment, the participants in the Combined group, on average, spoke with the fastest speech rate, while the Comparison had the slowest speech rate. It should be noted that the speech rate figures presented here are based on pruned discourse.

Figure 7.1

Speech Rate for Groups on the Same-Task Post-Test



7.2.1.2 Between-Groups Differences in Between-Clause Pausing on the Same-Task Post-Test

The results for post-hoc pairwise comparisons using Fischer's LSD tests in the frequency of between-clause pauses of the three groups on the Same-Task Post-Test are presented in Table 7.4. It can be seen that, when repeating the same task as on the Pre-Test, the students in both experimental groups paused less between clauses than those in the Comparison group. However, these differences did not reach statistical significance between any of the groups. Furthermore, the differences between the Comparison group and the two experimental groups (the Input group and Combined group) had only small effect sizes (p = .234, d = .28; and p = .196, d = .31, respectively), whereas the differences between the Input group and Combined group were negligible (p = .918, d = .03). These results indicate that the respective treatments did not result in significant between-groups differences in learners' conceptualization during their speech production.

Table 7.4

Groups Comparisons in Between-Clause Pausing on the Same-Task Post-Test

Dependent variable	Group comparison			Mean difference	Std. error	Sig.*	Cohen's *d***	95% CI for difference	
								Lower bound	Upper bound
Between-clause pausing[log10]	CpG	-	InG	0.052	0.043	0.234	0.28	-0.034	0.137
	CpG	-	CbG	0.056	0.043	0.196	0.31	-0.030	0.142
	InG	-	CbG	0.004	0.043	0.918	0.03	-0.081	0.090

Note. InG = Input group, CbG = Combined group, CpG = Comparison group.

[log10] Values represent Log10 transformation.

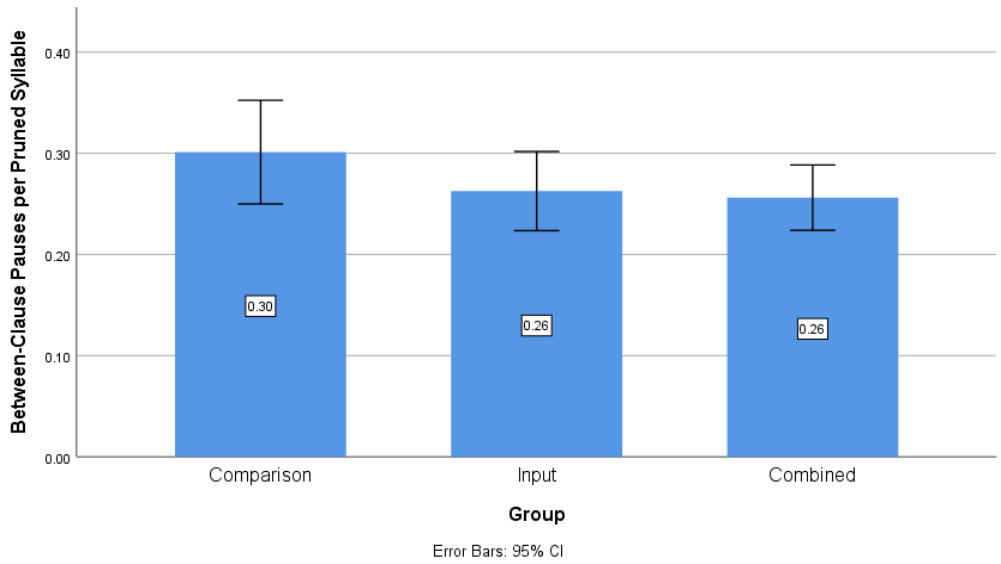
*The mean difference is significant at the .05 level.

** Cohen (1988) recommends *d* = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

Figure 7.2 displays untransformed means of frequency of between-clause pauses for groups on the Same-Task Post-Test with 95% confidence interval of the means. It can be seen that following the treatment, both experimental groups had comparably lower frequencies of between-clause pauses than the Comparison group.

Figure 7.2

Frequency of Between-Clause Pauses for Groups on the Same-Task Post-Test



7.2.1.3 Between-Groups Differences in Within-Clause Pausing on the Same-Task Post-Test

Similar to between-clause pausing, LSD post-hoc comparisons in the frequency of within-clause pauses on the Same-Task Post-Test indicate no statistically significant difference between any of the three groups (see Table 7.5). However, the Comparison group paused more in the middle of clauses than the two experimental groups. The differences reached small effect sizes in both cases, Comparison vs. Input (p = .089, d = .40) and Comparison vs. Combined (p = .119, d = .37). The effect sizes were null between the Input and the Combined group (p = .887, d = .04). From these results, it can be inferred that the treatments did not affect learners' encoding of language during post-test performances.

Table 7.5

Groups Comparisons in Within-Clause Pausing on the Same-Task Post-Test

Dependent variable	Group comparison	Mean difference	Std. error	Sig.*	Cohen's *d***	95% CI for difference	
						Lower bound	Upper bound
Within-clause pausing[sr]	CpG - InG	0.063	0.037	0.089	0.40	-0.010	0.135
	CpG - CbG	0.058	0.037	0.119	0.37	-0.015	0.130
	InG - CbG	-0.005	0.037	0.887	0.04	-0.078	0.067

Note. InG = Input group, CbG = Combined group, CpG = Comparison group.

[sr] Values represent square root transformation.

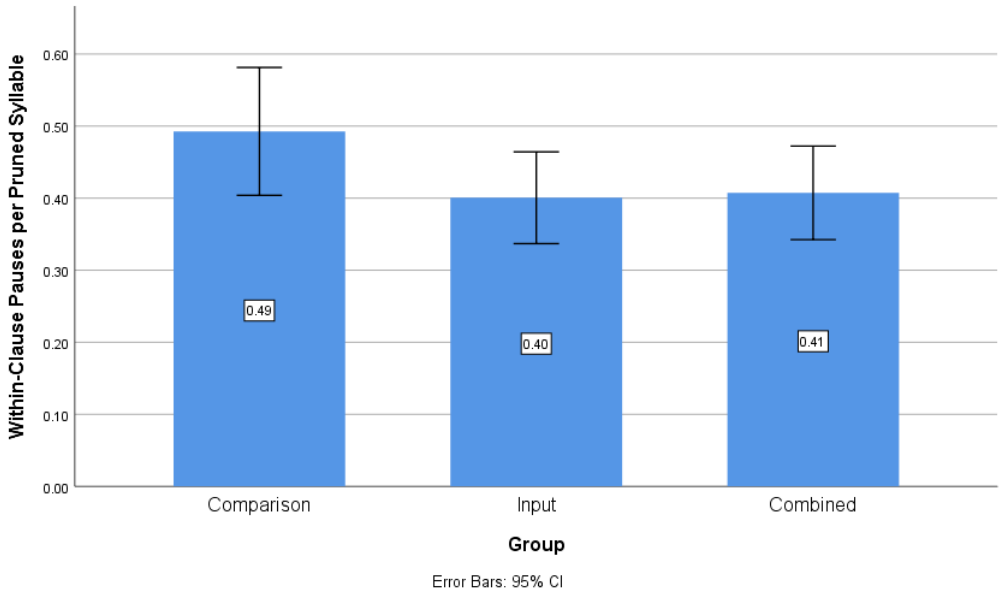
*The mean difference is significant at the .05 level.

** Cohen (1988) recommends d = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

Figure 7.3 illustrates the differences between groups in within-clause pausing during the performance of a repeated task, based on untransformed means. The Comparison group had the highest frequency of within-clause pauses of the three groups, whereas the scores of the two experimental groups were comparable.

Figure 7.3

Frequency of Within-Clause Pauses for Groups on the Same-Task Post-Test



7.2.1.4 Between-Groups Differences on Self-Repairs on the Same-Task Post-Test

Regarding the frequency of self-repairs (replacements and reformulations), post-hoc comparisons of group means on the Same-Task Post-Test reveal that the Comparison group was significantly different from the Input group with medium effect sizes (p = .002, d = .76), and from the Combined group with large effect sizes (p = .000, d = .1.06) (see Table 7.6). Meanwhile, the differences between two experimental groups were not statistically significant with a null effect size (p = .413, d = .19). Thus, task-based instruction, both the Input and Output conditions, appear to lead to a reduction in monitoring in comparison to non-task-based instruction.

Table 7.6

Groups Comparisons in Self-Repairs on the Same-Task Post-Test

Dependent variable	Group comparison	Mean difference	Std. error	Sig.*	Cohen's *d***	95% CI for difference	
						Lower bound	Upper bound
Repairs[sr]	CpG - InG	0.047	0.015	0.002*	0.76	0.019	0.076
	CpG - CbG	0.059	0.015	0.000*	1.06	0.030	0.088
	InG - CbG	0.012	0.015	0.413	0.19	-0.017	0.041

Note. InG = Input group, CbG = Combined group, CpG = Comparison group.

[sr] Values represent square root transformation.

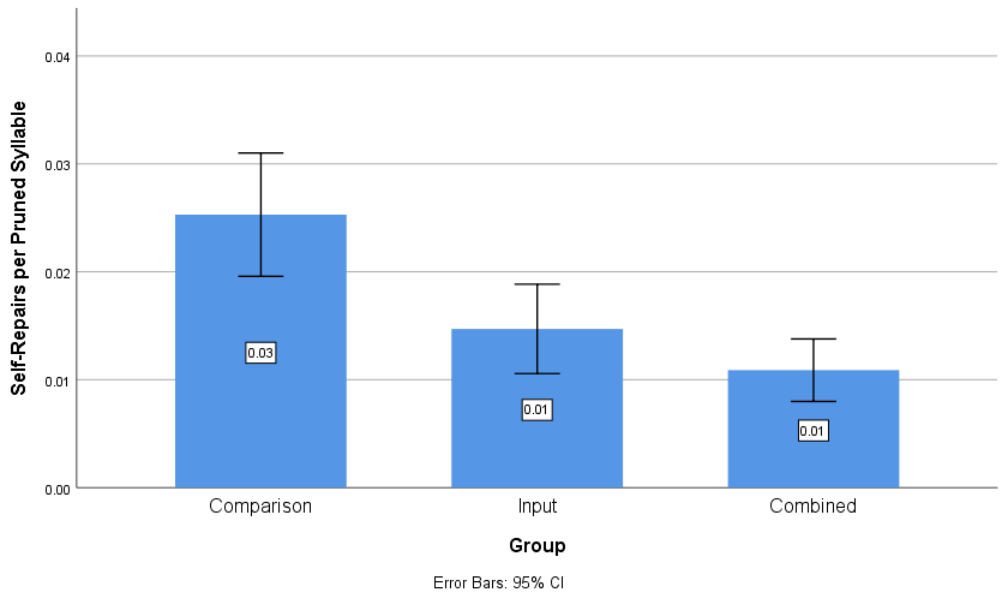
*The mean difference is significant at the .05 level.

** Cohen (1988) recommends *d* = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

The differences between groups in the frequency of self-repairs are illustrated in Figure 7.4, based on untransformed means. The figure shows that both the Input group and Combined group had lower frequencies of self-repairs than the Comparison group in the performance of a repeated task.

Figure 7.4.

Frequency of Self-Repairs for Groups on the Same-Task Post-Test



In summary, with regard to Research Question 1, the results show improvement in speech production capacity for both task-based instructional groups compared to the Comparison group when they performed a repeated task. Specifically, there were significant differences between the Combined group and the Comparison group in terms of speech rate (overall fluency) with a medium effect size and self-repair frequency (monitoring) with a large effect size on the Same-Task Post-Test. The Input group was also significantly different from the Comparison group with a medium effect size in respect of the frequency of self-repairs. Regarding between-clause pausing and within-clause pausing, both experimental groups exhibited positive trends with lower frequencies of between-clause pauses and within-clause pauses than the Comparison group, but not significantly so. There was no significant difference between the two experimental groups in any of the speech production measures.

7.2.2 Between-Groups Differences on the Parallel-Task Post-Test

As shown in Table 7.7, the univariate results indicate that when performing the Parallel-Task Post-Test, the three groups were not significantly different on any of the four dependent variables: speech rate, $F(2,99) = .613$, $p = .544$, $\eta_p^2 = .012$; frequency of between-clause pauses, $F(2,99) = 1.776$, $p = .175$, $\eta_p^2 = .035$; frequency of within-clause pauses, $F(2,99) = 1.176$, $p = .313$, $\eta_p^2 = .023$; and frequency of self-repairs, $F(2,99) = 2.904$, $p = .059$, $\eta_p^2 = .055$.

Table 7.7

Between-Group Differences in Speech Production Measures on the Parallel-Task Post-Test

Variable	Df	Mean square	F	Sig.	Partial eta squared
SR (sr)	2	0.015	0.613	0.544	0.012
BP_freq (lg10)	2	0.047	1.776	0.175	0.035
WP_freq (sr)	2	0.028	1.176	0.313	0.023
Rp_freq (sr)	2	0.012	2.904	0.059	0.055

Note. SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs, sr = square root transformation, lg10 = log10 transformation.

Follow-up post-hoc pairwise comparisons were examined using Fischer's LSD tests for speech rate, between-clause pausing, and within-clause pausing together, whereas Games-Howell tests were used for frequency of self-repairs because the assumption of homogeneity of variances was violated for this measure. Finally, Cohen's *d* values were calculated for each comparison between means as shown below.

7.2.2.1 Between-Groups Differences in Speech Rate on the Parallel-Task Post-Test

As can be seen from Table 7.8, post-hoc pairwise comparisons show no statistically significant differences in speech rate between any of the groups on the Parallel-Task Post-Test. However, there was a positive trend when the learners in both experimental groups maintained a higher speech rate than the Comparison group in the performance of a parallel task and did so with small effect sizes (Comparison vs. Input, $p = .320$, $d = .24$; Comparison vs. Output, $p = .363$, $d = .22$). However, the effect sizes were nil between the Input and the Combined group ($p = .933$, $d = .02$). Based on the magnitude of the effect sizes, it appears that there was no meaningful transfer of practice effects from either input-based or output-based task instruction to the performance of a parallel task in the context in which the study was conducted.

Table 7.8

Groups Comparisons in Speech Rate on the Parallel-Task Post-Test

Dependent variable	Group comparison	Mean difference	Std. error	Sig.*	Cohen's *d***	95% CI for difference	
						Lower bound	Upper bound
Speech rate [sr]	CpG - InG	-0.038	0.038	0.320	0.24	-0.113	0.037
	CpG - CbG	-0.035	0.038	0.363	0.22	-0.110	0.041
	InG - CbG	0.003	0.038	0.933	0.02	-0.072	0.078

Note. InG = Input group, CbG = Combined group, CpG = Comparison group.

[sr] Values represent square root transformation.

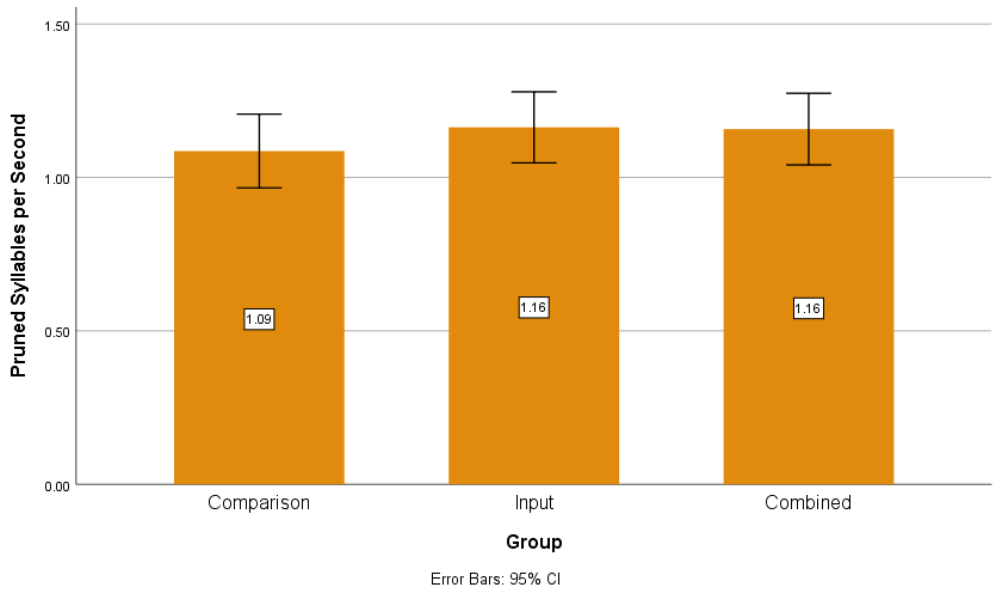
*The mean difference is significant at the .05 level.

**Cohen (1988) recommends *d* = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

Figure 7.5 presents the differences of the three groups on the Parallel-Task Post-Test, based on untransformed means. It can be seen that during the performance of a parallel task, the two experimental groups showed comparable speech rates that were slightly higher than the Comparison group, but not significantly so.

Figure 7.5

Speech Rate for Groups on the Parallel-Task Post-Test



7.2.2.2 Between-Groups Differences in Between-Clause Pausing on the Parallel-Task Post-Test

The results for post-hoc pairwise comparisons show that the differences in the frequency of between-clause pauses during the parallel-task performance did not reach statistical significance for any pairs in the groups (see Table 7.9). However, it can be seen from Table 7.9 that the students in the Comparison group tended to pause more between clauses than those in the Input and the Combined group with small effects sizes (*d* = .39, .40, respectively). The difference between the two experimental groups was negligible (*d* = .03). The pattern was similar to that of the Same-Task Post-Test. However, all three groups showed even lower frequencies of between-clause pauses on the Parallel-Task Post-Test than on the Same-Task Post-Test. This indicates that the treatments had non-significant, but positive effects on conceptualization when similar content was used.

Table 7.9

Groups Comparisons in Between-Clause Pausing on the Parallel-Task Post-Test

Dependent variable	Group comparison			Mean difference	Std. error	Sig.*	Cohen's *d***	95% CI for difference	
								Lower bound	Upper bound
Between-clause pausing ^{log10}	CpG	-	InG	0.067	0.039	0.093	0.39	-0.011	0.145
	CpG	-	CbG	0.061	0.039	0.122	0.40	-0.017	0.139
	InG	-	CbG	-0.006	0.039	0.89	0.03	-0.084	0.073

Note. InG = Input group, CbG = Combined group, CpG = Comparison group.

^{log10} Values represent Log10 transformation.

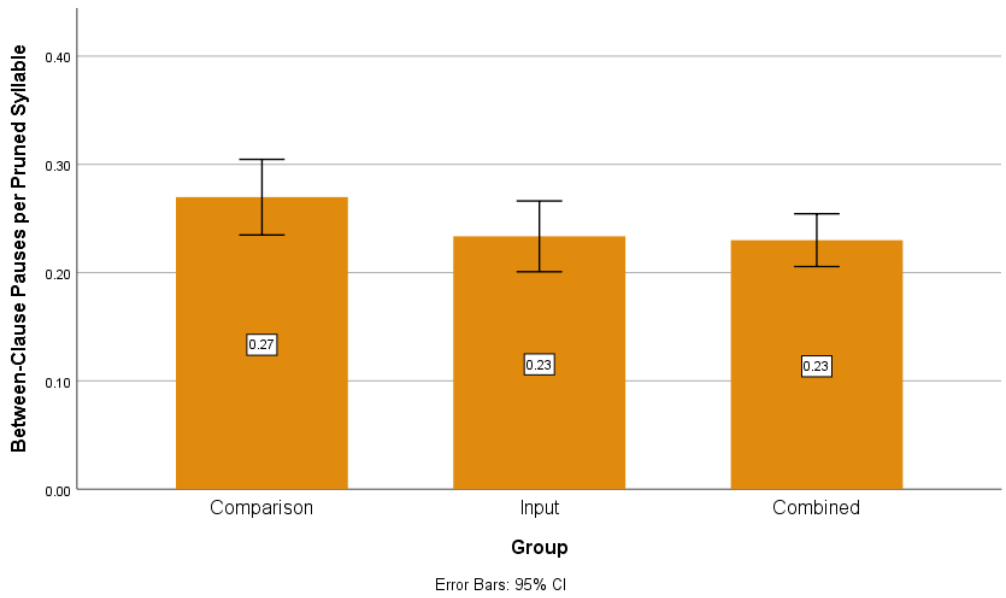
*The mean difference is significant at the .05 level.

**Cohen (1988) recommends *d* = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

The differences between groups are displayed in Figure 7.6, based on untransformed means. As shown in the figure, in the performance of a parallel task, the Comparison group had the highest rate of between-clause pausing, with the two experimental groups pausing between clauses at lower comparable rates.

Figure 7.6

Frequency of Between-Clause Pauses for Groups on the Parallel-Task Post-Test



7.2.2.3 Between-Groups Differences in Within-Clause Pausing on the Parallel-Task Post-Test

The results for post-hoc comparisons in the frequency of within-clause pauses on the Parallel-Task Post-Test, as presented in Table 7.10, reveal that there were no statistically significant differences between the three groups. However, the Comparison group had a nominally higher frequency of within-clause pauses than the two experimental groups. The difference between the Comparison group and the Input group was small (p = .133, d = .37) while the difference was negligible between the Comparison group and the Combined group (p = .577, d = .13). Interestingly, the Input group paused less than the Combined group on the Parallel-Task Post-Test with small effect sizes (p = .341, d = .23), while they paused at almost the same frequency on the Same-Task Post-Test. From these results, it can be assumed that the treatments did not significantly affect the learners' formulation (encoding) in the performance of a parallel task.

Table 7.10

Groups Comparisons in Within-Clause Pausing on the Parallel-Task Post-Test

Dependent variable	Group comparison	Mean difference	Std. error	Sig.*	Cohen's *d***	95% CI for difference	
						Lower bound	Upper bound
Within-clause pausing [sr]	CpG - InG	0.056	0.037	0.133	0.37	-0.017	0.130
	CpG - CbG	0.021	0.037	0.577	0.13	-0.053	0.094
	InG - CbG	-0.036	0.037	0.341	0.23	-0.109	0.038

Note. InG = Input group, CbG = Combined group, CpG = Comparison group.

[sr] Values represent square root transformation.

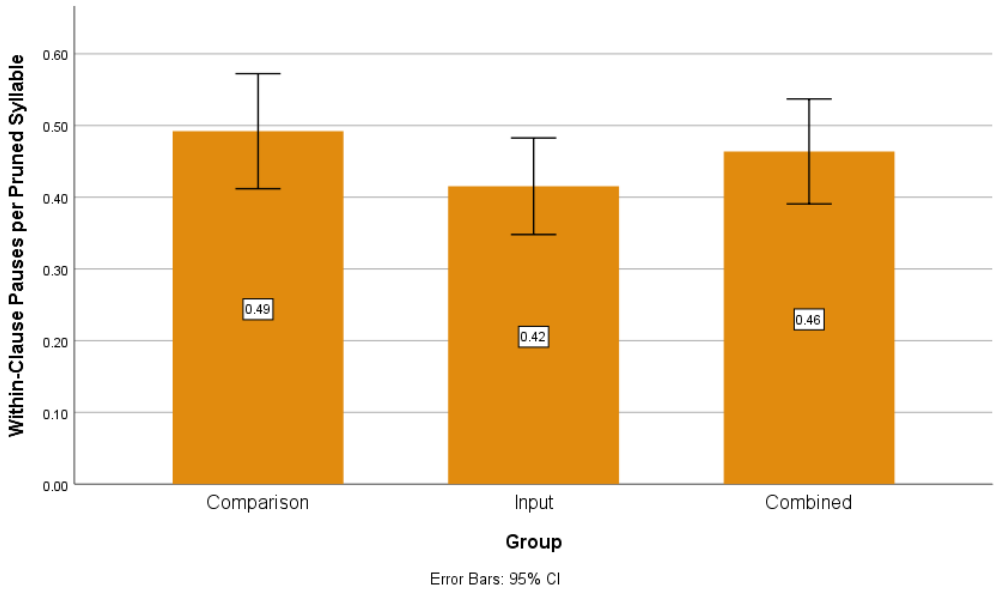
*The mean difference is significant at the .05 level.

**Cohen (1988) recommends d = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

Figure 7.7 illustrates these differences between groups for within-clause pausing based on untransformed means. It can be seen that when performing a parallel task, the Comparison group paused within clauses the most frequently of the three groups, while the Input group had the lowest scores.

Figure 7.7

Frequency of Within-Clause Pauses for Groups on the Parallel-Task Post-Test



7.2.2.4 Between-Groups Differences in Self-Repairs on the Parallel-Task Post-Test

While there was significant difference in the frequency of self-repairs on the Same-Task Post-Test, the results of Games-Howell post-hoc comparisons of group means in this measure revealed no statistically significant difference between any groups on the Parallel-Task Post-Test (see Table 7.11). However, the Cohen's *d* values did show that the Comparison group repaired more than the Input group and the Combined group with small to medium effect sizes (p = .197, d = .46; and p = .062, d = .56, respectively). Meanwhile, the difference between two experimental groups was nil (p = .745, d = .18). It appeared that treatment effects on reducing the need to monitor during task performance did not transfer to the performance of the same task type with new content.

Table 7.11

Groups Comparisons in Self-Repairs on the Parallel-Task Post-Test

Dependent variable	Group comparison	Mean difference	Std. error	Sig.*	Cohen's *d***	95% CI for difference	
						Lower bound	Upper bound
Repairs [sr] (Games-Howell)	CpG - InG	0.025	0.014	0.197	0.42	-0.010	0.060
	CpG - CbG	0.038	0.016	0.062	0.56	-0.002	0.077
	InG - CbG	0.012	0.017	0.745	0.18	-0.028	0.053

Note. InG = Input group, CbG = Combined group, CpG = Comparison group.

[sr] Values represent square root transformation.

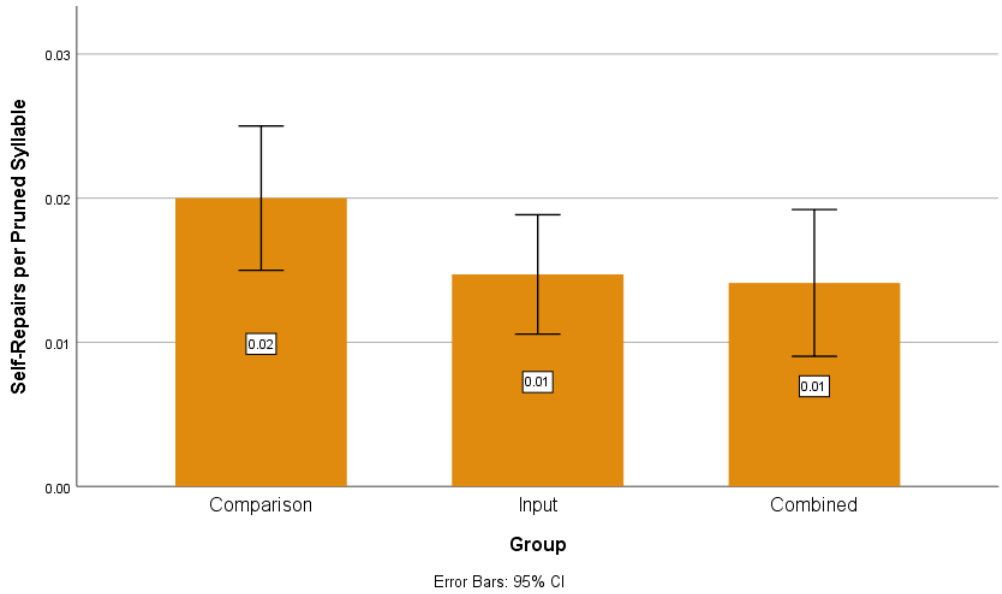
*The mean difference is significant at the .05 level.

**Cohen (1988) recommends *d* = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

The differences between groups in the frequency of self-repairs on the Parallel-Task Post-Test are presented in Figure 7.8 below, based on untransformed means. As shown in the figure, the two experimental groups had a comparable frequency of self-repairs, which was lower than that of the Comparison group, but not significantly so.

Figure 7.8

Frequency of Self-Repairs for Groups on the Parallel-Task Post-Test



In brief, with regard to Research Question 2, the results show no statistically significant differences on any speech production measures between the two task-based instruction groups and the Comparison group, as well as between the two task-based instruction groups on the Parallel-Task Post-Test. However, both experimental groups showed positive trends when performing a parallel task with higher speech rates, fewer between-clause and within-clause pauses, as well as lower rates of self-repairs than the Comparison group.

7.3 Results for Discriminant Function Analysis

The MANOVA was followed by discriminant function analysis to investigate the relationship between the dependent variables in the study and determine which measures were the best predictors of task-based treatment. The discriminant function analysis reveals two functions. The first function explained 87.7% of the variance (canonical $R^2 = .545$), whereas the second explained only 12.3% of the variance between groups (canonical $R^2 = .236$). In combination, these two discriminant functions significantly differentiated the groups, $\Lambda = .663$, $X^2(16) = 39.196$, $p = .001$. However, removing the first function indicated that the second function alone did not significantly differentiate the treatment groups, $\Lambda = .944$, $X^2(7) = 5.492$, $p = .600$. This means that the group differences shown by MANOVA can only be explained by the combination of two underlying dimensions. The canonical variate correlation coefficients for the functions (see Table 7.12) indicate the substantive nature of the eight variables in predicting group membership.

Table 7.12

Structure Matrix

Variable	Test	Function (*r*)*	
		1	2
Frequency of self-repairs	Same-Task Post-Test	0.642 (L)	0.466 (M)
Speech rate	Same-Task Post-Test	0.424 (M)	0.088
Frequency of self-repairs	Parallel-Task Post-Test	0.369 (M)	0.128
Frequency of within-clause pauses	Parallel-Task Post-Test	0.101	0.573 (L)
Frequency of within-clause pauses	Same-Task Post-Test	0.254	0.399 (M)
Frequency of between-clause pauses	Parallel-Task Post-Test	0.251	0.394 (M)
Speech rate	Parallel-Task Post-Test	0.147	0.233
Frequency of between-clause pauses	Same-Task Post-Test	0.207	0.225

Note. *Cohen (1988) recommends benchmarks of 0.10, 0.30, and 0.50 for interpreting *r* values as small, medium, and large effects respectively.

Table 7.12 shows predictor variables ordered by absolute size within the primary function. The correlations between outcomes and the discriminant functions indicated that the best predictor of group membership was frequency of self-repairs on the Same-Task Post-Test. It loaded highly onto both functions, with a large effect size on the first function (r = .642) and the medium effect size on the second (r = .466). The second important predictor was speech rate on the Same-Task Post-Test, which loaded with a medium effect size on the first function (r = .424), but had negligible impact on the second (r = .088). The third important discriminating variable was frequency of self-repairs on the Parallel-Task Post-Test, which also loaded with a medium effect size on the first function (r = .369) and a small effect on the second (r = .128).

Besides frequency of self-repairs on the Same-Task Post-Test, the variable that loaded with a medium effect size on the second function (r = .573), but had a small effect size on the first function (r = .101) was the frequency of within-clause pauses on the Parallel-Task Post-Test. The frequency of within-clause pauses on the Same-Task Post-Test also

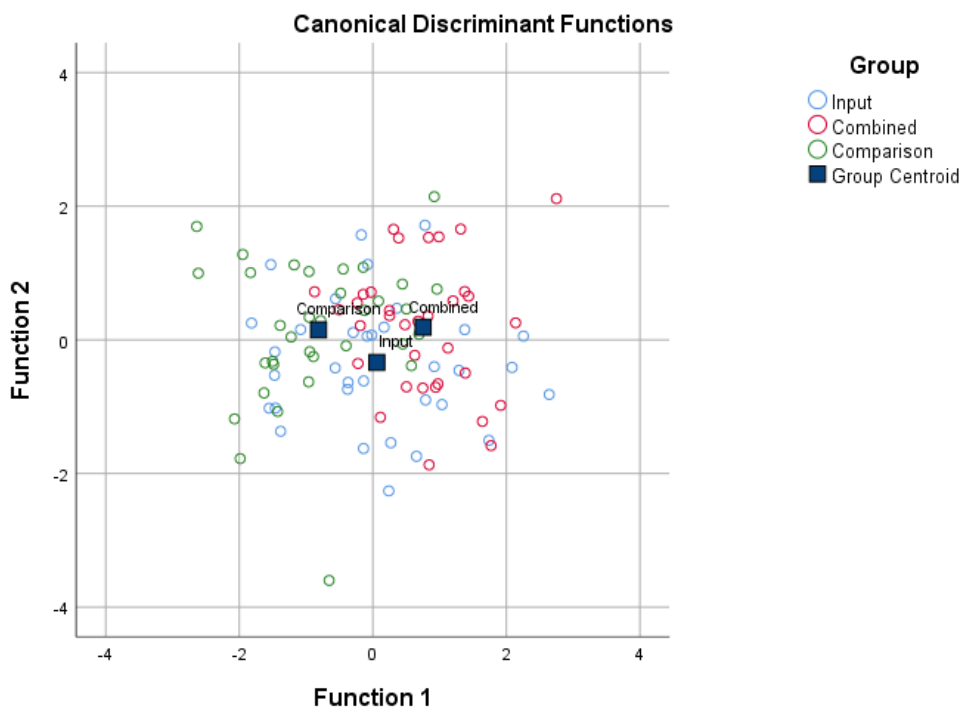
loaded highly with a medium effect size on the second function (*r* = .399) and with a small effect size on the first function (*r* = .254). Similarly, the frequency of between-clause pauses on the Parallel-Task Post-Test loaded with a medium effect size on the second function (*r* = .394) and with a small effect size on the first function (*r* = .251). Speech rate on the Parallel-Task Post-Test and frequency of between-clause pauses on the Same-Task Post-Test contributed the least to group separation (*r* = .147, *r* =.207 for the first function and *r* = .233, *r* = .225 for the second function respectively).

Because the first function explained most of the variance (87.7%, canonical R^2 = .545), it can be concluded that with regard to Research Question 3, the best predictors of group membership were frequency of self-repairs on the Same-Task Post-Test, followed by speech rate on the Same-Task Post-Test and frequency of self-repairs on the Parallel-Task Post-Test.

The group centroids and the discriminant function plot (Figure 7.9) showed that the first function discriminated the Comparison group from the two experimental groups, and the second function differentiated the Input group from the Output and the Comparison groups.

Figure 7.9

Discriminant Function Plot



Specifically the classification results showed accurate predictability of 59.8% for the two-function model. Of the 34 members of the Input group, 12 (35.3%) were correctly classified, nine (26.5%) appeared to be aligned with Combined group members, and 13 (38.2%) as Comparison group members. Of the 34 Combined group members, 24 (70.6%) were grouped correctly with four (11.8%) aligning with Input group members, and six (17.6%) with Comparison group members. Of the 34 Comparison group members, 25 (73.5%) were correctly classified, while two (5.9%) appeared to be better classified as Input group members, and seven (20.6%) as Combined group members. These results showed that the Output and the Comparison groups were the most accurately classified.

In brief, in response to Research Question 3, the results reveal that the group membership was best predicted by the frequency of self-repairs on the Same-Task Post-Test, followed by speech rate on the Same-Task Post-Test and frequency of self-repairs on the Parallel-Task Post-Test. The discriminant analysis suggests that the group separation could be best explained by the combination of two functions (two underlying dimensions). The two treatment groups can be discriminated from the Comparison group on Function 1. Membership in the control group as opposed to the treatment groups was best predicted by a lower speech rate on the post-tests as well as more between-clause pausing, within-clause pausing, and self-repairs.

7.4 Summary

In this chapter, the results of the differences between groups after treatment have been presented. The three groups were comparable in their speech production capacity at the time of the pre-test, and the differences between groups on the post-tests were attributed to the treatments. The results of the one-way MANOVAs reported in this chapter show the between-subjects effects of different task-based treatments on each post-test regarding learners' speech production capacity as measured by four dependent variables: speech rate (overall fluency), between-clause pausing (conceptualization), within-clause pausing (formulation), and self-repairs (monitoring). Table 7.13 summarizes all the differences between groups on the two post-tests by variables.

Table 7.13

Summary of Between-Groups Differences on Post-Tests

Variables	Same-Task Post-Test	Parallel-Task Post-Test
Speech rate	*Significant* InG > CpG (S) CbG > CpG (M)* InG < CbG (S)	*Not significant*
Between-clause pausing	*Not significant*	*Not significant*
Within-clause pausing	*Not significant*	*Not significant*
Repairs	*Significant* InG < CpG (M)* CbG < CpG (L)* InG = CbG (N)	*Not significant*

Note. InG = Input group, CbG = Combined group, CpG = Comparison group
N = Nil effect sizes, S = Small effect sizes, M = Medium effect sizes, L = Large effect sizes
*The mean difference is significant at the .05 level.

As noted earlier in this chapter, with regard to Research Question 1, the results of the one-way MANOVA show that there were significant differences between both task-based instructional groups and the Comparison group in terms of self-repairs on the Same-Task Post-Test. In addition, there were significant differences between the Combined group and the Comparison group in speech rate on the Same-Task Post-Test. The differences between the Input group and Combined group in all speech production measures on the Same-Task Post-Test were not significant.

In response to Research Question 2, there was no significant difference between any groups in any speech production measure on the Parallel-Task Post-Test.

The followed-up discriminant function analysis reveals the answer to Research Question 3. The frequency of self-repairs and speech rate on the Same-Task Post-Test, followed by frequency of self-repairs on the Parallel-Task Post-Test were the best predictors of group membership in the comparison group as opposed to the treatment groups.

Overall, the analyses reveal that different task-based instruction options had differential effects on learners' speech production capacity during the performance of a repeated task. The treatment effects, however, did not carry on to the performance of a

parallel task. The following chapter will discuss the changes within each group from the pre-test to the post-tests.

Chapter 8. Results for Within-Group Differences

This chapter presents the results for the changes within each group between the pre-test and the post-tests. The chapter first reports the results of the repeated-measures MANOVA for the Comparison group, followed by the Input group and the Combined group. The three research questions that this chapter addresses are:

RQ4. How does learners' speech production capacity change over time in the performance of a repeated task and a parallel task when they are involved in **narrow input instruction**?

RQ5. How does learners' speech production capacity change over time in the performance of a repeated task and a parallel task when they are involved in **input-based task instruction**?

RQ6. How does learners' speech production capacity change over time in the performance of a repeated task and a parallel task when they are involved in **combined input-output task-based instruction**?

8.1 Descriptive Statistics

Table 8.1 shows the descriptive statistics for each group across three tests in terms of the four speech production measures based on untransformed means (see Appendix M for full descriptive statistics of transformed means).

Table 8.1

*Descriptive Statistics for Speech Production Measures by Groups**

Group	Variable	Pre-test		Same-Task Post-Test		Parallel-Task Post-Test	
		Mean	SD	Mean	SD	Mean	SD
Comparison group	SR	1.09	0.34	1.08	0.33	1.09	0.34
	BP_freq	0.29	0.13	0.30	0.15	0.27	0.10
	WP_freq	0.45	0.22	0.49	0.25	0.49	0.23
	Rp_freq	0.02	0.02	0.03	0.02	0.02	0.01
Input group	SR	1.14	0.41	1.21	0.30	1.16	0.33
	BP_freq	0.28	0.13	0.26	0.11	0.23	0.09
	WP_freq	0.42	0.22	0.40	0.18	0.42	0.19
	Rp_freq	0.01	0.01	0.01	0.01	0.01	0.01
Combined group	SR	1.14	0.53	1.29	0.32	1.16	0.33
	BP_freq	0.30	0.13	0.26	0.09	0.23	0.07
	WP_freq	0.51	0.31	0.41	0.19	0.46	0.21
	Rp_freq	0.02	0.02	0.01	0.01	0.01	0.01

Note. SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs

*Values represent untransformed means.

As shown in Table 8.1, the Comparison group performed similarly across three tests with regard to speech rate, between-clause pausing and self-repairs. The frequencies of within-clause pauses on both post-tests were similar and higher than on the Pre-Test. The Input group performed better on the Same-Task Post-Test than on the Pre-Test with higher speech rate, and, lower frequency of both between-clause and within-clause pauses. However, when they performed a parallel task on the Parallel-Task Post-Test, the scores for speech rate decreased while the frequency of within-clause pauses increased to the same point as on the pre-test, and only between-clause pausing continued the declining trend. Frequency of self-repairs remained the same over three tests. Similarly, in the Combined group, the participants performed better on both post-tests than on the Pre-Test with higher speech rate, fewer between-clause and within-clause pauses, and fewer self-repairs.

However, the scores for the Same-Task Post-Test reflected better performance than the Parallel-Task Post-Test with regard to speech rate, between-clause and within-clause pausing. Frequency of self-repairs for the Combined group stayed stable on both post-tests. These trends were examined statistically and the results are presented in the following sections.

In order to specify whether different task-based instruction approaches significantly affected learners' speech production capacity over time, a repeated-measures MANOVA was performed on split-file data to examine the differences between the three test points within each group on each measure. Separate results for each group are presented below in response to Research Questions 4, 5, and 6.

8.2 Changes over Time for the Comparison Group

The results of Mauchly's Test of Sphericity show that the assumption of sphericity was met only for self-repairs, $\chi^2(2) = 1.070$, $p = .586$), but not for three other measures of fluency ($p > .05$): speech rate, $\chi^2(2) = 9.296$, $p = .010$); between-clause pausing, $\chi^2(2) = 7.141$, $p = .0028$); within-clause pausing, $\chi^2(2) = 8.218$, $p = .016$). Therefore, sphericity was assumed for self-repairs while Huynh-Feldt correction was used to interpret within-subject effects for speech rate, between-clause pausing, and within-clause pausing. The multivariate results for the Comparison group using Pillai's trace show that there was no statistically significant difference between the three tests, $F(8, 128) = 4.799$, $p = .070$, $\eta_p^2 = .105$). The results of univariate tests and post-hoc pairwise comparisons and Cohen's *d* values revealed more details on the differences across three tests on each measure as follows.

8.2.1 Speech Rate for the Comparison Group

The univariate test results using Huynh-Feldt correction indicate no significant difference in speech rate scores among three tests of the Comparison group, $F(1.67, 54.97) = .010$, $p = .980$, $\eta_p^2 = .000$. LSD post-hoc pairwise comparisons on speech rate, as shown in Table 8.2, also showed no significant difference with nil effect sizes for all pairs of test, (Pre-Test vs. Same-Task Post-Test, $p = .897$, $d = .02$; Pre-Test vs. Parallel-Task Post-Test, $p = .962$, $d = .01$; Same-Task Post-Test vs. Parallel-Task Post-Test, $p = .904$, $d = .01$). From these results, it is evident that the students in the Comparison group performed similarly across three tests in terms of speech rate.

Table 8.2

Speech Rate for the Comparison Group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Speech rate [sr]	1 - 2	0.003	0.024	0.897	0.02	-0.046	0.052
	1 - 3	0.001	0.025	0.962	0.01	-0.051	0.053
	2 - 3	-0.002	0.016	0.904	0.01	-0.034	0.03

Note. [sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d* = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.2.2 Between-Clause Pausing for the Comparison Group

The results of univariate tests using Huynh-Feldt show that there was no statistically significant difference between tests for between-clause pausing, F (1.67, 57.57) = 1.089, p = .336, η_p^2 = .032. LSD post-hoc pairwise comparisons further confirm no difference between pairs of tests, as shown in Table 8.3 (Pre-Test vs. Same-Task Post-Test, p = .773, d = .04; Pre-test vs. Parallel-Task Post-Test, p =.311, d= .15; Same-Task Post-Test vs. Parallel-Task Post-Test, p=.071, d= .19). The effect sizes in all cases were minimal. This indicates that participants in the Comparison group paused between clauses for conceptualizing ideas at the same rate across three tests.

Table 8.3

Between-Clause Pausing for the Comparison Group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Between-clause pausing ^{log10}	1 - 2	-0.008	0.028	0.773	0.04	-0.065	0.049
	1 - 3	0.026	0.025	0.311	0.15	-0.026	0.078
	2 - 3	0.034	0.018	0.071	0.19	-0.003	0.072

Note. ^{log10} Values represent Log10 transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d* = 0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.2.3 Within-Clause Pausing for the Comparison Group

The univariate test results using Huynh-Feldt indicate that there was no statistically significant difference between any of the three tests regarding within-clause pausing, F (1.70, 56.22) = 1.432, p = .247, η_p^2 = .042. LSD post-hoc pairwise comparisons, as shown in Table 8.4, show the same results with all test scores not being significantly different from each other, with negligible effect sizes (Pre-Test and Same-Task Post-Test, p = .239, d = .19; Same-Task Post-Test vs. Parallel-Task Post-Test, p=.839, d= .02), or with only small effect sizes (Pre-Test vs. Parallel-Task Post-Test, p =.158, d= .022). It seems that learners in the Comparison group made no significant improvement in terms of their formulation ability in the performance of the same or parallel task during the post-tests.

Table 8.4

Within-Clause Pausing for the Comparison Group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Within-clause pausing [sr]	1 - 2	-0.031	0.026	0.239	0.19	-0.084	0.022
	1 - 3	-0.035	0.024	0.158	0.22	-0.084	0.014
	2 - 3	-0.003	0.017	0.839	0.02	-0.037	0.03

Note. [sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d*=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.2.4 Self-Repairs for the Comparison Group

With Sphericity assumed the univariate test results show that there was no significant difference among three tests in terms of self-repairs, F (2, 66) = 2.734, p = .072, η_p^2 = .077. However, LSD post-hoc pairwise comparisons (see Table 8.5) together with Cohen's *d* values indicate that the difference between the Pre-Test and the Same-Task Post-Test was significant with a medium effect size (p = .026, d = .52). The differences between the other pairs were not significant, with small effect sizes (Pre-Test vs. Parallel-Task Post-Test, p =.291, d= .26; Same-Task Post-Test vs. Parallel-Task Post-Test, p =.205, d= .33). In other words, the participants in the Comparison group had higher frequency of self-repairs in both post-tests than in the Pre-Test, but only the increase between the Pre-Test and the Same-Task Post-Test was significant. This means that without task-based instruction, learners could produce even more reformulations and replacements on the second performance than on the first one, regardless of whether it was a repeated or a parallel task.

Table 8.5

Self-Repairs for the Comparison Group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Repairs [sr]	1 - 2	-.037*	0.016	0.026	0.52	-0.07	-0.005
	1 - 3	-0.018	0.017	0.291	0.26	-0.053	0.016
	2 - 3	0.019	0.015	0.205	0.33	-0.011	0.049

Note. [sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d*=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

In summary, the results of the repeated MANOVA for the Comparison group show that there was no significant effect for time. With regard to Research Question 4, it can be seen that when the learners received narrow reading and listening practice**,** their speech production capacity did not change significantly across three tests in terms of speech rate, between-clause pausing, and within-clause pausing, except for the observed increase in self-repairs.

8.3 Changes over Time for the Input Group

SPSS output of the repeated MANOVA for the Input group show non-significant Mauchly's Test of Sphericity for all four measures of fluency ($p > .05$): speech rate, $\chi^2(2) = 5.420$, $p = .067$); between-clause pausing, $\chi^2(2) = 5.273$, $p = .071$); within-clause pausing, $\chi^2(2) = 2.376$, $p = .305$); and self-repairs, $\chi^2(2) = 1.323$, $p = .516$). This indicates that variances of differences were equal and the assumption of sphericity was met. The multivariate results using Pillai's trace reveal a statistically significant difference for time, $F(8,128) = 2.545$, $p = .013$, $\eta_p^2 = .137$). This means that there were significant differences between the three tests done by the Input group. The results of univariate tests and post-hoc pairwise comparisons, together with Cohen's *d* values provided more insight into the differences among the three tests on each measure as follows.

8.3.1 Speech Rate for the Input Group

The univariate results show no significant difference between the Pre-Test and the two post-tests for speech rate, F (2, 66) = 1.975, p = 1.47, η_p^2 = .056. LSD post-hoc pairwise comparisons on speech rate, as presented in Table 8.6, show that the participants scored the highest in the Same-Task Post-Test, then lower in the Parallel-Task Post-Test. The effect sizes for differences were small between the Pre-Test and the Same-Task Post-Test (p = .086, d = .24), and negligible between other pairs (Pre-Test vs. Parallel-Task Post-Test, p = .397, d = .11; Same-Task Post-Test vs. Parallel-Task Post-Test, p = .179, d = .14). From these results, it appears that the learners' overall fluency as measured by speech rate did not change significantly following input-based task treatment.

Table 8.6

Speech Rate for the Input Group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Speech rate [sr]	1 - 2	-0.039	0.022	0.086	0.24	-0.085	0.006
	1 - 3	-0.018	0.021	0.397	0.11	-0.061	0.025
	2 - 3	0.021	0.016	0.179	0.14	-0.01	0.053

Note. [sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends d=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.3.2 Between-Clause Pausing for the Input Group

The results of univariate tests show that there were significant differences between the three tests with regard to between-clause pausing of the Input group, F(2, 66) = 7.457, p = .001, η_p^2 = .184. However, LSD post-hoc pairwise comparisons, as shown in Table 8.7, reveal that the difference between the Pre-Test and the Same-Task Post-Test was not significant (p = .157, d = .16). At the same time, there were significant decreases in frequencies of between-clause pauses with small effect sizes between the Pre-Test and the

Parallel-Task Post-Test (*p* =.002, *d*= .44), and between the two post-tests (*p*=.006, *d*= .28). This indicates that participants who received input-based task instruction improved on their conceptualization ability through pausing less between clauses in the performance of a repeated task and this gain transferred to the performance of a parallel task.

Table 8.7

Between-Clause Pausing for the Input Group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Between-clause pausing ^{log10}	1 - 2	0.029	0.020	0.157	0.16	-0.012	0.071
	1 - 3	0.079**	0.024	0.002	0.44	0.030	0.128
	2 - 3	0.049**	0.017	0.006	0.28	0.015	0.084

Note. ^{log10} Values represent Log10 transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d*=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.3.3 Within-clause pausing for the Input Group

The univariate results indicate no statistically significant difference between any of the three tests on within-clause pausing for the Input group, F (2, 66) = .371, p = .692, η_p^2 = .011. LSD post-hoc pairwise comparisons, as shown in Table 8.8, confirm this with nil effect sizes for all pairs of group means (Pre-Test vs. Same-Task Post-Test, p = .397, d = .09; Pre-Test vs. Parallel-Task Post-Test, p =.852, d= .02; Same-Task Post-Test vs. Parallel-Task Post-Test, p=.497, d= .07). Thus, it seems that input-based task instruction did not influence the learners' lexical and syntactic encoding, either when they repeated the same task or performed a parallel one.

Table 8.8

Within-Clause Pausing for the Input Group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Within-clause pausing [sr]	1 - 2	0.013	0.015	0.397	0.09	-0.018	0.045
	1 - 3	0.003	0.018	0.852	0.02	-0.033	0.040
	2 - 3	-0.010	0.014	0.497	0.07	-0.039	0.019

Note. [sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d*=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.3.4 Self-Repairs for the Input Group

The results show that there was no significant difference between any of the three tests in terms of self-repairs for the Input group, F (2, 66) = .035, p = .966, η_p^2 = .001. LSD post-hoc pairwise comparisons (see Table 8.9) reveal that the effect size for the difference between the Pre-Test and the Same-Task Post-Test was nil (p = .994, d = .000). Similarly, the difference in frequency of self-repairs in other pairs showed minimal effect sizes (Pre-Test vs. Parallel-Task Post-Test, p =.805, d= .05; Same-Task Post-Test vs. Parallel-Task Post-Test, p=.83, d= .05). Stable scores of the participants across three test points indicate that input-based task instruction did not influence their monitoring capacity as measured by the rate of repairs made to their initially encoded utterances.

Table 8.9

Self-Repairs for the Input Group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Repairs [sr]	1 - 2	0.000	0.015	0.994	0.00	-0.029	0.030
	1 - 3	-0.003	0.013	0.805	-0.05	-0.030	0.023
	2 - 3	-0.003	0.015	0.830	-0.05	-0.035	0.028

Note. [sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d*=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

Overall, the results of the repeated MANOVA for the Input group show a statistically significant time effect. However, the source of differences only came from the change in between-clause pausing. It seems that the other measures of fluency were not affected by the use of input-based tasks. In response to Research Question 5, it can be said that input-based task instruction had positive effects on learners' conceptualization, but did not significantly influence learners' ability to produce L2 speech in terms of overall fluency, formulation, and monitoring.

8.4 Changes over Time for the Combined group

The results of Mauchly's Test of Sphericity indicate that the assumption of sphericity was met for self-repairs, $\chi^2(2) = 4.723$, $p = .094$), but violated for three other measures of fluency ($p > .05$): speech rate, $\chi^2(2) = 20.685$, $p = .000$); between-clause pausing, $\chi^2(2) = 9.718$, $p = .008$); within-clause pausing, $\chi^2(2) = 11.343$, $p = .003$). Therefore, sphericity was assumed for self-repairs while a Greenhouse-Geisser correction was used for interpreting within-subject effects for speech rate, and Huynh-Feldt correction was used for between-clause pausing, and within-clause pausing.

The multivariate results for the Combined group using Pillai's trace show a statistically significant difference between the three tests, $F(8,128) = 4.799$, $p = .000$, $\eta_p^2 =$

.231). The results of univariate tests and post-hoc pairwise comparisons, together with Cohen's *d* values show more details on the differences among the three tests on each measure below.

8.4.1 Speech Rate for the Combined group

The univariate test results using Greenhouse-Geisser correction show that speech rate scores of the Combined group differed significantly across three tests, $F(2, 44.7) = 6.975$, $p = .007$, $\eta_p^2 = .171$. LSD post-hoc pairwise comparisons on speech rate, as presented in Table 8.10, show a significant increase with small effect sizes between the Pre-Test and the Same-Task Post-Test ($p = .004$, $d = .43$). However, the increase was not significant between the Pre-Test and the Parallel-Task Post-Test, and the effect sizes were negligible ($p = .406$, $d = .11$). The difference between the Same-Task Post-Test and the Parallel-Task Post-Test was a significant decrease with small effect sizes ($p = 0$, $d = .41$). From these results, it can be concluded that instruction combining both input-based and output-based tasks improved the learners' speech rate considerably in their repetition of the same task, but the effects declined when the learners performed a parallel task.

Table 8.10

Speech Rate for the Combined group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Speech rate [sr]	1 - 2	-0.084**	0.027	0.004	0.43	-0.140	-0.028
	1 - 3	-0.023	0.027	0.406	0.11	-0.078	0.032
	2 - 3	0.061**	0.013	0.000	0.41	0.034	0.088

Note.[sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d*=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.4.2 Between-Clause Pausing for the Combined group

The results of univariate tests using Huynh-Feldt indicate a statistically significant time effect for the Combined group in terms of between-clause pausing, F (2, 54.5) = 8.660, p = .001, η_p^2 = .208. LSD post-hoc pairwise comparisons further reveal that all the test means were significantly different from each other (see Table 8.11). Specifically, the participants' frequency of between-clause pauses decreased significantly with a small effect size between the Pre-Test and the Same-Task Post-Test (p = .03, d = .32). The participants' scores in the Parallel-Task Post-Test were also significantly lower than in the Pre-test with a medium effect size (p =.002, d= .56), and lower than in the Same-Task Post-Test with a small effect size (p=.024, d= .25). It is interesting that the learners paused less between clauses for conceptualizing ideas when they performed a parallel task than a repeated task. This indicates that participants who were involved in both the combined input-based and output-based task instruction improved considerably on their conceptualization during the performance of a repeated task and were able to transfer the gain to the performance of a parallel task.

Table 8.11

Between-Clause Pausing for the Combined group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Speech rate [sr]	1 - 2	0.058**	0.026	0.030	0.32	0.006	0.11
	1 - 3	0.098**	0.027	0.001	0.56	0.042	0.153
	2 - 3	0.039**	0.017	0.024	0.25	0.006	0.073

Note. [log10] Values represent Log10 transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends d=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.4.3 Within-Clause Pausing for the Combined group

The univariate test results using Huynh-Feldt show that the differences in within-clause pausing among three tests of the Combined group were approaching significance, F (2, 52.83) = 3.241, p = .057, η_p^2 = .089. However, LSD post-hoc pairwise comparisons, as shown in Table 8.12, indicate a significant decrease in the frequency of within-clause pausing between the Pre-Test and the Same-Task Post-Test with a small effect size (p = .036, d = .30). In the Parallel-Task Post-Test, the participants paused within clauses less than in Pre-Test with a minimal effect size (p =.572, d= .008), but the scores were significantly higher than those of the Same-Task Post-Test with a small effects size (p=.011, d= .027). This shows that combined input-output task-based instruction did have positive effects on learners' formulation stage during their repetition of the same task. However, the effects did not remain when they performed a parallel task.

Table 8.12

Within-Clause Pausing for the Combined group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Within-clause pausing [sr]	1 - 2	0.055**	0.025	0.036	0.30	0.004	0.105
	1 - 3	0.014	0.025	0.572	0.08	-0.037	0.066
	2 - 3	-0.040**	0.015	0.011	0.27	-0.071	-0.01

Note. [sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends d=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

8.4.4 Self-Repairs for the Combined group

The univariate test results with sphericity assumed for self-repairs indicate no significant difference between any of the three tests in terms of self-repairs for the Combined group, F (2, 66) = .649, p = .526, η_p^2 = .019. LSD post-hoc pairwise comparisons (see Table 8.13) together with Cohen's d values confirm that the difference between the

tests had nil or minimal effect sizes (Pre-Test vs. Same-Task Post-Test, *p* = .217, *d* = .27; Pre-Test vs. Parallel-Task Post-Test, *p* =.45, *d*= .20; Same-Task Post-Test vs Parallel-Task Post-Test, *p*=.857, *d*= .04). It appears that the combined input-output task-based instruction did not change learners' repairing behaviours to a noticeable degree in the performance of both the same task and a parallel one.

Table 8.13

Self-Repairs for the Combined group

Dependent variable	Test*	Mean difference	Std. error	Sig.**	Cohen's *d****	95% CI for difference	
						Lower bound	Upper bound
Repairs [sr]	1 - 2	0.018	0.014	0.217	0.27	-0.011	0.047
	1 - 3	0.015	0.02	0.450	0.20	-0.025	0.055
	2 - 3	-0.003	0.016	0.857	-0.04	-0.036	0.030

Note. [sr] Values represent square root transformation.

*1 = Pre-Test, 2 = Same-Task Post-Test, 3 = Parallel-Task Post-Test

**The mean difference is significant at the .05 level.

*** Cohen (1988) recommends *d*=0.2, 0.5, and 0.8 as benchmarks for small, medium, and large effect sizes, respectively.

In brief, the results of the repeated MANOVA for the Combined group show a statistically significant effect for time. The differences mainly came from the changes in speech rate on the Same-Task Post-Test, and between-clause pausing on both post-tests. For within-clause pausing and self-repairs, the combination of both input-based tasks and output-based tasks also had positive, but not significant effects. With regard to Research Question 6, the combined input-output task-based instruction significantly affected learners' speech production capacity in terms of overall fluency and conceptualization, but not formulation and monitoring.

8.5 Summary of Within-Group Differences by Variables

The results of the repeated MANOVA show differences across three tests for each group regarding learners' speech production capacity as measured by speech rate (overall fluency), between-clause pausing (conceptualization), within-clause pausing (formulation),

and self-repairs (monitoring). Specifically, learners in the Comparison group, who received narrow input through non-task-based reading and listening practice, did not make any significant change over the three tests. Learners in the Input group, who were involved in input-based tasks, significantly improved in terms of between-clause pausing only. However, learners in the Combined group, who were involved a combination of both input-based and output-based tasks, had significantly positive changes with regard to not only between-clause pausing but also speech rate. Table 8.14 summarizes all the differences across three tests within each group by variables.

Table 8.14

Summary of Within-Group Differences

Variables	Comparison group	Input group	Combined group
Speech rate	*Not significant*	*Not significant*	*Significant* Pre < Post A (S)* Pre < Post B (N) Post A > Post B (S)*
Between-clause pausing	*Not significant*	*Significant* Pre > Post A (N) Pre > Post B (M)* Post A > Post B (S)*	*Significant* Pre > Post A (S)* Pre > Post B (M)* Post A > Post B (S)*
Within-clause pausing	*Not significant*	*Not significant*	*Approaching significance* Pre > Post A (S)*
Repairs	*Not significant*	*Not significant*	*Not significant*

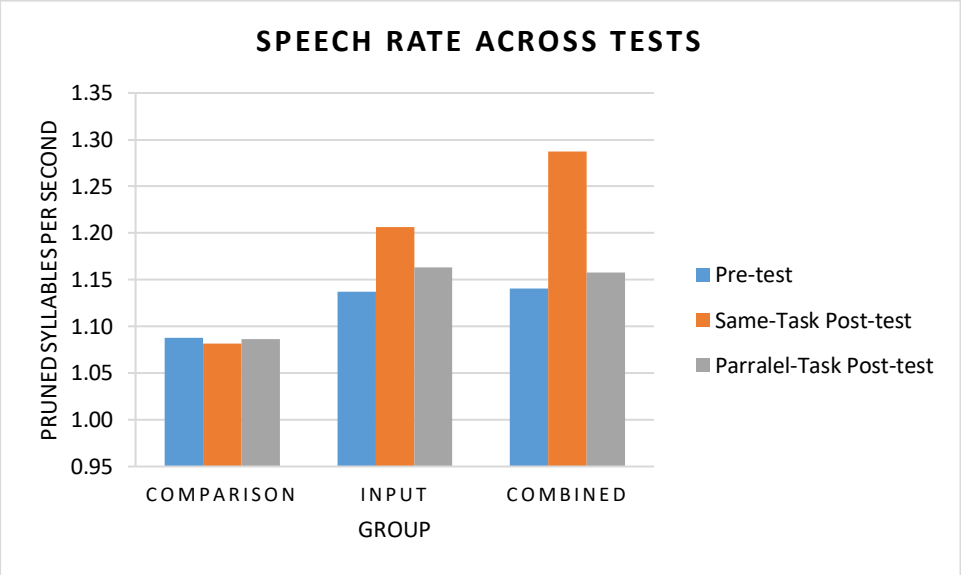
Note. Pre =Pre-Test, Post A = Same-Task Post-Test, Post B = Parallel-Task Post-Test
N = Nil effect sizes, S = Small effect sizes, M = Medium effect sizes, L = Large effect sizes
*The mean difference is significant at the .05 level.

In terms of speech rate, the two task-based instructional groups demonstrated similar patterns of changes across tests. Both Input group and Combined group scored higher in speech rate on the post-tests than on the pre-test, with a bigger increase on the Same-Task Post-Test than on the Parallel-Task Post-Test. However, only the improvement for the Combined group between the Pre-Test and the Same-Task Post-Test was significant whereas the changes for the Input group in both post-tests were not above chance levels at .05. In contrast, the speech rate for the Comparison group remained stable through all three

tests. Figure 8.1, based on untransformed means, illustrates the changes in speech rate across the tests for three groups.

Figure 8.1

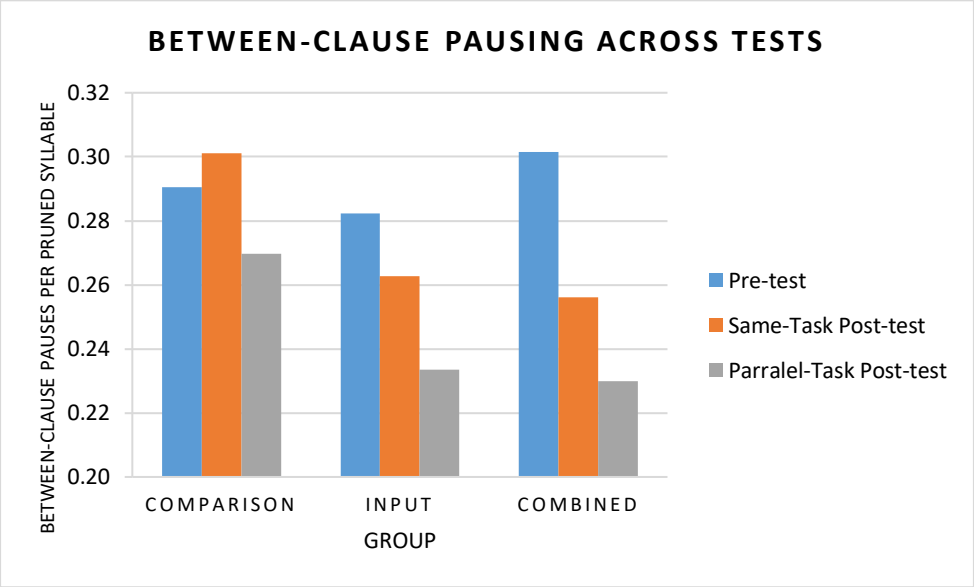
Changes across Tests in Speech Rate



With regard to between-clause pausing, both task-based instructional groups showed improvement following treatment. The Combined group showed significant decreases in both post-tests (the Same-Task-Post-Test and the Parallel-Task Post-Test). The Input group had lower frequencies of between-clause pauses in both post-tests, but only the decrease between the Pre-Test and the Parallel-Task Post-Test was significant. Both groups improved with greater effect sizes in the Parallel-Task Post-Test. The Comparison group had fluctuating frequencies of between-clause pauses across three tests, and the changes did not reach a significant level. These differences for groups over time are illustrated in Figure 8.2, based on the untransformed means. It can be seen that the frequencies of between-clause pauses of both Input group and Combined group decreased in the Same-Task Post-Test, and declined even further in the Parallel-Task Post-Test. This means that both input-based task instruction and combined input-output task-based instruction had positive effects on learners' conceptualization stage during L2 speech production when they repeated the same task, and the gains transferred to the performance of a parallel task.

Figure 8.2

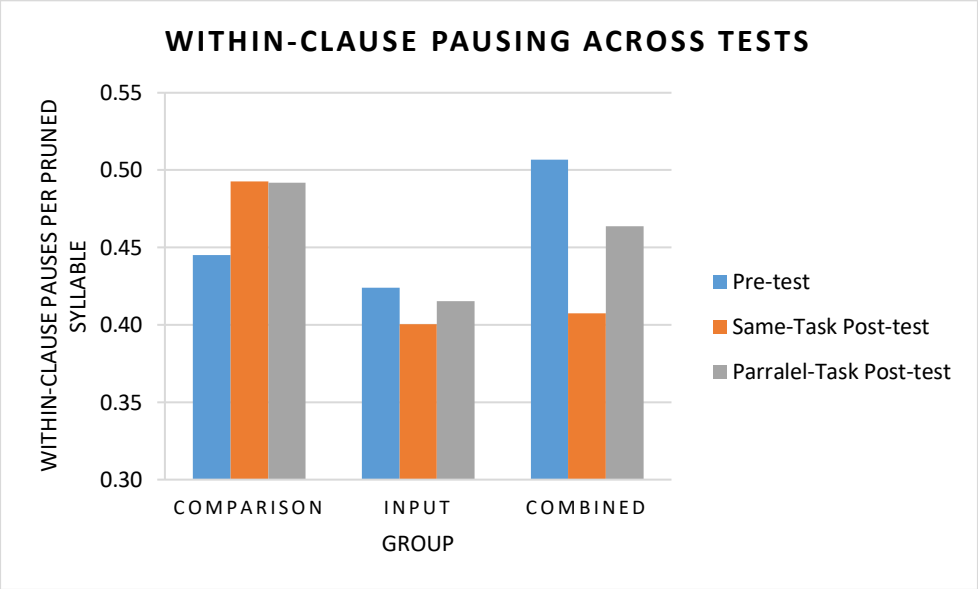
Changes across Tests in Between-Clause Pausing



Regarding within-clause pausing, the patterns of changes were similar for the two task-based instructional groups. Both the Input group and the Combined group had decreased within-clause pausing between the pre-test and both post-tests. However, only the decrease between the Pre-Test and Same-Task Post-Test for the Combined group was significant. In contrast, the Comparison group had similarly higher rates of within-clauses pausing in both post-tests than in the Pre-Test. This means that there was a significant improvement in learners' encoding of a repeated task for the Combined group, but the gains did not transfer to the performance of a parallel task. In contrast, treatments did not significantly change learners' within-clause pausing behaviour for the Input group and the Comparison group. Figure 8.3 visually summarizes the differences for within-clause pausing over time for each group based on untransformed means.

Figure 8.3

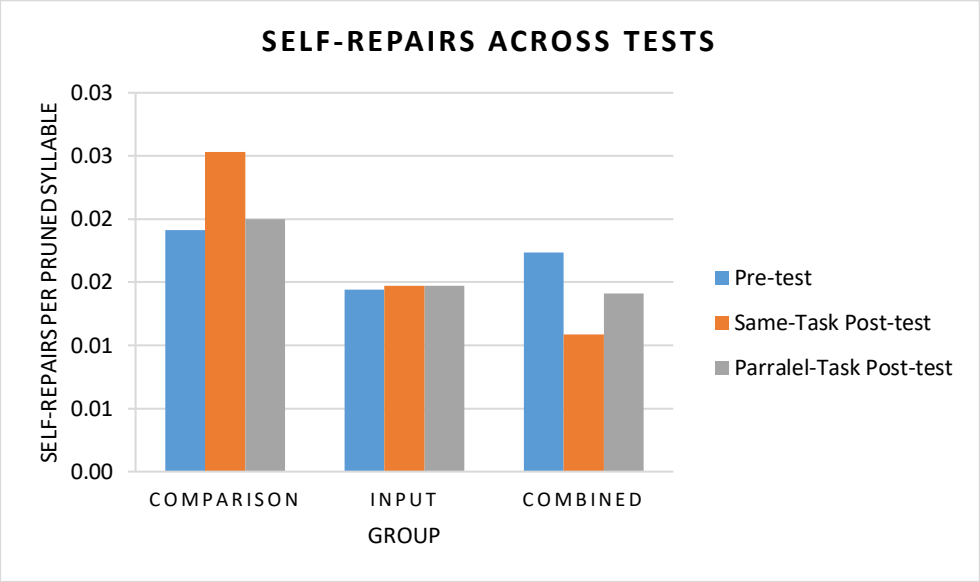
Changes across Tests in Within-Clause Pausing



In terms of self-repairs, the pattern of changes was different for each group and there was no overall significant time effect for self-repairs of any group. The Combined group had lower frequencies of self-repairs on both post-tests than on the Pre-Test, but not significantly so. The Input group had stable scores all over three tests. In contrast, the Comparison group had a significant increase in frequency of self-repairs between Pre-test and the Same-Task Post-Test, but the score decreased between the Same-Task Post-Test and the Parallel-Task Post-Test. Overall, different approaches to task-based instruction did not significantly affect learners' monitoring capacity as measured by frequency of self-repairs. The differences between the Pre-Test and two post-tests for each group are illustrated in Figure 8.4 based on the untransformed means.

Figure 8.4

Changes across Tests in Frequency of Self-Repairs



8.6 Summary

In this chapter, the results for within-group differences over time have been presented. The results of the repeated MANOVAs for the three groups show that there were differences in learners' speech production capacity over time when they were involved in task-based instruction versus non-task-based instruction. With regard to Research Question 4, the Comparison group's performances remained stable over three tests when the learners received narrow input through non-task-based reading and listening practice. With regard to Research Question 5, the Input group learners, who received input-based task practice, only changed for the better in terms of between-clause pausing. With regard to Research Question 6, the Combined group learners, who had opportunities for both input-based and output-based task practice, improved significantly in terms of speech rate and between-clause pausing. Overall, the analyses revealed that different task-based instruction approaches had different impacts on learners' speech production capacity as measured by various fluency measures. The following chapter brings together all these results for discussion.

Chapter 9. Discussion

This chapter interprets the results reported in Chapter 7 and Chapter 8. The first two sections discuss the effects of the different task-based instructional approaches on pre-test/post-test gains and transfer effects to novel tasks, relating the findings to previous ISLA research. Next, theoretical and pedagogical implications of the study are discussed. The chapter closes by pointing out some limitations to the study and making some suggestions for future research.

9.1 Instruction Effects on L2 Speech Production Capacity

The first research question asked whether the different forms of task-based instruction would improve learners' speech production capacity during their performance of a repeated task. The results reveal that there was significant improvement in all aspects of speech production for the group that received a combination of input-based and output-based task instruction.

9.1.1 Instruction Effects on L2 Overall Fluency

Results reveal that in terms of overall fluency only the gains made by the combined input-output task-based instructional group between the pre-test and the same-task post-test were statistically significant (see Section 8.4.1). This means that for the Vietnamese university-age EFL learners in the study, only task-based instruction that included opportunities to both receive task-based input and produce task-based output was effective for promoting speech processing capacity. These findings contrast with recent studies (e.g., Shintani, 2011, 2012a, 2012b) where it was reported that input-based instruction and output-based instruction had comparable effects on the development of beginning-level Japanese EFL learners' receptive and productive knowledge. In the present study, overall fluency gains between the pre-test and the post-test were only significant for learners that received *both* input *and* output opportunities and not for learners who received input only. This was true regardless of whether the input was provided through task-based instruction or narrow reading and listening. One explanation for the difference between the results of Shintani's research and the present study might be that she investigated the effects of input-based instruction and output-based instruction on the acquisition of specific

vocabulary and grammar knowledge, whereas the present study examined learners' fluency development in extended meaning-focused discourse. Moreover, it should be noted that the learners in the present study were false beginners in EFL while Shintani's research was with young children who were true EFL beginners. Furthermore, in Shintani's studies, there were abundant opportunities for output and the only way that the learners could not produce task-based output was when they directly imitated either their teacher or followed a script. Therefore, it could be argued that when learners can produce output based on their own L2 resources, regardless of their level, it is beneficial to their fluency development, and therefore, opportunities should be provided within the syllabus for this to occur.

In this respect, the results of the present study are consistent with earlier research which supports the use of both input and output opportunities in L2 instruction (e.g., De la Fuente, 2002; Gass & Selinker, 2008; Izumi, 2002; Shintani et al., 2013; Swain, 2000; Swain & Lapkin, 1995; Tanaka, 1999, 2001). Tanaka (1999, 2001), for example, found that for Japanese high-beginning EFL learners when production and comprehension practice was combined it was more effective for promoting performance, both within comprehension tasks and production tasks, than when each type of practice occurred alone. This was also true in terms of learners learning of both simple (psychological verbs, 2001) and complex grammar structures (relative clauses, 1999). The results of the present study and previous work such as those by Tanaka (1999, 2001) suggest that for Asian university-age EFL learners at the false beginning level, the ability to produce speech fluently might be best served by the combination of input-based and output-based tasks during TBLT instruction. It is also possible that the two forms of practice (input and output) are complementary in facilitating automatization (McLaughlin, 1987b, 1990) and that input alone may not be as effective in promoting automatic access to the language that learners bring to bear in completing a given task. While input-based practice might enhance learners' ability to comprehend the meaning of messages and establish form-meaning connections, output-based practice could reinforce the form-meaning connections of target structures through producing the language (Tanaka, 2001).

The positive, but non-significant improvement in overall fluency for learners who only received input-based tasks in the study is in line with some researchers' claim that comprehensible input is crucial, but it is not sufficient for language acquisition (Gass et al.,

2013; Long, 1985b; Swain, 1985). In contrast to theories which argue that input is either sufficient to drive SLA (Krashen, 1985 - see Chapter 2, Section 2.2.1) or equal to output, at least in the case of young beginning-level learners (Shintani, 2011, 2012a; Shintani & Ellis, 2010), the present study provides support for Swain's (1985, 1995, 2005) Output Hypothesis and Gass and Selinker's (2008) arguments concerning the role of output as opposed to input alone in SLA (see Section 2.2.3). It has been claimed that opportunities to produce output facilitate acquisition because by producing the language, the learners may notice their linguistic problems/gaps, test a hypothesis, modify their output, and take control over their own linguistic knowledge (i.e., automaticity, which is important for the development of fluency) (Gass & Selinker, 2008; Swain, 1985, 1995, 2005). In the present study, this was evidenced by the significant gains in overall fluency for those who got both input and output opportunities. It is also likely that when learners engaged in output-based tasks (retelling the stories in pairs), not only did they have opportunities for output, but also the exposure to oral input manipulated through verbal interaction with their interlocutor, and even the opportunities to get feedback and modified output, which are all claimed to be facilitative of learners' internal processing and acquisition (Ellis, 2008, 2015; Gass, 1997; Long, 1983, 1996; Mackey, 1999; Mackey & Goo, 2007; Oliver, 1998). However, learners' interactions during treatment were not investigated in the current study and thus deserve further examination.

9.1.2 Instruction Effects on Aspects of L2 Speech Production

In terms of between-clause pausing (conceptualization) and within-clause pausing (formulation), the results of the pre-test and the same-task post-test show that the combined input-output task-based group achieved a statistically significant decrease (i.e., improvement) on both measures (see Sections 8.4.2, 8.4.3). The group that received only input-based tasks also showed the pattern of changes in the same direction as the combined input-output task-based group, but there was no significant difference between their pre-test and post-test scores. In contrast, the group that received narrow input through reading and listening practice had higher frequencies of both between- and within-clause pauses on the post-tests than on the pre-test. The significant improvement in both between-clause and within-clause pausing frequency for the combined input-output task-based group can be seen as a likely indicator of the decreased need to plan content and

efficient linguistic encoding during L2 speech processing respectively. This means that for the Vietnamese university-age EFL learners in the study, only task-based instruction in which the learners had opportunities for both task-based input and output was effective for improving conceptualization and encoding processes during L2 speech processing capacity and increasing fluency.

With regard to self-repairs (monitoring), it will be remembered from Chapter 7 (Section 7.3) that a lower rate of self-repairs between the pre-test and the same-task post-test was the best predictor of learners who had received TBLT instruction as compared to those who engaged in non-TBLT instruction (narrow reading and listening). Specifically, self-repair behaviour decreased for the group that received a combination of input-based and output-based tasks, whereas it remained stable for the group that received input-based tasks only, and increased for the learners who received non-task-based narrow reading and listening (see Figure 8.4). The larger decrease in monitoring for the group that received integrated task-based instruction rather than input-based instruction only (either in the form of tasks or in the form of narrow reading and listening) indicated that the opportunity to produce task-based output helped these beginning-level learners to effectively automatize access to the L2 resources, and it also freed their attention from the need to monitor the language that they brought to bear on this task.

From these results, it can be concluded that only the combined input-output task-based treatment had significant effects on aspects of speech production. It is possible that the opportunities for both receiving input and producing the target language enabled learners to utilize the vocabulary and structures learned from previous narratives for subsequent task enactments better than the students who did not receive such meaningful combined practice. The significant improvement over time in terms of both between-clause pausing (content planning) and especially within-clause pausing (lexical and grammatical encoding) for the combined input-output task-based group in the present study could be explained through the proceduralization of content and linguistic knowledge by way of repeated use of certain task-related words, phrases and sentence structures through output practice (De Jong & Perfetti, 2011). While all speakers need to conceptualize content, the need to consciously encode this content in the L2 decreases with proficiency level as learners have progressively more automatized access to lexis and associated syntax. With

the changes in the learners' underlying knowledge, they are thus able to encode the language in the formulation module while organizing content in the conceptualization module.

Alternatively, the improvement for the combined input-output task-based group in all aspects of speech production could be attributed to the opportunities to use the language they have built up from parallel tasks of the same type when completing output-based tasks during treatment sessions. It is possible that the participants' familiarity with task content, task demands, task structures, and discourse types that was gained through exposure to parallel versions, together with their experience of the first performance of the same task has eased the cognitive demands, increased task-readiness, and freed up attentional capacity for more efficient parallel processing of conceptualization and formulation during subsequent performances of a repeated task (Bui & Huang, 2016; Bui, 2014; Bygate, 2001; Skehan & Foster, 1999).

The results of the present study are also in line with the Limited Attention Capacity hypothesis (Skehan, 1998a) which argues that due to limited attentional capacity, learners cannot attend to everything equally at the same time and the gains in one area may be at the expense of the others. In relation to Levelt's (1989) L1 speech production model, Skehan (2009) argues that L2 fluency is associated with access to the lexicon and syntax during the formulation stage. While L1 speakers or proficient L2 speakers who have automatic access to the mental lexicon can rely on parallel processing during the conceptualization and formulation of messages, low-proficiency speakers (as in the present study) need to resort to serial processing. Consequently, breakdowns often occur during task performances due to the lack of attention available for both conceptualizing and formulating the message to operate in parallel. The gains in overall fluency for the learners who received both input and output opportunities in the current study may be explained by the increased automatization during the formulation of messages which allowed parallel processing of the conceptualization and formulation stages (corresponding less between-clause and within-clause pausing) on the post-tests (McLaughlin, 1987b; Mitchell & Myles, 2004). Automaticity might be argued to be manifested in the greater decrease in self-repairs, between-clause pauses, within-clause pauses, and greater increase in speech rate in the combined input-output group than in the input-only groups. These results, together with previous research

(e.g., Lambert et al., 2020; Lambert et al., 2017; Skehan, 2009) suggest that combined input-based and output-based based tasks in the L2 curriculum can promote fluent L2 performance by easing attention demands for the conceptualization and formulation stages of speech production. For beginning-level EFL learners, in particular, combining the use of input-based and output-based tasks during TBLT instruction might be one way to facilitate parallel processing of L2 speech.

To summarize, the comparison of the pre-test and the same-task post-test scores on the repeated task for the three groups indicates that the combined input-output task-based instruction was more efficient in developing learners' speech production capacity than the instruction involving input only regardless of whether the input was task-based or through narrow reading and listening. The study thus shows that for beginning-level university-age learners in Asia who have had some previous English instruction in secondary school, a combination of input-based tasks and output-based tasks in the syllabus is likely to be the most effective means of promoting automaticity of the language that learners bring to bear on tasks. Input-based tasks alone may be insufficient for this purpose, even when learners are beginners and struggle with producing speech on tasks.

9.2 Transfer Effects on L2 Speech Production Capacity

The second research question asked whether different forms of task-based instruction would have differential transfer effects regarding learners' speech production capacity during their performance of a parallel task. The results show that learners in both task-based instructional groups paused significantly less between clauses on the Parallel-Task-Post-Test than on the Pre-test. This indicates that there were transfer effects with regard to between-clause pausing. For other measures, both task-based instructional groups also exhibited positive trends when performing a parallel task post-test with faster speech rates, fewer within-clause pauses, as well as lower rates of self-repairs than the group that received non-task-based narrow reading and listening. This means that for the Vietnamese beginning-level university-age EFL learners in the study, task-based instruction was more likely to result in transfer effects, especially in terms of conceptualization, during the performance of parallel tasks than merely exposure to narrow input through non-task-based reading and listening.

9.2.1 Transfer Effects on L2 Overall Fluency

In terms of overall fluency, although the practice of parallel narratives resulted in significant fluency gains for the combined input-output task-based group when the students performed a repeated task (see Section 9.1), these effects were not retained during their speech production of parallel task performances.

Given that the treatment for the combined input-output task-based group was the repetition of task type (parallel task repetition), the lack of fluency transfer effects in the performance of a parallel task of the present study mirror the results of previous studies (Bygate, 2001; De Jong & Perfetti, 2011) that also reported no transfer effects for parallel task repetition. Bygate (2001), for instance, found that practice of one task type did not benefit future performances of new versions of the same task type, while it did have positive impact on the performance of a repeated task. Similarly, De Jong and Perfetti (2011) found that both the same task repetition and parallel task repetition resulted in significant short-term improvement in fluency during the training period, but the effects transferred to the performance of a parallel-task post-test only for the same task repetition condition. This suggests that multiple same task repetitions might be necessary for the automatization of linguistic knowledge to occur before practice effects can carry over to the performance of a new task. For example, Lambert et al. (2020) provided evidence of fluency transfer for massed parallel task repetition of oral monologic tasks when used by intermediate-level university-age EFL learners in Japan. The authors argued that "learners were able to transfer fluency gains from one version of the task type to another" because they were familiarized with basic procedures of the task type and through massed practice of parallel task repetition, they became more flexible in encoding new content and language of the same task type with similar task structures and demands. However, Lambert et al.'s (2020) study was carried out with intermediate-level EFL learners, whereas the learners in the present study were false beginners in EFL. Another possible reason for the contrast with Lambert et al. (2020) might be that they employed a between-subjects design, whereas the present study employed a within-subject design to study transfer effects, which may allow more direct insight into the transfer of practice.

Alternatively, repetition intervals may also account for the lack of transfer in the present study. In Lambert et al. (2020), the participants performed parallel versions of the

task consecutively within a total treatment period of ten minutes, whereas the participants in the present study repeated parallel tasks on a weekly basis for three weeks and were assessed in the fourth week. It is thus possible that in Lambert et al.'s (2020) study the participants' memory of the task structures and procedures was still fresh, occurring within a short span of only 10 minutes. This short-term memory may have eased the cognitive demands on conceptualization and formulation related to basic schemata for the argumentative tasks, and thus freeing attentional resources to adapt new L2 resources to these schemata. Bui et al. (2019) referred to the differences in intervals between repetitions of a task as 'spacing effects' and suggested that immediate repetition is the most beneficial for speed fluency. In the present study, transfer effects may have been diluted due to the longer intervals between repetitions, leading to positive, but non-significant fluency gains. Thus, in order for practice effects to transfer to the performance of a novel version of the same task type, massed immediate parallel task repetition may be required. Future research is needed to investigate the conditions under which practice effects might transfer with beginning-level learners such as those in the present study. An interesting possibility would be to consider the impact of massed input-based and output-based parallel task repetition on learners at different levels of proficiency.

9.2.2 Transfer Effects on Aspects of L2 Speech Production

Interestingly, the present study found comparably significant transfer effects in terms of between-clause pausing for both task-based instructional groups. In other words, during the performance of a parallel-task post-test, learners in both TBLT conditions (i.e., input only and combined input-output) paused significantly less between clauses than they did on the pre-test. This finding lends support to the claim made by Skehan et al. (2016) that a more structured task is associated with less clause-boundary pausing for both native and non-native speakers. In the present study, the gains in between-clause pausing for both types of task-based instruction might be attributed to the increased familiarity with the narrative motifs and the task structures during their parallel task practice (Lambert et al., 2020). Knowing what to say and when to say it in completing the task might have reduced the need for frequent pauses at clause boundaries to plan ideas (Conceptualizer-linked macro-planning) and resulted in the observed decreases in the frequency of between-clause pauses for the TBLT groups (Skehan & Foster, 1999; Skehan et al., 2016).

In terms of within-clause pausing and self-repairs, the lack of transfer effects for the combined input-output task-based group in the present study is in line with the findings of Lambert et al. (2020) who found no significant improvement on these measures for parallel task repetition. They explained the difference between the four treatment groups (same task repetition, parallel task repetition, L1 planning, and L2 planning) by the amount of processing support for content, language, and practice, respectively that each treatment option could bring to the learners' performance of the assessment task. Based on previous research (De Jong, 2016b; Lambert et al., 2020), a possible explanation for the results of the present study could be that when the participants performed the new task version, they had little opportunity to rehearse and encode topic-specific language for the task like they did on the same task performance. The limited access to the needed lexis and structures would then increase the need for online encoding with corresponding more within-clause pauses to search for new topic-relevant expressions during the parallel task than the same task performance. Consequently, they needed to repair the initially encoded utterances more frequently than they did on the repeated task performance as was found in the study by Lambert et al. (2020). Thus, it is suggested that for low-proficiency learners, different types of processing support can be provided through task preparation options to assist learners to effectively access their available L2 resources for more efficient task performances.

In summary, the comparison of the pre-test and the parallel-task post-test results for the three groups shows that both task-based instructional groups exhibited more transfer effects in terms of conceptualization during the performance of parallel tasks than the group that received narrow reading and listening instruction. It might be speculated that either input-based or output-based practice of a task-type can increase the familiarity with task structures and task demands, which subsequently decreased the need to pause between clauses for conceptualization. However, the results of the present study overall indicate that the combined input-output task-based instruction was more effective in improving beginning-level university-age EFL learners' speech production capacity than the input-based instruction regardless of whether the input was provided through tasks or narrow reading and listening.

9.3 Theoretical Implications

The present study has some theoretical implications to contribute to the field of task-based language teaching and learning, as well as L2 speech production.

Firstly, the study fills the gap in previous studies on speech production. There have been a number of studies investigating the impact of task repetition or different planning options on learners' speech production (e.g., Lambert et al., 2020; Lambert et al., 2017). However, little is known about the potential role of task-based instruction on learners' speech production capacity. The present study addresses this gap by providing empirical evidence for the impact of two approaches to task-based language instruction: input-based tasks versus combined input-based and output-based tasks on EFL learners' speech production capacity. Specifically, the study shows that combining input-based and input-output-based task instruction can maximize low-proficiency EFL learners' development of speech production by providing opportunities for both receiving input and producing output. While these results do not support Shintani's (2011) claim that input-based task instruction is equal to output-based instruction, they lend support to Tanaka (1999, 2001) and Shintani et al.'s (2013) suggestions of combining the two types of instruction for better effects, at least for low-level university EFL learners in comparable contexts to the present study. Thus, this study reinforces the importance of both input and output not only in L2 learners' acquisition of grammar and vocabulary (e.g., De la Fuente, 2002; Gass & Selinker, 2008; Izumi, 2002; Shintani, 2011; Shintani et al., 2013; Swain, 2000; Swain & Lapkin, 1995; Tanaka, 1999, 2001), but also in the development of their L2 speech fluency, or their ability to produce fluent speech.

Secondly, the current study provides some support for the use of pause location and self-repairs in measuring speech production. The results of the present study are consistent with those reported in previous research about the measurement of L2 speech production (e.g., De Jong, 2016a; Götz, 2013; Kahng, 2020; Kormos, 2006; Lambert et al., 2020; Lambert et al., 2017). In these studies, it has been argued that between-clause pausing can be a measure of conceptualization (content planning), within-clause pausing represents formulation (encoding), and self-repairs reflects attention to monitoring during task performances. With regard to pause phenomena, previous research has provided evidence that the differences between native and non-native speakers, or between different

proficiency-level learners, are associated with pause location (within-clause as opposed to between-clause pausing) rather than pause frequency and pause duration. In previous research, within-clause pausing has varied with proficiency level, indicating that it is a measure of linguistic formulation or encoding, as increases in within-clause pausing directly reflect breakdowns during the formulation stage of L2 speech production (De Jong, 2016b; Kahng, 2014, 2018; Tavakoli, 2011). In contrast, although pauses between clauses can reflect many phenomena, they can be distinguished from within-clause pausing in that they are not associated with proficiency level and breakdowns in the formulation module of speech production. As between-clause pausing does not differ significantly between learners of different proficiency levels, it can be argued to generally reflect the differences related to conceptualization of messages and other cognitive processes that are not associated with the formulation module. The present study shows that repeated task-based practice seems to facilitate learners' automatization, which enables the move from serial to parallel processing of conceptualization and formulation stages and a corresponding lower frequency of between-clause pauses and within-clause pauses. The learners' automatized access to current L2 resources also means that more attention is freed from the need to monitor the language that they brought to bear on the task and a corresponding decrease in self-repairs.

Finally, the present study provides supporting evidence for the benefits that parallel task repetition or task-type practice can positively bring to learners' subsequent performances, at least when the same task is repeated. While the advantages of repeating the same task on both L2 oral and written performances have been evidenced in numerous studies (e.g., Ahmadian, 2011, 2012; Ahmadian et al., 2017; Ahmadian & Tavakoli, 2011; Bygate, 2001; De Jong & Perfetti, 2011; Fukuta, 2016; Lambert et al., 2017; Wang, 2014), the effects of parallel task repetition on learners' interlanguage development remain inconclusive (Bygate, 2001; Hunter, 2017; Lambert et al., 2020). Overall, the findings of the present study show the importance of task-type practice in developing learners' flexibility in encoding new content and language in task-based language teaching and learning, as well as in carrying real-world-related tasks (Bygate, 2001; Lambert et al., 2020). The benefits of task-type practice on fluency, in particular, can be enhanced through the use of repeating the same task. Prior experience of a task, coupled with knowledge of how to carry out the

task can bring about significant fluency improvement in learners' subsequence performances of the same task.

9.4 Pedagogical Implications

The findings of the present study have potential pedagogical implications. First of all, this study shows that for low-proficiency EFL learners combining input-based and output-based tasks can best promote learners' fluency development. When combined, the two forms of practice may be complementary in facilitating proceduralization of information processing through form-meaning connections, thus freeing up attentional resources for more fluent performances. However, it is the teacher who needs to choose appropriate tasks according to the learners' proficiency levels, balancing the amount of input and output in each lesson, and sequencing the tasks in order to best develop the learners' ability to speak fluently in real-time communication.

Second, reusing tasks can improve learners' ability to produce fluent speeches. Parallel task repetition and same task repetition of both input-based and output-based tasks can interact to boost learners' fluency development by reducing the demands of conceptualization, encoding, as well as attention to monitoring. On the one hand, same task repetition may increase learner's fluency by lexical and structural priming and enable the procedualization of the topic-specific lexis and structures, thereby freeing up learners' attentional resources for more efficient subsequent performances. On the other hand, repetition of parallel versions of the same task type would give learners the opportunity to strengthen form-meaning relationships and increase learners' adaptability and flexibility when handling new tasks with new content and language using their own linguistic resources. However, it is important that teachers choose the appropriate type of practice based on their learners' proficiency levels, educational contexts, and instructional objectives so as to foster language development (Bygate, 2001; Lambert et al., 2020). As Lambert et al. (2020) suggest, parallel task repetition may be more beneficial for experienced learners or those in contexts where English is used as a second language (ESL) and they have opportunities to use the language outside the classroom. Meanwhile, for less experienced learners or those who have limited opportunity to use the language outside the classroom, same task repetition may be a better option to enhance their confidence in using their current linguistic resources before engaging them in parallel task repetition. In terms of

lesson sequences, same task repetition can be employed in initial lessons to familiarize learners with a new task type, and provide them with more processing support and opportunities to gain “fluent access to known L2 resources” (Lambert et al., 2020, p. 27). At the later stage, when learners are able to perform the task at an acceptable level of fluency with automated access to current task-related L2 linguistic resources, parallel task repetition can be added to develop flexibility in processing new content and language associated with the same task type. It is necessary that the learners be made aware of the advantages of each type of practice and receive feedback about their performances so that they can improve in future enactments. It is also recommended that fluency development come along with gains in other areas of language proficiency in balanced instruction options.

9.5 Limitations and Future Research

First of all, the study is limited in that there was no output-only-based task instruction condition to explore how learners’ fluency might change when there were no input-based tasks. The inclusion of a group that received output-based tasks only might yield a more comprehensive data picture. However, it was impossible to recruit more than 102 participants with comparable backgrounds and proficiency levels at the time the data collection was conducted. In fact, it was planned to recruit 36 participants for each group (n = 108), but only 102 confirmed their availability. Second, the study is further limited by the lack of a delayed post-test to measure the stability of treatment effects. Data from a delayed post-test conducted after a longer interval would have answered the question if the treatment effects were durable in the longer term. Therefore, it is recommended that similar future studies include an output-based task condition and delayed post-tests for more comprehensive findings on how learners’ speech fluency may change when they are engaged in output-based tasks only and whether the treatment effects for each type of task-based instruction are maintained over a longer period of time. Additionally, the study could be replicated with a more diverse range of task types (e.g., opinion, description tasks) and proficiency levels in different instructional contexts. For instance, the results may be not similar if undertaken with learners in an ESL context with higher-proficiency levels. The findings of such studies would enable comparisons across research contexts and be of great value for future meta-analyses.

Third, the issue of test bias against the two experimental conditions might also be questionable. While the treatments involved both input-based and output-based tasks, the tests only deal with learners' production, but not comprehension. However, as the primary objective of the current study was to investigate the effects of two task-based instructional options on learners' speech production capacity, the investigation into how learners' comprehension ability is impacted is thus beyond the scope of the study and is left for future research. Even so, the inclusion of a comprehension test along with a production test in future studies may help avoid the test bias criticism.

Fourth, a fuller understanding of learners' pausing patterns and the nature of self-repairs would have been achieved if stimulated recalls or qualitative analysis of the transcripts had been conducted in addition to the quantitative results. However, due to the time-consuming nature of conducting temporal fluency analysis, using mixed methods is extremely challenging. Thus, future studies could incorporate quantitative data analysis with other instruments for more fine-grained findings – the results of the present study remain limited to quantitative results only. For example, qualitative analysis of transcripts could bring an in-depth understanding of the changes that take place in the cognitive processes underlying the participants' speech production. The use of stimulated recall as an introspective method in follow-up interviews could also be used to reveal deeper insight into the use of between-clause and within-clause pauses, and self-repairs that occur during the course of L2 speech production (Kahng, 2014).

Finally, the study is limited in that there was no control for individual variations in participants' L1 fluency behaviours. Previous research has provided evidence that that L1 and L2 fluency behaviours are correlated in many ways, for example, with respect to pausing frequency and speech rate (De Jong, 2016b; Derwing et al., 2009; Duran-Karaoz & Tavakoli, 2020; Kahng, 2014, 2020; Segalowitz, 2010). Although the sample size, group randomization, and the homogeneity of the participants had certain control over individual differences, the results might be different due to moderating effects of participants' varied L1 fluency behaviours such as pausing and use of fillers. Future research needs to take variations in L1 fluency into account in order to accurately measure L2 fluency because individual differences in L1 pausing behaviours, for example, may moderate the treatment effects.

9.6 Summary

This chapter has discussed the findings of the study with reference to previous research. The study shows that combining input-based and output-based task instruction appears to be the most beneficial approach for developing L2 speech fluency for low-proficiency EFL learners, at least in the context of the current study. Although the input-based task instruction also had positive trending impact on EFL learners' speech production capacity in the performance of repeated oral monologic tasks, the effects did not reach significance. It is also found that transfer effects were observed only in terms of between-clause pausing in the performance of a parallel task. The discussion of these findings is followed by theoretical implications, pedagogical implications, limitations, and recommendations for future research. The next chapter, the conclusion, will conclude the study with a summary of all previous chapters and drawing together the threads of this research.

Chapter 10. Conclusion

This final chapter provides a summary of the key aspects of the study. The first section presents the background and objectives of the study. A summary of the key findings and implications of the study is provided next. The chapter concludes with some final comments

10.1 Aims of the Study

The present study was conducted to address a specific issue regarding ELT in Vietnam, namely students' inability to communicate effectively after many years of learning. According to various researchers, (e.g., Hoang, 2013; Hoang, 2014; Le & Barnard, 2009; Mai & Iwashita, 2012), there are many reasons for this situation, with ineffective teaching and learning approaches being identified as the most important factor. It has been suggested that the use of task-based instructional approaches may counter this by helping to improve students' ability to use English for real-time communication in place of the existing English teaching approach that focuses on linguistic knowledge more than communicative competence (Hoang, 2008; Le, 2007). Hence, the present study was carried out to investigate the effects of different task-based instructional approaches on Vietnamese EFL students' speech production capacity. In particular, the study compared two types of input-based instruction with combined input-based and output-based instruction. The reason for this is that, on the one hand, previous research has found that effects for input-based instruction with beginning-level learners were comparable to output-based instruction (Shintani, 2011). If this proved to be the case in Vietnam, it would be of considerable practical value for curriculum renewal because classes tend to be large, which restricts opportunities for individual and group work speech production tasks. Furthermore, the instructional context is much better suited to input-based instruction than output-based instruction in terms of teachers' level of training, learners' background and preferences, and administrative constraints on teaching and testing. On the other hand, other studies (e.g., Shintani et al., 2013; Tanaka, 1999, 2001) have shown that input-based instruction and output-based instruction have distinct advantages for L2 acquisition. Thus, it has been suggested that combining the two types can maximize L2 learning. Therefore, the general aim of this study was to determine whether input-based task instruction was comparable to

combined input-output task-based instruction in terms of learners' ability to perform production tasks (Shintani, 2011) or whether combining two types of instruction was more effective for low-level EFL learners, in this case for Vietnamese university students (Shintani et al., 2013; Tanaka, 1999, 2001).

Following on from this general aim, the primary research objectives of the current study were to determine: (1) the effects of input-based task instruction versus combined input-output task-based instruction on EFL learners' speech production capacity in the performance of a repeated oral monologic task, and (2) whether or not any observed effects would transfer to the performance of a parallel task. These objectives were specified in six research questions. The first two research questions concerned the differences between treatment groups in the performance of a repeated and a parallel narrative task, followed by the third research question, which looked at the role of dependent variables in distinguishing the treatment groups. Research Questions 4, 5, and 6 sought to examine how learners' speech production capacity changed over time within each group.

10.2 Key Findings and Implications

Firstly, this study reveals that for beginning-level university-age EFL learners combining input-based tasks and output-based tasks in the syllabus was the option that produced significant pre-test to post-test gains in speech production capacity. Although the instruction that included input-based tasks alone also showed trends in the same direction as the combined input-output task-based group, the difference between the pre-test and the post-test did not reach statistical significance. It was likely that when combined, the two forms of task-based practice were complementary in promoting the automatization of different stages of L2 speech processing so that learners could efficiently process their speech on the targeted task type. It was the opportunity to produce the target language during output-based practice of parallel tasks that might have enabled the move from controlled to automatic processing by easing cognitive and attentional demands during the conceptualization and formulation stages of speech production. This was further enhanced by the learners' decreased need to monitor their output on subsequent task performances. In other words, the key to maximizing learners' speech processing capacity was the automatization of L2 language and skills, which was achieved through the combination of both input-based and output-based practice. Automatized access to available L2 linguistic

resources meant less attentional resources were needed for linguistic encoding, and parallel processing of both the Conceptualizer and the Formulator was enabled, which then facilitated fluency development (Anderson, 1980, 1983; McLaughlin, 1987b, 1990; Mitchell & Myles, 2004).

Secondly, the findings indicate that task-based instruction has more potential to result in transfer effects on low-proficiency EFL learners' speech production capacity than non-task-based narrow reading and listening instruction during subsequent performances of parallel tasks, but with different degrees of impact on each aspect. Specifically, while both forms of task-based instruction lead to significant gains on learners' pausing behaviours between clauses, the improvement on other measures exhibited positive trends from pre-test to post-test (faster speech rate, fewer within-clause pauses, and lower frequency of self-repairs), but not significantly so. It is speculated that the significant improvement in between-clause pausing for both TBLT groups was because of the familiarity with the story structure and task demands (Bygate, 2001; Skehan & Foster, 1999) that increased alongside with opportunities for task-based practice. This familiarity with a storyline can be argued to have led to a reduction in conceptual demands and a corresponding lower frequency of between-clause pauses. Meanwhile, the limited transfer effects on other measures of speech production might have been caused by the lack of automatized access to the language needed for the performance of a parallel task. It appeared that spaced practice rather than massed practice of parallel versions of tasks was insufficient for learners to keep the conceptual and linguistic schemata of the task type primed and active during the post-test, thus resulting in a lack of attentional resources being available for adopting new language to these structures. When learners could not resort to automatic processing of the general task schemata, they had to slow down their speech rate and pause more often within clauses to retrieve the language needed for task performance, and they also self-repaired the initially encoded utterances more frequently than they did during the repeated task performance. This at least provides one plausible explanation for why learners in both TBLT groups failed to transfer their fluency gains to a parallel task performance on the post-test.

Based on these findings, the study makes significant contributions to the field of ISLA in terms of both theory and pedagogy. Theoretically, the study fills a gap in previous

research about the impact of task-based instruction on learners' speech production capacity. In contrast to Shintani's (2011, 2012a, 2012b) claim that input-based instruction was effective for true beginner young learners, the present study shows that the combination of both input-based and output-based tasks during instruction is needed, particularly in this case, to improve beginning-level university-age EFL learners' fluency, and potentially in similar contexts. Second, the results provide further evidence confirming that in addition to pause frequencies, pause locations (between clauses vs. within clauses) and self-repairs can be useful measures of speech production. The study suggests that the combination of both input-based and output-based task practices might enable parallel processing of conceptualization and formulation stages and lead to corresponding lower frequencies of between-clause pauses and within-clause pauses. At the same time, automatized access to available L2 resources also seems to reduce the attention needed for monitoring the language that they brought to bear on the task with a corresponding decrease in self-repairs. Finally, the study reaffirms the importance of parallel task repetition on learners' subsequent task performances, especially on a repeated task performance by showing that task-type practice can increase learners' familiarity with task structures and develop their flexibility in encoding new content and language, thus improving overall fluency.

Pedagogically, the study confirms the important role of the combined use of input-based and output-based tasks in improving learners' speech production capacity - a critical issue in Vietnamese tertiary-level English instruction. While the gains for this type of instruction were significant between the pre-test and the post-test, the gains for both input-based task instruction as well as non-task-based narrow input instruction failed to produce significant gains in L2 speech processing capacity. The study thus points to the need for curriculum and pedagogical renewal in Vietnam to prepare teachers and learners so that they can utilize both input-based and output-based practice within the ELT curriculum. In addition, the results of the study suggest that opportunities might be needed for greater use of massed parallel task repetition within the syllabus. This is based on the findings showing that the one-lesson-per-week format that was used in this study proved to be insufficient for the transfer of speech processing gains to similar parallel versions of the target task type. Although more research is needed with beginning-level learners, previous research by

Lambert et al. (2020) suggests that massed parallel task repetition may be effective in promoting the transfer of speech processing gains by increasing learners' familiarity with task structure and task demands, thus allowing adaptability and flexibility during the formulation stage when performing new versions of a targeted task type in the L2 syllabus. Finally, the current study also suggests that more flexibility regarding task implementation should be built into the curriculum so that teachers can adjust the nature and amount of practice that they provide, depending on their specific educational context, instruction objectives, and the learners' proficiency level.

10.3 Final Comments

It is hoped that the present study has contributed to the existing body of empirical research on task-based learning and teaching. The study not only provides further evidence that tasks can be useful learning tools in second language teaching, it also demonstrates that task design and implementation factors can be linked to the development of different aspects of speech production. It is also expected that Bygate's (2016) call for studies that enable TBLT as a 'researched pedagogy' has been addressed to some extent by the current study reporting instructional procedures and relative results of using an input-based approach to TBLT in comparison with a combined input-output task-based approach. Finally, it is hoped that the study will provide some useful insights for teachers regarding the selection and sequencing of tasks during the course of TBLT so as to maximize learners' speech production capacity, thus contributing to the improvement of task-based English teaching and learning in Vietnam and similar EFL contexts.

References

Adams, R., & Newton, J. (2009). TBLT in Asia: Opportunities and constraints. *Asian Journal of English Language Teaching, 19*, 1-17.

Ahmadian, M. J. (2011). The effect of 'massed' task repetitions on complexity, accuracy and fluency: Does it transfer to a new task? *The Language Learning Journal, 39*(3), 269-280.

Ahmadian, M. J. (2012). Task repetition in ELT. *ELT journal, 66*(3), 380-382.

Ahmadian, M. J., Masouri, A., & Ghominejad, S. (2017). Language learners' and teachers' perceptions of task repetition. *ELT Journal, 71*(4), 467–477.

Ahmadian, M. J., & Tavakoli, M. (2011). The effects of simultaneous use of careful online planning and task repetition on accuracy, complexity, and fluency in EFL learners' oral production. *Language Teaching Research, 15*(1), 35-39.

Anderson, J. (1980). *Cognitive psychology and its implications*. Freeman.

Anderson, J. (1983). *The architecture of cognition*. Harvard University Press.

Anderson, J. (1985). *Cognitive psychology and its implications* (2nd ed.). Freeman.

Anderson, J. (2000). *Learning and memory: An integrated approach* (2nd ed.). John Wiley & Sons Inc.

Barnard, R., & Viet, N. G. (2010). Task-based language teaching (TBLT): A Vietnamese case study using narrative frames to elicit teachers' beliefs. *Language Education in Asia, 1*(1), 77-86.

Bernat, E. (2004). Investigating Vietnamese ESL learners' beliefs about language learning. *English Australia Journal, 21*(2), 40-54.

Bock, G. (2000). Difficulties in implementing communicative theory in Vietnam. *Teacher's Edition, 2*, 24-28.

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

Bui, G., Ahmadian, M. J., & Hunter, A. M. (2019). Spacing effects on repeated L2 task performance. *System, 81*, 1-13. https://doi.org/https://doi.org/10.1016/j.system.2018.12.006.

Bui, G., & Huang, Z. (2016). L2 fluency as influenced by content familiarity and planning: Performance, measurement, and pedagogy. *Language Teaching Research, 22*(1), 99-114. https://doi.org/https://doi.org/10.1177/1362168816656650

Bui, H. Y. G. (2014). Task readiness: Theoretical framework and empirical evidence from topic familiarity, strategic planning, and proficiency levels. In P. Skehan (Ed.), *Processing perspectives on task performance*. John Benjamins.

Bui, T. T. N., & Nguyen, H. T. M. (2016). Standardizing English for educational and socio-economic betterment- A critical analysis of English language policy reforms in Vietnam. In R. Kirkpatrick (Ed.), *English language education policy in Asia* (pp. 363-388). Springer. https://doi.org/https://doi.org/10.1007/978-3-319-22464-0_17

Bygate, M. (1996). Effects of task repetition: appraising the developing language of learners. In D. Willis & J. Willis (Eds.), *Challenge and change in language teaching* (pp. 134-146). 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* (Vol. 1, pp. 23-48). Pearson Longman.

Bygate, M. (2016). Sources, developments and directions of task-based teaching. *The Language Learning Journal, 44*(4), 381–400. https://doi.org/https://doi.org/10.1080/09571736.2015.1039566

Bygate, M., & Samuda, V. (2005). Integrative planning through the use of task repetition. In R. Ellis (Ed.), *Planning and Task Performance in Second Language* (pp. 37-74). John Benjamins.

Bygate, M., Skehan, P., & Swain, M. (2001). *Researching pedagogic tasks, second language learning, teaching and testing*. Longman.

Cao, P. T. H. (2018). Task-based language teaching: Affordances and challenges in TBLT implementation at the Vietnamese tertiary level. *The Journal of Asia TEFL, 15*(2), 510-515. https://doi.org/http://dx.doi.org/10.18823/asiatefl.2018.15.2.19.510

Caspino, B. K. (2005). Support for narrow listening libraries. *The International Journal of Foreign Language Teaching, 2005*, 2-9.

Chambers, F. (1997). What do we mean by fluency? *System, 25*(4), 535–544. https://doi.org/10.1016/S0346-251X(97)00046-8

Chang, A. (2009). Effects of narrow reading andlistening on L2 vocabulary learning: Multiple dimensions. *Studies in Second Language Acquisition, 41*(4), 769-794. https://doi.org/doi:10.1017/S0272263119000032

Cho, K. S., & Krashen, S. (1994). Acquisition of vocabulary from the Sweet Valley Kids series. *Journal of Reading, 37*, 662-667.

Cho, K. S., & Krashen, S. (1995). From Sweet Valley Kids to Harlequins in one year. *California English, 1*(1), 18-19.

Chu, H., & Oliver, R. (2017). Vietnamese TESOL teachers' cognitions and practices: Developing learner-centred teaching. In L. Gurzynski-Weiss (Ed.), *Expanding individual differences research in the interaction approach: Investigating learners, instructors and other interlocutors* (pp. 173-199). John Benjamins.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum.

Cohen, J. (1992). Statistical power analysis. *Current Directions in Psychological Science, 1*(3), 98-101. https://doi.org/10.1111/1467-8721.ep10768783

Colomé, A. (2001). Lexical activation in bilinguals' speech production: Language specific or language-independent? *Journal of Memory and Language, 45*, 721-736.

Corder, S. (1967). The significance of learners' errors. *International Review of Applied Linguistics, 5*, 161-170.

Costa, A. (2005). Lexical access in bilingual production. In J. Kroll & A. M. B. DeGroot (Eds.), *Handbook of bilingualism: Psycholinguistic perspectives*. Oxford University Press.

Council of Europe. (2001). Common European framework of reference for languages: Learning, teaching, assessment.

Cramer, E. M., & Bock, R. D. (1966). Multivariate analysis. *Review of Educational Research, 35*, 604–617. https://doi.org/https://doi.org/10.3102/00346543036005604

Davies, A. (2003). *Native speaker: Myth and reality*. Multilingual Matters.

De Bot, K. (1992). A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics, 13*, 1-24.

De Bot, K., & Schreuder, R. (1993). Word production and the bilingual lexicon. In R. Schreuder & B. Weltens (Eds.), *The bilingual lexicon* (pp. 191–214). Benjamins.

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.

De Jong, N. H. (2016a). Fluency in second language assessment. In D. Tsagari & J. Banerjee (Eds.), *Handbook of second language assessment* (pp. 203–218). De Gruyter Mouton.

De Jong, N. H. (2016b). Predicting pauses in L1 and L2 speech: The effects of utterance boundaries and word frequency. *IRAL, 54*(2), 113-132.

De Jong, N. H. (2018). Fluency in second language testing: Insights from different disciplines. *Language Assessment Quarterly, 15*(3), 237-254. https://doi.org/10.1080/15434303.2018.1477780

De Jong, N. H., & Bosker, H. R. (2013). Choosing a threshold for silent pauses to measure second language fluency. In *Proceedings of the 6th workshop on disfluency in spontaneous speech (DiSS)* (pp. 17-20).

De Jong, N. H., Steinel, M. P., Florijn, A. F., Schoonen, R., & Hulstijn, J. H. (2013). Linguistic skills and speaking fluency in a second language. *Applied Psycholinguistics, 33*, 1-24. https://doi.org/https://doi.org/10.1017/S0272263111000489

De la Fuente, M. (2002). Negotiation and oral acquisition of L2 vocabulary: The roles of input and output in the receptive and productive acquisition of words. *Studies in Second Language Acquisition, 24*, 81-112. https://doi.org/http://dx.doi.org/10.1017/S0272263102001043

DeKeyser, R. (1997). Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition, 19*, 195-221.

DeKeyser, R. (1998). Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on form in second language acquisition*. Cambridge University Press.

DeKeyser, R. (2015). Skill acquisition theory. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 94-112). Routledge.

Derwing, T. M. (2017). L2 fluency development. In S. Loewen & M. Sato (Eds.), *Routledge handbook of instructed second language acquisition*. Routledge.

Derwing, T. M., Munro, M. J., Thomson, R. I., & Rossiter, M. J. (2009). The relationship between L1 fluency and L2 fluency development. *Studies in Second Language Acquisition, 31*(4), 533-557. https://doi.org/10.1017/S0272263109990015

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language Learning, 54*(4), 655–680. https://doi.org/10.1111/j.1467-9922.2004.00282.x

Dörnyei, Z., & Kormos, J. (1998). Problem-solving mechanisms in L2 communication: A psycholinguistic perspective. *Studies in Second Language Acquisition, 20*, 349–385.

Doughty, C., & Long, M. H. (Eds.). (2003). *The handbook of second language acquisition*. Blackwell publishing.

Dupuy, B. (1999). Narrow listening: An alternative way to develop listening comprehension in the foreign language classroom. *System, 27*(3), 351-361.

Duran-Karaoz, Z., & Tavakoli, P. (2020). Predicting L2 fluency from L1 fluency behaviour: The case of L1 Turkish and L2 English speakers. *Studies in Second Language Acquisition*, 1-25. https://doi.org/10.1017/S0272263119000755

Ellis, N. (2003). Constructions, chunking and connectionism: The emergence of second language structure. . In C. J. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 63-103). Blackwell Publishing Ltd.

Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics, 16*, 409.

Ellis, R. (2003). *Task-based language learning and teaching*. Oxford University Press.

Ellis, R. (2005a). Planning and task-based research: Theory and research. In R. Ellis (Ed.), *Planning and Task-Performance in a Second Language* (pp. 3-34). John Benjamins.

Ellis, R. (2005b). Principles of instructed language learning. *System, 33*(2), 209-224. https://doi.org/https://doi.org/10.1016/j.system.2004.12.006

Ellis, R. (2006). The methodology of task-based teaching. *Asian EFL Journal, 8*(3), 19-45.

Ellis, R. (2008). *The study of second language acquisition* (Second ed.). Oxford University Press.

Ellis, R. (2009). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19*, 221-246.

Ellis, R. (2015). *Understanding second language aquisition*. Oxford University Press.

Ellis, R. (2017). Task-based language teaching. In S. Loewen & S. Masatoshi (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 108-125). Routledge. https://doi.org/10.4324/9781315676968.ch7

Ellis, R., & Barkhuizen, G. (2005). *Analysing learner language*. Oxford University Press.

Ellis, R., & He, X. (1999). The roles of modified input and output in the incidental acquisition of word meanings. *Studies in Second Language Acquisition, 21*, 285-301.

Ellis, R., & Heimbach, R. (1997). Bugs and birds: Children's acquisition of second language vocabulary through interaction. *System, 25*, 247–259.

Ellis, R., Heimbach, R., Tanaka, Y., & Yamazaki, A. (1999). Modified input and the acquisition of word meanings by children and adults. In R. Ellis (Ed.), *Learning a second language through interaction* (pp. 63-114). John Benjamins Publishing Company.

Ellis, R., Loewen, S., & Erlam, R. (2006). Implicit and explicit corrective feedback and the acquisition of L2 grammar. *Studies in Second Language Acquisition, 28*, 339–368.

Ellis, R., & Shintani, N. (2014). *Exploring language pedagogy through second language acquisition research*. Routledge.

Ellis, R., Tanaka, Y., & Yamazaki, A. (1994). Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning, 44*, 449–491.

Erlam, R. (2003). Evaluating the relative effectiveness of structured-input and output-based instruction in foreign language learning: Results from an experimental study. *Studies in Second Language Acquisition, 25 (4)*, 559-582. https://doi.org/10.1017/S027226310300024X

Field, A. (2009). *Discovering statistics using SPSS*. Sage.

Foster, P., Tonkyn, A., & Wigglesworth, G. (2000). Measuring spoken language: A unit for all reasons. *Applied Linguistics, 21*, 354 – 375.

Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research, 20*(3), 321-340.

Gass, S. M. (1997). *Input, interaction, and the second language learner*. Lawrence Erlbaum Associates.

Gass, S. M. (2003). Input and interaction. In C. Doughty & M. H. Long (Eds.), *The handbook of second language acquisition* (pp. 224–255). Basil Blackwell.

Gass, S. M., Behney, J., & Plonsky, L. (2013). *Second language acquisition: An introductory course* (4th ed.). Routledge.

Gass, S. M., & Mackey, A. (2015). Input, interaction, and output in second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 179-206). Routledge.

Gass, S. M., Mackey, A., & Pica, T. (1998). The role of input and interaction in second language acquisition Introduction to the special issue. *The Modern Language Journal, 82*, 299-307. https://doi.org/10.1111/j.1540-4781.1998.tb01206.x

Gass, S. M., & Selinker, L. (2000). *Second language acquisition: An introductory course* (2nd ed.). Lawrence Erlbaum Associates.

Gass, S. M., & Selinker, L. (2008). *Second language acquisition: An introductory course* (3rd ed.). Routledge.

Gass, S. M., & Varonis, E. (1994). Input, interaction and second language production. *Studies in Second Language Acquisition, 16*, 283–302.

George, D., & Mallery, P. (2010). *SPSS for Windows step by step: A simple guide and reference 17.0 Update* (10th ed.). Pearson.

Götz, S. (2013). *Fluency in native and nonnative English speech (Vol. 53)*. John Benjamins.

Gravetter, F., & Wallnau, L. (2014). *Essentials of statistics for the behavioral sciences* (8th ed.). Wadsworth.

Griffiths, R. (1990). Speech rate and NNS comprehension: A preliminary study in time-benefit analysis. *Language Learning, 40*(3), 311-336.

Hermans, D., Bongaerts, T., K., d. B., & Schreuder, R. (1998). Producing words in a foreign language: Can speakers prevent interference from their first language. *Bilingualism: Language and Cognition, 3*, 213–229.

Hien, T. T. T., & Loan, M. T. (2019). Current challenges in the teaching of tertiary English in Vietnam. In J. Albright (Ed.), *English tertiary education in Vietnam*. Routledge.

Hilton, H. (2008). The link between vocabulary knowledge and spoken L2 fluency. *Language Learning Journal, 36*(2), 153-166.

Hoang, V. V. (2008). Những yếu tố ảnh hưởng đến chất lượng đào tạo tiếng Anh không chuyên ở Đại học Quốc gia Hà Nội (Factors affecting the teaching quality for non-English major students at Ha Noi National University. *Tạp chí Khoa học ĐHQGHA (Journal of Science, Ha Noi National University), 24*, 22–37.

Hoang, V. V. (2010). The current situations and issues of teaching of English in Vietnam. *Ritsumeikan Language Culture Study, 22*(1), 7-18.

Hoang, V. V. (2013). TRAO DOI /DISCUSSION: The role of English in the internationalization of higher education in Vietnam. *VNU Journal of Foreign Studies, 29*(1), 72-80.

Hoang, Y. P. (2014). *The impact of PPP and TBLT on Vietnamse students' writing performance and self-regulation* Univeristy of KU Lueven]. Belgium.

Hunter, A. M. (2017). *Fluency development in the ESL classroom: The impact of immediate task repetition and procedural repetition on learners' oral fluency* St Mary's University]. Twickenham: London.

Iwasaki, J., & Oliver, R. (2003). Chat-line interaction and negative feedback. *Australian Review of Applied Linguistics, 17*, 60-73.

Izumi, S. (2002). Output, input enhancement, and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition, 24*, 541-577.

Kahng, J. (2014). Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning, 64*, 809-854.

Kahng, J. (2018). The effect of pause location on perceived fluency. *Applied Psycholinguistics, 39*, 569–591.

Kahng, J. (2020). Explaining second language utterance fluency: Contribution of cognitive fluency and first language utterance fluency. *Applied Psycholinguistics, 41*(2), 457-480. https://doi.org/10.1017/S0142716420000065

Kang, E. Y. (2015). Promoting L2 vocabulary learning through narrow reading. *RELC Journal, 46*(2), 165-179.

Kimura, H., & Ssali, V. (2009). The case for combining narrow reading and listening. *The Language Teacher, 33*(6), 9-13.

Kormos, J. (1999). Monitoring and self-repair in L2. *Language Learning, 49*, 303-342.

Kormos, J. (2006). *Speech production and second language acquisition*. Lawrence Erlbaum Associates.

Kormos, J. (2011). Speech production and the cognition hypothesis. In P. Robinson (Ed.), *L2 task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 39–60). John Benjamins.

Kormos, J., & Dénes, M. (2004). Exploring measures and perceptions of fluency in the speech of second language learners. *System, 32*, 145–164. https://doi.org/10.1016/j.system.2004.01.001

Krashen, S. (1981). The case for narrow reading. *TESOL Newsletter, 15*, 23.

Krashen, S. (1982). *Principles and practice in second language acquisition*. Pergamon.

Krashen, S. (1985). *The Input Hypothesis: Issues and implications*. Longman.

Krashen, S. (1996). The case for narrow listening. *System, 24*, 97-100.

Krashen, S. (2003). *Explorations in language acquisition and use: The Taipei lectures*. Heinemann.

Krashen, S. (2008). The Comprehension Hypothesis Extended. In T. Piske & M. Young-Scholten (Eds.), *Input Matters in SLA* (pp. 81-94). Multilingual Matters.

Krashen, S. (2014). *Applying the Comprehension Hypothesis: Some suggestions* 13th International Symposium and Book Fair on Language Teaching, Taipei, Taiwan.

La Heij, W. (2005). Selection processes in monolingual and bilingual lexical access. In J. Kroll & A. M. B. de Groot (Eds.), *Handbook of bilingualism. Psycholinguistic approaches*. Oxford University Press.

Lambert, C. (2020). Frameworks for using tasks in second language instruction In C. Lambert & R. Oliver (Eds.), *Using tasks in second language teaching: Practice in diverse contexts*. Multilingual Matters.

Lambert, C., Aubrey, S., & Leeming, P. (2020). Task preparation and second language speech production. *TESOL Quarterly*. https://doi.org/doi:10.1002/tesq.598

Lambert, C., & Kormos, J. (2014). Complexity, cccuracy, and fluency in task-based L2 research: Toward more developmentally based measures of second language acquisition. *Applied Linguistics, 35*(5), 607–614. https://doi.org/10.1093/applin/amu047

Lambert, C., Kormos, J., & Minn, D. (2017). Task repetition and second language speech processing. *Studies in Second Language Acquisition, 39*(1), 167-196. https://doi.org/10.1017/S0272263116000085

Lambert, C., & Robinson, P. (2014). Learning to perform narrative tasks in the L2 classroom: A semester-long study of L2 sequencing effects. In M. Baralt, R. Gilabert, & P. Robinson (Eds.), *Task sequencing and instructed second language learning* (pp. 207-230). Bloomsbury Academic.

Le, P. H. H., & Yeo, M. (2016). Evaluating in-service training of primary English teachers: A case study in Central Vietnam. *The Asian EFL Journal Quarterly, 18*(1), 34-51.

Le, V. C. (2007). A historical review of English language education in Vietnam. In Y. H. Choi & B. Spolsky (Eds.), *English education in Asia: History and policies* (pp. 167-180). Asia TEFL.

Le, V. C. (2011). *Form-focused instruction: A case study of Vietnamese teachers' beliefs and practices* University of Waikato]. New Zealand.

Le, V. C. (2012). A theory-practice divide: EFL teachers' and students' beliefs about grammar instruction. *English Australia Journal, 27*(2), 34-49.

Le, V. C. (2015). English language education innovation for the Vietnamese secondary school: The Project 2020. In B. Spolsky & K. Sung (Eds.), *Secondary school English education in Asia: From policy to practice* (pp. 182-200). Taylor & Francis.

Le, V. C., & Barnard, R. (2009). Curricular innovation behind closed classroom doors: A Vietnamese case study. *Teacher's Edition (Vietnam), 24*(2), 20-33.

Lee, J. (2000). *Tasks and communicating in language classrooms*. McGraw-Hill.

Lennon, P. (1990). Investigating fluency in EFL: A quantitative approach. *Language Learning, 40*, 387–417.

Lennon, P. (2000). The lexical element in spoken L2 fluency. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 25-42). University of Michigan Press.

Leow, R. P. (1993). To simplify or not to simplify. *Studies in Second Language Acquisition, 15*(3), 333-355.

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

Levelt, W. J. M. (1995). The ability to speak: From intentions to spoken words. *European Review, 3*(1), 13-23.

Levelt, W. J. M. (1999). Producing spoken language: A blueprint of the speaker. In C. Brown & P. Hagoort (Eds.), *Neurocognition of language* (pp. 83-122). Oxford University Press.

Loewen, S., & Sato, M. (2017). Instructed second language acquisition (ISLA): An overview. In S. Loewen & M. Sato (Eds.), *The Routledge handbook of instructed second language acquisition* (pp. 1-12). Routledge.

Long, M. (1981). Input, interaction, and second language acquisition. *Annals of the New York Academy of Sciences, 379*, 259-278.

Long, M. (1983). Native speaker/non-native speaker conversation and the negotiation of comprehensible input. *Applied Linguistics, 4*(2), 126-141. https://doi.org/https://doi.org/10.1093/applin/4.2.126

Long, M. (1985a). Input and second language acquisition theory. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 377-393). Newbury House.

Long, M. (1985b). A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and asssessing second language acquisition*. Multilingual Matters.

Long, M. (1996). The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of Second Language Acquisition* (pp. 413-468). Academic Press.

Long, M. (2016). In defense of tasks and TBLT: Nonissues and real issues. *Annual Review of Applied Linguistics, 36 (2016)*, 5-33. https://doi.org/10.1017/S0267190515000057 (Cambridge University Press)

Long, M., & Ross, S. (1993). Modifications that preserve language and content. In M. Tickoo (Ed.), *Simplification: Theory and Application* SEAMEO Regional Language Centre.

Loschky, L. (1994). Comprehensible input and second language acquisition: What is the relationship? *Studies in Second Language Acquisition, 16*, 303-323.

Lynch, T. (2001). Seeing what they mean: transcribing as a route to noticing. *English Language Teaching Journal, 55*(2), 124-132.

Lynch, T., & Maclean, J. (2000). Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research, 4*(3), 221-250.

Mackey, A. (1999). Input, interaction and second language development. *Studies in Second Language Acquisition, 21*, 557–587.

Mackey, A. (Ed.). (2007). *Conversational interaction in second language acquisition: A collection of empirical studies*. Oxford University Press.

Mackey, A., Abbuhl, R., & Gass, S. (2012). Interactionist approach. In S. Gass & A. Mackey (Eds.), *Routledge handbook of second language acquisition* (pp. 7-23). Routledge.

Mackey, A., & Goo, J. (2007). Interaction research in SLA: A meta-analysis and research synthesis. In A. Mackey (Ed.), *Conversational interaction in second language acquisition: A collection of empirical studies* (pp. 407–452). Oxford University Press.

Mackey, A., Oliver, R., & Leeman, J. (2003). Interactional input and the incorporation of feedback: An exploration of NS-NNS and NNS-NNS adult and child dyads. *Language Learning, 53*(1), 35–66.

Mai, N. K., & Iwashita, N. (2012). A comparison of learners' and teachers' attitudes toward communicative language teaching at two universities in Vietnam. *University of Sydney Papers in TESOL, 7*, 25–49.

Manh, L. D., Nguyen, H. T. M., & Burns, A. (2017). Teacher Language Proficiency and Reform of English Language Education in Vietnam, 2008-2020. In D. Freeman & L. LeDrean (Eds.), *Developing classroom English competence: Learning from the Vietnam experience* (pp. 84-94). IDP Australia.

Mason, B., & Krashen, S. (1997). Extensive reading in English as a foreign language. *System, 25*(1), 91-102.

McDonough, K., & Mackey, A. (2006). Responses to recasts: Repetitions, primed production, and linguistic development. *Language Learning, 56*(4), 693–720. https://doi.org/doi:10.1111/j.1467-9922.2006.00393

McLaughlin, B. (1987a). The monitor model: Some methodological considerations. *Language Learning, 28*, 309-332.

McLaughlin, B. (1987b). *Theories of second language learning*. Edward Arnold.

McLaughlin, B. (1990). Restructuring. *Applied Linguistics, 11*, 113-128.

Mitchell, R., & Myles, F. (2004). *Second language learning theories* (2nd ed.). Hodder Arnold.

MOET. (2003). *Quyết định về việc ban hành chương trình môn ngoại ngữ (Tiếng Anh, tiếng Pháp) trung học chuyên nghiệp (Decision on implementing foreign language curricula for English and French in higher education)*. Ministry of Education and Training.

MOET. (2008). *The project: Teaching and learning English language in the national education system from 2008 to 2020*. Education Publishing House.

Newton, J., & Bui, T. (2020). Low-proficiency learners and task-based language teaching. In C. Lambert & R. Oliver (Eds.), *Using tasks in second language teaching: Practice in diverse contexts*. Blue Ridge Summit: Multilingual Matters.

Nguyen, C., & Le, D. (2013). Communicative language testing: Do school tests measure students' communicative competence?FLLT Conference Proceedings,

Nguyen, T. H. (2002). Vietnam: Cultural background for ESL/EFL teachers. *Review of Vietnamese Studies, 2*(1), 1-6.

Nguyen, T. M. H. (2011). Developing EFL learners' intercultural communicative competence: A gap to be filled? In R. Nunn & S. Sivasubramaniam (Eds.), *From defining EIL Competence to Developing EIL learning* (Vol. 21, pp. 86-99). Asian EFL Journal Press.

Nguyen, V. T., & Mai, N. K. (2015). Responses to a language policy: EFL teachers' voices. *The European Journal of Social and Behavioural Sciences, 13*(2), 1830-1841.

Nobuyoshi, J., & Ellis, R. (1993). Focused communication tasks and second language acquisition. *ELT Journal, 47*, 203-210. https://doi.org/http://dx.doi.org/10.1093/elt/47.3.203

Nunan, D. (1989). *Designing tasks for the communicative classroom*. Cambridge University Press.

Nunan, D. (2004). *Task-based language teaching*. Cambridge University Press.

Nunan, D. (2006). Task-based language teaching in the Asia context: Defining 'task'. *Asian EFL Journal, 8(3), 8*(3), 12-18.

Oanh, D. T. H., & Hien, N. T. (2006). Memorization and EFL students' strategies at university level in Vietnam. *TESL-EJ, 10*(2), 1-21.

Oliver, R. (1995). Negative feedback in child NS/NNS conversation. *Studies in Second Language Acquisition, 18*, 459-481.

Oliver, R. (1998). Negotiation of meaning in child interactions. *Modern Language Journal, 82*, 372-386.

Oliver, R. (2002). The patterns of negotiation for meaning in child interactions. *Modern Language Journal, 86*, 97-111.

Oliver, R., & Bogachenko, T. (2018). Teacher perception and use of tasks in school ESL classrooms. In M. Bygate (Ed.), *Task-based language teaching as a researched pedagogy* (pp. 71-95). Johns Benjamins. https://doi.org/10.1075/tblt.12.04oli

Parker, K., & Chaudron, C. (1987). *The effects of linguistic simplification and elaborative modifications on L2 comprehension* The 21st Annual TESOL Convention

Pham, H. H. (2004). *Trained in the West, teaching in the East: Vietnamese teachers returning from TESOL courses abroad* University of Melbourne].

Phuong-Mai, N., Terlouw, C., & Pilot, A. (2005). Cooperative learning vs Confucian heritage cul-ture's collectivism: Confrontation to reveal some cultural conflicts and mismatch. *Asia Europe Journal, 3*(3), 403–419.

Pica, T. (1996). The essential role of negotiation in the communicative classroom *JALT Journal, 78*, 241-268.

Pica, T., Holliday, L., Lewis, N., & Morgenthaler, L. (1989). Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition, 11*, 63-90.

Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and using communicative tasks for second language instruction. In G. Crookes & S. Gass (Eds.), *Tasks and language learning: Integrating theory and practice* (pp. 9-54). Multilingual Matters.

Pica, T., Young, R., & Doughty, C. (1987). The impact of interaction on comprehension. *TESOL Quarterly, 21*, 737–758.

Pienemann, M. (1998). *Language processing and second language development: Processability Theory*. John Benjamins.

Pienemann, M. (2005). An introduction to Processability Theory. In M. Pienemann (Ed.), *Cross-linguistic aspects of Processability Theory* (pp. 1-60). John Benjamins.

Poulisse, N. (1997). Language production in bilinguals. In A. de Groot & J. Kroll (Eds.), *Tutorials in bilingualism. Psycholinguistic perspectives* (pp. 201-224). Lawrence Erlbaum Associates.

Poulisse, N. (1999). *Slips of the tongue: Speech errors in first and second language production*. Benjamins.

Poulisse, N., & Bongaerts, T. (1994). First language use in second language production. *Applied Linguistics, 15*, 36-57.

Prabhu, N. S. (1987). *Second language pedagogy*. Oxford University Press.

Préfontaine, Y. (2013). Perceptions of French fluency in second language speech production. *Canadian Modern Language Review/La Revue Canadienne Des Langues Vivantes, 69*(3), 324–348. https://doi.org/10.3138/cmlr.1748

Rassaei, E. (2012). The effects of input-based and output-based instruction on L2 development. *TESL-EJ, 16*(3).

Riazantseva, A. (2001). Second language proficiency and pausing: A study of Russian speakers of English. *Studies in Second Language Acquisition, 23*, 497–526.

Richards, J., Platt, J., & Weber, H. (1985). *Longman dictionary of applied linguistics*. Longman.

Richards, J., & Rodgers, T. (2001). *Approaches and methods in language teaching* (Vol. 2nd). Cambridge University Press.

Riggenbach, H. (1991). Towards an understanding of fluency: A microanalysis of nonnative speaker conversation. *Discourse Processes, 14*, 423–441.

Rodrigo, V., Krashen, S., & Gribbons, B. (2004). The effectiveness of two comprehensible-input approaches to foreign language instruction at the intermediate level. *System, 32*, 53-60. https://doi.org/http://dx.doi.org/10.1016/j.system.2003.08.003

Rose, R. (2017). Silent and filled pauses and speech planning in first and second language production. In R. Eklund & R. Rose (Eds.), *Proceedings of DiSS 2017, Disfluency in Spontaneous Speech* (pp. 49-52). Royal Institute of Technology (KTH).

Rossiter, M. J. (2009). Perceptions of L2 fluency by native and non-native speakers of English. *The Canadian Modern Language Review, 65*(3), 395-412. https://doi.org/10.3138/cmlr.65.3.395

Saito, K., Ilkan, M., Magne, V., Tran, M., & Suzuki, S. (2018). Acoustic characteristics and learner profiles of low, mid and high-level second language fluency. *Applied Psycholinguistics, 39*(3), 593–617. https://doi.org/10.1017/S0142716417000571

Samuda, V. (2009). Guiding relationships between form and meaning during task performance: The role of the teacher. In K. Van den Branden, M. Bygate, & J. M. Norris (Eds.), *Task-based language teaching: A reader* (pp. 379-400). John Benjamins.

Schmidt, R. (1990). The role of consciousness in second language learning. *Schmidt, R, 11*(2), 129-158.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *AILA Review, 11*, 11-26.

Schmidt, R. (2001). Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge University Press.

Schmitt, N., & Carter, R. (2000). The lexical advantages of narrow reading for second language learners. *TESOL Journal, 9*, 4-9.

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

Segalowitz, N. (2013). Fluency. In P. Robinson (Ed.), *The Routledge encyclopaedia of second language acquisition* (pp. 240–244). Routledge.

Segalowitz, N., & Freed, B. F. (2004). Context, contact, and cognition in oral fluency acquisition: Learning Spanish in at home and study abroad contexts. *Studies in Second Language Acquisition, 26*, 173–200.

Sharwood Smith, M. (1993). Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition, 15*, 165–179.

Shehadeh, A. (2002). Comprehensible output, from occurrence to acquisition: An agenda for acquisitional research. *Language Learning, 52*, 597–647. https://doi.org/doi:10.1111/1467-9922.00196

Shintani, N. (2011). A comparative study of the effects of input-based and production-based instruction on vocabulary acquisition by young EFL learners. *Language Teaching Research, 15*, 137–158. https://doi.org/10.1177/1362168810388692

Shintani, N. (2012a). Input-based tasks and the acquisition of vocabulary and grammar: A process-product study. *Language Teaching Research, 16*(2), 253 - 279. https://doi.org/10.1177/1362168811431378

Shintani, N. (2012b). Repeating input-based tasks with young beginner learners. *RELC Journal, 43*(1), 39-51. https://doi.org/10.1177/0033688212439322

Shintani, N. (2016). *Input-based tasks in foreign language instruction for young learners*. John Benjamins Publishing Company.

Shintani, N., & Ellis, R. (2010). The incidental acquisition of English plural -s by Japanese children in comprehension-based and production-based lessons. *Studies in Second Language Acquisition, 32*, 607-637.

Shintani, N., Li, S., & Ellis, R. (2013). Comprehension-based versus production-based grammar instruction: A meta-analysis of comparative studies. *Language Learning, 63*(2), 296-329.

Skehan, P. (1995). Analysability, accessibility, and ability for use. In G. Cook & B. Seidlhofer (Eds.), *Principle and Practice in Applied Linguistics*. Oxford University Press.

Skehan, P. (1996). A framework for the implementation of task-based instruction. *Applied Linguistics, 17*, 38-62.

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

Skehan, P. (1998b). Task-based instruction. *Annual Review of Applied Linguistics, 18*, 268-286.

Skehan, P. (2003). Task-based instruction. *Language Teaching Research, 36*(1), 1-14.

Skehan, P. (2009). Modelling L2 performance: Integrating complexity, accuracy, fluency and lexis. *Applied Linguistics, 30*, 510-532. https://doi.org/10.1093/applin/amp047

Skehan, P. (2014). The context for researching a processing perspective on task performance. In P. Skehan (Ed.), *Processing perspectives on task performance* (pp. 1-26). John Benjamins.

Skehan, P., & Foster, P. (1999). The influence of task structure and processing conditions on narrative retellings. *Language Learning, 49*(1), 93-120.

Skehan, P., Foster, P., & Shum, S. (2016). Ladders and snakes in second language fluency. *International Review of Applied Linguistics in Language Teaching, 54*(2), 97-111. https://doi.org/https://doi.org/10.1515/iral-2016-9992

Smith, B. (2005). The relationship between negotiated interaction, learner uptake, and lexical acquisition in task-based computer-mediated communication. *TESOL Quarterly, 39*, 33-58.

Smith, G. (2015). *Combining input-and output-based instruction in second language learning* Concordia University]. Montréal, Québec, Canada.

Subramaniam, G. (2008). Confronting Asian concerns in engaging learners to online education. *International Education Studies, 1*(4), 10-18.

Swain, M. (1985). Communicative competence: Some roles of comprehensible input and comprehensible output in its development. In S. Gass & C. Madden (Eds.), *Input in second language acquisition* (pp. 235-253). Newbury House Publishers.

Swain, M. (1995). Three functions of output in second language learning. In G. Cook & B. Seidhofer (Eds.), *Principle and practice in applied linguistics* (pp. 125-144). Oxford University Press.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning*. Oxford University Press.

Swain, M. (2005). The output hypothesis: Theory and research. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (pp. 471–483). Lawrence Erlbaum Associates.

Swain, M., & Lapkin, S. (1995). Problems in output and the cognitive processes they generate: A step towards second language learning *Applied Linguistics, 16*(3), 371-391.

Swan, M. (2005). Legislating by hypothesis: the case of task-based instruction. *Applied Linguistics, 26*, 376–401.

Tabachnick, B. G., & Fidell, L. S. (2005). *Using multivariate statistics* (5th ed.). Allyn & Bacon.

Tanaka, T. (1999). The effect of combination of comprehension and production practice in grammar instruction. *JACET Bulletin, 30*, 119-133.

Tanaka, T. (2001). Comprehension and production practice in grammar instruction: Does their combined use facilitate second language acquisition? *JALT Journal, 23*, 6-30.

Tavakoli, P. (2011). Pausing patterns: differences between L2 learners and native speakers. *ELT Journal, 65*(1), 71-79. https://doi.org/https://doi.org/10.1093/elt/ccq020

Tavakoli, P. (2016). Fluency in monologic and dialogic task performance: challenges in defining and measuring L2 fluency. *International Review of Applied Linguistics in Language Teaching, 54*(2), 133--150. https://doi.org/https://doi.org/10.1515

Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a L2* (pp. 239-277). John Benjamins.

Thi Tuyet, T. (2013). What shapes the passiveness in learning of Vietnamese students? *VNU Journal of Education Research, 29*(2), 71-83.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Harvard University Press.

Towell, R., Hawkins, R., & Bazergui, N. (1996). The development of fluency in advanced learners of French. *Applied Linguistics, 17*(1), 84–119. https://doi.org/10.1093/applin/17.1.84

Van Avermalet, P., Colpin, M., Van Gorp, K., Bogaert, N., & Van den Branden, K. (2006). The role of a teacher in task-based language teaching. In K. Van den Branden (Ed.), *Task-based language education: From theory to practice* (pp. 175-196). Cambridge University Press.

Van den Branden, K. (1997). Effects of negotiation on language learners' output. *Language Learning, 47*, 589-636.

Van den Branden, K. (2009). Mediating between predetermined order and complete chaos. The role of the teacher in task-based language education. *International Journal of Applied Linguistics, 19*(3), 264–285.

Van den Branden, K. (2016). The role of teachers in task-based language education. *Annual Review of Applied Linguistics, 36*, 164-181. https://doi.org/10.1017/S0267190515000070

Van den Branden, K., Bygate, M., & Norris, J. M. (2009). *Task-based language teaching: A reader*. John Benjamins.

VanPatten, B. (1996). *Input processing and grammar instruction in second language acquisition*. Ablex.

VanPatten, B. (2003). *From input to output: A teacher's guide to second language acquisition*. McGraw-Hill.

VanPatten, B. (2004). Input processing in second language acquisition. In B. Van Patten (Ed.), *Processing Instruction: Theory, research, and commentary* (pp. 5–32). Erlbaum.

VanPatten, B., & Williams, J. (Eds.). (2015). *Theories in second language acquisition: An introduction* (Second Edition ed.). Routledge.

Wang, Z. (2014). 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). John Benjamins.

White, L. (1987). Against comprehensible input: The Input Hypothesis and the development of L2 competence. *Applied Linguistics, 8*, 95-110.

Willis, J. (1996). *A framework for task-based learning*. Longman.

Wong, J. K. (2004). Are the learning styles of Asian international students culturally or contextually based. *International Education Journal, 4*(4), 154-166.

Ziegler, N. (2016). Synchronous computer-mediated communication and interaction. *Studies in Second Language Acquisition, 38*, 553 – 586.

Appendix A: Participant Information Sheet

HREC Project Number:	HRE2017-0249
Project Title:	The effects of task-based instruction on Vietnamese EFL learners' speech production capacity
Chief Investigator:	Dr. Craig Lambert
Student researcher:	Thi Huyen Thanh Do
Version Number:	1
Version Date:	28/MAR/2017

What is the Project About?

This study investigates how task-based instruction affects Vietnamese EFL learners' speech processing capacity. The research will employ a mixed-method approach, namely a quasi-experimental pre-test/post-test design. 108 EFL learners will participate in the study. The study is intended to contribute to our understanding of the development of communicative ability through task-based language teaching. Pedagogically, this study can provide direction for teachers about how to sequence tasks to suit learners' level of proficiency; particularly to develop their speech processing capacity.

Who is doing the Research?

The project is being conducted by Thi Huyen Thanh Do, as part of her doctoral degree at Curtin University, Australia. The results of the research will be used by the researcher to obtain a Doctor of Philosophy at Curtin University.

Why am I being asked to take part?

You have been asked to participate in this project because you have the condition we are researching. Your participation will involve three testing sessions (pre-test, and two post-test) and three treatment lessons with the researcher over an five-week period. If you agree to participate in this research, you must be willing to attend all the testing sessions and the treatment lessons and do so punctually. The research will take place at a convenient location at Hanoi University of Industry, Vietnam.

There will be no costs to you and you will not be paid for participating in this project. At the end of the study, we will give you up to 150000VND to reimburse for your travel or motorbike parking tickets while you attend appointments for this research project.

What will I have to do?

First, we will ask you to complete a background questionnaire to provide information about your age, gender and English learning experience for research purposes only. All information that you provide in the questionnaire will be kept confidential. You need to complete the questionnaire once before the research starts with relevant information. Please do not provide extra information that is not asked. You need to complete the questionnaire at once and hand it to Thanh as soon as you finish.

Next, we will stratify you randomly into one of the three homogeneous groups of 36: two experimental groups, and a comparison group. Neither you nor the researcher can choose which group you go in. All three groups will attend three stages: (1) pre-test (Week 1), (2) experimental treatment (Weeks 2, 3, 4) and (3) post-tests (Weeks 5).

During the treatment period, you will receive a 90-minute lesson per week for three weeks (4.5 contact hours of instruction total) which will be taught by the researcher. The experimental groups will involve either input-based task instruction or combined input-output task-based instruction whereas the comparison group will be engaged in narrow reading and listening activities with content unrelated to the target structure to ensure all of you get the same amount of instruction time.

All the testing and treatment sessions will involve oral narrative tasks using parallel versions of similar stories based on the same characters. You will be given time to look at the pictures adapted from Lambert and Robinson (2014) and think about the stories and a one-sentence prompt will be used to establish the narrative discourse context. You will then narrate the stories based on your own linguistic resources. These narratives in all testing and treatment sessions will be audio recorded for research purposes only and will not be used to evaluate or identify you personally in any way.

Are there any benefits to being in the research project?

If you participate in the research project, it will be a good opportunity to learn English and improve your communicative ability. With your participation, we hope this research will bring good results, which will contribute to our understanding of the development of communicative ability through task-based language teaching and enable us to develop better L2 educational programs

Are there any risks, side-effects, discomforts or inconveniences from being in the research project?

Apart from giving up your time, we do not expect that there will be any risks or inconveniences associated with taking part in this research.

We have been careful to make sure that the research procedures do not cause you any distress. But, if you feel anxious while participating in this research, please let us know. We will try to find a way to help you with it.

We may find new information about the risks and benefits of this research as we progress. If this happens we will let you know the new information and what it means to you. It may be that this new information means that you can no longer be in the study or you may choose to keep going or to leave the study. If this happens, you will be asked to sign a new consent form to let us know you understand any new information we have told you.

Who will have access to my information?

The information collected in this research will be re-identifiable (coded). This means that the stored information will be re-identifiable which means we will remove identifying information on any data or sample and replace it with a code. Only the research team have access to the code to match your name if it is necessary to do so. Any information we collect will be treated as confidential. The following people will have access to the information we collect in this research: the research team and, in the event of an audit or investigation, staff from the Curtin University Office of Research and Development

How information will be stored?

Electronic data will be password-protected and hard copy data (including audio tapes) will be in locked storage.

How long the information will be stored and what happens at the end of the storage period?

This needs to comply with the data management policy for Curtin University and any collaborating institutions. Therefore, the information we collect in this study will be kept under secure conditions at Curtin University for 7 years after the research has ended and then it will be destroyed.

How do you plan to discuss or publish the results?

The results of this research may be presented at conferences or published in professional journals. You will not be identified in any results that are published or presented.

Will you tell me the results of the research?

We are not able to send you any results from this research as we do not collect any personal information to be able to contact you. The results of the research will be made available in future publications.

Do I have to take part in the research project?

Taking part in a research project is voluntary. It is your choice to take part or not. You do not have to agree if you do not want to. If you decide to take part and then change your mind, that is okay, you can withdraw from the project. You do not have to give us a reason; just tell us that you want to stop. Please let us know you want to stop so we can make sure you are aware of any thing that needs to be done so you can withdraw safely. If you choose not

to take part or start and then stop the study, it will not affect your relationship with the University, staff or colleagues. If you choose to leave the study we will use any information collected unless you tell us not to.

What happens next and who can I contact about the research?

If you decide to take part in this research we will ask you to sign the consent form. By signing it is telling us that you understand what you have read and what has been discussed. Signing the consent indicates that you agree to be in the research project and have your health information used as described. Please take your time and ask any questions you have before you decide what to do. You will be given a copy of this information and the consent form to keep. After signing the consent form, you will then need to fill in the background questionnaire. At the start of the questionnaire there is a checkbox to indicate that you have understood the information provided here in this information statement.

Who can I contact about the research?

If you have any questions regarding your participation in this research project, do not hesitate to let us know. Please address your questions to Craig Lambert through email at craig.lambert@curtin.edu.au or Thi Huyen Thanh Do at t.do21@postgraduate.curtin.edu.au

Curtin University Human Research Ethics Committee (HREC) has approved this study (HREC number XX/XXXX). Should you wish to discuss the study with someone not directly involved, in particular, any matters concerning the conduct of the study or your rights as a participant, or you wish to make a confidential complaint, you may contact the Ethics Officer on (08) 9266 9223 or the Manager, Research Integrity on (08) 9266 7093 or email hrec@curtin.edu.au.

Appendix B: Consent Form

HREC Project Number:	HRE2017-0249
Project Title:	The effects of task-based instruction on Vietnamese EFL learners' speech production capacity
Principal Investigator:	Dr. Craig Lambert
Student researcher:	Thi Huyen Thanh Do
Version Number:	1
Version Date:	28/MAR/2017

- I have read the information statement version listed above and I understand its contents.
- I believe I understand the purpose, extent and possible risks of my involvement in this project.
- I voluntarily consent to take part in this research project.
- I have had an opportunity to ask questions and I am satisfied with the answers I have received.
- I understand that this project has been approved by Curtin University Human Research Ethics Committee and will be carried out in line with the National Statement on Ethical Conduct in Human Research (2007).
- I understand I will receive a copy of this Information Statement and Consent Form.
- I agree to be audio-recorded for the research.

Participant Name	
Participant Signature	
Date	

Declaration by researcher: I have supplied an Information Letter and Consent Form to the participant who has signed above, and believe that they understand the purpose, extent and possible risks of their involvement in this project.

Researcher Name	
Researcher Signature	
Date	

Note: All parties signing the Consent Form must date their own signature.

Appendix C: Background Questionnaire

Your information in this questionnaire will be kept confidential and used for research purpose only.

Please fill in the questions below:

Name: _____

Age: _____

Phone number: _____

Email address: _____

Gender: Male ☐ Female ☐

Nationality: _____

Province of origin: _____

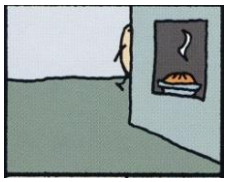
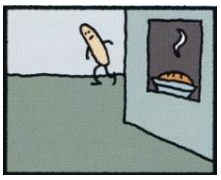
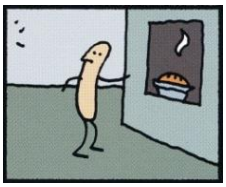
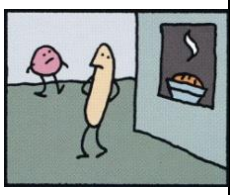
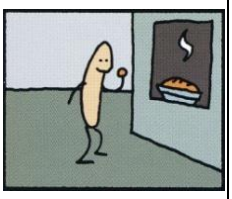
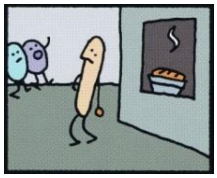
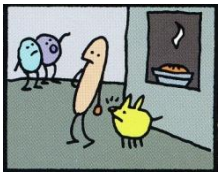
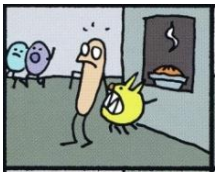
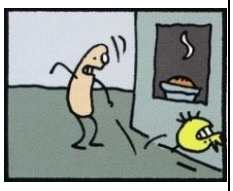
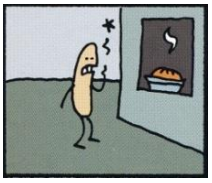
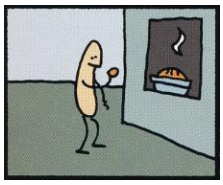
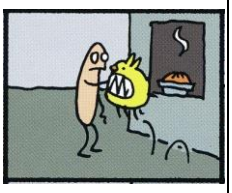
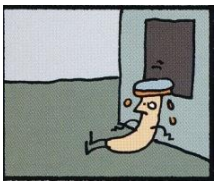
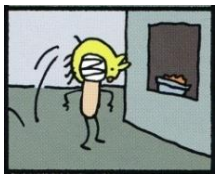
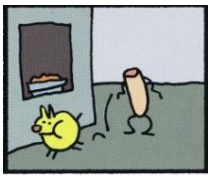
How old were you when you started learning English? _____ years old

How many years have you been learning English? _____ years

Appendix D: Pre-Test Materials

Look at the pictures and retell the story about Mr. I who gets in trouble for trying to steal some pie using your own English. Please start with the following sentence: *One day, Mr. I <u>was walking along</u> and he <u>passed</u> a fresh pie in someone's window.*

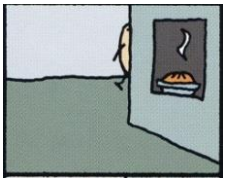
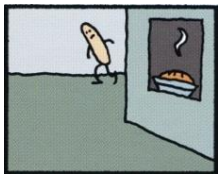
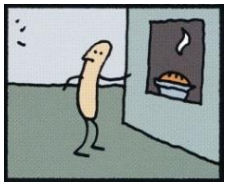
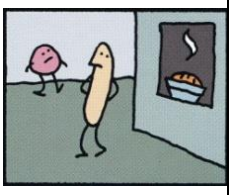
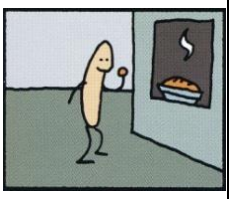
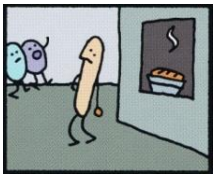
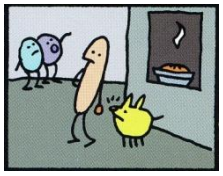
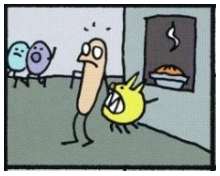
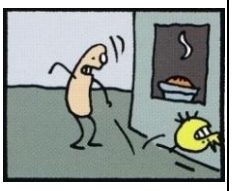
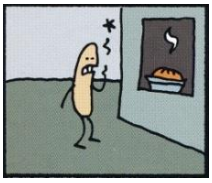
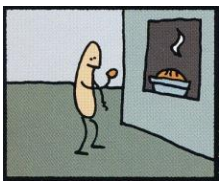
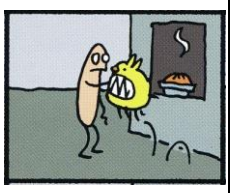
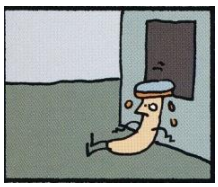
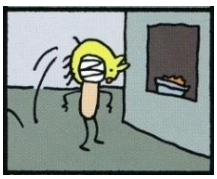
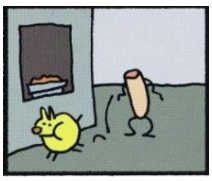
You have 3 minutes to prepare.

Part 1	1 	2 	3 	4 	5 
Part 2	6 	7 	8 	9 	10 
Part 3	11 	12 	13 	14 	15 

Appendix E: Post-Test Materials

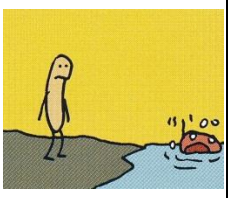
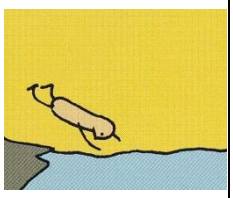
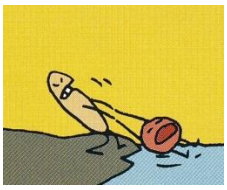
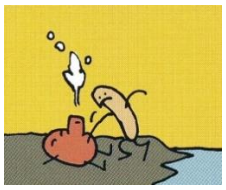
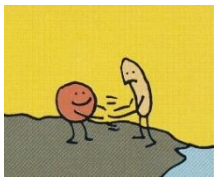
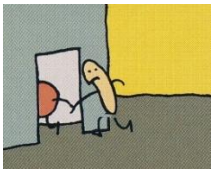
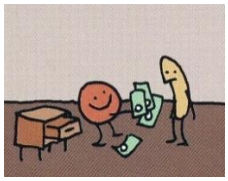
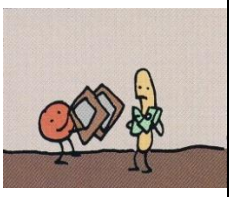
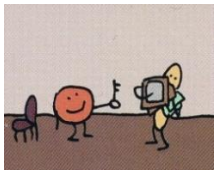
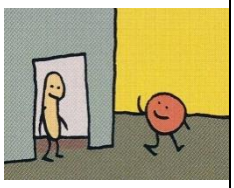
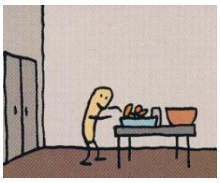
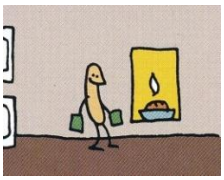
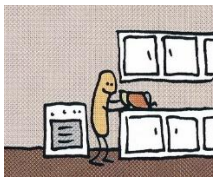
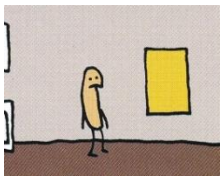
Task 1: Look at the pictures and retell the story about Mr. I who gets in trouble for trying to steal some pie using your own English. Please start with the following sentence: *One day, Mr. I was walking along and he passed a fresh pie in someone's window.*

You have 3 minutes to prepare.

Part 1	1 	2 	3 	4 	5 
Part 2	6 	7 	8 	9 	10 
Part 3	11 	12 	13 	14 	15 

Task 2: Look at the pictures and tell a story about another misadventure of Mr. I. Tell the story in English. Please start with the following sentence: One day, Mr. I was walking beside a lake and saw a man drowning.

You have 5 minutes to prepare.

Part 1	1 	2 	3 	4 	5 
Part 2	6 	7 	8 	9 	10 
Part 3	11 	12 	13 	14 	15 

Appendix F: Treatment Materials for Input Group

Input Group Lesson 1.1 (Story 2)

Pre-task 1 (5 minutes)

- First, listen to the teacher give a brief introduction about the story

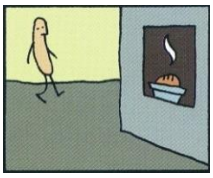
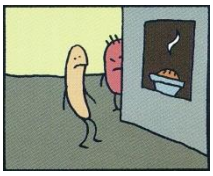
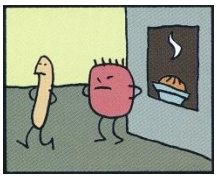
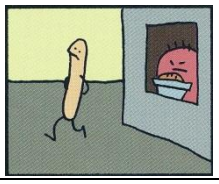
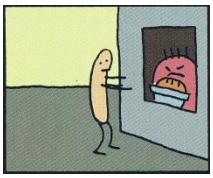
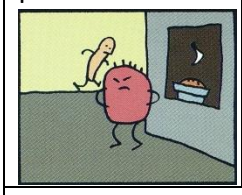
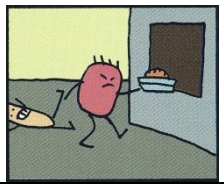
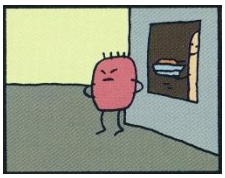
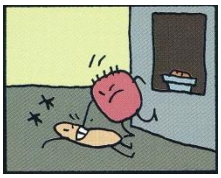
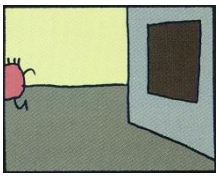
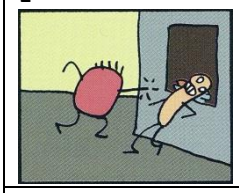
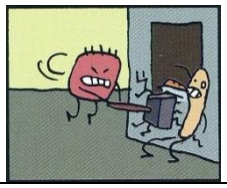
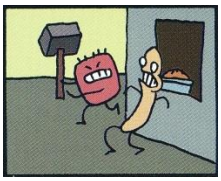
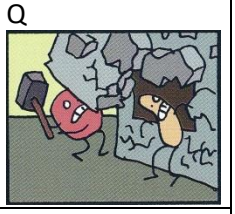
- Next listen to the recording of the story once to get some ideas what the story is about.

Task 1.1 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L__ 12_____ 13_____ 14_____ 15_____				

Now check the answers with your teacher. What is your score? ___ 12

Task 1.2 (15 minutes)

Read the following story and choose which of the underlined words you think is better to complete the story.

Put a tick (√) above the word(s) you choose.

One day, Mr. I was walking along and (1) passed/ was passing a fresh pie cooling in someone's window. He kept on walking past the house, but he (2) popped/ was popping his head back around the corner just to see the pie. He went to grab the pie, but the owner of the pie (3) looked/ was looking at him through the window, so he walked away until the owner (4) disappeared/ was disappearing. Later, he tried to come back to take the pie, (5) and/but the owner of the pie (6) appeared/ was appearing from around the corner, and Mr. I walked away again. The third time, while the owner of the pie was not paying attention, Mr. I ran behind the house, through the house, picked the pie up from the window, and (7) ran/ was running away. However, (8) while/because Mr. I was running, the owner of the pie saw this, ran after him, grabbed him by the legs, and (9) brought/ was bringing the pie back. Then, the owner of the pie punched Mr. I in the face repeatedly until he (10) hit/was hitting his hand against the wall. In which case, he went around the corner, grabbed a hammer to hit Mr. I. (11) Unfortunate/ Unfortunately, he smashed his own house instead. The house (12) collapsed/ was collapsing, crushing him and Mr. I, but leaving the pie unharmed.

Now listen to the story again and this time circle the word(s) that you hear.

Now check the answers with you teacher. What is your score? ___/12

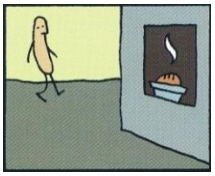
Task 1.3 (10 minutes)

Read this story and answer the questions below.

One day, Mr. I was walking along and passed a fresh pie cooling in someone's window. He kept on walking past the house, but he popped his head back around the corner just to see the pie. He went to grab the pie, but the owner of the pie was looking at him through the window, so he walked away until the owner disappeared. Later, he tried to come back to take the pie, but the owner of the pie appeared from around the corner, and Mr. I walked away again. The third time, while the owner of the pie was not paying attention, Mr. I ran behind the house, through the house, picked the pie up from the window, and ran away. However, while Mr. I was running, the owner of the pie saw this, ran after him, grabbed him by the legs, and brought the pie back. Then, the owner of the pie punched Mr. I in the face repeatedly until he hit his hand against the wall. In which case, he went around the corner, grabbed a hammer to hit Mr. I. Unfortunately, he smashed his own house instead. The house collapsed, crushing him and Mr. I, but leaving the pie unharmed.

Questions

1. How does the speaker open the story?

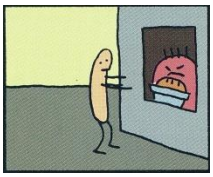


A. One day, Mr. I walked alone and he was passing a fresh pie cooling in someone's window.

B. One day, Mr. I walking alone and he passed a fresh pie cooling in someone's window.

C. One day, Mr. I was walking alone and he passed a fresh pie cooling in someone's window.

2. Which sentence best describes the picture below?

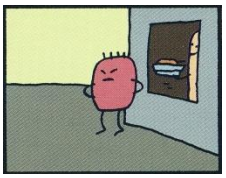


A. When he was going to grab the pie, the owner of the pie was looking at him through the window.

B. When he was go to grab the pie, the owner of the pie looked at him through the window.

C. When he went to grab the pie, the owner of the pie was looking at him through the window.

3. Which sentence best describes the picture below?

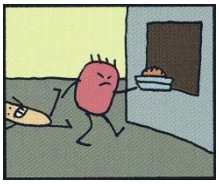


A. While the owner of the pie was not paying attention, Mr. I picked the pie up from the back of the window.

B. While the owner of the pie was not paying attention, Mr. I stood with his hands behind his back.

C. The owner of the pie stood with his hands behind his back until Mr. I was taking the pie and ran away.

4. Which sentence best describes the picture below?

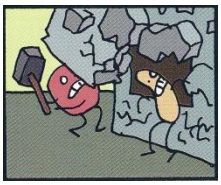


A. Mr. I was lying on the floor while the owner was eating the pie.

B. The owner of the pie grabbed Mr. I by the legs, and brought the pie back.

C. The owner of the pie was dancing with the pie when Mr. I brought the pie back.

5. Which sentence best describes the picture below?



A. Unfortunately, the owner of the pie smashed his own house, crushing himself and Mr. I in the collapsed house.

B. Unfortunately, Mr. I ran into the owner of the pie and smashed his house with a hammer.

C. In the end, while Mr. I was eating the pie, the owner of the pie hit him with a hammer.

Now check the answers with your teacher. What is your score? ___/5

Input Group Lesson 1.2 (Story 3)

Pre-task 2 (5 minutes)

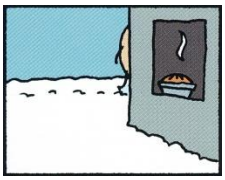
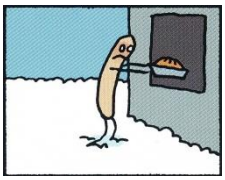
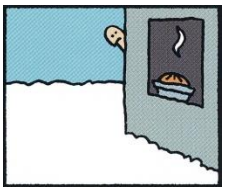
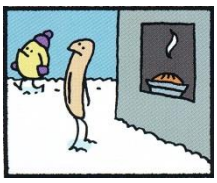
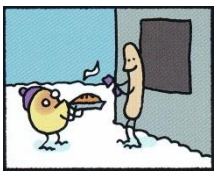
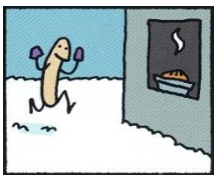
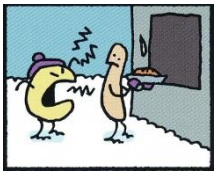
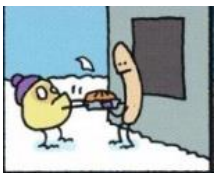
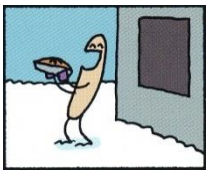
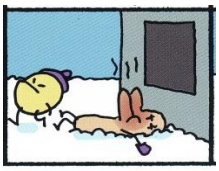
Listen to the teacher give a brief introduction about the story and listen to the recording of the story once to get some ideas what the story is about.

Task 2.1 (15 minutes)

- Look at the pictures for one minute and try to figure out what is in each picture.

- Listen to the recording of Story 1B again and sequence the pictures in each part of the story according to what you hear.

- Write a letter of the corresponding picture next to each number to show your answer. Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A__ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L__ 12_____ 13_____ 14_____ 15_____				

Now check the answers with your teacher. What is your score? ___ 12

Task 2.2 (15 minutes)

Read the story and guess which of the underlined words is better to complete the story.

Put a tick (√) above the word(s) you choose.

One snowy winter's day, Mr. I was walking along and he passed the pie in the window of a house. He (1) kept/ was keeping walking past, but he popped his head back around the corner just to see the pie. "Hang on, wait a minute" he (2) thought/ was thinking to himself "That's a hot pie and it's a very cold day!". So he came back through the snow and stood in front of the pie. There was no one there so he (3) decided/ was deciding to pick up the pie with both hands. "Oops, uh oh, that's a bit hot. It's a bit too hot. Ouch, it's very hot." Okay, he put it back on the shelf on the window. His hands were *incredibly* badly burnt. He (4) blew/ was blowing on his hands but it was not good enough, so he dug them into the snow to cool them down. (5) Then/ Before, he pulled his hands out of the snow but he still looked a bit disappointed. Suddenly, Mr I saw a yellow guy with snow gear walking past. "That's it, I've got a good idea" thought Mr. I, and (6) immediate/ immediately he ran off after the guy. Mr. I came back with a pair of gloves and a big smile on his face. With these gloves, he (7) managed/ was managing to pick up the pie. But *suddenly* the yellow guy came back over, and he looked very angry. While the yellow guy (8) shouted/ was shouting, Mr. I offered him some pie as some sort of consolation. The yellow guy considered the offer and said "Okay, yep, thank you, thank you very much, I'll have some pie."

However, as soon as the yellow guy took the pie from Mr I, he (9) realized/ was realizing that the pie was incredibly hot. His hands started burning too, so he (10) passed/ was passing the pie back to Mr. I, looking very unhappy. The yellow guy's hands were badly burnt, and Mr. I (11) smiled/ was smiling to himself. The yellow guy started shouting but Mr. I kept his cool. The yellow guy actually went away shouting. And Mr. I (12) laughed/ was laughing as he got the pie all to himself. He was about to take a big bite of the long-awaited pie and *suddenly* a big snow ball hit him in the face. He fell on the ground, and the pie (13) landed/ was landing on his stomach and started burning him. Mr. I started swearing and tried to lift the pie off his burnt stomach. However, the yellow guy (14) came/ was coming back and forced open Mr. I's mouth and stuffed the pie in. When his mouth started to burn, Mr. I struggled but the yellow guy was forced him to keep the pie in his mouth. Mr. I's lips were *badly* swollen and his arms and legs started falling. (15) Endly/ Finally, the yellow guy let him go but Mr. I was already dead. I guess that Mr. I got a taste of his own medicine.

Listen to the story again and circle the word(s) that you hear.
Now check the answers with your teacher. What is your score? ___/15

Task 2.3 (10 minutes)

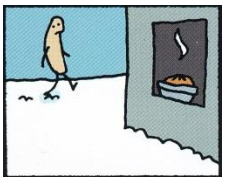
Read the story and answer the questions below.

One snowy winter's day, Mr. I was walking along and he passed the pie in the window of a house. He kept walking past, but he popped his head back around the corner just to see the pie. "Hang on, wait a minute" he thought to himself "That's a hot pie and it's a very cold day!". So he came back through the snow and stood in front of the pie. There was no one there so he decided to pick up the pie with both hands. "Oops, uh oh, that's a bit hot. It's a bit too hot. Ouch, it's very hot." Okay, he put it back on the shelf on the window. His hands were *incredibly* badly burnt. He blew on his hands but it was not good enough, so he dug them into the snow to cool them down. Then, he pulled his hands out of the snow but he still looked a bit disappointed. Suddenly, Mr I saw a yellow guy with snow gear walking past. "That's it, I've got a good idea" thought Mr. I, and immediately he ran off after the guy. Mr. I came back with a pair of gloves and a big smile on his face. With these gloves, he managed to pick up the pie. But *suddenly* the yellow guy came back over, and he looked very angry. While the yellow guy was shouting, Mr. I offered him some pie as some sort of consolation. The yellow guy considered the offer and said "Okay, yep, thank you, thank you very much, I'll have some pie."

However, as soon as the yellow guy took the pie from Mr I, he <u>realized</u> that the pie was incredibly hot. His hands started burning too, so he <u>passed</u> the pie back to Mr. I, looking very unhappy. The yellow guy's hands were badly burnt, and Mr. I <u>was smiling</u> to himself. The yellow guy started shouting but Mr. I kept his cool. The yellow guy actually went away shouting. And Mr. I <u>was laughing</u> as he got the pie all to himself. When he was about to take a big bite of the long-awaited pie, *suddenly* a big snow ball hit him in the face. He fell on the ground, and the pie <u>landed</u> on his stomach and started burning him. Mr. I started swearing and tried to lift the pie off his burnt stomach. However, the yellow guy <u>came</u> back and forced open Mr. I's mouth and stuffed the pie in. When his mouth started to burn, Mr. I struggled but the yellow guy was forced him to keep the pie in his mouth. Mr. I's lips were *badly* swollen and his arms and legs started falling. <u>Finall</u>y, the yellow guy let him go but Mr. I was already dead. I guess that Mr. I got a taste of his own medicine.

Questions

1. How does the speaker describe picture 1?



A. One snowy winter's day, Mr. I was walking along and he walked past a pie in the window of a house.

B. On a snowy winter's day, Mr. I was walking along and he walk past a pie in the window of a house.

C. One snowy winter's day, Mr. I walking along and he walked past a pie in the window of a house.

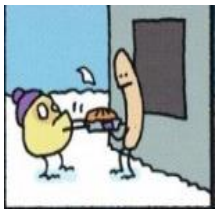
2. Which sentence describes the picture below?

A. There was no one there because he picked up the pie with both hands.

B. There was no one there so he picked up the pie with both hands.

C. There was no one there but he picked up the pie with both hands.

3. How does the speaker describe the picture below?



A. As soon as the yellow guy got the pie from Mr. I, he realized that the pie was incredibly hot.

B. As soon as the yellow guy get the pie from Mr. I, he realized that the pie was incredibly hot.

C. As soon as the yellow guy getting the pie from Mr. I, he realizes that the pie is incredibly hot.

4. Which sentence describes the picture below?

A. In spite of he was about to take a big bite of the long awaited pie, suddenly a big snow ball hit him in the face.

B. During he was about to take a big bite of the long awaited pie, suddenly a big snow ball hit him in the face.

C. When he was about to take a big bite of the long awaited pie, suddenly a big snow ball hit him in the face.

5. How does the speaker describe the last picture of the story?

A. Finally, the yellow guy lets him go but Mr. I was already dead.

B. Finally, the yellow guy let him go but Mr. I was already dead.

C. Finally, the yellow guy was let him go but Mr. I was already dead.

Now check the answers with your teacher. What is your score? ___/5

Input Group Lesson 2.1 (Story 4)

Pre-task 1 (5 minutes)

- First, listen to the teacher give a brief introduction about the story

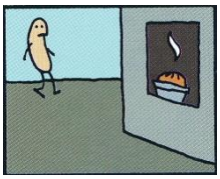
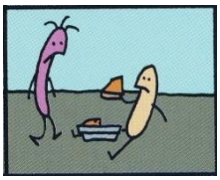
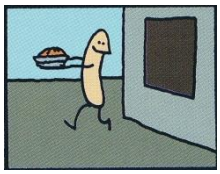
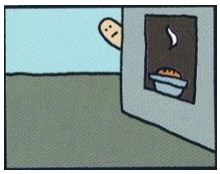
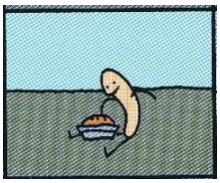
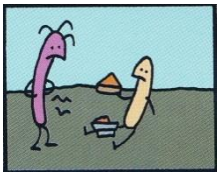
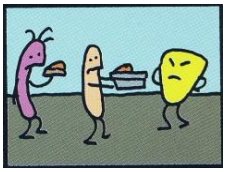
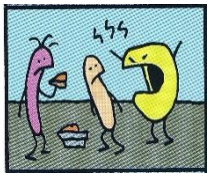
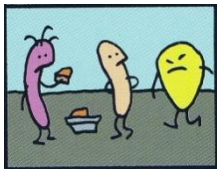
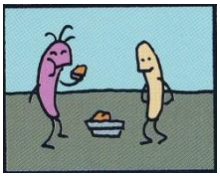
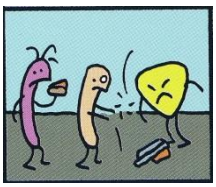
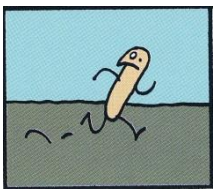
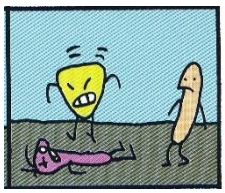
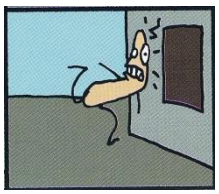
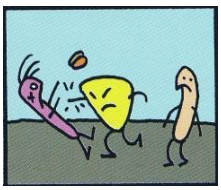
- Next listen to the recording of the story once to get some ideas what the story is about.

Task 1.1 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11__L__ 12_____ 13_____ 14_____ 15____				

Now check the answers with your teacher. What is your score? ___ 12

Task 1.2 (15 minutes)

Read the story and guess which of the underlined words is better to complete the story.
Put a tick (√) above the word(s) you choose.

One day, Mr. I (1) walked/was walking along and passed a fresh pie in someone's window. He kept on walking past the house, (2) but/ so he popped his head back around the corner to see the pie. He (3) noticed/ was noticing that no one was there. So he grabbed the pie and ran away quickly. When he was far away from the house, he (4) put/ was putting the pie on the ground, sat down and began to eat it. (5) As soon as/ Despite he took the first piece of pie and was about to eat it, however, a very thin man walked up. The man's stomach (6) growled/ was growling, and Mr. I thought he must be very hungry so Mr. I offered him a piece of the pie. The man (7) took/ was taking the pie and started eating it. (8) Sudden/ Suddenly, a very angry yellow man, who was the true owner of the pie appeared. He started shouting at Mr I and the thin man. Mr. I (9) realised/ was realising that he must be the rightful owner of the pie, so Mr. I (10) offered/ was offering to give him back the pie. But the man knocked it to ground and (11) punched/ was punching the hungry man in the face. He (12) knocked/ was knocking the man down on the ground and started beating him and jumping on him. Mr. I (13) managed/ was managing to escape during all of these commotions. However, as he (14) ran/ was running away, he looked back to see if anyone was following him and (15) ran/ was running into a wall knocking himself silly. This goes to show that crime doesn't pay.

Listen to the story again and circle the word(s) that you hear.
Now check the answers with you teacher. What is your score? ___/15

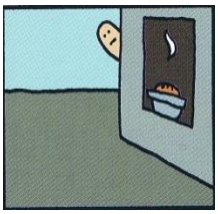
Task 1.3 (10 minutes)

Read the story and answer the questions below.

One day, Mr. I was walking along and passed a fresh pie in someone's window. He kept on walking past the house, but he popped his head back around the corner to see the pie. He noticed that no one was there. So he grabbed the pie and ran away quickly. When he was far away from the house, he put the pie on the ground, sat down and began to eat it. As soon as he took the first piece of pie and was about to eat it, however, a very thin man walked up. The man's stomach was growling, and Mr. I thought he must be very hungry so Mr. I offered him a piece of the pie. The man took the pie and started eating it. Suddenly, a very angry yellow man, who was the true owner of the pie, appeared. He started shouting at Mr. I and the thin man. Mr. I realised that he must be the rightful owner of the pie, so Mr. I offered to give him back the pie. But the man knocked it to ground and punched the hungry man in the face. He knocked the man down on the ground and started beating him and jumping on him. Mr. I managed to escape during all of these commotions. However, as he was running away, he looked back to see if anyone was following him and ran into a wall knocking himself silly. This goes to show that crime doesn't pay.

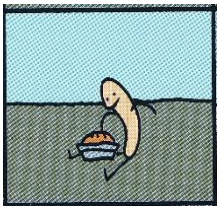
Questions

1. How does the speaker describe the picture below?



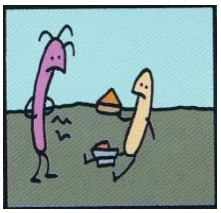
A. He walking past the house, but he popped his head back around the corner to see the pie.
B. He kept on walking past the house, but he popped his head back around the corner to see the pie.
C. He kept on walking past the house, but he popping his head back around the corner to see the pie.

2. Which sentence best describes the picture below?



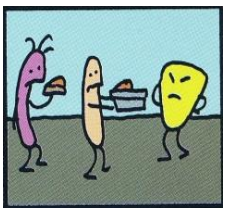
A. When he goes far away from the house, he put it on the ground, sat down and began to eat it.
B. When he was far away from the house, he was putting it on the ground, sitting down and began to eat it.
C. When he was far away from the house, he put it on the ground, sat down and began to eat it.

3. Which sentence best describes the picture below?



A. The man's stomach was growling, so Mr. I offered him a piece of the pie.
B. When the man's stomach growling, Mr. I offered him a piece of the pie.
C. Because the man's stomach growled, so Mr. I offered him a piece of the pie.

4. How does the speaker describe the picture below?

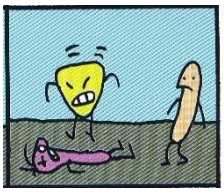


A. Mr. I realising that he must be the rightful owner of the pie, so Mr. I offered to give him back the pie.

B. Mr. I realised that he must be the rightful owner of the pie, because Mr. I offered to give him back the pie.

C. Mr. I realised that he must be the rightful owner of the pie, so Mr. I offered to give him back the pie.

5. Which sentence best describes the picture below?



A. While he knocked the man down on the ground and started beating him and jumping on him.
B. He knocked the man down on the ground and started beating him and jumping on him.
C. He was knocking the man down on the ground and started beating him and jumped on him.

Now check the answers with you teacher. What is your score? ___/5

Input Group Lesson 2.2 (Story 5)

Pre-task 2 (5 minutes)

- First, listen to the teacher give a brief introduction about the story

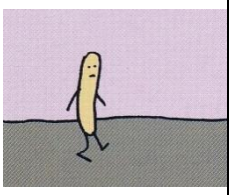
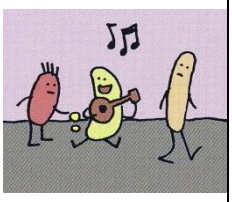
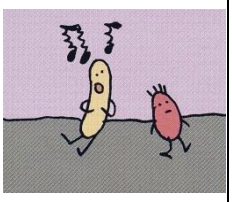
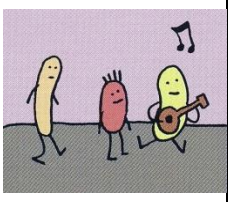
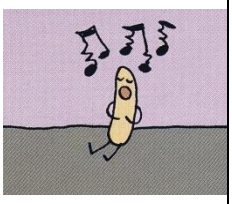
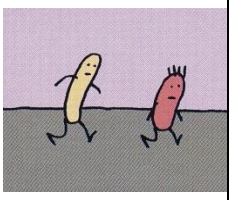
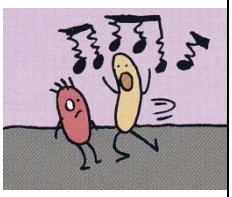
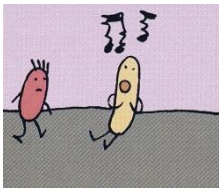
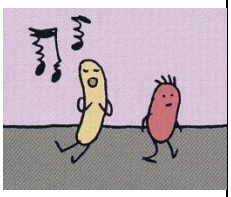
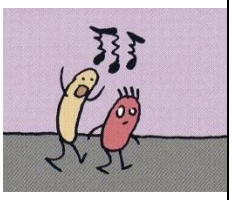
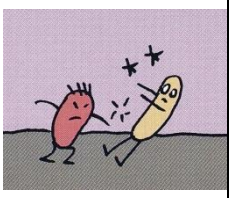
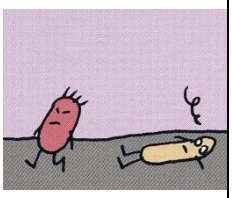
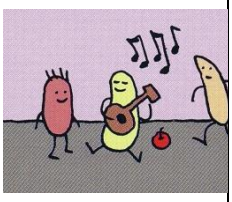
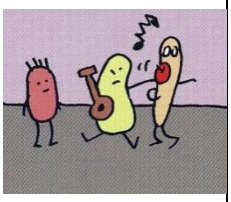
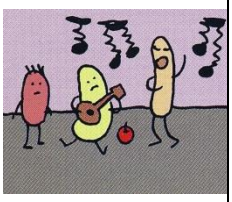
- Next listen to the recording of the story once to get some ideas what the story is about.

Task 2.1 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F___ 7_____ 8_____ 9_____ 10____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11__L__ 12_____ 13_____ 14_____ 15____				

Now check the answers with your teacher. What is your score? ___ 12

Task 2.2 (15 minutes)

Read the story and guess which of the underlined words is better to complete the story.

Put a tick (√) above the word(s) you choose.

One day, Mr. I (1) walked/was walking alone, and he came across a street singer. He watched and (2) noticed/ was noticing that the singer received some money for his singing. Mr. I was hungry and (3) thought/ was thinking this would be a great way to get some money of his own to buy his own food. So he (4) sat/ was sitting down and tried to sing as well. (5) However/ Although, when the man who had given the street singer money before came past Mr. I, he just (6) walked/ was walking by without giving Mr. I anything. Mr. I was very surprised, so he (7) chased / was chasing after him, passed him, and sat down in front of him and tried to sing for him again. (8) But/ Therefore the man just walked past and ignored him completely. When Mr. I saw this, he (9) ran/ was running up behind the man and (10) started/ was starting singing even louder so that the man could not help but hear him. The man tried to escape from him but Mr. I chased after him, blocked his way, and (11) sang/ was singing very loudly in the hope of getting some money from him. (12) Unfortunate/ Unfortunately, the man got angry and punched him. Mr I was knocked down on the ground and the man got away. Later, when Mr. I got up and collected himself, he (13) came/ was coming across the street singer and the man again. This time he tried joining in with them, but as he (14) sang/ was singing, the street singer jammed an apple down his throat to get him to shut up. (15) In the end/ At the end, Mr. I got his apple but this goes to show that no one likes a copycat.

Listen to the story again and circle the word(s) that you hear.

Now check the answers with you teacher. What is your score? ___/15

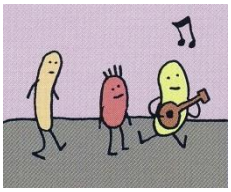
Task 2.3 (10 minutes)

Read the story and answer the questions below.

One day, when Mr. I was walking alone, he came across a street singer. He watched and noticed that the singer received some money for his singing. Mr. I was hungry and thought this would be a great way to get some money of his own to buy his own food. So he sat down and tried to sing as well. However, when the man who had given the street singer money before came past Mr. I, he just walked by without giving Mr. I anything. Mr. I was very surprised, so he chased after him, passed him, and sat down in front of him and tried to sing for him again. But the man just walked past and ignored him completely. When Mr. I saw this, he ran up behind the man and started singing even louder so that the man could not help but hear him. When Mr. I was singing louder and louder, the man tried to escape from him. This time Mr. I chased after him, blocked his way, and sang very loudly again with the hope to get some money from him. The man tried to escape from him but Mr. I chased after him, blocked his way, and sang very loudly in the hope of getting some money from him. Unfortunately, the man got angry and punched him. Mr I was knocked down on the ground and the man got away. Later, when Mr. I got up and collected himself, he came across the street singer and the man again. This time he tried joining in with them, but as he was singing, the street singer jammed an apple down his throat to get him to shut up. In the end, Mr. I got his apple but this goes to show that no one likes a copycat.

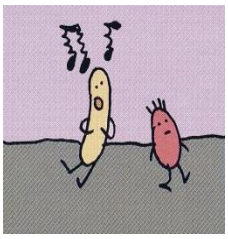
Questions

1. How does the speaker open the story?



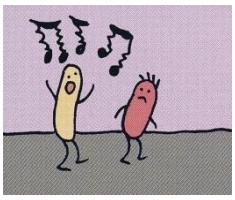
A. One day, Mr. I walking alone when came across a street singer.
B. One day, when Mr. I was walking alone, he came across a street singer.
C. One day, when Mr. I walked alone, he was coming across a street singer.

2. How does the speaker describe the picture below?



A. When the man who given the street singer the money came past, he just walked by without giving Mr. I anything.
B. When the man who was giving the street singer the money came past, he just walked by without giving Mr. I anything.
C. When the man who had given the street singer the money came past, he just walked by without giving Mr. I anything.

3. Which sentence best describes the picture below?

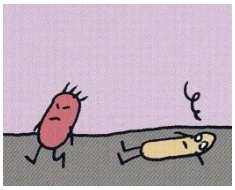


A. When Mr. I was singing louder and louder, the man tried to escape from him.

B. During Mr. I was singing louder and louder, the man tried escaping to escape from him.

C. When Mr. I sang louder and louder, the man was trying to escape from him.

4. Which sentence best describes the picture below?

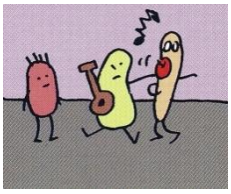


A. Mr. I was knocked down on the floor and the man got away.

B. Mr. I was knocking down on the floor so the man got away.

C. Mr. I was knocked down on the floor because the man got away.

5. How does the speaker describe the last picture of the story?



A. When he is singing, the street singer jam an apple down his throat.

B. As he sang, the street singer was jamming an apple down his throat.

C. As he was singing, the street singer jammed an apple down his throat.

Now check the answers with your teacher. What is your score? ___/5

Input Group Lesson 3.1 (Story 6)

Pre-task 1 (5 minutes)

- First, listen to the teacher give a brief introduction about the story

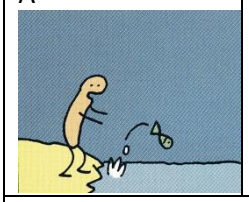
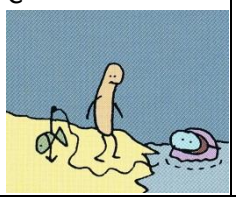
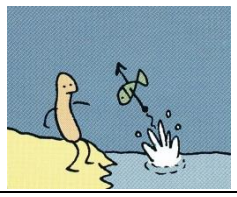
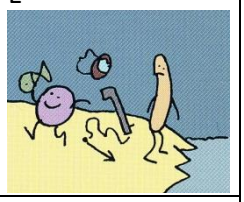
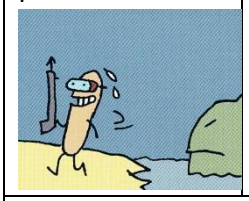
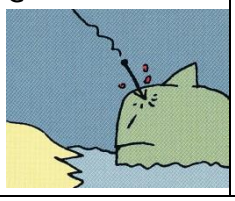
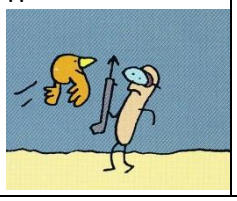
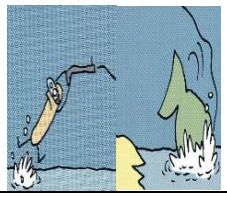
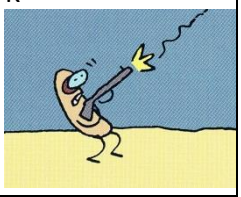
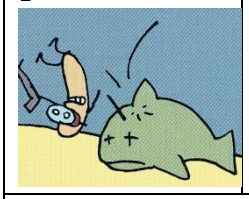
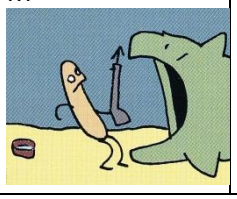
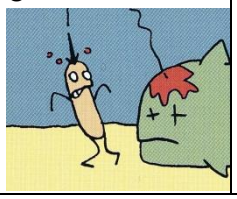
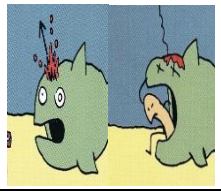
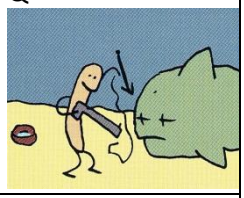
- Next listen to the recording of the story once to get some ideas what the story is about.

Task 1.1 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L___ 12_____ 13_____ 14_____ 15____				

Now check the answers with your teacher. What is your score? ___ 12

Task 1.2 (15 minutes)

Read the story and guess which of the underlined words is better to complete the story.

Put a tick (√) above the word(s) you choose.

One day, when Mr. I (1) stood/ was standing on the beach, he saw a fish jump. He tried reaching out for the fish because he was hungry, (2) but/and he couldn't catch it. Then suddenly the fish was harpooned and landed on the beach. From the sea, a man with a snorkel mask appeared. The man walked out of the sea and (3) jettisoned/ was jettisoning his scuba mask and his harpoon gun, but carried the fish away. Mr. I (4) thought/ was thinking this could be a good way to catch a fish of his own (5) so/ because he picked up his snorkel mask and the harpoon gun, and went fishing. When he (6) met/ was meeting face-to-face with a giant whale, however, he lost all of his confidence, turned around and ran away. Then, a bird flew past him, and he got an idea. He (7) thought/ was thinking the bird was an easier target than the whale so he shot at it instead. (8) However/ Therefore, he missed the bird and (9) hit/ was hitting the whale in the head with the harpoon. The whale dove into the sea, and Mr. I was dragged along as well.

The whale jumped up and down in the sea and finally (10) landed/ was landing on the beach, knocking itself out. At the same time, Mr. I crashed onto the beach next to the harpooned whale. As the whale appeared to be dead, Mr. I (11) took/ was taking off his mask and removing the harpoon when the whale suddenly opened his mouth wide and (12) gobbled/ was gobbling Mr. I up. From inside, Mr. I shot the harpoon through the roof of the whale's mouth and (13) managed/ was managing to climb out. He thought he was finally safe, but (14) unfortunate/ unfortunately, the harpoon he shot out of the whale's head (15) came/ was coming back down to earth and hit Mr. I. This goes to show that is is not good to copy everyone you see.

Listen to the story again and circle the word(s) that you hear.

Now check the answers with you teacher. What is your score? ___/15

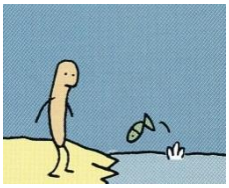
Task 1.3 (10 minutes)

Read the story and answer the questions below.

One day, when Mr. I was standing on the beach, he saw a fish jump. He tried reaching out for the fish because he was hungry, but he couldn't catch it. Then suddenly the fish was harpooned and landed on the beach. From the sea, a man with a snorkel mask appeared. The man walked out of the sea and jettisoned his scuba mask and his harpoon gun, but carried the fish away. Mr. I thought this would be a good way to catch a fish of his own. So he picked up his snorkel mask and the harpoon gun, and went fishing. When he met face-to-face with a giant whale, however, he lost all of his confidence, turned around and went away. Then, a bird flew past him, and he got an idea. He thought the bird was an easier target than the whale. So he shot at it instead. However, he missed the bird and hit the whale in the head with a harpoon. The whale dove into the sea, and Mr. I was dragged along as well. The whale jumped up and down in the sea and finally landed on the beach, knocking itself out. At the same time, Mr. I crashed onto the beach next to the harpooned whale. As the whale appeared to be dead, Mr. I was taking off his mask and removing his harpoon when the whale suddenly opened his mouth wide and gobbled Mr. I up. From inside, Mr. I shot the harpoon through the roof of the whale's mouth and managed to climb out. He thought he was finally safe but unfortunately, the harpoon he shot out of the whale's head came back down to earth and hit Mr. I. This goes to show that it is not good to copy everyone your see.

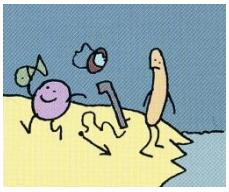
Questions

1. How does the speaker describe the picture below?



A. One day, Mr. I was stand on the beach when saw a fish jump.
B. One day, Mr. I stood on the beach and seeing a fish jump.
C. One day, when Mr. I was standing on the beach, he saw a fish jump.

2. Which sentence describes the picture below?

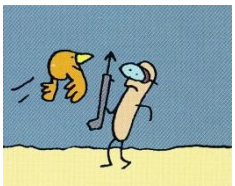


A. The man jettisoned his scuba mask and his harpoon gun, but carried the fish away.
B. The man was jettisoning his scuba mask and his harpoon gun, but carrying the fish away.
C. The man was jettisoned his scuba mask and his harpoon gun, but carried the fish away.

3. How does the speaker describe the picture below?

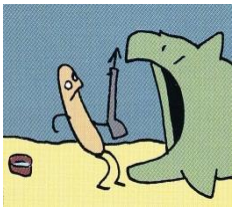
A. Until he met face-to-face with a giant whale, he lost all of his confidence, turned around and walked away.
B. When he met face-to-face with a giant whale, he lost all of his confidence, turned around and walked away.
C. When he met face-to-face with a giant whale, he was losing all of his confidence, turned around and walked away.

4. How does the speaker describe the picture below?



A. He thought the bird was an easier target than the whale.
B. He thought the bird is an easier target than the whale.
C. He thinking the bird was an easier target than the whale.

5. How does the speaker describe the picture below?



A. When the whale was suddenly opening his mouth wide and gobbled Mr. I up, Mr. I took off his mask and removed his harpoon.
B. Mr. I was taking off his mask and removing his harpoon because the whale suddenly opened his mouth wide and gobbled Mr. I up.
C. Mr. I was taking off his mask and removing his harpoon when the whale suddenly opened his mouth wide and gobbled Mr. I up.

Now check the answers with you teacher. What is your score? ___/5

Input Group Lesson 3.2 (Story 7)

Pre-task 2 (5 minutes)

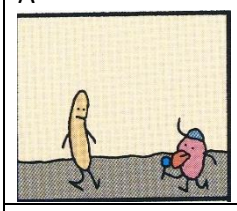
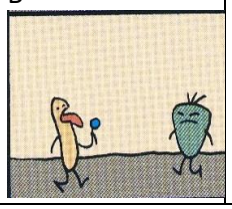
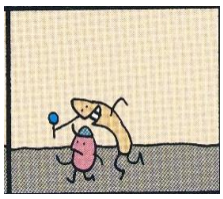
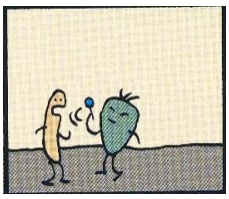
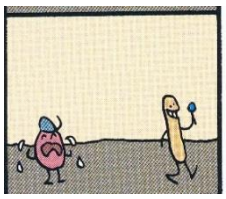
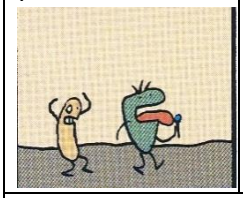
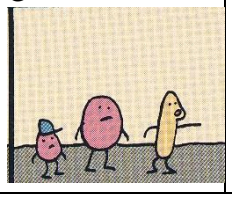
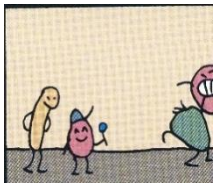
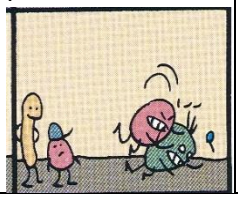
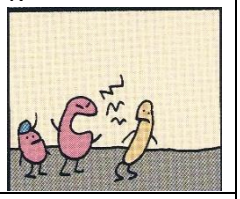
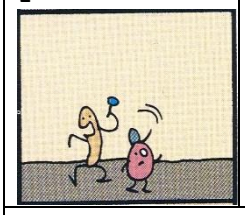
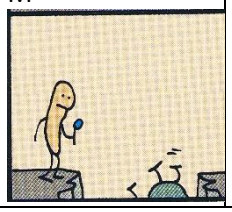
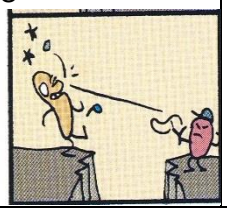
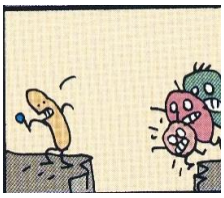
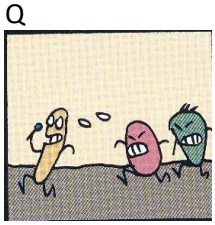
Listen to the teacher give a brief introduction about the story and listen to the recording of the story once to get some ideas what the story is about.

Task 2.1 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L___ 12_____ 13_____ 14_____ 15____				

Now check the answers with your teacher. What is your score? ___ 12

Task 2.2 (15 minutes)

Read the story and guess which of the underlined words is better to complete the story.

Put a tick (√) above the word(s) you choose.

One day, Mr. I (1) walked/was walking along and came across the boy with a lollipop. The boy (2) enjoyed/ was enjoying sucking his lollipop. But Mr. I (3) thought/ was thinking it looked very good so he started following behind the boy and took it away from him. The boy started shouting and crying (4) but/due to Mr I ran away with the lollipop. As Mr. I was enjoying the lollipop, he (5) ran/ was running into a green triangular man who grabbed the lollipop away from him in turn. (6) Although/ Despite Mr. I started shouting, stomping and stomping his feet, the green man (7) walked/ was walking away with the lollipop, leaving him there alone. (8) Sudden/Suddenly, the boy's mother came up from behind and started shouting at him: "Why do you steal my son's lollipop?". Mr. I (9) explained/explaining that the green guy had taken it from him. So the son's mother (10) went/ was going over, grabbed the other man and started beating him. Meanwhile the little boy got the lollipop back and was watching the fight. When he (11) wasn't look/ wasn't looking, Mr. I grabbed the lollipop away and made a run for it. Both the boy's mother and the green man (12) came/ was coming chasing after Mr. I. When they arrived at a large gorge, Mr. I made a jump for it and successfully (13) reached/ was reaching the other side while the others fell into the gorge all except the little boy. Mr. I thought he was safe but as soon as he (14) saw/was seeing the little boy, the boy shot him with a slingshot so he (15) dropped/ was dropping the lollipop into the gorge. In the end, no one got the lollipop.

Listen to the story again and circle the word(s) that you hear.

Now check the answers with you teacher. What is your score? ___/15

Task 2.3 (10 minutes)

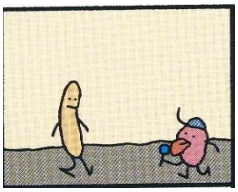
Read the story and answer the questions below.

One day, Mr. I was walking along and came across the boy with a lollipop. The boy was enjoying sucking his lollipop. But Mr. I thought it looked very good so he started following behind the boy and took it away from him. The boy started shouting and crying but Mr I ran away with the lollipop. As Mr. I was enjoying the lollipop, he ran into a green triangular man who grabbed the lollipop away from him in turn. Although Mr. I started shouting, stomping and stomping his feet, the green man walked away with the lollipop, leaving him there alone. Suddenly, the boy's mother came up from behind and started shouting at him: "Why do you steal my son's lollipop?". Mr. I explained that the green guy had taken it from him. So the son's mother went over, grabbed the other man and started beating him.

Meanwhile the little boy got the lollipop back and was watching the fight. When he wasn't looking, Mr. I grabbed the lollipop away and made a run for it. Both the boy's mother and the green man came chasing after Mr. I. When they arrived at a large gorge, Mr. I made a jump for it and successfully reached the other side while the others fell into the gorge all except the little boy. Mr. I thought he was safe but as soon as he saw the little boy, the boy shot him with a slingshot so he dropped the lollipop into the gorge. In the end, no one got the lollipop.

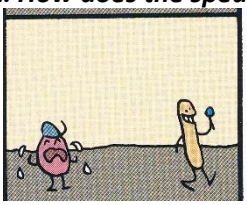
Questions

1. How does the speaker describe the picture below?



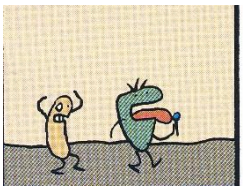
A. Mr. I walking along and came across the boy with a lollipop.
B. Mr. I walked along and come across the boy with a lollipop.
C. Mr. I was walking along and came across the boy with a lollipop.

2. How does the speaker describe the picture below?



A. The boy started shouting and crying but Mr. I kept walking away with the lollipop.
B. The boy started shouting and crying so Mr. I kept walking away with the lollipop.
C. The boy started shouting and crying although Mr. I kept walking away with the lollipop.

3. Which sentence best describes the picture below?

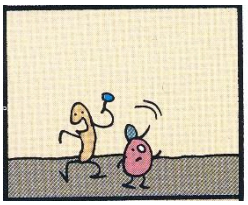


A. Because Mr. I start shouting, stomping his feet, the green man was walking away with the lollipop.

B. Although Mr. I started shouting, stomping his feet, the green man walked away with the lollipop.

C. Despite Mr. I started shouting, stomping his feet, the green man walked away with the lollipop.

4. Which sentence best describes the picture below?

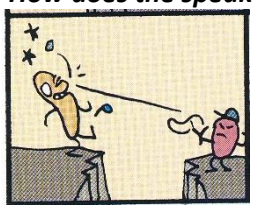


A. While the little boy was watching the fight, Mr. I grabbed the lollipop again and made a run for it.

B. During the little boy was watching the fight, Mr. I grabbed the lollipop again and made a run for it.

C. While the little boy watching the fight, Mr. I grabbed the lollipop again and made a run for it.

5. How does the speaker describe the last picture of the story?



A. While Mr. I saw the little boy, the boy was shooting him with a slingshot.

B. Because Mr. I saw the little boy, the boy shoot him with a slingshot.

C. As soon as Mr. I saw the little boy, the boy shot him with a slingshot.

Now check the answers with you teacher. What is your score? ___/5

Appendix G: Treatment Materials for Combined group

Combined group Lesson 1A (Student A, Story 2-3)

Pre-task 1.1 (5 minutes)

- First let's discuss the theme of the Story

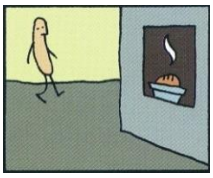
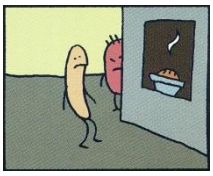
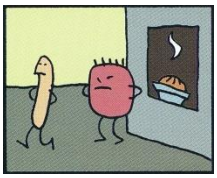
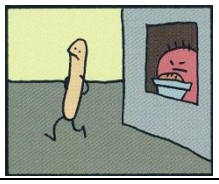
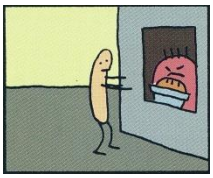
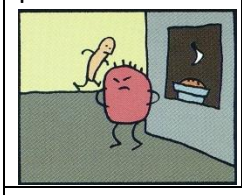
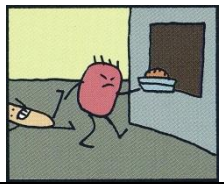
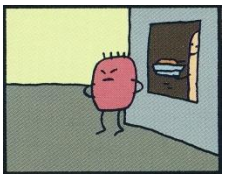
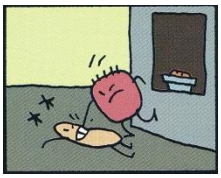
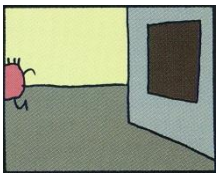
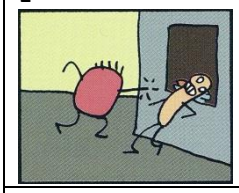
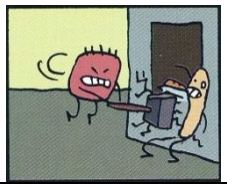
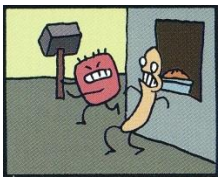
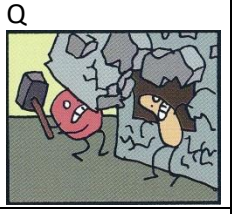
- Next listen to the recording to get some ideas about the story.

Task 1.1 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L__ 12_____ 13_____ 14_____ 15_____				

Now check the answers with your teacher. What is your score? ___ 12

Pre-task 1.2 (5 minutes)

- First let's discuss the theme of the Story

- Next listen to the recording to get some ideas about the story.

Task 1.2 (15 minutes)

Listen to the recording of the story again and identify the key events in the story (i.e., what happened in each step). You can take some notes, if you wish.

……..

……..

……..

……..

……..

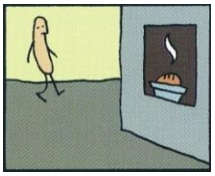
Task 2 (10 minutes)

Read this story and answer the questions below.

One day, Mr. I was walking along and passed a fresh pie cooling in someone's window. He kept on walking past the house, but he popped his head back around the corner just to see the pie. He went to grab the pie, but the owner of the pie was looking at him through the window, so he walked away until the owner disappeared. Later, he tried to come back to take the pie, but the owner of the pie appeared from around the corner, and Mr. I walked away again. The third time, while the owner of the pie was not paying attention, Mr. I ran behind the house, through the house, picked the pie up from the window, and ran away. However, while Mr. I was running, the owner of the pie saw this, ran after him, grabbed him by the legs, and brought the pie back. Then, the owner of the pie punched Mr. I in the face repeatedly until he hit his hand against the wall. In which case, he went around the corner, grabbed a hammer to hit Mr. I. Unfortunately, he smashed his own house instead. The house collapsed, crushing him and Mr. I, but leaving the pie unharmed.

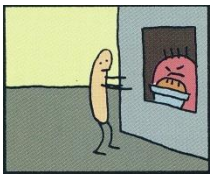
Questions

2. How does the speaker open the story?



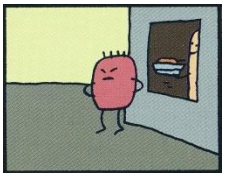
A. One day, Mr. I walked alone and he was passing a fresh pie cooling in someone's window.
B. One day, Mr. I walking alone and he passed a fresh pie cooling in someone's window.
C. One day, Mr. I was walking alone and he passed a fresh pie cooling in someone's window.

2. Which sentence best describes the picture below?



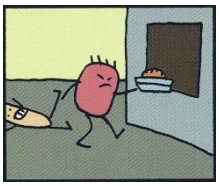
A. When he was going to grab the pie, the owner of the pie was looking at him through the window.
B. When he was go to grab the pie, the owner of the pie looked at him through the window.
C. When he went to grab the pie, the owner of the pie was looking at him through the window.

3. Which sentence best describes the picture below?



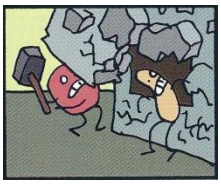
A. While the owner of the pie was not paying attention, Mr. I picked the pie up from the back of the window.
B. While the owner of the pie was not paying attention, Mr. I stood with his hands behind his back.
C. The owner of the pie stood with his hands behind his back until Mr. I was taking the pie and ran away.

4. Which sentence best describes the picture below?



A. Mr. I was lying on the floor while the owner was eating the pie.
B. The owner of the pie grabbed Mr. I by the legs, and brought the pie back.
C. The owner of the pie was dancing with the pie when Mr. I brought the pie back.

5. Which sentence best describes the picture below?



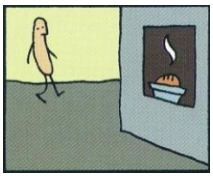
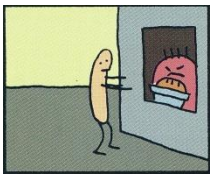
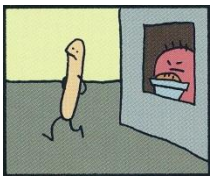
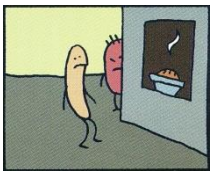
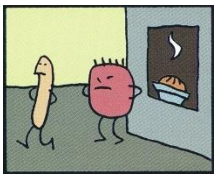
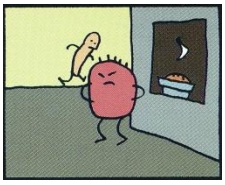
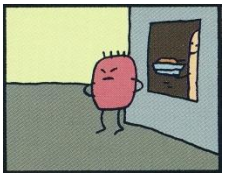
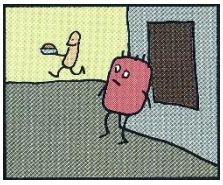
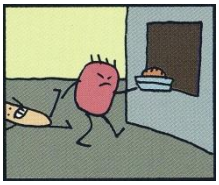
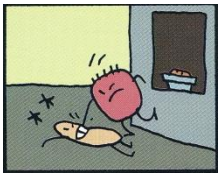
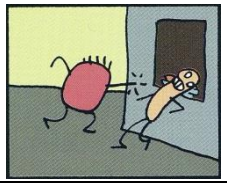
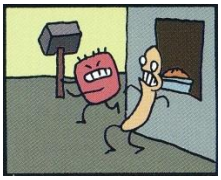
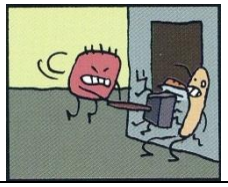
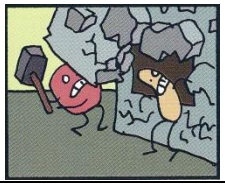
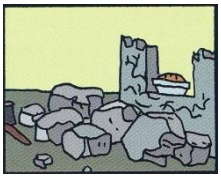
A. Unfortunately, the owner of the pie smashed his own house, crushing himself and Mr. I in the collapsed house.
B. Unfortunately, Mr. I ran into the owner of the pie and smashed his house with a hammer.
C. In the end, while Mr. I was eating the pie, the owner of the pie hit him with a hammer.

Now check the answers with your teacher. What is your score? ___/5

Task 3.1 (10 minutes)

Look at the pictures and retell the story about Mr. I to your partner using your own English. Please start with the following sentence: ***One day, Mr. I <u>was walking along</u> and he <u>passed</u> a fresh pie cooling in someone's window.***

You have 3 minutes to prepare. Do not use your notes or the written version of the story when you do this.

Part 1	1 	2 	3 	4 	5 
Part 2	6 	7 	8 	9 	10 
Part 3	11 	12 	13 	14 	15 

Task 3.2 (10 minutes)

Listen to your partner retell Story 2 and check if he/ her understood it correctly.

Task 3.3 (10 minutes)

Listen to your partner tell Story 3. Make notes about the main events.

……

……

……

……

……

Task 3.4 (10 minutes)

Next, retell Story 3 to show that you understood it correctly.

Combined group Lesson 1B (Student B, Story 2-3)

Pre-task 1.1 (5 minutes)

- First let's discuss the theme of the story

- Next listen to the recording to get some ideas about the story.

Task 1.1 (15 minutes)

Listen to the recording of the story again and identify the key events in the story (i.e., what happened in each step). You can take some notes, if you wish.

……..

……..

……..

……..

……..

……..

……..

……..

Pre-task 1.2 (5 minutes)

- First let's discuss the theme of the Story

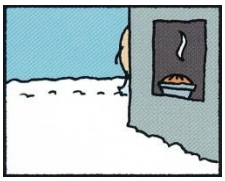
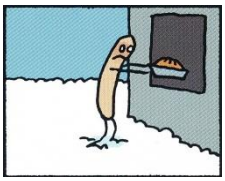
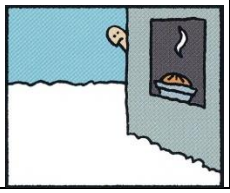
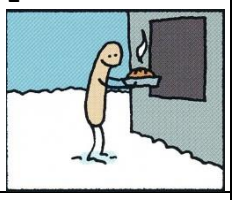
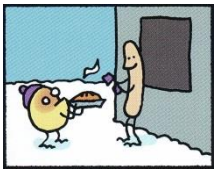
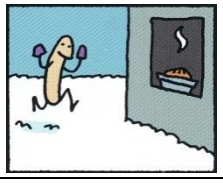
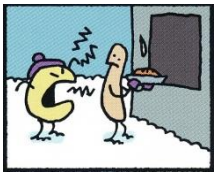
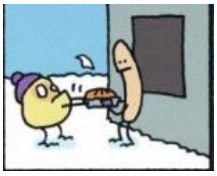
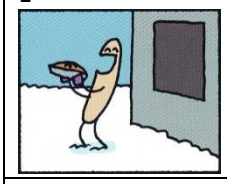
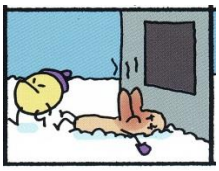
- Next listen to the recording to get some ideas about the story.

Task 1.2 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A__ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L__ 12_____ 13_____ 14_____ 15_____				

Now check the answers with your teacher. What is your score? ___ 12

Task 2 (10 minutes)

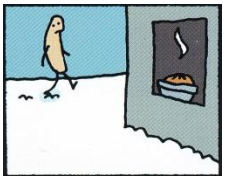
Read the story and answer the questions below.

One snowy winter's day, Mr. I was walking along and he passed the pie in the window of a house. He kept walking past, but he popped his head back around the corner just to see the pie. "Hang on, wait a minute" he thought to himself "That's a hot pie and it's a very cold day!". So he came back through the snow and stood in front of the pie. There was no one there so he decided to pick up the pie with both hands. "Oops, uh oh, that's a bit hot. It's a bit too hot. Ouch, it's very hot." Okay, he put it back on the shelf on the window. His hands were *incredibly* badly burnt. He blew on his hands but it was not good enough, so he dug them into the snow to cool them down. Then, he pulled his hands out of the snow but he still looked a bit disappointed. Suddenly, Mr I saw a yellow guy with snow gear walking past. "That's it, I've got a good idea" thought Mr. I, and immediately he ran off after the guy. Mr. I came back with a pair of gloves and a big smile on his face. With these gloves, he managed to pick up the pie. But *suddenly* the yellow guy came back over, and he looked very angry. While the yellow guy was shouting, Mr. I offered him some pie as some sort of consolation. The yellow guy considered the offer and said "Okay, yep, thank you, thank you very much, I'll have some pie."

However, as soon as the yellow guy took the pie from Mr I, he <u>realized</u> that the pie was incredibly hot. His hands started burning too, so he <u>passed</u> the pie back to Mr. I, looking very unhappy. The yellow guy's hands were badly burnt, and Mr. I <u>was smiling</u> to himself. The yellow guy started shouting but Mr. I kept his cool. The yellow guy actually went away shouting. And Mr. I <u>was laughing</u> as he got the pie all to himself. When he was about to take a big bite of the long-awaited pie, *suddenly* a big snow ball hit him in the face. He fell on the ground, and the pie <u>landed</u> on his stomach and started burning him. Mr. I started swearing and tried to lift the pie off his burnt stomach. However, the yellow guy <u>came</u> back and forced open Mr. I's mouth and stuffed the pie in. When his mouth started to burn, Mr. I struggled but the yellow guy was forced him to keep the pie in his mouth. Mr. I's lips were *badly* swollen and his arms and legs started falling. <u>Finall</u>y, the yellow guy let him go but Mr. I was already dead. I guess that Mr. I got a taste of his own medicine.

Questions

1. How does the speaker describe picture 1?



A. One snowy winter's day, Mr. I was walking along and he walked past a pie in the window of a house.

B. On a snowy winter's day, Mr. I was walking along and he walk past a pie in the window of a house.

C. One snowy winter's day, Mr. I walking along and he walked past a pie in the window of a house.

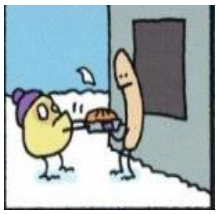
2. Which sentence describes the picture below?

A. There was no one there because he picked up the pie with both hands.

B. There was no one there so he picked up the pie with both hands.

C. There was no one there but he picked up the pie with both hands.

3. How does the speaker describe the picture below?



A. As soon as the yellow guy got the pie from Mr. I, he realized that the pie was incredibly hot.

B. As soon as the yellow guy get the pie from Mr. I, he realized that the pie was incredibly hot.

C. As soon as the yellow guy getting the pie from Mr. I, he realizes that the pie is incredibly hot.

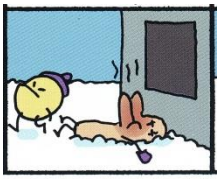
4. Which sentence describes the picture below?

A. In spite of he was about to take a big bite of the long awaited pie, suddenly a big snow ball hit him in the face.

B. During he was about to take a big bite of the long awaited pie, suddenly a big snow ball hit him in the face.

C. When he was about to take a big bite of the long awaited pie, suddenly a big snow ball hit him in the face.

5. How does the speaker describe the last picture of the story?



A. Finally, the yellow guy lets him go but Mr. I was already dead.

B. Finally, the yellow guy let him go but Mr. I was already dead.

C. Finally, the yellow guy was let him go but Mr. I was already dead.

Now check the answers with your teacher. What is your score? ___/5

Task 3.1 (10 minutes)

Listen to your partner tell Story 2. Make notes about the main events.

……

……

……

……

……

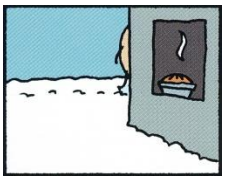
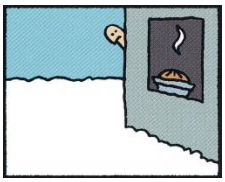
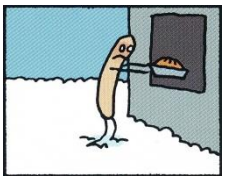
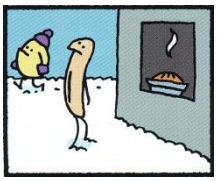
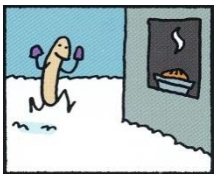
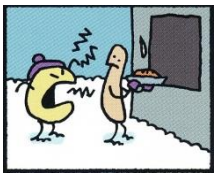
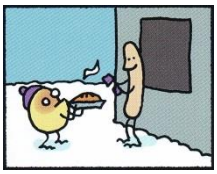
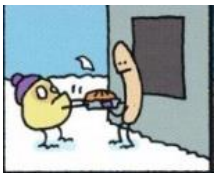
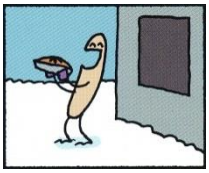
Task 3.2 (10 minutes)

Next, retell Story 2 to show that you understood it correctly.

Task 3.3 (10 minutes)

Look at the pictures and retell the story about Mr. I to your partner using your own English. Please start with the following sentence: ***One snowy winter's day, Mr. I was walking along and he passed the pie in the window of a house.***

You have 3 minutes to prepare. Do not use your notes or the written version of the story when you do this.

Part 1	A 	B 	C 	D 	E 
Part 2	F 	G 	H 	I 	K 
Part 3	L 	M 	O 	P 	Q 

Task 3.4 (10 minutes)

Listen to your partner retell Story 3 and check if s/he understood it correctly.

Combined group Lesson 2A (Student A, Story 4-5)

Pre-task 1.1 (5 minutes)

- First let's discuss the theme of Story 4

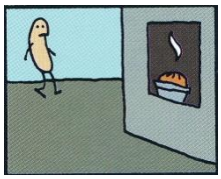
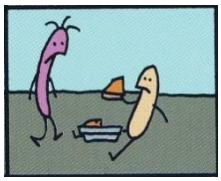
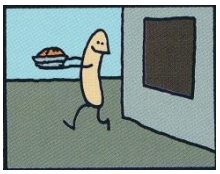
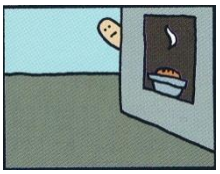
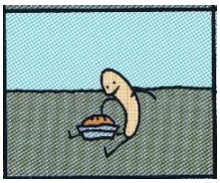
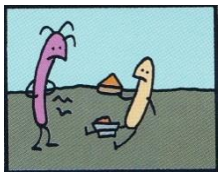
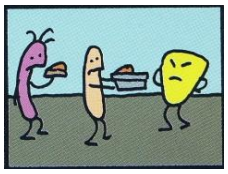
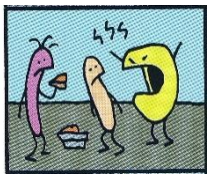
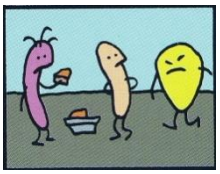
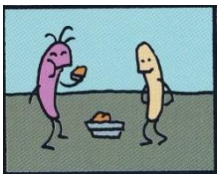
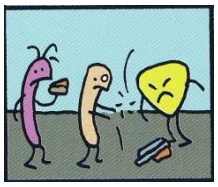
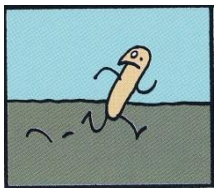
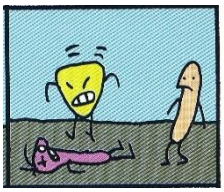
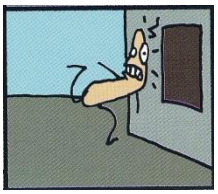
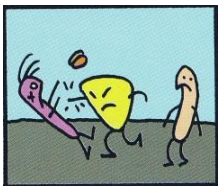
- Next listen to the recording to get some ideas about the story.

Task 1.1 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11__L__ 12_____ 13_____ 14_____ 15____				

Now check the answers with your teacher. What is your score? ___ 12

Pre-task 1.2 (5 minutes)

- First let's discuss the theme of Story 5

- Next listen to the recording to get some ideas about the story.

Task 1.2 (15 minutes)

Listen to the recording of the story again and identify the key events in the story (i.e., what happened in each step). You can take some notes, if you wish.

…….

……..

……..

……..

……

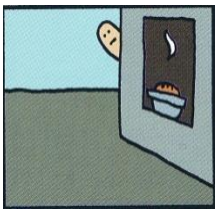
Task 2 (10 minutes)

Read the story and answer the questions below.

One day, Mr. I was walking along and passed a fresh pie in someone's window. He kept on walking past the house, but he popped his head back around the corner to see the pie. He noticed that no one was there. So he grabbed the pie and ran away quickly. When he was far away from the house, he put the pie on the ground, sat down and began to eat it. As soon as he took the first piece of pie and was about to eat it, however, a very thin man walked up. The man's stomach was growling, and Mr. I thought he must be very hungry so Mr. I offered him a piece of the pie. The man took the pie and started eating it. Suddenly, a very angry yellow man, who was the true owner of the pie, appeared. He started shouting at Mr. I and the thin man. Mr. I realised that he must be the rightful owner of the pie, so Mr. I offered to give him back the pie. But the man knocked it to ground and punched the hungry man in the face. He knocked the man down on the ground and started beating him and jumping on him. Mr. I managed to escape during all of these commotions. However, as he was running away, he looked back to see if anyone was following him and ran into a wall knocking himself silly. This goes to show that crime doesn't pay.

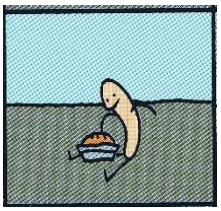
Questions

1. How does the speaker describe the picture below?



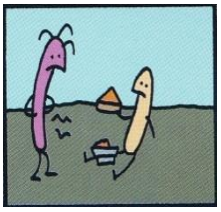
A. He walking past the house, but he popped his head back around the corner to see the pie.
B. He kept on walking past the house, but he popped his head back around the corner to see the pie.
C. He kept on walking past the house, but he popping his head back around the corner to see the pie.

2. Which sentence best describes the picture below?



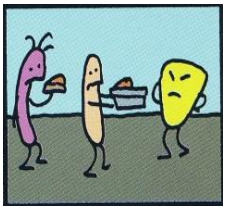
A. When he goes far away from the house, he put it on the ground, sat down and began to eat it.
B. When he was far away from the house, he was putting it on the ground, sitting down and began to eat it.
C. When he was far away from the house, he put it on the ground, sat down and began to eat it.

3. Which sentence best describes the picture below?



A. The man's stomach was growling, so Mr. I offered him a piece of the pie.
B. When the man's stomach growling, Mr. I offered him a piece of the pie.
C. Because the man's stomach growled, so Mr. I offered him a piece of the pie.

4. How does the speaker describe the picture below?

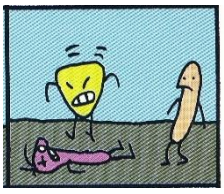


A. Mr. I realising that he must be the rightful owner of the pie, so Mr. I offered to give him back the pie.

B. Mr. I realised that he must be the rightful owner of the pie, because Mr. I offered to give him back the pie.

C. Mr. I realised that he must be the rightful owner of the pie, so Mr. I offered to give him back the pie.

5. Which sentence best describes the picture below?



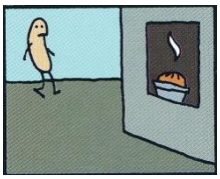
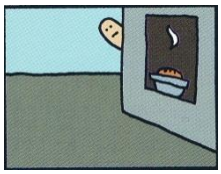
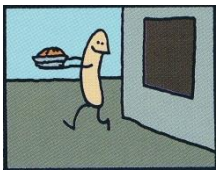
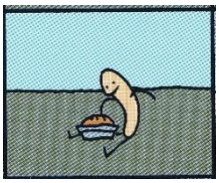
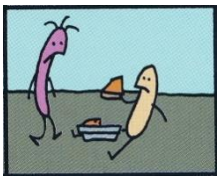
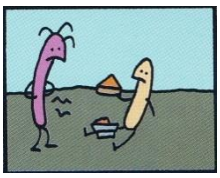
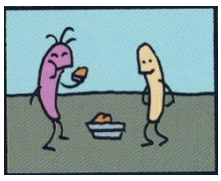
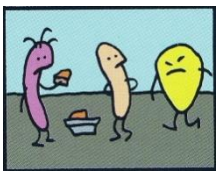
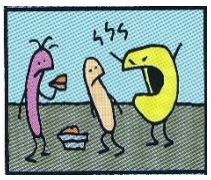
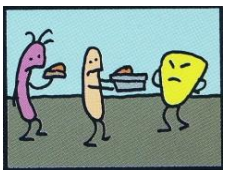
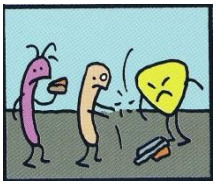
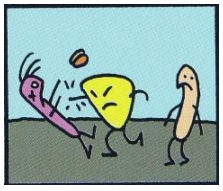
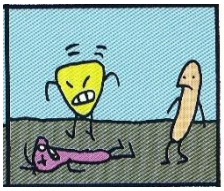
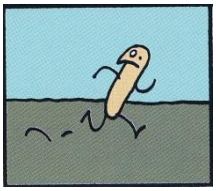
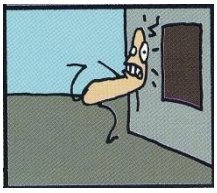
A. While he knocked the man down on the ground and started beating him and jumping on him.
B. He knocked the man down on the ground and started beating him and jumping on him.
C. He was knocking the man down on the ground and started beating him and jumped on him.

Now check the answers with you teacher. What is your score? ___/5

Task 3.1 (10 minutes)

Look at the pictures and retell the story about Mr. I who gets in trouble for stealing some pie to your partner using your own English. Please start with the following sentence: ***One day, Mr. I <u>was walking along</u> and <u>passed</u> a fresh pie cooling in someone's window.***

You have 3 minutes to prepare. Do not use your notes or the written version of the story when you do this.

Part 1	A 	B 	C 	D 	E 
Part 2	F 	G 	H 	I 	K 
Part 3	L 	M 	O 	P 	Q 

Task 3.2 (10 minutes)

Listen to your partner retell Story 4 and check if he/ her understood it correctly.

Task 3.3 (10 minutes)

Listen to your partner tell Story 5. Make notes about the main events.

……..

……..

……

……..

Task 3.4 (10 minutes)

Next, retell Story 5 to show that you understood it correctly.

Combined group Lesson 2B (Student B, Story 4-5)

Pre-task 1.1 (5 minutes)

- First let's discuss the theme of Story 4

- Next listen to the recording to get some ideas about the story.

Task 1. 1 (15 minutes)

Listen to the recording of the story again and identify the key events in the story (i.e., what happened in each step). You can take some notes, if you wish.

……..

……..

……..

……..

……..

……..

……..

Pre-task 1.2 (5 minutes)

- First let's discuss the theme of Story 5

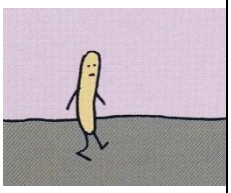
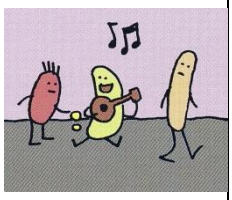
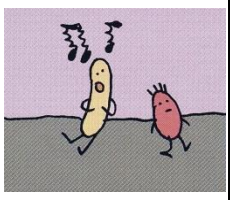
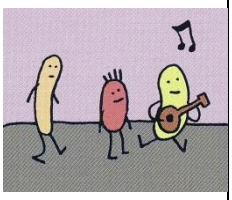
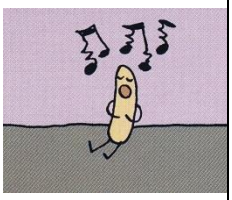
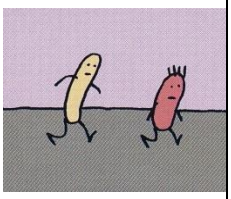
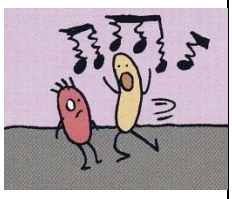
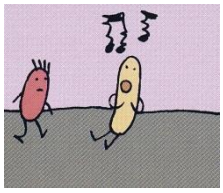
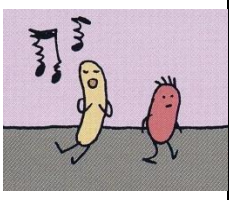
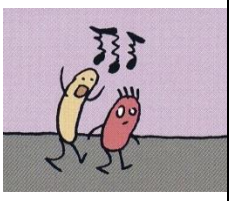
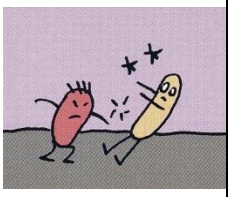
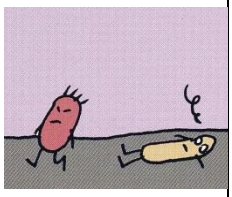
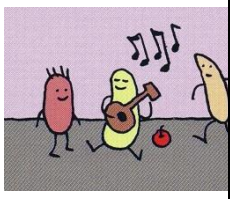
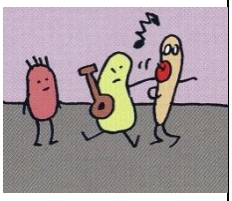
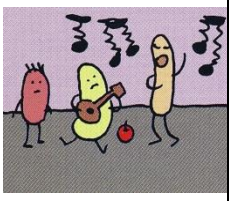
- Next listen to the recording to get some ideas about the story.

Task 1.2 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F___ 7_____ 8_____ 9_____ 10____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11__L__ 12_____ 13_____ 14_____ 15____				

Now check the answers with your teacher. What is your score? ___ 12

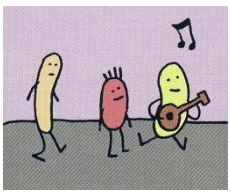
Task 2 (10 minutes)

Read the story and answer the questions below.

One day, when Mr. I was walking alone, he came across a street singer. He watched and noticed that the singer received some money for his singing. Mr. I was hungry and thought this would be a great way to get some money of his own to buy his own food. So he sat down and tried to sing as well. However, when the man who had given the street singer money before came past Mr. I, he just walked by without giving Mr. I anything. Mr. I was very surprised, so he chased after him, passed him, and sat down in front of him and tried to sing for him again. But the man just walked past and ignored him completely. When Mr. I saw this, he ran up behind the man and started singing even louder so that the man could not help but hear him. When Mr. I was singing louder and louder, the man tried to escape from him. This time Mr. I chased after him, blocked his way, and sang very loudly again with the hope to get some money from him. The man tried to escape from him but Mr. I chased after him, blocked his way, and sang very loudly in the hope of getting some money from him. Unfortunately, the man got angry and punched him. Mr I was knocked down on the ground and the man got away. Later, when Mr. I got up and collected himself, he came across the street singer and the man again. This time he tried joining in with them, but as he was singing, the street singer jammed an apple down his throat to get him to shut up. In the end, Mr. I got his apple but this goes to show that no one likes a copycat.

Questions

1. How does the speaker open the story?

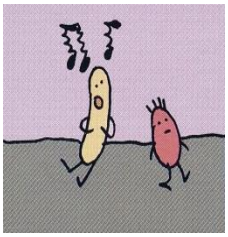


A. One day, Mr. I walking alone when came across a street singer.

B. One day, when Mr. I was walking alone, he came across a street singer.

C. One day, when Mr. I walked alone, he was coming across a street singer.

2. How does the speaker describe the picture below?

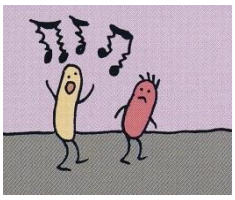


A. When the man who given the street singer the money came past, he just walked by without giving Mr. I anything.

B. When the man who was giving the street singer the money came past, he just walked by without giving Mr. I anything.

C. When the man who had given the street singer the money came past, he just walked by without giving Mr. I anything.

3. Which sentence best describes the picture below?

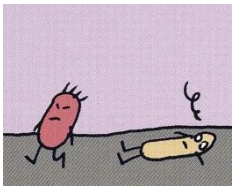


A. When Mr. I was singing louder and louder, the man tried to escape from him.

B. During Mr. I was singing louder and louder, the man tried escaping to escape from him.

C. When Mr. I sang louder and louder, the man was trying to escape from him.

4. Which sentence best describes the picture below?

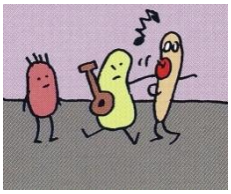


A. Mr. I was knocked down on the floor and the man got away.

B. Mr. I was knocking down on the floor so the man got away.

C. Mr. I was knocked down on the floor because the man got away.

5. How does the speaker describe the last picture of the story?



A. When he is singing, the street singer jam an apple down his throat.

B. As he sang, the street singer was jamming an apple down his throat.

C. As he was singing, the street singer jammed an apple down his throat.

Now check the answers with your teacher. What is your score? ___/5

Task 3.1 (10 minutes)

Listen to your partner tell Story 4. Make notes about the main events.

……

……

……..

……

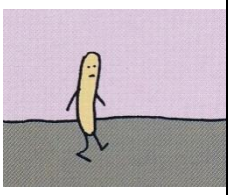
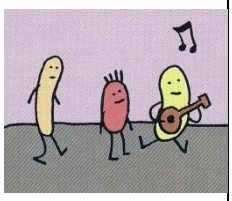
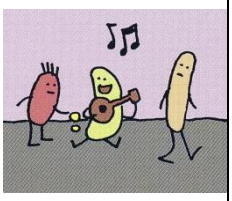
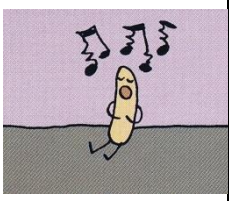
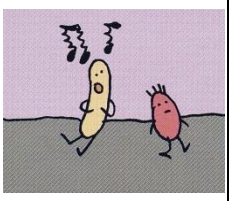
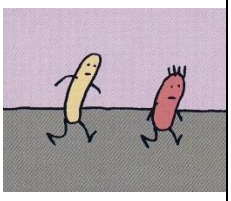
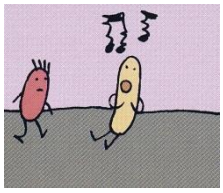
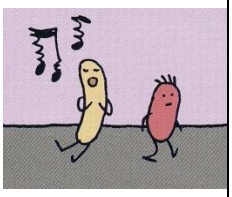
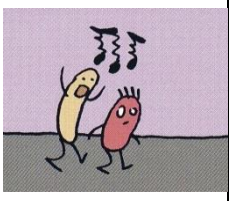
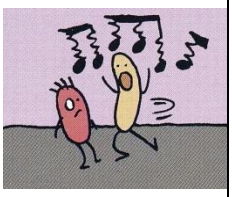
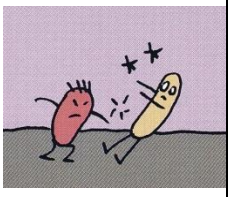
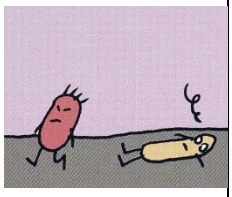
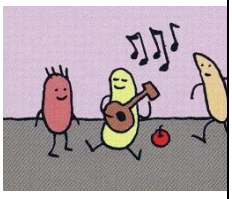
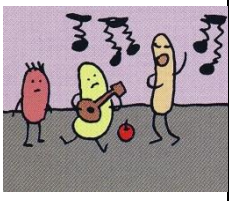
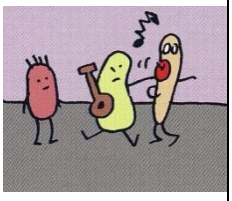
Task 3.2 (10 minutes)

Next, retell Story 4 to show that you understood it correctly.

Task 3.3 (10 minutes)

Look at the pictures and retell the story about Mr. I who gets in trouble for trying to get some money for his singing. Tell the story in English. Please start with the following sentence: ***One day, Mr. I <u>was</u> walking along and <u>came across</u> a street singer.***

You have 3 minutes to prepare. Do not use your notes or the written version of the story when you do this.

Part 1	A 	B 	C 	D 	E 
Part 2	F 	G 	H 	I 	K 
Part 3	L 	M 	O 	P 	Q 

Task 3.4 (10 minutes)

Listen to your partner retell Story 5 and check if he/ her understood it correctly.

Combined group Lesson 3A (Student A, Story 6-7)

Pre-task 1.1 (5 minutes)

- First let's discuss the theme of Story 6

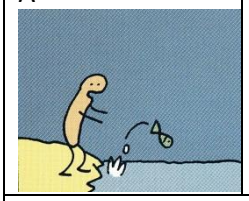
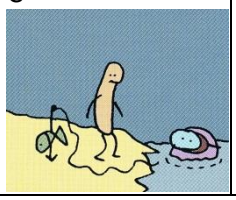
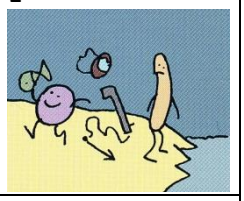
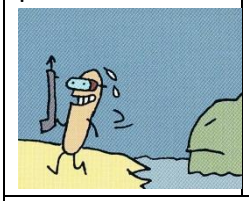
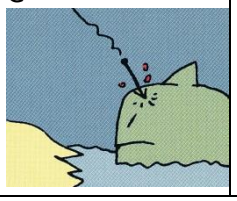
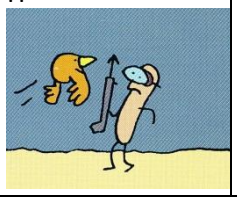
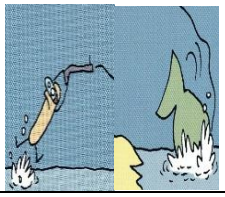
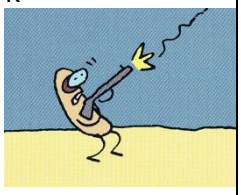
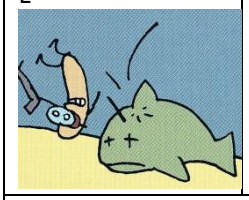
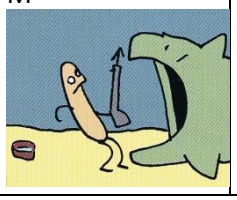
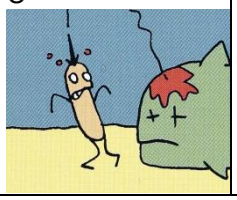
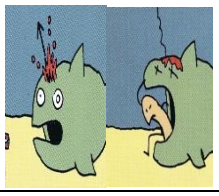
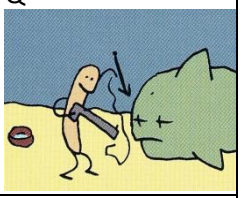
- Next listen to the recording to get some ideas about the story.

Task 1.1 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L___ 12_____ 13_____ 14_____ 15____				

Now check the answers with your teacher. What is your score? ___ 12

Pre-task 1.2 (5 minutes)

- First let's discuss the theme of Story 7

- Next listen to the recording to get some ideas about the story.

Task 1.2 (15 minutes)

Listen to the recording of the story again and identify the key events in the story (i.e., what happened in each step). You can take some notes, if you wish.

……

……

……

……

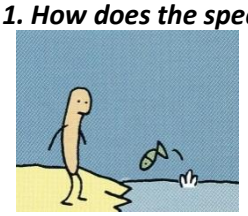
Task 2 (10 minutes)

Read this story and answer the questions below.

One day, when Mr. I was standing on the beach, he saw a fish jump. He tried reaching out for the fish because he was hungry, but he couldn't catch it. Then suddenly the fish was harpooned and landed on the beach. From the sea, a man with a snorkel mask appeared. The man walked out of the sea and jettisoned his scuba mask and his harpoon gun, but carried the fish away. Mr. I thought this would be a good way to catch a fish of his own. So he picked up his snorkel mask and the harpoon gun, and went fishing. When he met face-to-face with a giant whale, however, he lost all of his confidence, turned around and went away. Then, a bird flew past him, and he got an idea. He thought the bird was an easier target than the whale. So he shot at it instead. However, he missed the bird and hit the whale in the head with a harpoon. The whale dove into the sea, and Mr. I was dragged along as well. The whale jumped up and down in the sea and finally landed on the beach, knocking itself out. At the same time, Mr. I crashed onto the beach next to the harpooned whale. As the whale appeared to be dead, Mr. I was taking off his mask and removing his harpoon when the whale suddenly opened his mouth wide and gobbled Mr. I up. From inside, Mr. I shot the harpoon through the roof of the whale's mouth and managed to climb out. He thought he was finally safe but unfortunately, the harpoon he shot out of the whale's head came back down to earth and hit Mr. I. This goes to show that it is not good to copy everyone your see.

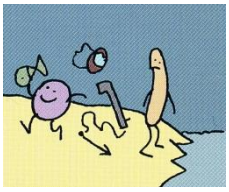
Questions

1. How does the speaker describe the picture below?



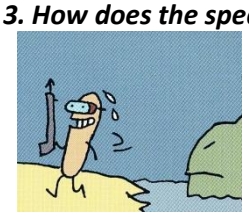
A. One day, Mr. I was stand on the beach when saw a fish jump.
B. One day, Mr. I stood on the beach and seeing a fish jump.
C. One day, when Mr. I was standing on the beach, he saw a fish jump.

2. Which sentence describes the picture below?



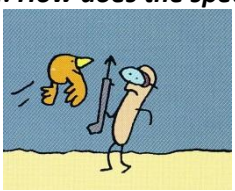
A. The man jettisoned his scuba mask and his harpoon gun, but carried the fish away.
B. The man was jettisoning his scuba mask and his harpoon gun, but carrying the fish away.
C. The man was jettisoned his scuba mask and his harpoon gun, but carried the fish away.

3. How does the speaker describe the picture below?



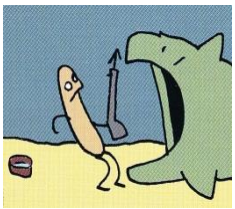
A. Until he met face-to-face with a giant whale, he lost all of his confidence, turned around and walked away.
B. When he met face-to-face with a giant whale, he lost all of his confidence, turned around and walked away.
C. When he met face-to-face with a giant whale, he was losing all of his confidence, turned around and walked away.

4. How does the speaker describe the picture below?



A. He thought the bird was an easier target than the whale.
B. He thought the bird is an easier target than the whale.
C. He thinking the bird was an easier target than the whale.

5. How does the speaker describe the picture below?



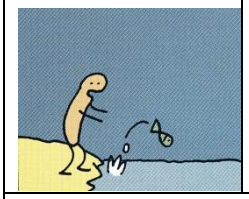
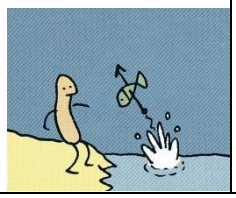
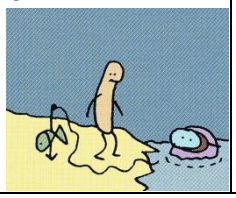
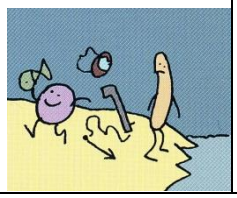
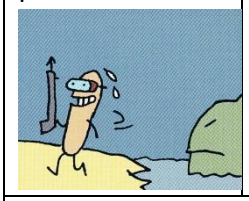
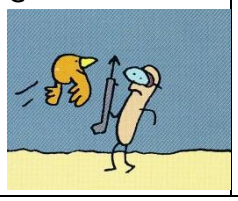
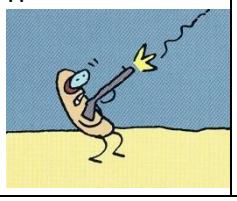
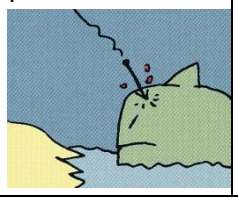
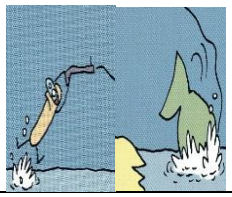
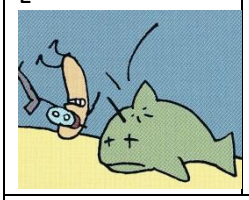
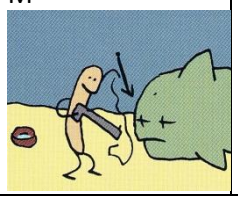
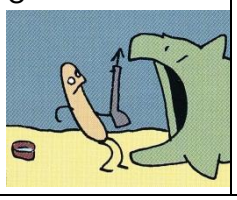
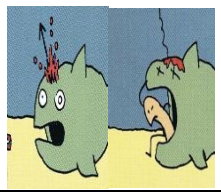
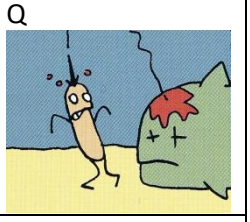
A. When the whale was suddenly opening his mouth wide and gobbled Mr. I up, Mr. I took off his mask and removed his harpoon.
B. Mr. I was taking off his mask and removing his harpoon because the whale suddenly opened his mouth wide and gobbled Mr. I up.
C. Mr. I was taking off his mask and removing his harpoon when the whale suddenly opened his mouth wide and gobbled Mr. I up.

Now check the answers with you teacher. What is your score? ___/5

Task 3.1 (10 minutes)

Look at the pictures and retell the story about Mr. I who gets in trouble with harpoon fishing. Tell the story in English. Please start with the following sentence: ***One day, Mr. I was standing on the beach and saw a fish jump.***

You have 3 minutes to prepare. Do not use your notes or the written version of the story when you do this.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6____ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11____ 12_____ 13_____ 14_____ 15_Q__				

Task 3.2 (10 minutes)
Listen to your partner retell Story 6 and check if he/ her understood it correctly.
Task 3.3 (10 minutes)

Listen to your partner tell Story 7. Make notes about the main events.

……
……
……
……
……

Task 3.4 (10 minutes)
Next, retell Story 7 to show that you understood it correctly.

Combined group Lesson 3B (Student B, Story 6-7)

Pre-task 1.1 (5 minutes)

- First let's discuss the theme of Story 6

- Next listen to the recording to get some ideas about the story.

Task 1.1 (15 minutes)

Listen to the recording of the story again and identify the key events in the story (i.e., what happened in each step). You can take some notes, if you wish.

……

……

……

……

Pre-task 1.2 (5 minutes)

- First let's discuss the theme of Story 7

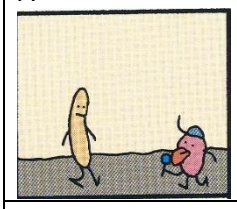
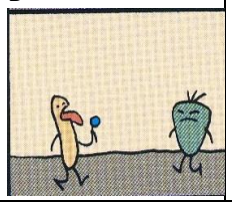
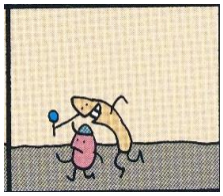
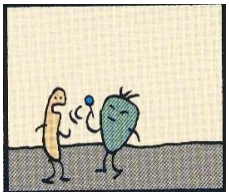
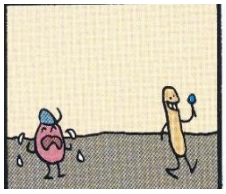
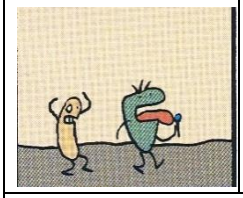
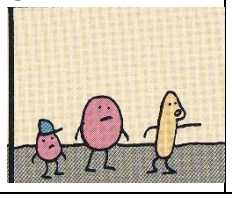
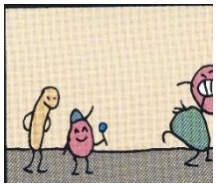
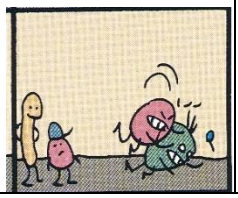
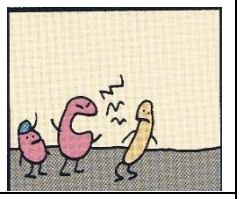
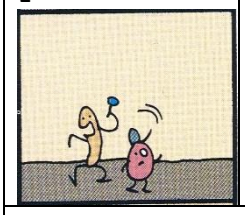
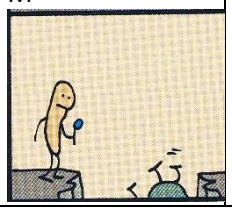
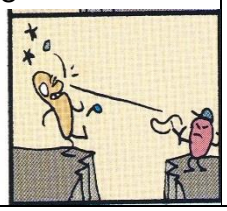
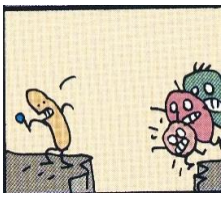
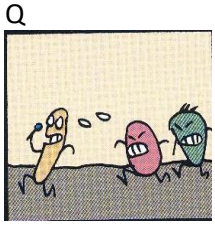
- Next listen to the recording to get some ideas about the story.

Task 1.2 (15 minutes)

- Now look at the pictures for one minute and try to figure out what is in each one.

- Listen to the recording of the story again and then sequence the pictures in each part of the story according to what you hear.

- Write the letter corresponding to each picture next to Numbers 2-15 to show the order of the pictures. The first one in each part of the story, which is Number 1, 6 and 11 are already done for you.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L___ 12_____ 13_____ 14_____ 15_Q__				

Now check the answers with your teacher. What is your score? ___ 12

Task 2 (10 minutes)

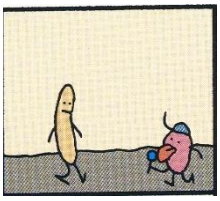
Read the story and answer the questions below.

One day, Mr. I was walking along and came across the boy with a lollipop. The boy was enjoying sucking his lollipop. But Mr. I thought it looked very good so he started following behind the boy and took it away from him. The boy started shouting and crying but Mr I ran away with the lollipop. As Mr. I was enjoying the lollipop, he ran into a green triangular man who grabbed the lollipop away from him in turn. Although Mr. I started shouting, stomping and stomping his feet, the green man walked away with the lollipop, leaving him there alone. Suddenly, the boy's mother came up from behind and started shouting at him: "Why do you steal my son's lollipop?". Mr. I explained that the green guy had taken it from him. So the son's mother went over, grabbed the other man and started beating him.

Meanwhile the little boy got the lollipop back and was watching the fight. When he wasn't looking, Mr. I grabbed the lollipop away and made a run for it. Both the boy's mother and the green man came chasing after Mr. I. When they arrived at a large gorge, Mr. I made a jump for it and successfully reached the other side while the others fell into the gorge all except the little boy. Mr. I thought he was safe but as soon as he saw the little boy, the boy shot him with a slingshot so he dropped the lollipop into the gorge. In the end, no one got the lollipop.

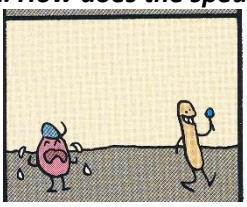
Questions

1. How does the speaker describe the picture below?



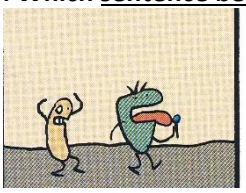
A. Mr. I walking along and came across the boy with a lollipop.
B. Mr. I walked along and come across the boy with a lollipop.
C. Mr. I was walking along and came across the boy with a lollipop.

2. How does the speaker describe the picture below?



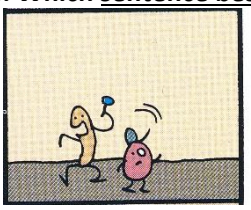
A. The boy started shouting and crying but Mr. I kept walking away with the lollipop.
B. The boy started shouting and crying so Mr. I kept walking away with the lollipop.
C. The boy started shouting and crying although Mr. I kept walking away with the lollipop.

3. Which <u>sentence best describes</u> the picture below?



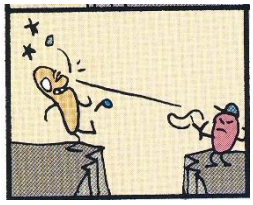
A. Because Mr. I start shouting, stomping his feet, the green man was walking away with the lollipop.
B. Although Mr. I started shouting, stomping his feet, the green man walked away with the lollipop.
C. Despite Mr. I started shouting, stomping his feet, the green man walked away with the lollipop.

4. Which <u>sentence best describes</u> the picture below?



A. While the little boy was watching the fight, Mr. I grabbed the lollipop again and made a run for it.
B. During the little boy was watching the fight, Mr. I grabbed the lollipop again and made a run for it.
C. While the little boy watching the fight, Mr. I grabbed the lollipop again and made a run for it.

5. How does the speaker describe the last picture of the story?



A. While Mr. I saw the little boy, the boy was shooting him with a slingshot.
B. Because Mr. I saw the little boy, the boy shoot him with a slingshot.
C. As soon as Mr. I saw the little boy, the boy shot him with a slingshot.

Now check the answers with you teacher. What is your score? ___/5

Task 3.1 (10 minutes)

Listen to your partner tell Story 3A. Make notes about the main events.

……
……
……
……
……

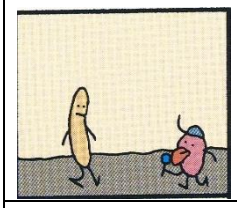
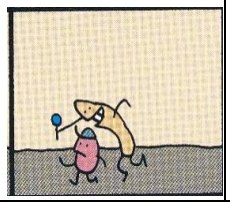
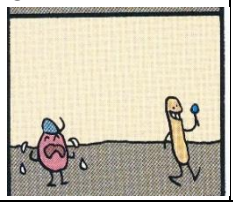
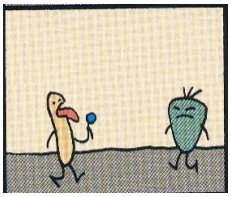
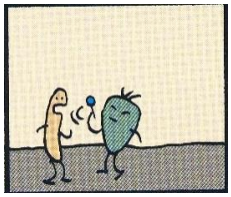
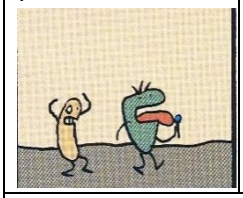
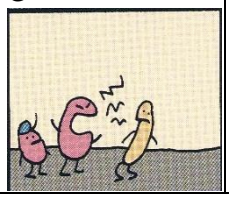
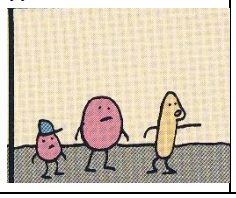
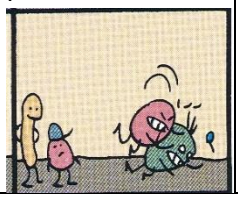
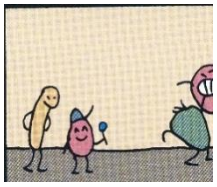
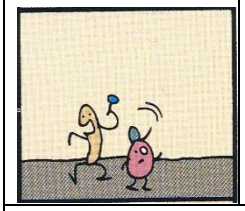
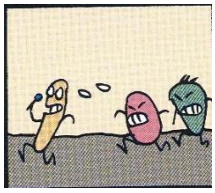
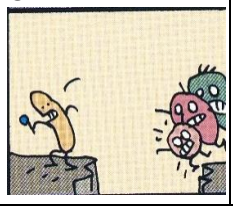
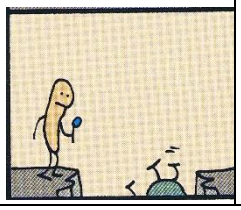
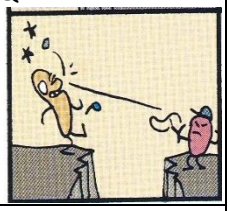
Task 3.2 (10 minutes)

Next, retell Story 3A to show that you understood it correctly.

Task 3.3 (10 minutes)

Look at the pictures and retell Story 3B about Mr. I who gets in trouble for trying to get a lollipop. Tell the story in English. Please start with the following sentence: ***One day, Mr I <u>was walking</u> along and <u>came</u> across the boy with a lollipop.***

You have 3 minutes to prepare. Do not use your notes or the written version of the story when you do this.

Part 1	A 	B 	C 	D 	E 
	Your order: 1_A_ 2_____ 3_____ 4_____ 5_____				
Part 2	F 	G 	H 	I 	K 
	Your order: 6_F__ 7_____ 8_____ 9_____ 10_____				
Part 3	L 	M 	O 	P 	Q 
	Your order: 11_L___ 12_____ 13_____ 14_____ 15____				

Task 3.4 (10 minutes)

Listen to your partner retell Story 3B and check if he/ her understood it correctly.

Appendix H: Transcription and Coding Conventions

|	AS-unit boundaries
::	clause boundaries
(0.0)	length of silent pauses within clauses (in seconds)
(0.0)	length of silent pauses between clauses (in seconds)
<…>	dysfluencies (false starts, hesitations and repetitions)
[…]	repairs (reformulations and replacements)

Appendix I: Example of a Coded Transcript

IG#13-STPT, Total sample time: 176.78''

He kept go passed :: and he (0.42) <turn> turn back to just to see the pie (0.47) eh |

it's a delicious pie (0.44) eh |

suddenly have a noise from his back (0.51) eh |

the Pink guy (0.51) eh is coming (0.47) eh |

Mr. I <put> (0.4) put (0.28) his hand on his back :: (0.42) uh to make pink guy don't care <this> (0.7) uh this pie (0.42) eh (0.42) eh |

when <mis-> the Pink guy <is> (0.73) eh (0.82) <is to> (0.61) <is> go away :: (0.37) eh Mr. I (0.42) uh take (0.28) uh [took] a piece of this pie (0.37) uh |

<He> (0.36) <he prepare to> (0.73) he (0.35) prepare (0.28) to (0.47) eh (1.18) try [trying] <this> (0.6) this pie (0.38) eh (0.76) |

Suddenly have <two> (0.69) two blue guy (0.53) eh is coming (0.49) eh |

he put (0.48) a piece of pie eh on his back (0.53) Eh |

they talk (0.57) [talked] lot of story (0.42) :: and they don't know about the appear of this pie (0.45) eh uh (0.72) |

<there was a> (0.31) when they are (0.45) talking :: (0.48) eh (0.61) there was an (0.53) eh (0.99) [there was a] (0.56) green dog (0.49) eh <appear> (0.54) appear :: (0.33) and (0.5) It sniff (0.37) the (0.37) piece of pie on his (0.48) [on Mr. I's hand] (0.49) eh (0.53) eh (0.68) |

and then (0.34) the blue guy (0.42) eh go away (0.39) eh |

the (0.67) green dog (0.48) eh bit (0.57) eh the (0.48) Mr. I's hand. (0.44) |

Oh! He fell very hurt :: (0.58) and (0.39) Mr. I struggle with the dog (0.41) |

He try to <threw it> threw it out (0.58) eh (0.44) eh |

<the (0.38) left> (0.7) <the left> (0.47) eh the left hand of Mr. I (0.48) feel very (0.97) sore (0.45) eh |

he (0.47) eh use the right hand to (0.76) uh take a (0.59) piece of pie again (0.43) |

<and the> (1.01) and the Green dog is (0.57) eh (0.45) come back (0.5). |

<It (0.57) have> (0.38) <it> (0.43) it was very angry (0.44) :: and it (0.5) eh bit the (1.29) <right> right (0.26) Mr. I's hand (0.59) eh (0.31) |

<he> (0.39) he fell very dizzy (0.51) |

and he fall out on the (0.49) land (0.28) eh |

the pie (0.26) was (0.25) uh (0.82) eh felt on the his head (0.55) eh |

the (0.58) green dog (0.59) eh (1.48) jump (0.38) into his head :: and (0.28) bit his head (0.5) uh |

in the end of the story (0.26) Mr. I (0.25) lost (0.68) his head |

Appendix K: Example of an Unpruned Transcript

IG#13-STPT, Total sample time: 176.78’’

He kept go passed and he turn turn back to just to see the pie eh it’s a delicious pie. Eh suddenly have a noise from his back eh the Pink guy eh is coming eh Mr. I put put his hand on his back uh to make pink guy don’t care this uh this pie eh eh when mis- the Pink guy is eh is to is go away eh Mr. I uh take uh took a piece of this pie uh. He he prepare to he prepare to eh try trying this this pie eh. Suddenly have two two blue guy eh is coming eh he put a piece of pie eh on his back. Eh they talk talked lot of story and they don’t know about the appear of this pie eh uh there was a, when they are talking eh there was an eh there was a green dog eh appears appear and It sniff the piece of pie on his on Mr. I’s hand eh eh and then the blue guy eh go away eh the green dog eh bit eh the Mr. I’s hand. Oh! He fell very hurt and Mr. I struggle with the dog. He try to threw it threw it out eh eh the left the left eh the left hand of Mr. I feel very sore eh he eh use the right hand to uh take a piece of pie again and the and the Green dog is eh comeback. It have it it was very angry and it eh bit the right right Mr. I’s hand eh he he fell very dizzy and she fall out on the land eh the pie was uh eh felt on the his head eh the green dog eh jump into his head and bit his head uh in the end of the story Mr. I lost his head.

Appendix L: Example of a Pruned Transcript

IG#13-STPT, Total sample time: 176.78’’

He kept go passed and he turn back to just to see the pie it’s a delicious pie. Suddenly have a noise from his back the Pink guy is coming Mr. I put his hand on his back to make pink guy don’t care this pie when the Pink guy go away Mr. I took a piece of this pie. He prepare trying this pie. Suddenly have two blue guy is coming he put a piece of pie on his back. they talked lot of story and they don’t know about the appear of this pie, when they are talking there was a green dog appear and It sniff the piece of pie on Mr. I’s hand and then the blue guy go away the green dog bit the Mr. I’s hand. Oh! He fell very hurt and Mr. I struggle with the dog. He try to threw it out the left hand of Mr. I feel very sore he use the right hand to take a piece of pie again and the Green dog is come back. It was very angry and it bit the right Mr. I’s hand he felt very dizzy and she fall out on the land the pie was felt on the his head the green dog jump into his head and bit his head in the end of the story Mr. I lost his head.

Appendix M: Descriptive Statistics

Descriptive Statistics for the Comparison Group

Test	Variable	N Statistic	Range Statistic	Minimum Statistic	Maximum Statistic	Mean		Std. Deviation Statistic	Variance Statistic	Skewness		Kurtosis	
						Statistic	Std. Error			Statistic	Std. Error	Statistic	Std. Error
Pre-test	SR_sr	34	0.67	0.73	1.4	1.032	0.027	0.157	0.025	0.423	0.403	0.159	0.788
	BP_freq_lg10	34	0.7	-0.85	-0.15	-0.573	0.030	0.175	0.031	0.402	0.403	-0.094	0.788
	WP_freq_sr	34	0.74	0.32	1.05	0.650	0.026	0.154	0.024	0.591	0.403	0.448	0.788
	Rp_freq_sr	34	0.3	0	0.3	0.112	0.014	0.083	0.007	-0.068	0.403	-0.673	0.788
Same-Task Post-Test	SR_sr	34	0.65	0.77	1.42	1.028	0.027	0.157	0.025	0.393	0.403	-0.063	0.788
	BP_freq_lg10	34	0.67	-0.82	-0.15	-0.564	0.033	0.190	0.036	0.618	0.403	-0.816	0.788
	WP_freq_sr	34	0.71	0.42	1.13	0.681	0.030	0.173	0.03	0.587	0.403	-0.268	0.788
	Rp_freq_sr	34	0.26	0	0.26	0.149	0.010	0.057	0.003	-0.518	0.403	1.501	0.788
Parallel-Task Post-Test	SR_sr	34	0.8	0.71	1.51	1.030	0.027	0.159	0.025	0.572	0.403	1.107	0.788
	BP_freq_lg10	34	0.68	-0.96	-0.28	-0.599	0.028	0.165	0.027	-0.226	0.403	-0.281	0.788
	WP_freq_sr	34	0.73	0.37	1.1	0.684	0.027	0.157	0.025	0.58	0.403	0.396	0.788
	Rp_freq_sr	34	0.26	0	0.26	0.130	0.010	0.057	0.003	-0.465	0.403	1.284	0.788

Note. SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs, sr = square root transformation, lg10 = log10 transformation

Descriptive Statistics for the Input Group

Test	Variable	N Statistic	Range Statistic	Minimum Statistic	Maximum Statistic	Mean		Std. Deviation Statistic	Variance Statistic	Skewness		Kurtosis	
						Statistic	Std. Error			Statistic	Std. Error	Statistic	Std. Error
Pre-test	SR_sr	34	0.76	0.73	1.48	1.050	0.032	0.189	0.036	0.386	0.403	0.074	0.788
	BP_freq_lg10	34	0.67	-0.89	-0.21	-0.587	0.031	0.180	0.032	0.364	0.403	-0.706	0.788
	WP_freq_sr	34	0.7	0.32	1.01	0.631	0.028	0.162	0.026	0.44	0.403	-0.106	0.788
	Rp_freq_sr	34	0.22	0	0.22	0.101	0.011	0.065	0.004	-0.414	0.403	-0.701	0.788
Same-Task Post-Test	SR_sr	34	0.58	0.72	1.3	1.089	0.024	0.142	0.02	-0.712	0.403	0.406	0.788
	BP_freq_lg10	34	0.7	-0.96	-0.26	-0.616	0.030	0.177	0.031	0.183	0.403	-0.347	0.788
	WP_freq_sr	34	0.5	0.42	0.93	0.618	0.024	0.138	0.019	0.617	0.403	-0.543	0.788
	Rp_freq_sr	34	0.2	0	0.2	0.101	0.012	0.068	0.005	-0.547	0.403	-1.128	0.788
Parallel-Task Post-Test	SR_sr	34	0.7	0.73	1.43	1.068	0.026	0.152	0.023	0.250	0.403	0.304	0.788
	BP_freq_lg10	34	0.7	-1.05	-0.35	-0.666	0.030	0.176	0.031	-0.113	0.403	-0.414	0.788
	WP_freq_sr	34	0.61	0.39	1	0.628	0.025	0.147	0.022	0.335	0.403	0.005	0.788
	Rp_freq_sr	34	0.22	0	0.22	0.105	0.011	0.062	0.004	-0.535	0.403	-0.368	0.788

Note. SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs, sr = square root transformation, lg10 = log10 transformation

Descriptive Statistics for the Combined group

Test	Variable	N Statistic	Range Statistic	Minimum Statistic	Maximum Statistic	Mean		Std. Deviation Statistic	Variance Statistic	Skewness		Kurtosis	
						Statistic	Std. Error			Statistic	Std. Error	Statistic	Std. Error
Pre-test	SR_sr	34	0.9	0.68	1.58	1.042	0.041	0.238	0.056	0.655	0.403	-0.316	0.788
	BP_freq_lg10	34	0.91	-1.1	-0.19	-0.562	0.034	0.199	0.04	-0.485	0.403	0.428	0.788
	WP_freq_sr	34	0.94	0.24	1.18	0.678	0.038	0.220	0.049	0.214	0.403	-0.372	0.788
	Rp_freq_sr	34	0.24	0	0.24	0.107	0.013	0.078	0.006	-0.096	0.403	-1.013	0.788
Same-Task Post-Test	SR_sr	34	0.63	0.79	1.42	1.126	0.025	0.143	0.021	0.008	0.403	0.097	0.788
	BP_freq_lg10	34	0.66	-0.96	-0.3	-0.621	0.028	0.165	0.027	-0.363	0.403	-0.478	0.788
	WP_freq_sr	34	0.71	0.35	1.06	0.623	0.024	0.139	0.019	0.518	0.403	1.94	0.788
	Rp_freq_sr	34	0.17	0	0.17	0.089	0.009	0.055	0.003	-0.688	0.403	-0.62	0.788
Parallel-Task Post-Test	SR_sr	34	0.55	0.78	1.33	1.065	0.027	0.156	0.024	0.043	0.403	-1.24	0.788
	BP_freq_lg10	34	0.56	-1	-0.44	-0.660	0.025	0.144	0.021	-0.554	0.403	-0.599	0.788
	WP_freq_sr	34	0.58	0.4	0.98	0.664	0.027	0.156	0.024	0.068	0.403	-0.842	0.788
	Rp_freq_sr	34	0.24	0	0.24	0.092	0.013	0.076	0.006	-0.052	0.403	-1.26	0.788

Note. SR = Speech rate, BP_freq = Frequency of between-clause pauses, WP_freq = Frequency of within-clause pauses, Rp_freq = Frequency of self-repairs, sr = square root transformation, lg10 = log10 transformation