

University of Warwick institutional repository: <http://go.warwick.ac.uk/wrap>

**A Thesis Submitted for the Degree of PhD at the University of Warwick**

<http://go.warwick.ac.uk/wrap/59397>

This thesis is made available online and is protected by original copyright.

Please scroll down to view the document itself.

Please refer to the repository record for this item for information to help you to cite it. Our policy information is available from the repository home page.

The Focus-on-Form Effects of Strategic and On-Line  
Planning: An Analysis of Japanese Oral Performance  
and Verbal Reports

by

Ryo Nitta

A thesis submitted in partial fulfillment of the requirements for the  
degree of  
Doctor of Philosophy in Applied Linguistics

University of Warwick, Centre for English Language Teacher Education

January 2007

## **Declaration**

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgement has been made in the text.

Ryo Nitta

January 2007

## Acknowledgements

Throughout my PhD years, innumerable people supported me directly and indirectly. Without their generous and professional support, this thesis would never have been completed. Among the many people to whom I am indebted, first of all, I wish to express deepest thanks to my supervisors, Dr Sheena Gardner and Dr Annamaria Pinter. Their insightful comments and professional guidance kept this work on the right track throughout the entire process of my doctoral work. I have learned a lot of important things from them, not only as a researcher but also as a teacher. The time I have spent with them at their offices and at their homes with their families constitutes the most essential part of my life in Warwick. I have also had the support and encouragement of many teachers and colleagues, and particularly I would like to express my sincere thanks to Dr Ema Ushioda and Mr Peter Brown for their continuous support and warm encouragement. I am also grateful to Dr Pauline Foster, Prof Miyoko Kobayashi, Prof Peter Skehan, Dr Martin Bygate and Dr Lourdes Ortega for their helpful discussions and suggestions. My discussions with them inspired me immensely and guided my thinking over the last couple of years. My special thanks also go to the following organizations for their financial support: the Graduate School of the University of Warwick for the Warwick Postgraduate Research Fellowship, Universities UK for the Overseas Research Studentship and the International Office of the University of Warwick for the Japan Postgraduate Scholarship. Without these, I would never have started and completed the course. Last but not least, I would like to express my gratitude to my wife and my parents for their ceaseless help and patience.

## Abstract

Within the framework of task-based language learning, there has been much research on planning, under the premise that learners' language would be enhanced in planned conditions. However, the underlying mechanisms of this rationale have not been fully explored. To develop the present understanding, this study aims to explore the nature of planning and the psycholinguistic mechanisms of its effects on L2 performance.

Earlier planning research has tended to focus on 'strategic planning' (i.e., a period of time given prior to a task), suggesting that it may improve learners' language in terms of fluency and complexity but not always in accuracy (e.g., Crookes, 1989, Foster & Skehan, 1996). In response to this, Yuan and Ellis (2003) propose 'on-line planning' (i.e., on-line processing pressure is lessened to allow active formulation and monitoring) and show its positive effect on accuracy as well as complexity. Building on these previous studies, the purpose of this research is to investigate the different form-focused effects between strategic and on-line planning.

The study takes a process-product approach to planning by using a quantitative analysis of oral performance and a qualitative analysis of post-task verbal reports, prompted by stimulated recall, under non-planning, strategic planning and on-line planning conditions. The analysis of the performance of twenty-seven Japanese learners of English (grouped as high vs. low proficiency levels) demonstrates the positive effects of strategic planning on complexity and those of on-line planning on complexity and accuracy. Most importantly, different planning effects on specific accuracy measures were observed between different proficiency groups – verb forms in the low-proficiency and articles in the high-proficiency group.

To complement the results of the performance analysis, the examination of verbal reports presents participants' planning processes. To support the improvement in accuracy in on-line planning, the analysis reveals that pressured conditions (i.e., non-planning and strategic planning) made participants prioritize meaning over form; on the other hand, on-line planning tended to push them into more complex structures while maintaining certain attention to accuracy.

Drawing on pedagogical considerations of focus-on-form instruction, this thesis argues that strategic planning and on-line planning have different degrees of form-focused effects. In particular, on-line planning, beyond a simple improvement of accuracy, would increase consciousness of form and bring L2 learners to deeper, syntactic processing. It is suggested that some kind of on-line planning would be useful for developing learners' abilities of syntactic formulation.

# Table of Contents

Declaration .....	ii
Acknowledgements .....	iii
Abstract .....	iv
Table of Contents.....	v
List of Tables .....	x
List of Figures .....	xi
List of Appendices.....	xii
List of Abbreviations.....	xiii
INTRODUCTION.....	1
CHAPTER ONE: A PSYCHOLINGUISTIC RATIONALE FOR SECOND LANGUAGE	
SPEECH PRODUCTION .....	7
<b>1.1 Introduction</b> .....	7
<b>1.2 Speaking in a First Language: Levelt’s speech production model</b> .....	8
<b>1.3 Differences between L1 and L2 Speech Production</b> .....	12
<b>1.4 Second Language Production for Learning: Swain’s output hypothesis</b> ...	14
1.4.1 <i>Fluency Function</i> .....	16
1.4.2 <i>Noticing Function</i> .....	17
1.4.3 <i>Hypothesis-Testing Function</i> .....	18
1.4.4 <i>Metalinguistic Function</i> .....	20
1.4.5 <i>Beyond the Output Hypothesis</i> .....	21
<b>1.5 Cognitive Processes of L2 Speech</b> .....	24
1.5.1 <i>Working Memory</i> .....	25
1.5.2 <i>Consciousness and Noticing</i> .....	26
1.5.3 <i>Attentional Capacity and Allocation</i> .....	29
1.5.4 <i>Monitoring</i> .....	31
1.5.5 <i>Automaticity</i> .....	34
1.5.6 <i>Lexicalized Language</i> .....	37
<b>1.6 Dual-Mode Perspective for L2 Processing</b> .....	38
1.6.1 <i>Rule-Based and Exemplar-Based Processing</i> .....	39
1.6.2 <i>Controlled and Automatic Processing</i> .....	42
<b>1.7 Conclusion</b> .....	45
CHAPTER TWO: A PEDAGOGIC APPROACH TO SECOND LANGUAGE	
SPEAKING .....	46

<b>2.1 Introduction</b> .....	46
<b>2.2 A Cognitive Approach to Task-Based Pedagogy</b> .....	47
2.2.1 <i>Defining ‘Task’</i> .....	47
2.2.2 <i>Goals in Task-Based Research</i> .....	50
<b>2.3 Investigating the Planning Effects on L2 Performance</b> .....	52
2.3.1 <i>Types of Planning</i> .....	53
2.3.2 <i>Strategic Planning</i> .....	55
2.3.3 <i>On-Line Planning</i> .....	58
2.3.4 <i>Interaction between Proficiency and Planning</i> .....	62
<b>2.4 Implementing Focus-on-Form Pedagogy</b> .....	64
2.4.1 <i>Form-Focused Instruction</i> .....	64
2.4.2 <i>Defining Focus-on-Form</i> .....	66
<b>2.5 Integrating Planning and Focus-on-Form</b> .....	72
2.5.1 <i>Mechanisms of Planning Effects on L2 Performance</i> .....	72
2.5.2 <i>Focus-on-Form Effects on Task Performance in On-Line Planning</i> .....	75
<b>2.6 Conclusion</b> .....	77
CHAPTER THREE: METHODOLOGY .....	79
<b>3.1 Introduction</b> .....	79
<b>3.2 Research Objectives</b> .....	80
<b>3.3 Research Hypotheses and Questions</b> .....	80
<b>3.4 Participants</b> .....	83
3.4.1 <i>Learning Background</i> .....	83
3.4.2 <i>Length of Residence</i> .....	84
3.4.3 <i>Making Different Proficiency Groups</i> .....	85
<b>3.5 Research Design</b> .....	87
3.5.1 <i>Research Stages</i> .....	88
3.5.2 <i>Research Procedures</i> .....	89
3.5.3 <i>Planning Order and a Rehearsal Task</i> .....	90
<b>3.6 Tasks</b> .....	91
3.6.1 <i>Monologic Tasks</i> .....	91
3.6.2 <i>Task Difficulty</i> .....	93
<b>3.7 Instructions</b> .....	96
3.7.1 <i>Obligatory Occasion of the Past Tense Form</i> .....	96
3.7.2 <i>Limitations of Production Time</i> .....	97
3.7.3 <i>On-Line Planning Instruction</i> .....	98
<b>3.8 Task Performance Analysis</b> .....	100
3.8.1 <i>Three Goals of Task-Based Research: Interaction of fluency, complexity and accuracy</i> .....	100

3.8.2 <i>Defining Fluency</i> .....	102
3.8.2.1 FLUENCY MEASURES (1): TEMPORAL PHENOMENA .....	104
3.8.2.2 FLUENCY MEASURES (2): DYSFLUENT MARKERS .....	105
3.8.3 <i>Defining Complexity</i> .....	111
3.8.3.1 COMPLEXITY MEASURES (1): SYNTACTIC COMPLEXITY ...	113
3.8.3.2 COMPLEXITY MEASURES (2): DISCOURSE ORGANIZATION DEVICES .....	119
3.8.4 <i>Defining Accuracy</i> .....	121
3.8.4.1 ACCURACY MEASURES (1): ERROR-FREE CLAUSES.....	122
3.8.4.2 ACCURACY MEASURES (2): CORRECT VERB FORMS AND ARTICLE .....	124
3.8.5 <i>Criteria for Error Identification</i> .....	125
3.8.6 <i>Summary of Interlanguage Measures</i> .....	126
<b>3.9 Verbal protocol analysis</b> .....	127
3.9.1 <i>Performance and Competence-Oriented Research</i> .....	127
3.9.2 <i>Types of Verbalization</i> .....	129
3.9.3 <i>Criticisms of Retrospection</i> .....	131
3.9.4 <i>Stimulated Recall Methodology</i> .....	132
3.9.5 <i>Use of the First Language</i> .....	134
3.9.6 <i>Degree of Verbalization Structure</i> .....	135
<b>3.10 Conclusion</b> .....	135
CHAPTER FOUR: TASK PERFORMANCE ANALYSIS .....	138
<b>4.1 Stages of Statistical Analysis</b> .....	138
<b>4.2 Results of Statistical Analysis</b> .....	140
4.2.1 <i>Basic Descriptive Statistics</i> .....	140
4.2.2 <i>Factor Analysis</i> .....	141
4.2.2.1 SAMPLE SIZE .....	142
4.2.2.2 PRELIMINARY ANALYSIS .....	143
4.2.2.3 RESULTS OF FACTOR ANALYSIS .....	144
4.2.3 <i>Repeated Measures MANOVA</i> .....	148
4.2.3.1 REPEATED MEASURES VS. INDEPENDENT MEASURES .....	148
4.2.3.2 RESULTS OF MANOVA .....	149
4.2.4 <i>Factorial ANOVAs</i> .....	151
4.2.4.1 UNIVARIATE TEST OF WITHIN-PARTICIPANT EFFECT .....	151
4.2.4.2 OVERALL RESULTS OF ANOVA .....	152
4.2.4.3 RESULTS OF ANOVAS IN TWO PROFICIENCY GROUPS .....	155
<b>4.3 Implications of Task Performance Analysis</b> .....	160
4.3.1 <i>Interaction between Proficiency and Performance</i> .....	160



4.3.1.1 INTERACTION BETWEEN FLUENCY AND PROFICIENCY ....	160
4.3.1.2 INTERACTION BETWEEN ACCURACY AND PROFICIENCY.	162
4.3.2 <i>Different Form-Focused Effects of Strategic and On-Line Planning</i> .....	163
4.3.2.1 PRIORITIZING FLUENCY IN NON-PLANNING.....	164
4.3.2.2 ACHIEVING BALANCED PERFORMANCE IN STRATEGIC PLANNING.....	166
4.3.2.3 CHALLENGING THE EXISTING LEVEL OF INTERLANGUAGE IN ON-LINE PLANNING .....	169
4.3.3 <i>Trading-off between Fluency, Complexity and Accuracy</i> .....	172
4.3.3.1 FLUENCY AND ACCURACY .....	173
4.3.3.2 FLUENCY AND COMPLEXITY.....	175
4.3.3.3 COMPLEXITY AND ACCURACY .....	176
<b>4.4 Conclusion</b> .....	177
CHAPTER FIVE: VERBAL REPORT ANALYSIS.....	179
<b>5.1 Introduction</b> .....	179
<b>5.2 Methodology</b> .....	180
5.2.1 <i>Research Questions for the Analysis of Verbal Protocols</i> .....	180
5.2.2 <i>Interview Procedure</i> .....	181
5.2.3 <i>Semi-Structured Interview Questions</i> .....	182
5.2.4 <i>Data Analysis Method</i> .....	183
5.2.5 <i>Limitations</i> .....	184
<b>5.3 General Overview of Verbal Reporting</b> .....	185
5.3.1 <i>Conceptualizing Processes</i> .....	188
5.3.2 <i>Lexical Operations</i> .....	189
5.3.3 <i>Formulating Processes</i> .....	190
5.3.4 <i>Monitoring Processes</i> .....	190
<b>5.4 Content Analysis of Verbal Reporting</b> .....	191
5.4.1 <i>Non-Planning Process</i> .....	192
5.4.1.1 AVOIDING PRODUCTION .....	192
5.4.1.2 LEXICAL PRIORITIZATION IN NON-PLANNING .....	195
5.4.1.3 MEANING PRIORITIZATION IN NON-PLANNING .....	199
5.4.1.4 SUMMARY.....	200
5.4.2 <i>Strategic Planning Process</i> .....	201
5.4.2.1 OVERVIEW OF PLANNING STRATEGIES .....	201
5.4.2.2 TRANSFERRING THE PLANNED PRODUCTION.....	204
5.4.2.3 A GAP BETWEEN THE PLANNED AND PRODUCED LANGUAGE .....	205
5.4.2.4 FOCUSING ON FORM AND MEANING IN THE STRATEGIC	

PLANNING PRODUCTION.....	208
5.4.2.5 SHIFT OF ATTENTION TO FORM IN THE STRATEGIC PLANNING PRODUCTION .....	213
5.4.2.6 SUMMARY.....	215
5.4.3 <i>On-Line Planning Process</i> .....	217
5.4.3.1 ENGAGING IN THE FORMULATING PROCESS AND ACCESSING GRAMMATICAL KNOWLEDGE .....	218
5.4.3.2 SELF-MONITORING OF THE LANGUAGE PROCESSING: REASONS FOR PAUSING .....	223
5.4.3.3 PLACE OF PAUSING .....	231
5.4.3.4 SUMMARY.....	232
<b>5.5 Conclusion</b> .....	233
CHAPTER SIX: DISCUSSION.....	236
<b>6.1 Introduction</b> .....	236
<b>6.2 Integrating the Findings of the Task Performance and Verbal Protocol Analysis</b> .....	236
6.2.1 <i>The Reduction Strategies in Non-Planning</i> .....	238
6.2.2 <i>Compensatory Strategies in Strategic Planning and On-Line Planning</i> ..	241
6.2.3 <i>Shifting into Deeper Syntactic Mode</i> .....	248
<b>6.3 Learner Proficiency</b> .....	252
6.3.1 <i>The Relationship between Planning and Proficiency</i> .....	253
6.3.2 <i>Some Psycholinguistic Evidence of Different Proficiencies: Automatization and lexicalization</i> .....	255
6.3.3 <i>Psycholinguistic and Developmental Constraints on Focus-on-Form</i> ....	259
6.3.4 <i>Strategic Attentional Allocation in Different Proficiencies</i> .....	263
<b>6.4 Conclusion</b> .....	268
CHAPTER SEVEN: CONCLUSIONS.....	270
<b>7.1 Introduction</b> .....	270
<b>7.2 Summary of the study</b> .....	270
<b>7.3 Implications for task-based pedagogy</b> .....	272
<b>7.4 Limitations of the study</b> .....	276
REFERENCES.....	280
APPENDICES.....	309

## List of Tables

Table 3.1: Planning Order x Stories .....	91
Table 3.2: Grading the Difficulty of the Three Tasks (Adapted from Ellis, 2003:228) .....	94
Table 3.3: Complication – Resolution in the Three Tasks.....	95
Table 3.4: Summary of the Interlanguage Measure for the Task Performance Analysis .....	127
Table 3.5: Types of Protocol Reports (Based on Jourdenais, 2001:355).....	130
Table 4.1: Summary of Statistical Methods Applied.....	139
Table 4.2: Descriptive Statistics .....	140
Table 4.3: Results of Factor Analysis (NP) .....	145
Table 4.4: Results of Factor Analysis (SP).....	145
Table 4.5: Results of Factor Analysis (OP) .....	146
Table 4.6: Results of Repeated measures MANOVA.....	150
Table 4.7: Results of Univariate Test.....	151
Table 4.8: Results of ANOVAs in All Participants.....	152
Table 4.9: Results of ANOVAs in the <u>High</u> Proficiency Group .....	155
Table 4.10: Results of ANOVAs in the <u>Low</u> Proficiency Group .....	157
Table 5.1: Summary of Themes Generated from the Retrospective Interviews.	187
Table 5.2: Reasons for Avoidance of Production in NP .....	193
Table 5.3: Planning Strategies .....	202
Table 6.1: Summary of Performance Features and Underlying Psychological Processes .....	237
Table 6.2: Reduction Strategies (Poullisse, 1993:161, based on Færch & Kasper, 1983:52-3).....	238
Table 6.3: NP Production and Speaker [C7]’s Commentary on Thought Processes .....	239
Table 6.4: Three Types of Compensatory Strategies (Based on Poullisse, 1993: 180-183) .....	242
Table 6.5: SP Production and Speaker [C7]’s Commentary on Thought Processes. .....	245
Table 6.6: On-Line Planning Production and Speaker [C7]’s Commentary on Thought Processes .....	246
Table 3-E.1: Summary of the Task Performance of the Pilot Study.....	315

## List of Figures

Figure 1.1: Levelt's (1989:9) Speech Production Model .....	9
Figure 1.2: Output and L2 Development (Swain & Lapkin, 1995:388/Izumi, 2003:187) .....	22
Figure 2.1: Types of Task Planning (Ellis, 2005:4).....	54
Figure 2.2: Types of Focus-on-Form Instruction .....	71
Figure 2.3: The Effects of Planned Condition on L2 Performance.....	74
Figure 3.1: Research Design .....	88
Figure 3.2: Theorizing Dimensions of Performance (Skehan & Foster, 2001:190) .....	101
Figure 4.1: Percentage of Correct Verb Forms.....	158
Figure 4.2: Percentage of Correct Articles .....	159
Figure 4.3: 'Waiting for a Bus' Task .....	164
Figure 5.1: Research Design .....	181

## List of Appendices

Appendix: Chapter 3 .....	310
Appendix 3-A: Participant’s Consent Form .....	310
Appendix 3-B: Pre-Task Questionnaire.....	311
Appendix 3-C: Results of Global Ratings .....	312
Appendix 3-D: Examples of Speeches by High and Low Proficiency Speakers..	313
Appendix 3-E: Summary of the Pilot Study.....	314
Appendix 3-F: Story-Telling Tasks .....	318
Appendix 3-G: Vocabulary and Structures for Tasks .....	320
Appendix 3-H: Instructions for Planning Implementations .....	322
Appendix: Chapter 5 .....	334
Appendix 5-A: Examples of Reports on Speech Production Processes.....	324
Appendix 5-B: Examples of Reports on Planning Strategies.....	329

## **List of Abbreviations**

ANOVA: Analysis of Variance

BICS: Basic Interpersonal Communication Skills

CALP: Cognitive/Academic Language Proficiency

CLT: Communicative Language Teaching

EFL: English as a Foreign Language

ELT: English Language Teaching

ESL: English as a Second Language

FFI: Form-Focused Instruction

ID: Individual Differences

IL: Interlanguage

LOR: Length of Residence

LRE: Language-related episode

LSS: Lexicalized Sentence Stem

LTM: Long-Term Memory

L1: First Language

L2: Second Language

MANOVA: Multiple Analysis of Variance

MEXT: Ministry of Education, Culture, Sports, Science and Technology, Japan

MLR: Mean Length of Run

NNS: Non-Native Speakers

NP: Non-Planning

NS: Native Speakers

OP: On-Line Planning

SLA: Second Language Acquisition

SP: Strategic Planning

SPSS: Statistical Package for Social Studies

SR: Speech Rate

STM: Short-Term Memory

TBLT: Task-Based Language Teaching

TL: Target Language

WM: Working Memory

# INTRODUCTION

I have long been obsessed with the thought of why learning a second language (L2) is such a difficult task. In the same way as most Japanese learners do, I started learning English at junior high school at the age of twelve. Since then, I had devoted a massive amount of time to studying English throughout years, but I still could not feel sure of my English level. From this unsatisfactory experience, I realized the limitations of learning English only in a school curriculum and dreamed of learning the language in English speaking countries. When I had the first opportunity to study abroad, I believed that an English immersion environment would surely improve my English.

However, this ideal learning condition could not satisfy my expectations; my English improved to a certain extent, but I still could not be confident of my English. I (and probably many other Japanese learners too) believed that we could gain L2 fully through natural-learning environments, but this experience heartlessly shattered my illusion. That is, living in an English speaking country could not be a sufficient condition to achieve a competent English proficiency which would allow me to express my opinions and feelings thoroughly. Since this bitter experience, I have been thinking of why I could not gain a sufficient level of English even in a natural learning context; and if being immersed in a target language (TL) is not enough, what is the necessary condition to support one's learning?

This study originally comes from this fundamental concern of learning L2 and the limitations of natural learning. In order to pursue this far-reaching question, I focus on a particular group of learners in a particular context: 'Japanese learners of English in a study abroad context'; because they are also equivalent of my learning profile. This type of learner appears to have a rather unique status, reflecting a



combination of formal classroom settings in the native country and natural learning in the target language (TL) country, but the importance of this special learning condition is not too overemphasized in today's globalized society. It is true to say that more and more Japanese people study abroad after spending substantial time in formal instruction. According to the statistics of Japan's Ministry of Education, Culture, Sports, Science and Technology (MEXT), the number of Japanese who study abroad each year has been steadily increasing from 18,066 in 1983 to 76,464 in 2001. They are immersed in English using environments, but such seemingly idealistic conditions do not necessarily produce a sufficient amount of improvement in their L2.

The research into study abroad programmes depicts the effectiveness of developing fluency in the natural environment and limitations in formal aspects of TL. Freed (1995:27) suggests that learners on such programmes appear to speak at a faster rate with fewer dysfluent markers, while 'at least for more advanced learners, that significant changes do not take place within the study abroad context [in the structural accuracy of their language]'. This means that the first period of immersion may contribute to an improvement in L2, particularly in terms of 'fluency', but the more advanced learners become, the weaker the impact of learning environment, particularly in terms of 'accuracy'. That is, the natural interaction seems difficult to break through a barrier of the linguistically stabilized state (Selinker & Han, 2001).

If natural contexts do not provide sufficient conditions for learning, the next question concerns what kinds of pedagogical interventions are beneficial. Among a number of possibilities, I will pay attention to 'planning' implementations as one of the promising pedagogic interventions, overcoming the limitations of natural learning and achieving more balanced development of fluency and accuracy.

Through a growing body of research on the effects of task planning in both

experimental and classroom settings, second language acquisition (SLA) researchers have developed an understanding of planning implementation. It has been argued that pre-task planning time is likely to bring a considerable improvement in learners' production in the short-term, and push up the level of IL in the long-term (Foster, 1996). The effects of planning time on learners' performance have been investigated from various angles. The literature tells us that under planned conditions, learners attempt to use more varied and complex language at a faster rate with fewer dysfluent features, but the effect on accuracy is inconsistent; that is, whether pre-task planning time (so-called 'strategic planning') leads to more accurate language seems less certain than the other performance dimensions (Crookes, 1989; Foster & Skehan, 1996). To disentangle such complexity of planning effects, further investigation into the cognitive processes underlying performance is needed. Despite an increasing interest in the 'pre-task' planning processes (e.g., Foster & Skehan, 1999; Kawauchi, 2005; Ortega, 1999, 2005; Sangarun, 2005), attempts to specify how planning helps L2 production 'on-line' have been neglected. If the case is that planning promises to upgrade the quality of language and ultimately leads to a significant improvement in proficiency, then what is the underlying mechanism of this process? For what reasons are planning effects on accuracy often marginalized?

To answer these questions, I draw particular attention to recent developments in 'on-line planning' research (Ellis & Yuan, 2005; Ortega, 1999; Skehan & Foster, 2005; Yuan & Ellis, 2003; Wendel, 1997), which extend our concept of planning to 'on-task' planning processes. Whereas strategic planning is a prospective influence on performance, learners may engage in a form of 'planning-as-regrouping' on-task, as they exploit time to enable them to 'think ahead' (Skehan & Foster, 2005). Defining on-line planning as 'careful production', Yuan and Ellis (2003) argue that this is likely to increase the level of accuracy as well as that of

complexity at the expense of fluency. Skehan and Foster (2005) attempted to identify the on-line planning components by investigating learners' performance, finding that particular performance features display on-line planning indices.

Building on these recent attempts, this thesis starts by defining on-line planning as 'careful speech production' following Yuan and Ellis, but ultimately aims to inquire into the construct of on-line planning in terms of the underlying cognitive processes. A main inquiry of this study is into on-line planning, but its distinctiveness is only illuminated by comparison with other planning conditions. Hence the study specifically investigates (1) whether strategic and on-line planning have differential form-focused effects upon L2 oral performance through the task performance analysis and (2) what cognitive processes operate in planning production through the verbal report analysis.

This two-way research design will show various characteristic features of performance and underlying cognitive processes in strategic planning and non-planning as well as on-line planning. From the findings of the task performance analysis, the study argues that both strategic and on-line planning lead to increased 'focus-on-form' (Long, 1991) but to different extents with different proficiency levels. The analysis of verbal reporting suggests that different planning conditions lead to different psycholinguistic operations, but learners' proficiencies also influence their attentional allocation in various speech processes. An examination of the effects of different planning conditions would be of significance to our understanding of task-based pedagogy.

The rest of this introductory chapter presents an overview of the issues which will be central in the following chapters. Based on the theories of psycholinguistics and SLA, the study mainly focuses on L2 speech processing and learning by

considering ‘planning’ within a task-based framework (Robinson, 2001a, 2001b; Skehan, 1996a, 1998, 2003). To understand these aspects, the next chapter (Chapter One) is primarily concerned with the issues of L2 processing for speech production, but also considers matters related to L2 representation. More specifically, I initially take Levelt’s (1989) monolingual speech model in order to grasp fundamental speech processes. Then, I focus on Swain’s (e.g., 1985, 1995) output hypothesis in order to consider how output production leads to learning. The theory argues that output tends to lead to syntactic processing more than comprehension, which is significant for sound L2 development. However, this does not mean that any output condition guarantees automatic attention to form aspects. To understand more fine-grained accounts of L2 speech processing and to search for more robust conditions to trigger syntactic processing, the chapter reviews some key cognitive concepts (e.g., working memory, attention, monitoring, automatization and lexicalized language), and finally attempts to construct an L2 speech model as a basis for the following discussion.

Building on this, Chapter Two approaches pedagogic issues by taking up ‘planning’ (particularly, strategic planning and on-line planning) and ‘focus-on-form’ interventions. Based on the understanding of cognitive processes of L2 production, I shall further account for more detailed planning effects by reviewing L2 task planning studies.

Having reviewed the theoretical accounts, Chapter Three describes various methodological issues of the study. Firstly, the chapter proposes research hypotheses and questions arising from previous planning studies. In addition, I will give an account of participants, research design, tasks and planning instructions in some detail. As a tool to test the effects of different planning conditions, I consider task performance analysis by reviewing various IL variables in the task-based and SLA literature. In contrast to this hypothesis-testing approach, the next part looks at verbal

report analysis combined with stimulated recall in order to explore learners' cognitive processes in different planning conditions.

The subsequent two chapters reveal the findings of the statistical analysis (Chapter Four) and those of the verbal report analysis (Chapter Five). The primary purpose of these chapters is to present the results in each research stage in terms of different planning conditions and different proficiency levels. The statistical analysis suggests focus-on-form effects in both strategic and on-line planning conditions, but to different extents; that is, there appear more positive (but not statistically significant) form-focused effects in on-line planning. It also shows significant interaction between different planning types and different proficiency groups. Not only to complement the quantitative results in the performance analysis but also to investigate the underlying reasons for those results, the verbal report analysis was conducted, showing from learners' reports significant psycholinguistic operations brought about by different planning conditions.

The purpose of Chapter Six ('Discussion') is to integrate the two different types of data and to reconsider the original question: how might reliable conditions to trigger syntactic-processing be achieved? The main discussion concerns not only the different types of planning but also learner proficiency in order to understand wider pedagogical implications from the present research.

Summarizing the main issues illuminated in the previous chapters, Chapter Seven concludes the study, suggesting some implications for task-based pedagogy and research, and discussing the limitations of the research. By combining this hypothesis-testing approach investigating the effects of on-line planning and strategic planning on L2 performance with an exploratory approach investigating the underlying cognitive operations, the study aims to uncover the effective engagement of planning implementation in task-based language teaching.

# CHAPTER ONE: A PSYCHOLINGUISTIC RATIONALE FOR SECOND LANGUAGE SPEECH PRODUCTION

## 1.1 Introduction

As widely recognized, speaking in L2 is not equivalent to speaking in L1, but this recognition does not mean that the reasons for the difficulty of L2 speech have been fully understood among researchers. It is also widely recognized (particularly by discourse analysts) that the ‘product’ of speaking is different from that of writing in a number of respects (linguistically such as lexical density and the use of nominalization, and pragmatically such as the degree of formality and contextualization; e.g., Carter & McCarthy, 1995; Hatch, 1992). However, understanding the ‘process’ underlying the ‘product’ of speaking, or what kinds of internal psycholinguistic processes bring about different language features from writing remains unclear. Based on the theories of psycholinguistics and SLA, this chapter aims to consider distinct L2 speech processes as a first step to understanding the necessary conditions to facilitate L2 learning beyond the limitations of natural learning.

As a basis of L2 processing, I start by looking at Levelt’s (1989) L1 speech production model in order to clarify the similarities between L1 and L2. Levelt’s model provides not only fundamental speech stages of a language but also vital distinctions between L1 and L2. Focusing on the distinctive nature of L2, the next section takes up Swain’s (1985, 1995, 1998) output hypothesis as one of the widely accepted L2 production models, explaining how output production leads to learning. This is also important in terms of listing the essential cognitive features of L2

production. Thus, this section considers the key cognitive factors hindering efficient L2 production in terms of an explicit knowledge account involving ‘working memory’, ‘noticing’, ‘attentional allocation’, ‘monitoring’, and in terms of an implicit knowledge account involving ‘automaticity’ and ‘lexicalized language’. The accounts of the two distinctive knowledge systems suggest the importance of a balanced use of the two dimensions for efficient transmission and sound development of L2. With the cognitive constructs differentiating L2 from L1 in mind, the last section integrates various cognitive features and adopts a dual-mode processing model as a frame of reference for discussing L2 production and learning throughout the thesis.

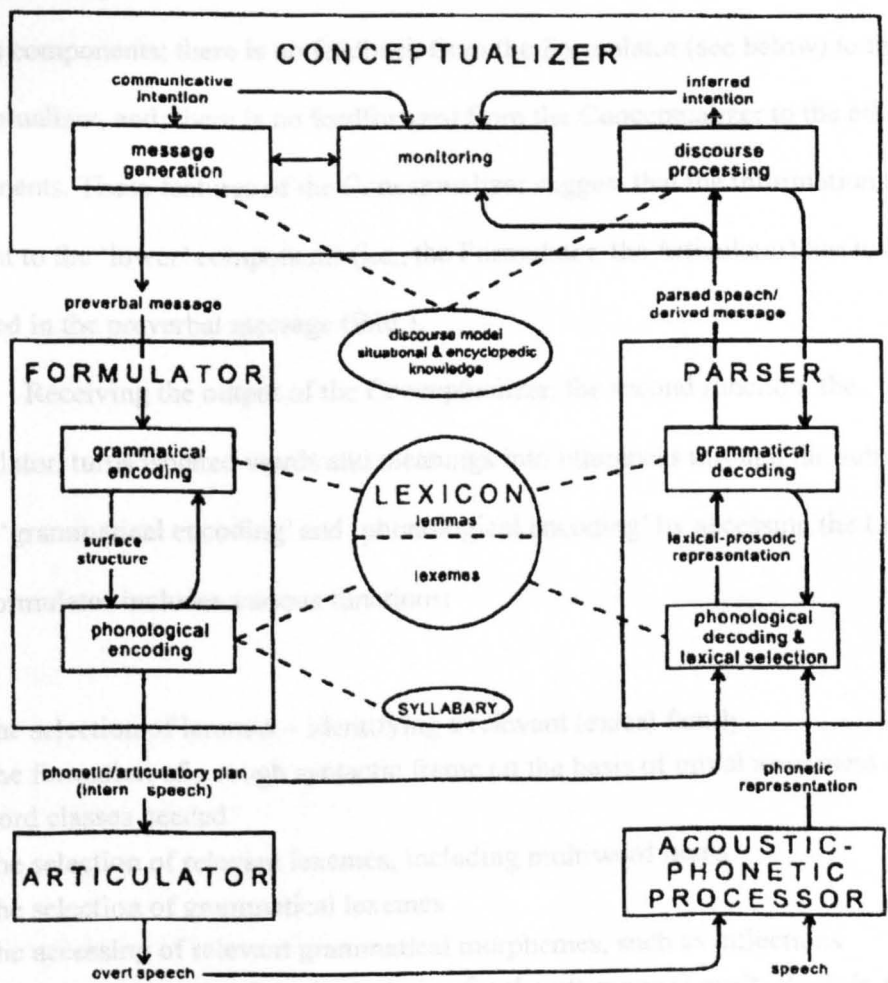
## **1.2 Speaking in a First Language: Levelt’s speech production model**

To understand the mechanisms of L2 speech production, it is useful to see those of L1 in the first place, because they seem to share a number of similarities at fundamental levels. This section discusses Levelt’s (1989) model in some detail, searching for significant implications for our understanding of L2 oral production.

One of the most important indications drawn from the current understanding in the field of psycholinguistics is that speech production is made through a series of separate stages, each devoted to a single level of linguistic analysis (e.g., Carroll, 2004, Dell, 1986, Fromkin, 1971, Garrett, 1982 & Levelt, 1989). Among various candidates, the present study adopts Levelt’s (1989) monolingual model as a basis, not only because this is constructed by integrating significant findings in the preceding psycholinguistic studies, but also because this is widely accepted as the most influential in SLA (e.g., Bygate, 2001, 2002; de Bot, 1992, 1996; Izumi, 2003) as well as psycholinguistics (e.g., de Bot, 2002; de Bot & Kroll, 2002; Field, 2003; Libben, 1997; Scovel, 1998).

Similar to other speech production models (e.g., Fromkin, 1971; Garret, 1982; Dell, 1986), this has distinctive stages comprising the ‘Conceptualizer’, the ‘Formulator’ and the ‘Articulator’. In Figure 1.1, circles represent declarative knowledge and squares represent procedural knowledge. Each component is operating independently, often referred to as the modularity principle.

**Figure 1.1:** Levelt’s (1989:9) Speech Production Model



In the Conceptualizer, ‘communicative intentions are turned into something that can be expressed in human language’ (de Bot & Kroll, 2002:135). This stage can



be subdivided into (1) the processes of planning the content and (2) those of the form of the message. The former (so-called ‘macro-planning’) plays a role in selecting the information which may realize the communicative goals concerning the kinds of relationships, identity, formality, the expected patterns of oral discourse and content knowledge (Bygate, 2002). The latter (so-called ‘micro-planning’) includes propositionalization of the event to be expressed, the perspective taken and certain language-specific decisions (de Bot, 2002). According to de Bot (2002:289), there are three crucial aspects of the Conceptualizer: there is no external unit controlling the various components; there is no feedback from the Formulator (see below) to the Conceptualizer, and; there is no feedforward from the Conceptualizer to the other components. These features of the Conceptualizer suggest that the information that is relevant to the ‘lower’ components (i.e., the Formulator, the Articulator) has to be included in the preverbal message (ibid.).

Receiving the output of the Conceptualizer, the second function, the Formulator, turns isolated words and meanings into utterances through the substages called ‘grammatical encoding’ and ‘phonological encoding’ by accessing the Lexicon. The Formulator includes various functions:

- The selection of lemmas – identifying a relevant lexical family
- The formation of a rough syntactic frame on the basis of initial awareness of the word classes needed
- The selection of relevant lexemes, including multiword items
- The selection of grammatical lexemes
- The accessing of relevant grammatical morphemes, such as inflections
- The preparation of a phonological plan for the utterance (Levelt, Roelofs & Meyer, 1999, cited in Bygate, 2002:31)

With respect to the Lexicon, there are two levels of lexical items: ‘lemma’ and

'lexeme'. The former concerns a lexical entry's meaning and syntax, which are used to generate appropriate phrase structures, while the latter includes morphological and phonological representations. Lemma information is retrieved before lexeme information. There are several stages of growth for word selection and alteration. Firstly, particular lexical items are activated by matching the meaning part of the lemma with the semantic information in the preverbal message transferred from the Conceptualizer. Secondly, this semantic and syntactic information in the lemma contributes to the formation of 'surface structure'. Then, the surface structure being formed, the morpho-phonological information in the lexeme is activated and encoded. Finally, the phonological encoding changes the information into a form of phonetic plan.

The formulated information (i.e., the output of the Formulator) is translated into sounds or 'overt speech' in the Articulator. The information is processed and temporarily stored in such a way that the phonetic plan can be fed back to the speech comprehension system (de Bot, 2002:290). According to Bygate (2002:32), articulation is generally a relatively automated phase, allowing the execution of speech with minimal conscious attention, while nonetheless open to some degree of action monitoring.

All aspects of speech production can involve the function of 'monitoring'. Not only can the accuracy of the formulating language but also the message generation and articulation be constantly monitored, whether or not the generated information matches the speaker's intention. It is also important to mention that the degree of on-line monitoring is subject to various contextual and cognitive conditions. That is, monitoring can play a significant role in some conditions, but cannot work efficiently in other conditions.

As will be discussed, the direct application of the model to L2 speech is not

possible, but it provides a general picture of the human speech system. The macro-structure of the Conceptualizer – Formulator – Articulator seems common to L2 speech. On the other hand, there are different components underlying each stage. For example, high degree of automaticity and its lexical dependency make it possible to process the language efficiently, but the same extents of these functions are not expected in the L2 system. The next section pays particular attention to such differences between L1 and L2.

### **1.3 Differences between L1 and L2 Speech Production**

In contrast to the studies of monolingual speech production model, those of bilingual speakers are preliminary, but some progress can be seen in the field of bilingualism. The researchers have traditionally and perennially asked the question of whether bilinguals' language systems are interlinked or separate in the same mind. Initiated by this fundamental question, a growing body of studies have attempted to construct bilingual models of speech production (e.g., Green, 1986, 1993; Grainger & Dijkstra, 1992; de Bot, 1992; de Bot & Schreuder, 1993; Poulisse & Bongaerts, 1994), in accordance with the recent development of L1 speech production models, particularly by Levelt. That is, most bilingual models of speech production also seem to imply hierarchical speech processing stages, as argued by McLaughlin, Rossman and McLeod (1983). Speaking is seen as a complex cognitive task with hierarchical task structure, and each component requires more or less attention, depending on how well learned it is.

Despite fundamental similarities between L1 and L2, more consideration of bilingual speech production is necessary in order to explain distinctive L2 features.

Poulisse (1997) suggests three certain characteristics of L2 production: (1)

incompleteness of L2 knowledge, (2) a lack of automaticity in procedures, and (3) a possibility of mixing two languages. Among the three factors, she argues that the first two differences can be accounted for by monolingual models of speech production, saying that the L2 lexicon contains fewer and less specified lexical items, and processing is to some extent 'serial' rather than 'parallel' to allow sufficient attention to be devoted to the different steps (Poulishse, 1997). On the other hand, the third characteristic proposes the necessity to modify the monolingual models, because they do not assume L2-specific phenomena such as 'code-switching' between two languages. That is, most bilingual speech production models are constructed to deal with the third characteristic, L1 traces in L2, essentially related to the question of 'whether word forms and concepts in two languages are represented independently or integrated within unitary lexical and conceptual memory systems' (Kroll & de Groot, 1997:169).

Despite much attention to this characteristic in bilingualism, it is also important to consider the first two L2 specific features, i.e., incompleteness of L2 knowledge and a lack of automaticity in processing procedures, because how we could overcome these problems seems to have been a central concern in the field of SLA as well as language teaching. The fact that a high degree of automaticity allows 'parallel' and 'incremental' processing (Levelt, 1989) blurs this essential nature in monolinguals. Therefore, the activities that usually require L1 speakers' continual attention are only message generation and monitoring, and the grammatical and phonological encoding of a message are largely automatic (Poulishse, 1997). In contrast, incomplete automaticity in L2 speech makes the processing speed slower and, as a result, tends to clarify the existence of the hierarchical stages of speech production and linguistic problems in a more observable manner. That is, the lack of automaticity simply makes us assume serial, step-by-step processing, and slow serial

processing allows the speaker to replenish the resources needed to carry out nonautomatic, attention-demanding processes (Poulisse, 1997). These fundamental deficiencies seem to include significant reasons for the difficulty of attaining a high level of L2 performance.

In the next section, I shift my attention to an L2 production model widely recognized in SLA. Unlike adult monolingual speech, L2 production seems inextricably related to the development of the system, as learners' language has not been completed. For this aim, I take up Swain's output hypothesis to understand fundamental issues of L2 production and grasp a general picture of how production contributes to learning.

#### **1.4 Second Language Production for Learning: Swain's output hypothesis**

Many SLA models were based on the processes of acquiring input and to a lesser extent on the aspect of production (e.g., Ellis, 1994, 1997; Færch & Kasper, 1986; Gass, 1988; VanPatten, 1990, 1996). At first sight, it is not very obvious why 'production' contributes to learning of an L2, because this is an activity which does not give new input but promotes using previously learned linguistic items. Therefore, it is important to consider, in the first place, why using the existing linguistic knowledge is requisite or at least facilitative for L2 development.

Swain's conceptualization of the output hypothesis originally came from the observations of French immersion classrooms in Canada, where students were given abundant opportunities of comprehending L2 input. This environment was then believed to contribute to an ideal approach to language teaching taken from Krashen's (e.g., 1982, 1985) theory of 'comprehensible input'. However, research on such

immersion programmes indicated that the students were likely to gain high receptive skills (e.g., listening) and communicative fluency but often failed to gain grammatical accuracy (e.g., Allen, Swain, Harley & Cummins, 1990).

From these observations, Swain (1985, 1995, 1998) has argued that, in input, L2 learners do not necessarily engage in syntactic processing, which plays a significant part in developing their IL; on the other hand, in production (both speaking and writing), L2 learners are 'pushed' into making their output more precise, coherent and appropriate, and through these efforts they discover what they can and cannot do. As a result of this attempt, learners' consciousness moves 'from the semantic, open-ended, non-deterministic, strategic processing prevalent in comprehension to the complete grammatical processing needed for accurate production' (Swain, 1995:128). An important aspect of the concept is to notice the importance of encouraging learners to move from the semantic processing to the grammatical processing within the contextualized setting, because learners cannot rely on external cues and general world knowledge in production in the same way as in comprehension (Gass, 1988). That is, output is seen not only as a product of acquisition but also as an active component in the overall acquisition processes (Izumi, 2003).

Since its first introduction (Swain, 1985), Swain and her colleagues have elaborated and constructed a whole picture of the output hypothesis (e.g., Swain, 1995, 1998; Kowal & Swain, 1994; Swain & Lapkin, 1995), identifying four functions of output: fluency, noticing, hypothesis-testing and metalinguistic functions. The following parts address each function in some detail in order to understand the concept more fully and the link between output and learning.

### *1.4.1 Fluency Function*

The first, most obvious function is that output gives opportunities for automatizing L2 language use. This ‘fluency function’ is not very surprising because it has been practised in the traditional language classroom, and also successfully achieved even in the early immersion classrooms, despite a limited amount of comprehensible output. We should not underestimate the role of practising because many teachers (Hopkins & Nettle, 1994; Larsen-Freeman, 2003) and material writers (Ellis, 2002; Nitta & Gardner, 2005) continue using some kinds of practising task. Since the influential Monitor Theory by Krashen (e.g., 1985), which argues that there is no interface or transfer between ‘learned’ and ‘acquired’ knowledge, the roles of production practice had been rather negatively viewed. However, practice associated with ‘automaticity’ has gradually drawn much attention in SLA, deriving from the idea, as Segalowitz (2003) argues, that extended practice under particular conditions and circumstances aims to develop automaticity for L2 performance. It has been argued that practice via input-based instruction will only serve to develop learners’ ability to comprehend the target structure, not to produce it (Ellis, 1999:67). That is, to comprehend input, learners need practice comprehending input, and to produce, they need practice producing (DeKeyser & Sokalski, 1996). De Bot (1996) argues that ‘acquisition’ in SLA refers to the changes in the processing of existing knowledge (i.e., the acquisition of procedural knowledge) as well as the acquisition of new linguistic knowledge (i.e., acquisition of declarative knowledge), suggesting that the locus of the effects of output must be in the transition of declarative knowledge to procedural knowledge. In brief, learning implies the qualitative development from controlled processing to automatic processing (ibid.:546). I will return to the issues of automaticity of L2 later in this chapter (1.5.5) because it is obvious that this function

plays a significant role in achieving efficient transmission of L2, which is one of the distinctive features of L2 pointed out by Poulisse (1997).

#### *1.4.2 Noticing Function*

The second function is called 'noticing/triggering' (or 'consciousness-raising').

Drawing on the study by Schmidt (e.g., 1990; Schmidt & Frota, 1986) arguing that *noticing* linguistic forms is indispensable for successful learning, Swain (1998) points out three types of noticing: (1) 'noticing a form' in TL due to the frequency or salience of the linguistic features; (2) 'noticing the gap' between the target form and the IL; (3) 'noticing a hole' between what learners can say and what they want to say. Through the noticing experiences, 'learners will turn to others, or to their own linguistic resources and work out a solution; or they will be primed to notice it in future input' (Kowal & Swain, 1994:75). Simple noticing of a problem does not mean its solution, but the awareness of this may lead to more attention to relevant information in the input, given incentives to solve the problem (de Bot, 1996:551).

Following the study by Swain and Lapkin (1995) that output not only leads to noticing the gaps in IL knowledge but also facilitates learning by triggering various internal processes conducive to SLA, Izumi and his colleagues (Izumi, 2002; Izumi & Bigelow, 2000; Izumi, Bigelow, Fujiwara & Fearnow, 1999) investigated whether output would alter the learners' subsequent input-processing and promote L2 development. SLA literature postulated that noticing can be triggered by visual input (e.g., Sharwood-Smith, 1993) as well as output. Thus, Izumi (2002) devised a study to investigate the effects of visual input and output opportunities respectively, in addition to an integrated condition of both mediums, on noticing and acquisition. The results showed that, despite positive effects of input enhancement on noticing of the target



forms in the input, the output group revealed superior learning of the form to the input group in both production and comprehension tests. From this, Izumi (2002) argues that input enhancement may have caused mere recirculation, rehearsal at a relatively shallow processing level, which led the learners to experience only a short-term retention of the noticed form; while the output treatment triggered deeper and more elaborate processing of the form, which led them to establish a more durable memory trace. This noticing function is widely recognized as an essential process of incorporating input into the IL system. I will further consider this aspect later by referring to the relevant literature (1.5.2).

### *1.4.3 Hypothesis-Testing Function*

As much discussed in the SLA literature, noticing is an essential but not sufficient process to SLA. That is, noticed items need to be further restructured in order to be incorporated into the IL system (McLaughlin, 1990). Swain and Lapkin (1995) argue that noticing a problem ‘pushes’ the learner to modify his/her output, which may be sometimes forced into syntactic processing. The third, ‘hypothesis-testing function’ is closely related to this process. Being directly related to the notion of comprehensible output, this function proposes that learners can judge the comprehensibility and linguistic well-formedness of their IL utterances against feedback obtained from their interlocutors (Izumi *et al.*, 1999). In being actively involved in production, learners will have more opportunities to check the correctness of their current understanding. Referring to Tarone and Liu (1995), Swain (1998) mentions that when the learner needs to produce output beyond the current IL system, the hypothesis-testing function pushes the limits of the system to make it handle the task. Similarly, Skehan (1996a) postulates that L2 learners need to engage in an attempt to produce complex as well as

accurate language, which is also expected to lead to stretching IL.

Much SLA research has been interested in this hypothesis-testing triggered by output production, but most studies have only looked at an aspect of 'other initiation' in giving opportunities of modified output when comprehension of the message was not clear (e.g., Nobuyoshi & Ellis, 1993; Pica, 1988; Pica, Holliday, Lewis & Morgenthaler, 1989). Following the output hypothesis proposal that learners may engage in this function internally as well as externally, Shehadeh (1999, 2001) examined 'self-initiated' modified output, reflecting the speaker's noticing that the intended message was not successfully transmitted or an ill-formed utterance was made without an explicit indication by the interlocutor. Shehadeh (2001) found that both self- and other-initiated modified output were given as opportunities to engage in restructuring the currently existing system, but the number of self-initiated instances was significantly greater than the number of other-initiated instances in many interactional contexts. Following these observations, Shehadeh (2001:450) argues that the interactional contexts that encourage repair by learners are more conducive to L2 learning than those which provide them with model utterances in the TL and invite them to confirm. This implies that learners need both time and opportunities for self-initiated and completed repairs of their messages, which have often been neglected in classroom settings (ibid.:451). This function suggests that 'deeper' psycholinguistic operations may be important for successful learning. In addition, Swain's studies indicate opportunities of self-correction triggered by learners' own production. It is thus important to consider how self-initiated repairs are made possible by output opportunities.

#### 1.4.4 Metalinguistic Function

The hypothesis-testing function which triggers reflection on learners' own language use is called the 'metalinguistic function'. Reflecting on language leads to learners' awareness of forms, rules, and form-function relationships if the context of production is communicative in nature. A series of studies by Swain have focused on this function (e.g., Swain, 1998; Swain & Lapkin, 1995, 2001; Kowal & Swain, 1994), because such metatalk is, she believes, likely to give rise to 'language-related episodes' (LREs). An LRE is defined as 'any part of a dialogue where students talk about language they are producing, question their language use, or other- or self-correct their language production' (Swain & Lapkin, 2001:104). As suggested by the definition, LRE primarily presupposes a by-product arising from student-student or student-teacher interactions, but it may be that, in addition to such an *inter*-learner advantage, learners are engaged in *intra*-learner processing as well. Learners' recognition of problems may be triggered not only by external feedback (e.g., obtained from a teacher and students) but also by internal feedback. That is, stimulated by their own effort to produce accurate language use and monitoring their own output, they are likely to notice their own linguistic problems and a hole between what they want to say and what they can say. This is the function by 'monitoring' which allows attention to be given to the well-formedness and appropriateness of the production outcome (Izumi, 2003:184), occurring overtly (i.e., external feedback) and covertly (i.e., internal feedback) in speech processes (Levelt, 1989). The metalinguistic function may not be evident in usual communication, but is considered to bring about significant learning effects. It is most common in dialogue, in which a teacher or more advanced learners help weaker learners. However, it may be possible that similar sorts of thought promote learning, for example, by accessing stored

explicit knowledge. The present study does not incorporate interaction between learners, but ‘internal’ metalinguistic reflection is taken into account.

#### *1.4.5 Beyond the Output Hypothesis*

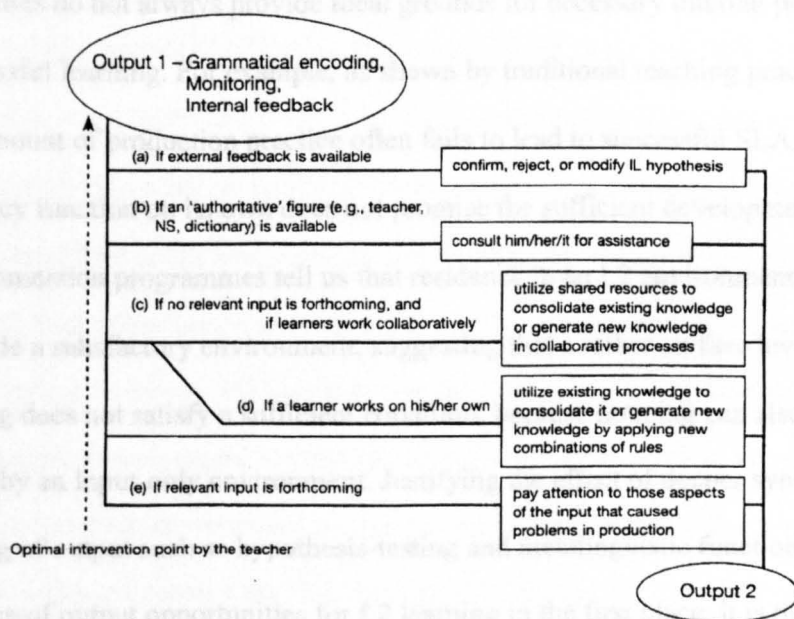
As mentioned so far, it is probably true to suggest that the output hypothesis is built on the central issues in SLA theories, involving ‘automaticity’, ‘noticing’, ‘restructuring’, ‘monitoring’ and facilitative roles of ‘explicit knowledge’. These notions were briefly mentioned, but, to make progress in our understanding of SLA, cognitive theories of such processes are in need of greater specification. Thus, I further address these key cognitive issues in the next section.

Before further examination of these issues, it is essential to point out that the four functions are distinctive but also interrelated as a series of cognitive activities, or ‘chains of psycholinguistic processes’ (Izumi, 2003:187), triggered by an attempt to make the output more comprehensible. This point is also important in terms of connecting these functions to the basic speech stages proposed by Levelt (1989). Figure 1.2 accounts for the stages of SLA, or how output production leads to learning.

Through output production, focused attention to specific production processes stimulates the development of connections in memory (de Bot, 1992). If the output matches the existing internal system, then the connection will be strengthened; if not, the information that there is something wrong with the existing system may be fed back. Solving the problem, formulating a hypothesis and testing it in the subsequent occasion of output (or comparing it in the subsequent input), is likely to bring about, internally or externally, metalinguistic reflections. In either case of strengthening or rejecting the current system, the speaker should be aware of this process for effective L2 development. There is no guarantee that the output will lead

to a noticing of the mismatches, but at least there is much less chance for the speaker to bypass syntactic processing in the course of production than comprehension. As briefly reviewed above, the findings of the research by Izumi (2002) are particularly important in emphasizing the essential nature of output, forcing L2 learners into profound syntactic processing, beyond the noticing level of processing. Output thus serves as a useful means to promote the interaction between learner internal factors (including selective attention and developing L2 competence) and environmental factors (input, interaction, and pedagogical intervention), or the interaction within the learners themselves for internal metalinguistic reflection (Izumi, 2003).

**Figure 1.2:** Output and L2 Development (Swain & Lapkin, 1995:388/Izumi, 2003:187)



What is particularly important in relation to the present discussion is the strong connection between language development and output production. As claimed by Swain, learners are likely to engage in syntactic processing when producing the target items rather than merely comprehending them. However, output production

does not always guarantee learners' automatic shift of attention to formal aspects. Often only an insufficient amount of attention is left for manipulation of form in language processing, because the meaning aspect is more readily prioritized over form (VanPatten, 1996). A challenge in instructional SLA is to establish form (and function) – meaning mapping in communicative situations. With reference to this, a series of studies by Lyster (1998a, 1998b, Lyster & Ranta, 1997) give a clearer picture of Swain's output hypothesis by showing that learners' output or self-initiated repairs seem more promisingly to guarantee that learners pay attention to form and connect it to meaning within communicative contexts. For successful L2 development, learners need to engage in not only 'negotiation-of-meaning', but also in more pedagogically-oriented discourse of 'negotiation-of-form'.

As implied in the statements so far, it should also be stressed that output opportunities do not always provide ideal grounds for necessary internal processing for successful learning. For example, as shown by traditional teaching practices, even a great amount of production practice often fails to lead to successful SLA, suggesting that fluency function on its own does not promise the sufficient development. Also, French immersion programmes tell us that residence in an L2 environment alone does not provide a satisfactory environment, suggesting that a rather surface level of processing does not satisfy a sufficient condition, because noticing can also be triggered by an input-only environment. Justifying the effect of deeper syntactic processing of output such as hypothesis-testing and metalinguistic functions and the importance of output opportunities for L2 learning in the first place, it is then important to ask more specifically, what psycholinguistic operations contribute to language learning. All the issues mentioned in this section are a matter of concern in the following section. To understand the nature and more detailed processing features of L2 output production, the discussion will shift to specific cognitive processes in L2

speech production.

## **1.5 Cognitive Processes of L2 Speech Production**

Having looked at the fundamental problems of L2 oral production and the link between output and learning, it is necessary to explore more precise mechanisms underlying such difficulties and processing. In the following discussion, I assume that the difficulty of L2 speech fundamentally lies in the heavy burden on processing the language. In the traditional language classroom, learners learn L2 with the rules of a pedagogic grammar and try to apply these rules while speaking. However, this often fails, because we can manipulate and control, 'online', only a very small amount of information in our working memory (Hulstijn, 2002). Much recent SLA theory refers to the notion that one of the difficulties of L2 production comes from 'limited working memory capacity,' reflecting the fact that learners are generally less efficient in terms of processing the L2 (Cook, 1997). Among the three characteristics of L2 production pointed out by Poulisse (1997), incompleteness of L2 knowledge leads to loading a considerable burden on on-line processing, because structural construction devours more attention than selection of appropriate, pre-fabricated chunks. To compensate for this processing problem, not only increasing the level of automaticity in particular processing stages but also allocating attention to the less automatized processes are considered necessary for improving L2 performance. This section thus depicts the issues concerning the 'oft-invoked SLA notion of 'capacity' constraints' (Robinson, 2003:664) by limiting the discussion to the issues concerning working memory, the roles of consciousness, noticing, attentional allocation, monitoring, automaticity and lexicalized language. The first four aspects concern an 'explicit knowledge' account, while the last two are more related to an 'implicit knowledge'

account. Both distinctive aspects complement each other and play essential roles in L2 speech production; balanced development of both systems is indispensable for successful L2 performance.

### *1.5.1 Working Memory*

The issue of 'working memory' (WM) is central in the information-processing view of L2 performance, because it is in this site that ongoing language processing takes place. Speakers need to make a number of decisions for production within a limited amount of time. The problem concerning WM is the limitation of its capacity, despite its significant function in language practice. In the case of L2, incompleteness of L2 knowledge and a lack of automaticity place more burdens on WM than in L1, finally making it more difficult to maintain the level of efficient performance. To understand the functions which occur in L2 oral production, this section starts by describing the essential characteristics of WM as a foundation for other cognitive processes.

Since the introduction of the modal view of memory advocated by Atkinson and Shiffrin (1968), it has been widely recognized that the human memory system can be functionally distinguished between perceptual/sensory memory, short-term/working memory, and long-term/episodic and semantic memory. The fundamental features of these memory systems are that the information in short-term memory (STM) is accessible quickly and effortlessly, but the amount of information is limited and decays rapidly. Therefore, the information in STM needs to be transferred to long-term memory (LTM), in which the information is more durable, but retrieval from this is slower and effortful. The study of STM was developed into a model of WM of Alan Baddeley (1986), by distinguishing 'active' measures from 'passive' measures of STM. WM capacity reflects both on-going processing and storage



elements, and differs from traditional conceptions of static STM in which the emphasis was on the ability to passively store bits of random information (Harrington & Sawyer, 1992).

Despite a number of different features of individual WM models among different researchers (Miyake & Shah, 1999), most memory researchers acknowledge that WM is within STM (Nairne, 1996, cited in Robinson, 2003:659), and WM is part of LTM in a currently heightened state of activation (Robinson, 2003). The difference between STM and WM is characterized by the extent to which attention is provided; that is, only peripheral attention is in STM while focal attention is given in WM, triggering maintenance rehearsal (i.e., data-driven, instance-based processing) or elaborative rehearsal (i.e., conceptually-driven, schema-based processing) (Hulstijn, 2001). This conceptualization is important for understanding why processing already stored knowledge leads to learning, as output production argues; because new information encoding processes have continual access to (activated) IL mental representations from LTM (Doughty, 2001), which seem to strengthen particular lexical items and/or promote the extent of automaticity. The WM capacity is also important for the functions of accessing explicit knowledge and monitoring in order to maintain and/or elaborate production.

The issue of WM capacity is closely related to the extent of consciousness and efficiency of focal attention allocation. These related issues affect learners' performance and language development. So, issues concerning attention will be discussed in the next section.

### *1.5.2 Consciousness and Noticing*

Following the assumption that the amount of information that can be processed by

WM is constrained, it is 'attention' that controls what information is kept active in WM and is retrieved from LTM. The roles of attention and those of the memory system are closely correlated, because the focus of attention is a subset of WM, which is part of LTM in a currently heightened state of activation; that is, attention is a process for which memory provides structure and constraint (Robinson, 2003:63).

However, understanding the notion of attention is not simple, because it includes a variety of mechanisms interacting in complex ways. To simplify the complex nature of attention, I assume that attention operates at three general stages of information processing, following Robinson (2003): (1) auditory and visual information intake and processing (i.e., consciousness and noticing); (2) central control and decision-making functions, such as allocation of attention to competing task demands (see 1.5.3) and automatization (see 1.5.5); and (3) response execution and monitoring via sustained attention (see 1.5.4). This section considers the aspect of consciousness and noticing.

Since a study by Corder (1967), much concern has been traditionally paid to the issues of how 'input' is transformed into 'intake'. In mobilizing this internal change, 'conscious awareness' is considered to play a key role. There has been much discussion on the role of consciousness in the human attentional system, particularly among the theorists informed by cognitive psychology. One of the reasons for inconsistency seems a terminological problem concerning consciousness. As 'consciousness' has several senses in everyday language, conflicts of opinion are often due as much to its vagueness as to substantive disagreement in scientific discussion (White, 1982, cited in Schmidt, 1990:131). Despite such terminological difficulties, Schmidt (1990) regards the term, 'consciousness', as a useful one because 'it ties together such related concepts as attention, short term memory, controlled vs.

automatic processing, and serial vs. parallel processing'. Schmidt (1990:131) argues that 'conscious processing is a necessary condition for one step in the language learning process, and is facilitative for other aspects of learning'.

To approach this complex concept, Schmidt (1990) disentangles the notion by carefully distinguishing several senses of consciousness: 'consciousness as awareness', 'consciousness as intention' and 'consciousness as knowledge'. Among these categories, the first category, 'consciousness as awareness', is important in the present discussion. It primarily involves three levels: 'perception', 'noticing (focal awareness)' and 'understanding'. While 'perceptions' (associated with subliminal learning) are not necessarily conscious, the other levels, 'noticing' and 'understanding', require some sort of awareness. What is noticed is the item chosen from competing perceptual stimuli surrounding us. Thus, noticing is 'the level at which stimuli are subjectively experienced' (ibid.:132), which may be (but not always) available for verbal report. On the other hand, 'understanding' is considered a more profound level of consciousness, which can be attained by analyzing a noticed item and comparing it with what we have noticed on other occasions. With reference to the link to the memory system, Schmidt (1993:213) argues that '[n]oticing is related to rehearsal within working memory and the transfer of information to long-term memory, to intake, and to item learning' while '[u]nderstanding is related to the organization of material in long-term memory, to restructuring, and to system learning'.

The issue of 'understanding' is important in order to consider explicit knowledge, but the current discussion focuses on whether consciousness at a more fundamental level (i.e., noticing) is necessary for language learning. From his experience of learning Portuguese in Brazil and the analysis of the diary notes he kept during his learning (Schmidt & Frota, 1986), Schmidt (1990) found a close

connection between noticing and emergence in production; all input do not become intake but 'intake is that part of the input that the learner's notices' (p.139). The issue of consciousness, more specifically to what extent consciousness is necessary for L2 learning, is important in language pedagogy because it has directly influenced our thoughts of how teachers teach and how learners learn an L2.

The noticing hypothesis is not unchallenged. For example, Tomlin and Villa (1994) raised a doubt about the concept developed by Schmidt as a 'coarse-grained, limited-capacity system', and proposed more detailed and elaborated ideas about attention (i.e., 'alertness', 'orientation' and 'detection'). Nevertheless, the concept has been very influential in SLA and has contributed to building up many other theoretical frameworks. The importance of conscious awareness for language learning can be accompanied by the fact that a variety of similar terms such as 'focus-on-form' (Long, 1991), 'consciousness-raising' (Rutherford, 1987; Sharwood-Smith, 1981) and 'input-processing' (VanPatten, 1990) sprung up around the same period.

### *1.5.3 Attentional Capacity and Allocation*

The roles of attention are not limited only to noticing sensory stimuli in the perceptual encoding level, but also extend to decision-making functions in the central processing level. Due to the limitation of attentional capacity, it is necessary that attention should be 'selective'. Thus, one of the significant functions of central control is the allocation of attentional capacity to task demands, or how an L2 speaker's attention is distributed to the different levels of production stages. Referring back to Levelt's model (Figure 1.1), as a consequence of the top-down processing of speech production, more attention can be allocated to the higher planning ranks (e.g., the Conceptualizer), and the lower planning processes (e.g., the Formulator) tend to suffer

from a lack of attention to control, in spite of the fact that the lower level processes are also essentially requisite for production. Garman (1990:372) explains this tendency in terms of the memory system:

[In contrast to relatively long memory for conceptual aspects] Memory for specific grammatical forms is much shorter, and we more frequently encounter a point in output where one's choice of how to continue an utterance may be in doubt because memory for the initial form has been lost.

In L1 speech processing, neglecting to pay 'attention' to the lower processes in the Formulator and the Articulator do not usually lead to a serious problem, because these processes are largely automatized, and conscious attention to control is not always necessary. Thus, native speakers (NS) could pay most attention to higher processes, such as the coordination of intentions (de Bot, 1996:550). According to Levelt (1989:22), because speech is normally produced at a rate of about two to three words per second, selecting from the many tens of thousands of words in the mental lexicon, there is just no time to consciously decide on a word. Thus, formulating and articulating are 'underground processes' that are largely impenetrable to executive control (ibid.). On the other hand, for L2 speakers, this lack of attention to the lower planning stages would be problematic, because morphosyntactic processes are not sufficiently automatized. It is also important to mention that there should also be further distinctive differences of proficiency within non-native speakers (NNS); i.e., while beginners need to pay more attention to 'lower' level planning, advanced learners could pay more attention to 'higher' level planning within a time constraint. Therefore, if learners, despite the lack of attention, attempt to complete the formulation stage, then they suffer from serious processing time limitations, because recruiting further attention to control the language formulation stage simply requires

more time (Hulstijn & Hulstijn, 1984).

#### *1.5.4 Monitoring*

In the limited attentional capacity view, an increase of task complexity would deplete attention, and thus negatively affect learners' performance. Although it is inconclusive whether the attentional resource capacity is limited in single or multiple levels (see Robinson, 2001a, 2001b, 2003), it is probably true to suggest that 'effortful' attention is a limited resource (Kihlstrom, 1984, cited in Schmidt, 2001:12). Robinson (2003:651) mentions that sustained attention to an activity over time is a central notion in studies of vigilance, energetic states, and the causes of decline in performance on a task, which is also acknowledged by Skehan and Foster (2005) for the difficulty of sustaining the quality of L2 performance. In the field of SLA, failure to sustain attention to a communicative task can be identified as one cause of decline in self-repairs and monitoring of output (Robinson, 2003:652). Thus, monitoring can be regarded as one of the significant processes in L2 performance. The following pages refer to the functions of monitoring by referring to the differences between L1 and L2.

With respect to the places where monitoring functions, Levelt (1989) devises 'the perceptual loop theory', in which the speech comprehension system (i.e., parser) is used for attending to one's own speech as well as that of others in order to avoid reduplication of knowledge. According to Levelt (1989:471), the roles of monitoring can be epitomized into two pathways, i.e., perceptions of self-produced overt speech (i.e., external loop) and internal speech (i.e., internal loop). Speech production is not a one-way transmission of messages but a self-regulating process with a feedback loop to ensure that each previous stage of output was accurate (Scovel, 1998). For the roles

of monitoring, Levelt (1989:470) demonstrates:

When we listen to the speech of others we can discern deviant sound form, deviant morphology, and deviant syntax...the same mechanism is involved in monitoring one's own internal or overt speech.

In spite of possible similarities between L1 and L2 monitoring (Kormos, 1999), there seem not to be precisely equivalent definitions of the term, monitoring, between L1 and L2 traditions, because monitoring in L2, based on Krashen's line of argument, has not been handled in such a sensitive manner. Krashen (1982) argued that monitoring can only be effective under certain limited circumstances (e.g., no time constraints, focus on form rather than focus on meaning situations), but it might be difficult to judge whether his statement implies either pre-production (i.e., covert) monitoring, post-production (i.e., overt) monitoring, or even both phases.

To my best knowledge, the first attempt to construct the psycholinguistic model for the L2 monitoring process is the one by Morrison and Low (1983). Criticizing Krashen's view of monitoring (emphasizing 'learned' formal rule applications) as 'too narrow in scope', Morrison and Low (1983) argue that monitoring should be regarded as a more comprehensive ability. Engaging in monitoring, adult L2 users pay active attention to the form of their utterances on a number of different levels such as 'lexis, discourse, and truth value, as well as syntax, without benefits of audience feedback, all in the same stretch of discourse' (ibid.:236), and 'differences in customary level of attention correlate with characteristic differences in the learner's acquired system' (p. 245). In addition to these wider roles, their definition clearly states the inclusion of both covert (pre-articulatory) and overt (post-articulatory) monitoring; as a result of the former function, L2 speakers may engage in hesitation, while as a result of the latter function, they may engage in

false-starts and self-corrections (Morrison & Low, 1983).

An important implication of these statements is that there is a distinction between a weak version and a strong version of monitoring; the former is, as presumably Krashen intended, simply a kind of quality control mechanism to edit out 'silly mistakes', while the latter gives monitoring a more central role in the acquisition process, 'in mediating between conflicting internal rules coinciding with different stages in the evolution of the learner's internal grammar' (ibid.:244).

It has been argued that there seem to be more or less corresponding processing stages between L1 and L2, but it is necessary to consider the primary reason for qualitatively different phenomena of L2 speech (such as distinguishably slower speech rate, noticeable amount of grammatical errors) from L1 speech production. In a similar vein, there are similar functions of L2 monitoring to those of L1, but there are also specific reasons for hindering effective functioning of L2 monitoring. Doughty (2001:214) maintains that, in addition to the regular monitoring function, processing for language learning includes a special kind of monitoring involving cognitive comparisons of the intention, the input, and the output, as reviewed in the output hypothesis (see 1.4). For a promising clue to this question, Kormos (1999:312) mentions:

Because in L2 speech processing a considerably lower number of processes are automatized, and in turn, they require more attention than L1 encoding mechanism, it is likely that these limitations will have more noticeable effects in monitoring in L2.

In both L1 and L2, monitoring requires conscious attention such as 'a fairly high degree of control' or 'a more or less conscious decision-making process' (Morrison & Low, 1983:245), but, from the degree of automaticity of language formulation, it is



not difficult to imagine that L2 speakers tend to suffer from maintaining sufficient attention to the monitoring process more seriously. Thinking about the incomplete nature of linguistic competence, monitoring should play a more significant role in L2 (Crookes, 1991:116), but it seems that L2 speakers, especially at an elementary level, often drain the attentional resources in the linguistic formulation, and as a result, neglect to allocate them to the monitoring process. In line with this, the fact that the error rate in L1 is astonishingly low (for example, Garnham *et al.* [1982], reported in Levelt [1992:2], found 86 errors of lexical selection in their spoken corpus of 200,000 words, and 105 other slips of the tongue) underscores the considerably lower importance of monitoring L1 formulation. Accordingly, NSs could allocate more attention to monitoring the other levels, especially the Conceptualizer. From these indications, speech monitoring can be considered vital in various functions of L2 performance.

### *1.5.5 Automaticity*

So far, the discussion has focused on the explicit knowledge account, but this does not mean the implicit knowledge side is insignificant. Rather, following the discussion that speaking is rapid and parallel processing (see 1.2) and only limited information can be held under conscious or effortful attention (see 1.5.3), the speech production process largely needs to take place ‘implicitly’. Although production practice to promote automatization has been criticized due to its strong association with behaviourism (e.g., Ellis, 1991; VanPatten, 1996), there has also been a growing body of research to support the value of automatization for effective L2 performance. This concerns a fundamental question prevailing in SLA, whether adult language learning takes similar steps to L1 acquisition or other cognitive skills (DeKeyser & Sokalski,

1996). Those who insist on an essential role of automatization view L2 use as a similar process to other cognitive skills, such as typing a keyboard and driving a car. It is inconclusive whether production practice plays a key role in developing L2, but the need for certain levels of automatization seems widely acknowledged, because of limited attentional capacity. Therefore, this section looks at issues concerning automaticity in L2 and its function in terms of limited attention management.

In the language classroom, the question of whether production practice should be given – and if so how – has always been a matter of concern (e.g., Ellis, 1993; Hopkins & Nettle, 1994; Nitta & Gardner, 2005). Some demonstrate that certain forms of practice in a certain sequential pattern are beneficial for L2 learning (e.g., DeKeyser, 1998), while others (e.g., VanPatten, 1996) insist on the inefficiency of production practice on the whole. This line of discussion is inextricably linked to the issue of automatization, as almost all automaticity proposals imply that extended practice develops automatization (Segalowitz, 2003). As pointed out by Segalowitz (2003), automaticity should not be used as simply synonymous with fast processing, because fast non-automatic conditions are possible. Rather, in the process of automatization, ‘fluent robust habit of incorrect speech’ (Segalowitz, 2003:387) is also likely to occur, as a result of a controlled process becoming automatic before it is native-like. That is, as automatic processes are outside the attentional control of L2 learners, non-targetlike structures are likely to remain in the learner’s IL, giving rise to a stable but erroneous construction (Mitchell & Myles, 1998). Therefore, automaticity can be rather interpreted as a significant qualitative change in the way that processing is carried out with the quantitative increase of automaticity, and to understand this, it is requisite to focus more on the co-development of automatic processing and attention (Segalowitz, 2003).

Among a number of theories to support the importance of automaticity in the

skill-acquisition view, the most widely accepted model in SLA is Anderson's Adaptive Control of Thought (ACT) theory (e.g., DeKeyser, 1998, 2001; Johnson, 1996; Towell & Hawkins, 1994). Since its first introduction in the 1970s, this has been repeatedly revised (e.g., Anderson, 1983, 1993, Anderson & Lebiere, 1998), but the fundamental principle remains unchanged. The theory essentially poses two different types of knowledge, declarative (i.e., 'knowledge about facts and things') and procedural knowledge (i.e., 'knowledge about how to perform various cognitive activities'), and claims that skill can be acquired through three stages: (1) declarative knowledge, (2) compilation/proceduralization of knowledge and (3) automatizing or fine-tuning of procedural knowledge. First of all, any kind of behaviour can be performed in principle by using general production rules to retrieve relevant chunks of knowledge from declarative knowledge in LTM and assembling them in WM. Through repeatedly retrieval, chunks of declarative knowledge become incorporated into the production rule through the processes of compilation and proceduralization. Then the rule can operate faster and with less risk of error, bypassing retrieval of information from long-term declarative memory. Also, in the process of fine-tuning, the skill will become more selective through the processes of generalization, discrimination and strengthening (DeKeyser, 2001:132-3). As a result of this 'automatization' process, learners could attain 'automaticity', enabling them to carry out tasks with speed and ease (ibid.:125).

The issue of automaticity has been traditionally recognized in SLA (e.g., Swain's output hypothesis; see 1.4.1) and language teaching, particularly in terms of developing the fluency function (Fillmore, 1979). As repeatedly mentioned, this is also important in terms of the limited attentional capacity, because the more knowledge is automatized, the more attention can be devoted to other processes.

### *1.5.6 Lexicalized Language*

In addition to such issues as limited WM capacity and insufficient automaticity of processing, quantity of lexicalized items can also be considered as a crucial component differentiating L2 from L1. There is a close relation between the level of automaticity and the amount of lexicalized items; ‘automatization reflects a transition from algorithm-based performance to memory-based performance’ (Logan, 1988:493, cited in DeKeyser, 2001:134). As suggested by de Bot (1992:5), Levelt’s monolingual model has the lexis-precedence tendency, because the characteristics of lexical items, which are retrieved first, determine the application of grammatical and phonological rules.

An exclusive lexis-based nature is crucial to accounting for ‘incremental’, ‘parallel’ and ‘automatized’ characteristics in L1 speech production (Levelt, 1989). It is true to suggest that a large stock of lexicalized items (also called ‘memorized sequences’, ‘formulaic chunks’ ‘lexicalized language’, ‘ready-made chunks’; see Wray [2000, 2002] for the summary of terms used in the literature), can be one of the keys to reducing the processing burden of encoding the preverbal message and attending to other aspects of speech processing in order to push up L2 proficiency (e.g., Lennon, 1998; Nattinger & DeCarrico, 1992; Pawley & Syder, 1983, 2000; Sinclair, 1991; Wray, 2002; Weinert, 1995). It is important to mention that what ‘lexicalized language’ means here does not limit to a single word level but includes longer fixed phrases, which Pawley and Syder (1983) call ‘lexicalized sentence stems’ (LSS). More proficient speakers are able to keep track of where they are, syntactically, as they incorporate fully or partially fixed sequences (Foster, Tonkyn & Wigglesworth, 2000:356). However, the practical problem is that adult learners are highly unlikely to take in such a great amount of lexicalized items as the ordinary NSs store, which

amount at least to several hundreds of thousands (Pawley & Syder, 1983:213).

Related to the role of lexicalized language, Foster (2001) investigates the extent to which the NSs and NNSs exploited lexicalized sequences during a discussion task, and whether giving planning time before the task would affect the degree to which they were used. The results show that NS's language was composed of a much higher proportion of identified lexicalized sequences. In addition, NNSs' use of these sequences was unchanged by planning, while NSs' use of them was considerably reduced by planning. In line with the present argument, Foster's (2001) study suggests a differential gap of lexicalized items between NSs and NNSs as well as NSs' strategy of relying on formulaic sequences in order to maintain the level of fluency in pressured conditions. Foster (2001:80) argues that '[o]ur reliance on ready-to-use 'chunks' of one kind or another is a useful processing strategy, enabling normal speech to be produced at seamless speed'.

Foster's study lends support to the idea that a lack of L2 fluency is partly due to the incomplete nature of lexicalized items in the L2 lexicon. Thus, it is necessary to build up the amount of lexicalization of L2 to develop the efficient transmission of L2. However, it is also noticeable that for adult NNS, achieving the same amount of lexicalization as NSs is intimidating. Therefore, it is important to consider how we could achieve a good level of L2 development not only by relying on lexicalized items but also other functions of a language system.

## **1.6 Dual-Mode Perspective for L2 Processing**

This chapter has covered a number of distinctive cognitive factors for understanding L2 speech production. The next challenge is to integrate these complex issues and to construct a unified picture of L2 oral production. The rest of this chapter, therefore,

attempts to build a psycholinguistically-oriented L2 production model, a dual-mode perspective for L2 processing, by synthesizing the above-mentioned cognitive processes. This will be a significant basis for understanding the planning research in the subsequent chapters.

### *1.6.1 Rule-Based and Exemplar-Based Processing*

Understanding the function of lexicalized language in L2 is important, because it concerns one of the underlying potential differences between L1 and L2 indicated by Poulisse (1997; see 1.3). This ‘memory-based’ approach accounts for the mechanism of efficient L1 acquisition well, but does not seem to reflect the reality that most adult L2 learners are finally unable to reach such a level equipped with the same amount of lexical items as in L1. Despite a limited amount of lexicalized items, some learners are able to efficiently transmit their messages and accomplish the communicative purpose successfully. On the other hand, the argument that L2 development only comes from speeding up controlled processes under attentional supervision in WM (Robinson, 1997:226) does not explain the failure of traditional teaching approaches to provide successful communicative competence. Thus, it seems inappropriate to adopt too extreme positions but rather to take an eclectic attitude and keep a balance between these proposals.

To compensate for insufficient amounts of lexicalized items, L2 speakers, more often than not, need to improvise language by accessing the explicit, rule-based system. Drawing on the ‘accessibility’ and ‘analysability’ distinction (Widdowson, 1989), Skehan (1995:97) explains the mechanisms of the two distinctive systems:

What seems to be the default is that accessibility [i.e., access to the

exemplar-based system] has greater priority, but given that such a system, not inherently focusing on rules, may hit problems, it is possible to ‘shift down’ to a more rule-governed mode of processing, closer to the analyst’s model, as the need arises.

As pointed out by Bygate (1998), this processing can be regarded as one of the unique functions of L2; because ‘second language learners are likely to manage the two processes less seamlessly, compared with L1 speakers’ fluent balance between the two processes’ (p. 28). This dual-system account of speech production has been increasingly gaining theoretical importance among a number of researchers. For example, similar to Sinclair’s (1991) ‘open choice principle’ and ‘idiom principle’, Wray (2002:14) also proposes a dual-system model of analytic and holistic processing:

Analytic processing entails the interaction of words and morphemes with grammatical rules, to create, and decode, novel, or potentially novel, linguistic material. Holistic processing relies on prefabricated strings stored in memory. The strategy preferred at any given moment depends on the demands of the material and on the communicative situation, and so, importantly, holistic processing is not restricted to only those strings which *cannot* be created or understood by rule, such as idioms.

Combining this mechanism with Levelt’s monolingual model (see 1.2), L2 speech processing can be explained in the following way. Like L1, L2 speakers also prioritize the exemplar-based system in order to economize the processing burden. The fact that collocations are even more frequent in spoken than in written corpora (Butler, 1995, cited in N. Ellis, 2001:45) suggests the importance of formulaic items particularly under time constraints. Thus, receiving the output from the Conceptualizer (i.e., preverbal speech), learners attempt to search for appropriate lexicalized items in

lemmas in the first place. However, it is not uncommon to be in a situation of lacking appropriate lexical resources. In such a case, as a second resort, they are forced to improvise appropriate structures by accessing the rule-based system. Due to limited amounts of lexicalized items, L2 speakers more frequently need to step aside from the smooth processing and construct language from scratch.

L2 speakers utilize the rule-based system, not only to compensate for the exemplar-based system, but also to complexify their IL. In Levelt's model, lexemes comprising morpho-phonological information can partly play the role of the rule-based system, but its role in L2 speech processing is, beyond such lower processing levels, considered more dynamic and essential for constructing a global syntactic structure. Skilled L2 learners can 'switch between the two modes to take account of whatever processing demands are most pressing' (Skehan, 1996a:42) by considering various surrounding factors (e.g., fluency should be paramount in casual talk, or accuracy and complexity should be more emphasized in formal presentation) and anticipating the forthcoming situation. More importantly, access to the rule-based system might reduce the speed of speech delivery, but it plays a significant role in the 'competence' stretching point of view, because 'on occasions where rule-based systems are used for the generation of language, the products of such activity can themselves become exemplars and then retrieved and used as exemplars on subsequent occasions' (Skehan, 1996a:43).

Interestingly, this dichotomy between exemplar-based and rule-based systems is similar to the distinction between a 'creative faculty' and a 'critical faculty' in L2 monitoring, as defined by Morrison and Low (1983). The former 'assembles strings of language for private consumption or for articulation as utterances, while the latter 'gives awareness of what has been created, making it possible to check, either before or after articulation, for the frequent slips of the tongue, grammatical errors, social



infelicities and other deviations from intention that characterize normal speech'

(Morrison & Low, 1983:228). Morrison and Low (1983:246) argue:

Under optimum conditions, i.e. where demands on the system do not exceed the ability to cope, we may assume that the user has fairly good control over output. Under less than optimum conditions however (and this is probably the normal situation) demands on the system may constantly threaten to exceed capacity, and thus the user will be at best only barely in control.

One of the monitoring functions is to foster a balanced condition of L2 use, which tends to be lopsided in the exemplar-based processing in natural interaction.

### *1.6.2 Controlled and Automatic Processing*

Related to the distinction between rule-based and the exemplar-based processing, it is also important to consider the balance between controlled processing, which 'occurs slowly and functions serially' and automatic processing, which 'occurs rapidly and in parallel form' (Ellis, 2003:144). Becoming more automatic enhances overall performance by leaving relatively more time available for slower, centrally controlled processing to make important decisions (Segalowitz, 2003:387). Depending largely on controlled processes, learners tend to pay more attention to grammatical structures rather than transmitting message content (Ellis, 2003:144).

The distinction of controlled vs. automatic processing is closely correlated to the explicit vs. implicit knowledge distinction, because 'whereas implicit knowledge is typically highly proceduralized, allowing rapid access, explicit knowledge is available only via controlled processing' (Ellis, 2003:105). Since one of the difficulties in speaking is to formulate intended ideas under certain communication pressure, explicit knowledge needs to be transformed into implicit knowledge; as a

result of this, a greater amount of automatic processes are at one's disposal with increasing expertise in L2. Skill-learning theories account for this transformation by positing that the mastery of implicit knowledge is derived from the achievement of explicit knowledge, contributing to establishment of the groundwork for language acquisition (DeKeyser, 1998:59; see 1.5.5). Without this training, smooth and efficient transmission of messages is difficult, because learners always have to engage in controlled processing in order to access explicit knowledge in LTM.

This account is compatible with the recent trend of placing more weight on implicit knowledge than explicit knowledge in SLA. Hulstijn (2002:210) suggests that explicit knowledge is especially helpful in situations allowing careful monitoring of the information to be understood and produced without time pressure. Ellis (2004:245) also postulates that explicit knowledge is typically accessed through controlled processing when L2 learners experience some kind of linguistic difficulty in the use of L2. In a similar line of discussion, Wray (2002:10) demonstrated that 'although we have tremendous capacity for grammatical processing, this is not our only, nor even preferred way of coping with language input and output.' Following these statements, it is conceivable that the exemplar-based system associated with implicit knowledge is usually prioritized and the rule-based system associated with explicit knowledge tends to be secondary in spontaneous speech.

Following this view, it is also conceivable that the more explicit knowledge changes into implicit knowledge, the more proficient learners become. Ultimately and ideally, learners have to be equipped with only implicit knowledge for effective communication. The conventional view of skill learning assumed that learners automatize factual rule knowledge by applying and practising the rules until the entire procedure can be executed without conscious control (Hulstijn, 2002:211). However, this is not the case, not only because it is extremely difficult for adult learners to

achieve automaticity in a great part of linguistic processing (Poullisse, 1997), but also because explicit knowledge seems to play an important role even in advanced levels. Bialystok (1978) investigated the performance of grammaticality judgements under time pressure and more relaxed conditions with 317 learners of French, showing that when they had to make more detailed judgements about what part of the sentence was problematic or what rule was violated, time pressure made a difference. From this research, she concluded that L2 learners make their grammaticality judgements on the basis of implicit knowledge and switch to the use of explicit knowledge when more fine-grained decisions are required. Although this research was conducted in a decontextualized condition, it can be inferred that L2 learners might change gear between controlled and automatic processing and selectively and most efficiently use different levels of knowledge to meet the needs of particular situations. That is, they tend to economize the processing by mainly using implicit knowledge in rather simple conditions (i.e., grammaticality judgements), but they may engage in effortful processing using explicit knowledge in more cognitively demanding conditions. As pointed out by Schmidt (1992:366), 'skilled performance requires a balance between the speed of automatic processing and the goal directedness of controlled processing', including specific control processes such as rehearsal, search, planning, monitoring, and decision making of all kinds, as well as selective attentional allocation. That is, for effective language use, even in advanced levels, there should be some sort of controlled processing with sufficient attention in L2 speech. To sum up, it is important that L2 learners balance both controlled and automatic modes of processing for successful communication, although the extent of each greatly vary depending on conditions (e.g., the time available and familiarity with the subject matter [Foster, 2001]) as well as their proficiency levels.

## 1.7 Conclusion

Building on Levelt's monolingual speech model and Swain's output hypothesis, this chapter observed various cognitive factors influencing L2 oral performance. It is true to suggest that most discussion primarily focused on the Formulator in Levelt's model, because this stage of speech processing embodies qualitative and quantitative distinction between NS and NNS speech processing. De Bot (1996:548) observes that 'in principle the production system is available and that the generation of communicative intentions, which precedes language independently, poses no problems'. From a psycholinguistic point of view, it is the remaining steps in the Formulator that we should pay particular attention to in order to comprehend the L2 speaker's language. In line with this argument, it is a great challenge for SLA research to establish conditions to push learners into paying attention to the formulating processes, which is largely unproblematic for NSs.

Bygate (1998:29) claims that one's level of skill is conceptualized as a 'capacity for performance', and this can be influenced by 'contexts of operation'. Thus, the next chapter considers pedagogic interventions and how they influence the L2 processing. To investigate a pedagogic approach, the present study will take up 'task planning' and reveal more detailed functions of the dual-mode processing system. This implementation is likely to encourage learners to use the rule-based system, which tends to be neglected in everyday language use, to the fullest extent.

# CHAPTER TWO: A PEDAGOGIC APPROACH TO SECOND LANGUAGE SPEAKING

## 2.1 Introduction

Following the account of the output hypothesis and the cognitive processes of L2 production, this chapter considers pedagogic interventions of how balanced processing can be achieved. In what Kumaravadivelu (1994) calls the ‘post-method’ era, there are a considerable number of pedagogic approaches under the umbrella of Communicative Language Teaching (CLT). Among various candidates, I focus on ‘a subject of keen contemporary interest’ (Johnson, 2001:194), so-called task-based language teaching (TBLT). Within this framework, I pay particular attention to the two pedagogic traditions of ‘task planning’ and ‘focus-on-form’, not only because much related research (but not all) has focused on L2 oral production and its development in these areas, but also because both proposals have been developed within the field of SLA in close connection with psycholinguistic perspectives. The following section starts by accounting for a cognitive approach to TBLT, and then reviews the two specific pedagogic interventions. Through reviewing these teaching traditions, this chapter searches for the answers to the following questions: what do previous planning studies tell us?; how does planning contribute to L2 development?; what is the relationship between planning and focus-on-form? Understanding the issues concerning the two SLA theories, the last section attempts to integrate both approaches by referring to the significant cognitive processes.

## 2.2 A Cognitive Approach to Task-Based Pedagogy

Because the area of TBLT has been growing rapidly and extensively in both language pedagogy (e.g., Brown, Anderson, Shillcock, & Yule, 1984; Edwards & Willis, 2005; Nunan 1989, 2005; Prahbu, 1987; Willis, 1996) and SLA (e.g., Bygate, 1999, 2000, 2005; Bygate, Skehan & Swain, 2001; Candlin, 1987; Ellis, 2000, 2003; Long & Crookes, 1992; Skehan 1996a, 1998, 2003; Robinson 2001a, 2001b), it is necessary, in the first place, to clarify the perspective I take throughout the thesis. I start by defining the meaning of ‘task’ in the present pedagogic and research contexts, then account for the information-processing approach to task-based research as a theoretical foundation of this study.

### 2.2.1 Defining ‘Task’

The generic meaning of ‘task’ includes any types of classroom activity such as mechanical grammar exercises and practice activities (Willis, 1996), but what this term implies in the current trend is limited to ‘communicative task’. According to Rubdy (1998), in the early 1980s, the term ‘task’ deliberately replaced ‘communicative activity’ in applied linguistics and gained the specific meaning both in language teaching syllabus design and SLA research. As this account suggests, an interest in ‘task’ derives from CLT (Bygate, 2005), which had originally emerged in response to the failure of traditional language teaching.

Although the original meaning of ‘task’ covers wide categories, the present study adopts more focused guiding principles. Skehan (1998) summarizes the main task characteristics, following Candlin (1987), Nunan (1989) and Long (1989):

- meaning is primary;
- there is some communication problem to solve;
- there is some sort of relationship to comparable real-world activities;
- task completion has some priority;
- the assessment of the task is in terms of outcome.

These characteristics reflect the important themes of CLT, which puts a high value on the development of the ability to use appropriate language in authentic, communicative contexts. The first, ‘meaning-focused’, characteristic clarifies the separation of TBLT from explicit form-focused instruction (FFI; see 2.4.1 below). In this sense, tasks appear synonymous with ‘pure communicative activities’, but it is noticeable that primacy of meaning does not mean an absolute rejection of any FFI. TBLT does not include explicit types of FFI in the sense that traditional approaches do, but some sort of ‘focus-on-form’ can be one of the central objectives of TBLT (Long, 2000; Skehan, 2003). That is, the extent to which a learner attends to the message or code when undertaking tasks is not seen as categorical but variable and probabilistic (Ellis, 2003). It is thus possible that, despite their primary focus-on-meaning orientation, tasks aim to create the opportunities to switch learners’ attention to form in some conditions.

With ‘communication problem’ as the second characteristic, ‘task’ is primarily different from ‘exercise’. For example, the former assumes that linguistic skills are to be developed through engaging in communicative activities while the latter views linguistic skills as a pre-requisite for learning communicative abilities (Ellis, 2000). Furthermore, as pointed out by Widdowson (1998), tasks should be concerned with ‘pragmatic meaning’ rather than ‘semantic meaning’. That is, exercises generally focus on linguistic form and semantic meaning, evaluating the outcome in terms of conformity to the code, while tasks focus more on propositional

content and pragmatic communicative meaning, evaluating the performance in terms of whether the communicative goal has been achieved (Ellis, 2000). Moreover, task completion has some priority over the correctness of produced language. Doing a task to achieve a particular communicative goal will be far more likely to lead to increased fluency and natural acquisition than controlled exercises (Willis, 1996). This emphasis on communicative values is certainly regarded as one of the essential characteristics of TBLT.

The third principle implies 'genuineness' of language use. The meaning of task authenticity is slightly different between researchers in terms of the extent to which pedagogic tasks should conform to 'real' communication. One extreme is the definition by Long (1985; Long & Crookes, 1992). For Long, 'task' should incorporate 'the hundred and one things people do in everyday life, at work, at play, and in between' (Long, 1985:89) such as painting a fence and making an airline reservation, so that tasks are regarded as complex approximations of 'real-world' needs (Long & Crookes, 1992). In contrast to such 'situational authenticity', Skehan's standpoint is not so strict in that it tolerates indirect as well as direct relation to real world language. Skehan (2003:3) argues that 'it is the reaction of the learner or the research subject which is key, rather than the fact that a particular task may have, at some time, been used by native speakers (possibly in a different context)'. Taking similar view, Ellis (2003) demonstrates that the kind of language behaviour elicited by language learning tasks resembles the kinds of communicative behaviour that arises from performing real-world tasks.

This authentic learning account takes beyond the view regarding tasks as simply a tool to provide real communicative experience. Guariento and Morley (2001) suggest that the importance of task authenticity lies not only in its 'genuineness' but also in providing opportunities to control linguistic knowledge under 'real operating



conditions'. Following on from this view, tasks are not necessarily the mirror reflecting a real activity but involve opportunities to develop language processing capacity needed for real world language use.

Such pedagogic authenticity leads to the issue of task outcome. Ellis (2003) suggests that it is important to distinguish 'outcome' from 'aim', because the outcome of a task can be judged in terms of content, while the aim refers to the pedagogic purpose of the task. Learners may achieve the former without the latter, but 'the real purpose of the task is not that learners should arrive at a successful outcome but that they should use language in ways that will promote language learning' (ibid.:8). Therefore, tasks need to be devised to elicit language performance leading to language learning (but usually not specifying the exact language) through the process of achieving the designed outcome.

### *2.2.2 Goals in Task-Based Research*

Through the task defining process above, the present approach to 'task' has been specified, diverging from traditional approaches involving the procedures of explicit FFI and exercises in decontextualized contexts. Identification of these task characteristics is a significant starting point, but there are still ambiguities in theoretical perspective. The following section narrows down the focus by taking an information-processing approach.

As underlying task principles, Bygate (1999) suggests that a pedagogic task consists of a set of 'rubrics' (Bachman & Palmer, 1996) and input which promotes learners' language use involving language features and processing. Following Breen (1989), Ellis (2000) uses the term, 'task-as-workplan', to represent a similar concept, typically involving (1) some input (i.e., information that learners are required to

process and use) and (2) some instruction relating to what outcome the learners are supposed to achieve. A general principle in this view is that ‘a task is a device that guides learners to engage in certain types of information-processing that are believed to be important for effective language use and/or for language acquisition from some theoretical standpoint’ (Ellis, 2000:197).

A particularly noticeable point implied in these statements by Bygate and Ellis is an inseparable relationship between language and processing; that is, language use is not possible without language processing. It is important to note that learning involves both learning of language content and learning of processing (Bygate, 1999). The role of task lies in ‘shaping language development by mobilizing, stretching and refining current interlanguage (IL) resources’ (Samuda 2001:121). In other words, tasks should be based not only on such linguistic insights as the nature of knowledge of language but also on psycholinguistic insights such as to the processes involved in its use (Johnson 1979, cited in Bygate 2000:186).

Based on this processing-as-learning perspective, Skehan (1996a, 1998) proposes a cognitive approach which focuses on the psycholinguistic processes typically engaged in when learners do tasks. The underlying principle is that attentional resources are limited, thus learners need to allocate their attention to particular aspects of performance (see Chapter One), which can be affected by task characteristics and task conditions.

There is a dilemma that using, for example, a fluency-developing type of task usually leads to neglecting the learning of linguistic forms while the accuracy-oriented type easily fails to develop fluent use of a language. In order to achieve a balanced development, consideration of any impact of particular task characteristics and some kinds of implementation are important. Skehan and Foster (1999:98) specify the three goals of task-based research:

1. How balance may be achieved among different performance areas;
2. How task characteristics can influence performance and influence balance among the goals; and
3. How task conditions can influence performance and influence balance among the goals.

As tasks should prioritize meaning in principle, it may be necessary to incorporate learners' attention to formal aspects by particular task characteristics and conditions. That is, some specific task-characteristics may influence learners' performance variable, while the influence of the task conditions can affect the language outcome. The former, the design features of tasks have been traditionally the main concern in much SLA and language teaching research, and researched in various angles, for example, types of interaction in different unfocused tasks (e.g., Pica *et al.*, 1989) and the ways to achieve the learning of target structures in focused tasks (e.g., Fotos & Ellis, 1990; Loschky & Bley-Vroman, 1991). The present study pays more attention to the latter issue, task condition.

### **2.3 Investigating the Planning Effects on L2 Performance**

Among various ways of establishing task conditions, the present study focuses on one of the pre-emptive activities, 'planning', which is one of the most active areas in task-based research. Skehan (1996a:53) postulates:

...the general purpose of pre-emptive activities is to increase the chance that some restructuring will occur in the underlying language system, and that either new elements will be incorporated, or that some re-arrangement of existing elements will take place.

This first section considers the meaning of ‘planning’ in the current SLA research context and the ways influencing L2 performance and development. In the following sections, I review previous planning studies and the relationship between planning and learner proficiency.

### *2.3.1 Types of Planning*

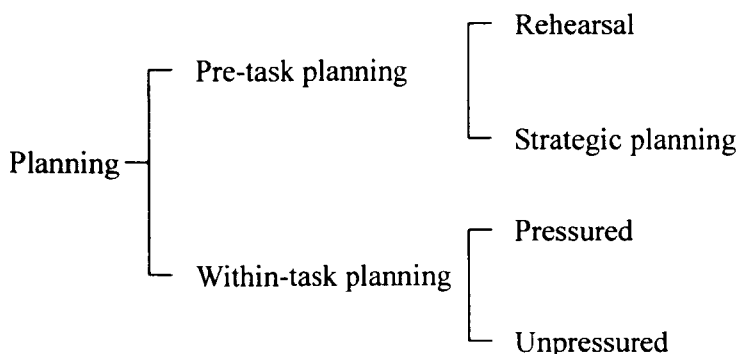
Beyond the everyday sense of *planning*, it is necessary to understand that, before making any utterance, we engage in some sort of planning process. In a recent paper on task planning, Ellis (2005:3) describes the nature of planning:

All spoken and written language use, even that which appears effortless and automatic, involves planning. That is speakers and writers have to decide what to say/write and how to say/write it. Planning is essentially a problem solving activity; it involves deciding what linguistic devices need to be selected in order to affect the audience in the desired way.

Certainly, our daily life consists of a number of effortless and automatic productions, seemingly made without *planning*, but the reason for this impression is largely a matter of consciousness. Even in a highly automated production, there is always some sort of planning process involved, such as constructing a message and selecting an appropriate expression.

Because there are several senses of planning in the research tradition, it is necessary to identify which type of planning researchers particularly focus on in their research. Figure 2.1 summarizes various types of planning, which have been developed in the task-based framework.

**Figure 2.1:** Types of Task Planning (Ellis, 2005:4)



The first principal distinction is whether planning particularly occurs before an assigned task (i.e., pre-task planning) or during task performance (i.e., within-task planning). L2 planning studies have been traditionally interested in the effects of pre-task planning, probably because of its pedagogical orientation. Pre-task planning can be subdivided into ‘rehearsal’ and ‘strategic planning’; ‘rehearsal’ provides an actual performance opportunity before the subsequent performance as preparation (e.g., Bygate, 1996, 2001; Bygate & Samuda, 2005; Gass, Mackay, Alvarez-Torres & Fernandez-Garcia 1999; Lynch & Maclean, 2000, 2001; Pinter, 2005, 2007). The other, ‘strategic planning’ (SP) gives learners the opportunity to use a period of time (e.g., ten minutes) *strategically*.

The within-task planning position is also further divided into ‘pressured’ and ‘unpressured’. The premise that every spoken and written production involves a planning phase (Ellis, 2005) suggests that we cannot establish the complete ‘non-planning’ condition in a strict sense, but it is possible to minimize the degree of planning ‘on-line’ methodologically by putting speakers under some sort of pressure. In real world language use, speakers usually suffer from communication pressure from interlocutors, and this often makes them unable to use their competence to its

full potential, finally ending up with simplified and erroneous production. Thus, it can be assumed that pressured conditions tend to limit their potentiality, while unpressured conditions give speakers more comfortable access to their linguistic repertoire. The former condition can be called 'non-planning' (or more strictly 'minimal planning' [Crookes, 1989]), which is often used as a controlled condition in planning research. On the other hand, the latter can be particularly labeled 'on-line planning' (Wendel, 1997), 'careful on-line planning' (Ellis, 2005) or 'on-task planning' (Bygate, 2005), in contrast to 'off-line' in strategic planning.

### 2.3.2 *Strategic Planning*

Among the various types of planning, the present thesis focuses on 'strategic planning' and 'on-line planning'. This section reviews the previous studies involving strategic planning, and then the next section concerns those involving on-line planning.

In the early period of planning research, Crookes (1989) investigated the relative effects of planning conditions (10-minute strategic planning vs. non-planning) on IL variation (i.e., complexity, accuracy, lexical variety and discourse). The result showed that the participants produced significantly more complex language, but did not clearly identify a significant improvement of accuracy.

Following Crookes (1989), Foster and Skehan (1996) investigated the effects of planning variables on oral performance (fluency, complexity, and accuracy). There is a significant shift from exclusive focus on planning *product* to more focus on planning *process* by operationalizing detailed (i.e., instruction of how to conduct planning is given) and undetailed planning conditions (i.e., no instruction is given). Consistent with Crookes (1989), the results showed that planning had an influence on

fluency and complexity but mixed results with accuracy; that is, in terms of accuracy, undetailed planners, unexpectedly, outperformed detailed planners as well as non-planners in all tasks. In response to this result, Foster and Skehan (1996) proposed trade-off effects between complexity and accuracy, arguing that 'the goals of complexity and accuracy compete for limited information processing resources and that what is achievable depends on the precise allocational decisions made' (p. 320). That is, it can be assumed that the detailed planners used more complex language at the expense of accuracy, while the less ambitious inclination of the undetailed planners produced more accurate language.

Following on from these early planning studies, there have been a number of other studies conducted from various perspectives. For example, Mehnert (1998) investigated the effects of different amounts of planning time (1, 3, 5 and 10 minutes). The analysis of task performance suggests that fluency and lexical density increases as a function of planning time, while the results of complexity and those of accuracy are not consistent. That is, complexity was significantly higher only for the 10-minute planning, and accuracy improved with only 1-minute planning. From the findings, Mehnert (1998:104) observes that 'any gains in accuracy and complexity are not simultaneously achieved', following Foster and Skehan (1996). That is, when provided with planning time, the first priority was accuracy, while when provided with a greater amount of planning time, attentional resources were allocated to complexity.

To explore planning processes, Foster and Skehan (1999) examined different sources of planning (teacher-led, solitary, group-based) and different foci (language vs. content). The teacher-led condition showed greater accuracy effects, while the solitary planning condition had greater influence on complexity, fluency and turn length. On the other hand, the group-based planning did not produce any positive effects, and the

language vs. content planning conditions led to only little effects.

A significant development in recent planning research is an interest in the process-oriented view. A clear example is the studies by Ortega (1999, 2005) that examine the process of L2 planning through a retrospective method. Drawing on the information-processing theories that planning may lessen the cognitive load and free up attentional resources at the micro levels of speech production, Ortega (1999) hypothesized that the conscious attention of L2 speakers may shift to formal aspects of the language, and thus strategic planning could enhance learners' attention to form without directing them specifically to attend to the code. However, she also suggested that 'attention to form cannot be assumed as a guaranteed by-product of pretask planning opportunity' (ibid.:136). Many other factors (e.g., the communicative requirements of the task, learners' predisposition towards communication or accuracy, and learner proficiency) might have an influence on accurate performance (Ortega, 1999, 2005; also see Sangarun [2005] for an analysis of strategic planning processes through a introspective method).

To sum up, the most noticeable point drawn from the strategic planning literature is a clear influence of planning on 'complexity' and 'fluency' but limited influence on 'accuracy'. In theory, it can be hypothesized that planning opportunities lead to better performance in all the three components by freeing up the cognitive capacity, but in practice the impact on accuracy is fairly limited. Such observations are in accordance with those in Canadian French immersion studies (e.g., Allen, *et al.*, 1990); advanced learners have good command of fluent, complex language use and a wide range of vocabulary but still repeatedly fail to produce accurate utterances, suggesting that accuracy is a final gigantic step for L2 speakers. As a natural consequence of this inextricable problem, it is necessary to advocate a particular intervention to increase the level of accuracy, assuming, as this research does



throughout, that this is both a desirable and an achievable goal.

Another observable point is that most studies have been rather lopsided towards the *product* aspect of planning through the task performance analysis, and an investigation into the planning *process* remains largely untouched, except in a highly speculative manner. Foster and Skehan (1996) took the first step into an exploration of this aspect, but their distinction of detailed and undetailed planning seems simplistic. An exception is the introspective/retrospective methods by Ortega (1999, 2005) and Sangarun (2005), but these studies focus only on the strategic planning processes during a certain period of time, not the on-task processes. Thus, further exploration of this aspect is needed for our understanding of the processes of planning.

### 2.3.3 On-Line Planning

To understand the nature of planning and the difficulty of accuracy-improvement, it is useful to take 'on-line planning' into account. With reference to the mixed results of accuracy in strategic planning research, Yuan and Ellis (2003) emphasize that it is important to distinguish between strategic planning and on-line planning. This argument is primarily based on the study by Wendel (1997) that the trade-off is more relevant to the relationship between fluency and accuracy, and whether or not learners attend to fluency or accuracy depends on the type of planning. It cannot be emphasized too strongly that, in spite of such focus on *off-line* (i.e., strategic) planning, it is likely that the participants in the previous strategic planning studies (see 2.3.2) were also implicitly engaged in *on-line* planning, assuming that the degree of on-line planning engagement might yield different effects on accuracy results.

However, it is noticeable that focusing on this methodological issue might give us a clue to solving the complex nature of accuracy in L2 speech production. One

of the possible procedures to improve accuracy is to implement this on-line planning process in a research design. As noted by Wendel (1997), this distinction is not only a methodological convenience, but also involves an important psychological orientation of a participant towards a task with implications for processing capacity, attentional resources and language production.

In a similar line of argument, Hulstijn and Hulstijn (1984) investigated the influence of time pressure (present or absent) and focus of attention (on information or on grammar) on the correct use of two Dutch word order rules in 32 adult L2 learners. The findings are that focus of attention on grammar resulted in more correct utterances, while the time pressure did not directly contribute to correctness of their grammar. From the results, Hulstijn and Hulstijn (1984) concluded that time in itself is not a necessary condition for successful self-correction but focus on form generally requires time in order to bring about successful self-correction. That is, erroneous production is the result of insufficient involvement into gramaticization, which is often hindered by a lack of on-line processing time.

To illustrate the effects of time pressure on task performance, it is important to look at the study by Ellis (1987), investigating the differences of adult learners' use of past tense between planned (written and oral) and unplanned (oral) conditions. In the first task, the participants were asked to write a story based on a given picture in one hour (planned written). Then, they were asked to record an oral version of the story without looking at the written version in the second condition (planned oral). Finally, given two more minutes, the participants were asked to record an oral version of the second picture (less planned oral). As pointed out by Yuan and Ellis (2003), this three-way condition (less planned oral, planned oral and planned written) includes the distinction of strategic planning and on-line planning in an implicit manner, because the writing process allows the learners to monitor the ongoing composition process in

terms of both content and language. Taking into account this implicit link with on-line planning, Ellis's findings are striking; accuracy on the regular past form declined in the unplanned condition (Task 3), while the correctness of the irregular past form was more or less consistent between the planned (Task 1 & 2) and unplanned conditions (Task 3). That is, regular past tense was improved in the on-line planning condition, giving evidence that the learners were able to access rule-based knowledge, while at the same time, exemplar-based knowledge was unaffected.

Following Ellis (1987), Yuan and Ellis (2003) investigate the effects of strategic and on-line planning on L2 oral performance with Chinese college students in China. The results show that, in fluency measures, the strategic planning group spoke the fastest and reformulated the least, while the on-line planning group spoke the slowest, reformulated and repeated the most. In complexity measures, strategic planning had a positive influence on grammatical complexity and also greater lexical variety, while on-line planning led to greater grammatical complexity but not lexical variety. In accuracy measures, the on-line planning group had the highest results, while the non-planning group was the lowest.

To summarize these studies, it can be assumed that careful on-line planning has a marked effect on learners' oral and written production, but in a different way from strategic planning. The results also support the on-line planning rationale and account for the mixed results of accuracy yielded by the strategic planning research; because learners with limited amount of lexicalized items and procedural ability in L2 might experience problems in formulating the language, they may be able to compensate for this lack of knowledge by monitoring their production using explicit knowledge.

The findings of Yuan and Ellis (2003) have made a significant impact on the progress of task planning research, but a more detailed examination of the effects of

on-line planning is needed. Skehan and Foster (2005) give a slightly different definition of the concept. They regard on-line planning as ‘a measure of how much speakers regroup in real-time as they modify what is formulated in their utterance’ (ibid.:214), and attempt to identify measures reflecting speakers’ on-line planning engagement. The study provides surprise information mid-task to examine any impact on the nature of performance, assuming that there may be more evidence of on-line planning among participants after surprise information is presented compared to participants who simply carry on doing a task with the same information as before. Factor analysis shows a set of indicators entirely consistent with the construct of on-line planning; a group of measures such as mid-clause pauses, filled pauses, length of run, reformulation and false-starts (but, importantly, not end-of-clause pauses) work together and ‘link particular features of disfluency to the ways on-line planning has been characterised psycholinguistically’ (ibid.:205).

However, such performance measures cannot be straightforwardly linked to a range of psycholinguistic processes which implicate on-line attention to speech such as snatching time in order to plan the language for what will be said and monitoring ongoing performance (Skehan & Foster, 2005). The on-line planning cluster can be observable from speakers’ performance, but these underlying cognitive thought processes, or reasons for committing such dysfluent features cannot be identified from the analysis of task performance. As mentioned above, there have been some attempts to investigate learners’ planning process during strategic planning (e.g., Ortega, 1999, 2005; Sangarun, 2005), but it is also necessary to investigate learners’ ‘on-line’ planning processes.

Another noteworthy point is the use of interactive tasks in Skehan and Foster (2005) rather than monologic tasks, which were used by Yuan and Ellis (2003). Interactive tasks seem to give slightly different effects of on-line planning from

monologic tasks because ‘one contributor may have time pressures eased while the other speaker holds the floor’ (Skehan & Foster, 2005:196). This point makes us realize a possibility that on-line planning research of interactional tasks could be methodologically more difficult, because speakers may more frequently engage in on-line planning without obvious linguistic features (e.g., unfilled pauses) while the interlocutor is holding the floor. I will further consider the different impacts of monologic and dialogic tasks on learners’ performance in Chapter Three (3.6.1).

### *2.3.4 Interaction between Proficiency and Planning*

Related to planning implementation, another significant factor which may influence performance is learner proficiency. Many previous planning studies only investigated a single proficiency group. Some studies (e.g., Crookes, 1989) include a wide range of proficiency levels, but did not examine the effects of proficiency in the analysis.

There have been a few attempts to inquire into the relationship between planning and proficiency levels. In a testing situation, Wigglesworth (1997) conducted research on the influence of planning time (one minute) on L2 performance between different proficiency levels (28 high vs. 28 low-proficiency candidates) in different types of task (i.e., ‘summary of conversation’ – ‘picture description-comparison’ – ‘telephone answering-machine message’ – ‘general discussion questions’). She found that the planning time only helped the more highly proficient learners to produce more complex language (i.e., subordinate clauses) and more accurate language (i.e., verb morphology) in more cognitively demanding tasks, while clear planning effects on low-proficiency candidates were not found. From this result, Wigglesworth (1997:120) tentatively observed that ‘for the high-proficiency candidates, planning time may be beneficial in a situation where the cognitive load required of the task is

heavy' while 'the low-proficient level candidates do not appear to benefit from planning time' (p. 104). This observation led to the assumption that the high-proficiency candidates may be focusing on the form and complexity of their linguistic output, while the low-proficiency candidates are focusing on content (ibid.).

In addition, Ortega (1999) approached the relationship between planning and proficiency by looking at qualitative data collected in retrospective interviews. From the retrospective findings in 1995 and 1999 studies involving different proficiency levels of speaker (low and advanced respectively), Ortega (1999) speculated that planning may be more likely to have an effect on the performance of learners with higher levels of proficiency. However, she also identified that low proficiency learners seem to be able to engage in effective and varied strategy use in strategic planning, although they may need more attention and time on planning content and lexis as a first priority.

To explore the effects of proficiency factors, Kawauchi (2005) devised 'planning activities' in a series of tasks comprising unplanned story-telling task – planning activities – questionnaire – planned story-telling task with three different proficiency levels (low-intermediate, high-intermediate and advanced). In the planning stage, learners were given one of the following activities: writing a draft, rehearsing, or reading a model L2 input. Although the study did not show any significant statistical differences in the effects of the three planning activities on performance, proficiency factors revealed positive effects on fluency, complexity and accuracy in all the three groups; the greatest gains for high-intermediate were in fluency (i.e., the number of words and percentage of repetitions) and complexity (i.e., clauses per t-units, words per t-units, number of subordination and number of types); and the largest gains for low-intermediate were in accuracy (i.e., the percentage of correct past tense). The findings are different from the positive planning effects on

higher proficiency learners found in Wigglesworth (1997), suggesting that sufficient proceduralized knowledge allowed advanced learners to handle the task well even in the unplanned condition (i.e., ceiling effects), while planned conditions enabled the high-intermediate learners to access declarative knowledge because of their limited proceduralized knowledge (Kawauchi, 2005:162).

Given the inconsistent results in Wigglesworth (1997), Ortega (1999) and Kawauchi (2005), further qualitative and quantitative investigation into the effects of planning in learner of varying proficiencies is necessary.

## **2.4 Implementing Focus-on-Form Pedagogy**

So far, this chapter has mentioned that planning implementation could enhance learners' performance by releasing on-line processing pressure. A key to this functioning is that learners' attention should be drawn to form *as well as* meaning. The investigation of the learners' attention to form in communicative settings is particularly called 'focus-on-form', and ultimately the present thesis searches for the link between planning and focus-on-form. Before looking into the issues of focus-on-form, I will describe more general, 'form-focused instruction' (FFI) with a view to knowing the position of the concept in a broad category. Because the term, focus-on-form, has been differently defined among researchers, I attempt to classify the definitions of focus-on-form first, and then specify the focus-on-form position in the present discussion.

### *2.4.1 Form-Focused Instruction*

Ellis (2001:1-2) defines FFI as any planned or incidental instructional activity that is

intended to induce language learners to pay attention to linguistic forms, ranging from traditional grammar teaching to a communicative approach with some grammar reference. Following this definition, those who take the most explicit and elaborated form-focused position are regarded as upholders of the traditional method (e.g., explicit rule provision, repetition of models and transformation exercises), which is, despite much criticism in the history of SLA, still dominant in many countries (Thornbury, 1999; Holliday, 1994). An alternative position which involves less explicitness and elaboration but predetermined grammar structures is widely recognized as a ‘consciousness-raising’ approach (e.g., Rutherford, 1987; Rutherford & Sharwood-Smith, 1985; Sharwood-Smith, 1981), which takes certain communicative-oriented task forms but involves certain grammar targets involving various forms such as grammar consciousness-raising tasks (Fotos & Ellis, 1990), structured communication tasks (Loschky & Bley-Vroman, 1990) and focused communication tasks (Nobuyoshi & Ellis, 1993). This type of approach promotes learners’ intake of linguistic items and the restructuring of existing knowledge by taking account of their psycholinguistic processes, but this is still considered as a teacher-centered position in terms of pre-setting the teaching items against the natural order hypothesis (Dulay, Burt & Krashen, 1982).

At the opposite end of the continuum is ‘focus-on-form’ (Long, 1991), in Ellis’s definition of FFI, involving ‘incidental’ activity including a communicative approach with some grammar reference. Central to this view is the claim that learners’ attention to form needs to arise *incidentally* from on-going communication in order to facilitate form-meaning-use mapping. In the post-Krashen era, the importance of FFI is well justified and the question of what particular approach is most effective has been a major concern. Norris and Ortega (2000, 2001) examined 49 FFI research articles published between 1980 and 1998 and found that explicit types of instruction



are more effective than implicit types, and that ‘focus-on-form’ and discrete-point grammar teaching (so-called ‘focus-on-formS’; see 2.4.2 below) interventions result in equivalent and large effects. The concept of focus-on-form has attracted much attention in SLA, because of its firm psycholinguistic foundations (e.g., Long, 1991, 2000), but the essential features that supposedly distinguish focus-on-form and focus-on-formS instructional approaches have been inconsistently operationalized (Norris & Ortega, 2001:203). Therefore, the following section specifically looks at the issues of focus-on-form involving its theoretical definition, pedagogical classification of the focus-on-form types, and psycholinguistic effects of this particular intervention.

#### *2.4.2 Defining Focus-on-Form*

There is a considerable body of research on the focus-on-form model constructed by Long and associates. The original motivation for the concept is to reject the dichotomous options of focus-on-formS (teaching linguistic forms isolated from communication, such as Grammar-Translation Method and Audiolingualism) and focus-on-meaning (no overt focus on linguistic forms at all, such as the natural approach and content-based learning) widely employed in foreign language classrooms, and compromise them to raise the ultimate level of attainment. The next section first looks at the focus-on-form and focus-on-formS classification. Then, more specific classifications are considered: ‘planned’ – ‘incidental’, and ‘pre-emptive’ – ‘reactive’ focus-on-form. Through these classifications, I specify the focus-on-form position applied in the present study.

**Focus-on-form versus focus-on-formS.** In his seminal paper (Long, 1991, replicated as Long, 2001), the term focus-on-form is defined as overtly drawing learners’

attention to linguistic elements as they arise incidentally in lessons whose overriding focus is on meaning, or communication. As Long and Robinson (1998:23) explain: 'focus on form often consists of an occasional shift of attention to linguistic code features – by the teacher and/or one or more students – triggered by perceived problems with comprehension or production', but 'not isolated from communication' (Doughty & Williams 1998a:197). Presumably, focus-on-form is a deliberate realization of what successful L2 learners have experienced; through natural or classroom interaction, they may notice the significant differences between input (i.e., TL) and their current knowledge (i.e., IL); it follows that learners' IL, activated by this focal attention, may start restructuring. That is, focus-on-form is an attempt to encourage such essential processes of SLA more actively.

A key to this theorization is its conformity to a learner's psycholinguistic readiness to acquire an L2, because focus-on-form teaching puts emphasis on learners' 'built-in syllabus' (Corder, 1967; Lightbown, 1998), which should be discovered and solved within meaning-focused situations. According to Pienemann's (1985) Teachability Hypothesis, because there is a gap between the teacher's and learner's syllabus, teaching particular linguistic items (focus-on-formS) often fails to lead to integration in the learners' own system. On the other hand, as many Canadian immersion studies show, natural learning and content-based learning (focus-on-meaning) tend to contribute to the development of fluency but not sufficiently to that of accuracy. The latter problem has made us realize the importance of paying conscious attention to linguistic forms, but simple return to the former position is not favoured. The solution that has attracted much attention is that learners are encouraged to notice their own current problems through meaningful language use, because teachable items can only be identified through learnable items (Long, 2001:184). Thus, in this option, 'which forms are targeted, and when is determined by

the learner's developing language system, not by a predetermined external linguistic description' (Long, 2001:185).

**Planned versus incidental focus-on-form.** As suggested by Ellis (2001:15), Long's original definition of focus-on-form essentially includes 'incidental' characteristics. However, other focus-on-form studies do not necessarily regard this as an essential feature, and the definition has been further extended to 'planned' as well as 'incidental'. In the planned option, the instruction will be intensive, in the sense that learners will have the opportunity to attend to a single preselected form many times, while the incidental option will lead to extensive instruction, because a range of linguistic forms (grammatical, lexical, phonological, pragmatic) are likely to arise as candidates for attention (Ellis, 2001). Taking these things into account, Doughty and Williams (1998a) slightly modified Long's definition, involving three features of focus-on-form: (1) the need for learner engagement with meaning to precede attention to the code; (2) the importance of analyzing learners' linguistic needs to identify the forms that require treatment; and (3) the need for the treatment to be brief and obtrusive. In any planned and incidental options, the primary focus of learners is on meaning, and the subsequent needs to analyze linguistic codes trigger the focus-on-form processing. In the case of incidental focus-on-form, this shift of attention is fundamentally natural, while the planned option needs to devise 'natural', 'essential' and 'useful' characteristics in tasks (Loschy & Bley-Vromann, 1990) as if learners voluntarily notice such linguistic problems.

**Pre-emptive versus reactive focus-on-form.** In addition to the above elaboration of the concept, incidental focus-on-form can be further subdivided into two options: 'pre-emptive' and 'reactive'. In the former option, a teacher or a learner takes time out

from a communicative activity to initiate attention to a form aspect that is perceived to be problematic in the use of the form or difficulty with message comprehension (Ellis, 2001). An important characteristic of this option is the teacher's and learners' brief switches from language user to language learner (Ellis, 2001:22-3). In this classification, therefore, the separation of form-focused teaching and communicative involvement seems rather clear. So, it is possible to devise a very explicit, deductive explanation of language in the focus-on-form instruction, which can be regarded as a simple combination of traditional form-focused method and communicative activities. On the other hand, the reactive option is often recognized in the form of negative feedback, which is given by teachers in response to learners' errors. This differs from the pre-emptive option in drawing learners' attention to form in the case of error-detection. This type of focus-on-form has been widely investigated either in implicit forms, specifically known as 'recasts' (e.g., Long, Inagaki & Ortega, 1998) or more explicit form, such as 'metalinguistic feedback'.

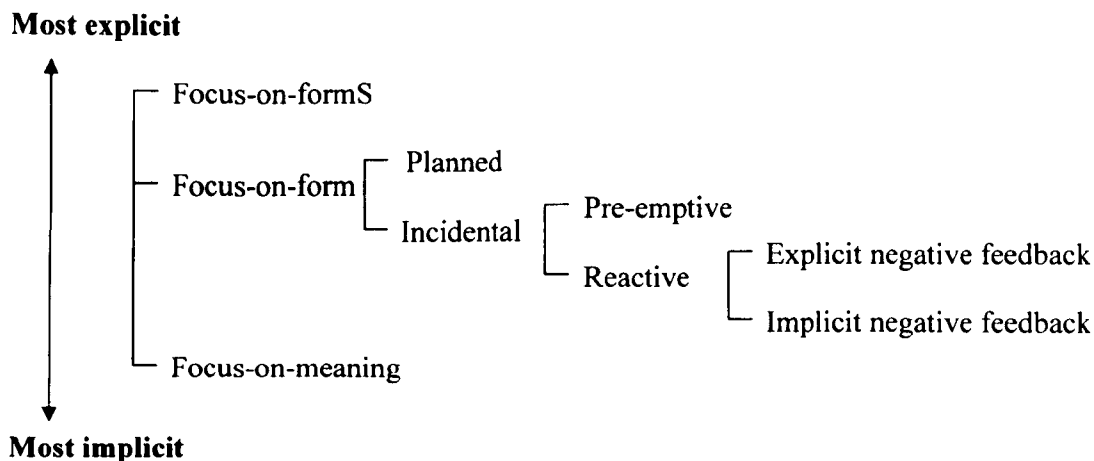
As briefly mentioned above (1.4.5), Lyster (1998a, 1998b; Lyster & Ranta, 1997) investigated some reactive focus-on-form, concluding that explicit types of focus-on-form are more effective in resulting in learners' uptake, which is important for successful L2 development. That is, learners need to be pushed to produce the linguistic items which they already know, ideally in a form of learner-generated repair. In doing so, they could pay more attention to the problematic linguistic aspect than merely being given a model form (i.e., recasts). Drawing on van Lier (1988), Lyster and Ranta (1997) pay attention to the distinction in classroom discourse between a conversational function and a didactic function. The former involves the 'negotiation of meaning', characterized as 'exchanges between learners and their interlocutors as they attempt to resolve communication breakdowns and to work toward mutual comprehension' (Pica, *et al.*, 1989:65). According to the Interaction Hypothesis, the

relatively implicit use of interactional moves provides opportunities of negotiation of meaning, which is believed to facilitate L2 development (Long, 1996). On the other hand, the latter involves the so-called 'negotiation of form', the provision of corrective feedback that encourages self-repair involving accuracy and precision and not merely comprehensibility, using elicitation, metalinguistic clues, clarification requests and repetition of error (Lyster & Ranta, 1997). Lyster (1998a) doubted the effectiveness of recasts, because this technique tends not to interrupt the flow of communication, and so there is a danger that L2 learners may not readily notice target-nontarget mismatches in the interactional input. That is, recasts, similar to input provision, seem to bring only limited amount of attention to form, whereas the other forms of corrective feedback (i.e., negotiation of form) tend to need more attention to form without blocking the flow of communication. Underlying these principles, the study by Lyster and Ranta (1997) shows that neither recasts nor explicit correction led to any peer- or self-repair, because they already provide correct forms to learners. In contrast, other ways of corrective feedback such as elicitation, metalinguistic clues, clarification requests and repetition of errors not only led to higher rates of uptake but also led to more peer- and self-repairs. This is because the retrieval of TL knowledge that results in self-repair following a teacher's metalinguistic clue or elicitation move requires more attention to the analysis of target-nontarget mismatches than does repetition of a teacher's recast or explicit correction (Lyster, 1998b:185-6). Therefore, Lyster and Ranta (1997) distinguished these interactional moves as negotiation of form from recasts and explicit correction, which are classified as mere rephrasings of students' utterances.

Although it is inconclusive whether explicit or implicit negative feedback is more effective for SLA, at least the series of studies by Lyster emphasize the importance of considering the degree of explicitness within the realm of

focus-on-form and the difficulty of balancing between the two extremes of most explicit and most implicit conditions. Because both the most explicit type (e.g., deductive rule explanation) and the most implicit type (e.g., natural learning) are unlikely to lead to successful L2 learning, what is required is to identify the most workable option among the intermediate positions, or ‘focus-on-form’. The problem is that the most appropriate level of explicitness is not absolute but highly dependent on external factors such as the teaching contexts (e.g., ESL vs. EFL) and internal factors such as proficiency levels and individual differences (ID) of learner types (Robinson, 2002; Skehan, 1989, 2002). Therefore, what is more favourable for teachers is to select the most effective type of focus-on-form from such options in individual teaching situations, rather than searching for the best generic approach. Figure 2.2 summarizes the focus-on-form options observed in this section.

**Figure 2.2:** Types of Focus-on-Form Instruction



In search of the optimal condition for the present study combined with planning implementations, my position comprises the incidental option rather than planned, because planning instruction does not include any specific target items. With

regard to the pre-emptive vs. reactive option, it is possible to incorporate the former in teaching, but the present study does not intend to provide this sort of ‘stepping aside’ stage. It is thus considered that planning opportunities provide conditions to elicit incidental/reactive focus-on-form chances. Moreover, as any negative feedback is not expected by interlocutors (through the use of monologic tasks), focus-on-form may occur ‘internally’, not ‘externally’, in the present study.

## **2.5 Integrating Planning and Focus-on-Form**

Having overviewed the issues of task planning and focus-on-form, the rest of this chapter attempts to integrate these two different research traditions. In the following section, I will argue that planning implementations are likely to produce form-focused effects on learners’ language by referring to the psycholinguistic processes, which were considered in Chapter One.

### *2.5.1 Mechanisms of Planning Effects on L2 Performance*

Having looked at previous strategic/on-line planning studies and focus-on-form principles, this section considers how planning psycholinguistically influences learners’ performance. With respect to the relationship between ‘task’ and ‘focus-on-form’, it is important to mention the essential condition to induce focus-on-form. Doughty and Williams (1998b:4) demonstrate that ‘meaning and use must already be evident to the learner at the time that attention is drawn to the linguistic apparatus needed to get the meaning across’. The principle of ‘task’ satisfies this feature. Since the task has a meaning focus and the decision of what to produce is left to the learner, ‘meaning and use must be evident’ at the time of production. As the

meaning processing aspect is completed, the learner's attention tends to be drawn to the form aspect (VanPatten, 1996). If the intended production is unproblematic, the learner may not focus on the form or strengthen it, but if the item is not fully internalized yet, his/her attention may be directed to form, so noticing his/her inability to utilize this item.

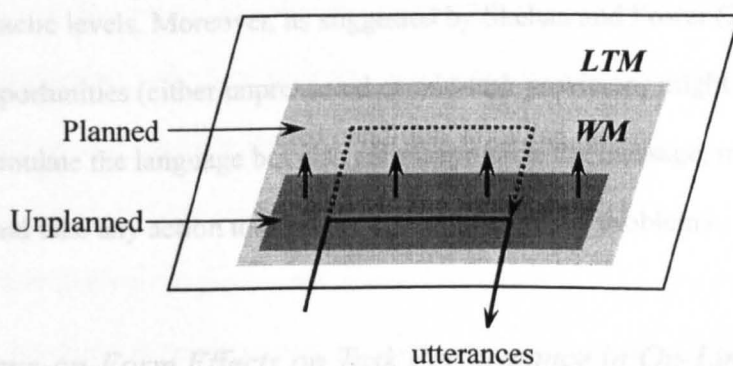
In addition to the relationship between task and focus-on-form, 'planning' seems to encourage learners' attention to form. Spontaneous speech tends to focus on conceptual processing and neglects morpho-syntactic processing due to the limitations of cognitive capacity and processing time. Pursuing this relationship of human cognition and L2 performance, it is conceivable that as a result of relaxing time pressure by planning, learners can more comfortably access pre-lexicalized and/or less automatized items. These items are not under their full control, because they are not developmentally ready to incorporate them into their IL (Lightbown, 1998; Pienemann, 1985; Pienemann & Johnston, 1987).

In spite of this prematurity, if sufficient processing time is given, it is likely that learners are able to use the language more correctly and appropriately by accessing grammatical knowledge in their LTM (Hulstijn & Hulstijn, 1984; Bialystok, 1978). For example, learners may be able to correctly answer a question about a particular grammar point in grammatical judgement tests, but this ability does not necessarily guarantee their performance in communicative situations. Access to declarative knowledge requires considerable time to search for and employ it. Taking into account that speakers' minds are pressured within the limited WM capacity, depleting attention and time is extremely problematic, particularly in speaking. As pointed out by much SLA research (e.g., Swain's [e.g., 1985] 'pushed output'; see 1.4), the opportunities to attempt to produce the language beyond the present comfortable state of control is crucial to developing linguistic competence beyond the



state of stabilization. As far as these processing problems are concerned, having planning time prior to a task can be one possible implementation to stretch the upper limit of IL by providing the optimal condition to access not-yet-internalized items and to elicit more upgraded language use in LTM. Figure 2.3 illustrates that the capacity of WM can be expanded in the planned condition, allowing the speaker to access a wider range of linguistic knowledge in LTM. The figure shows that only a limited amount of knowledge stored in LTM can be accessed, while accessible knowledge can be expanded in the planned condition.

**Figure 2.3:** The Effects of Planned Condition on L2 Performance



Applying the findings of strategic and on-line planning research to Levelt's (1989; see 1.2) speech production model, it is conceivable that strategic planning opportunities are mainly related to the Conceptualizer rather than the Formulator, as suggested by the results of limited accuracy improvement in much strategic planning research. That is, strategic planners are mainly involved in the overall planning of 'the direction and phrases of the discourse' (Bygate, 2001:25) through accessing the general knowledge store.

On the other hand, careful on-line planning is considered to bring learners not only into the Conceptualizer but also into the Formulator by releasing them from

on-line processing pressure. Following the function of WM (e.g., Baddeley, 1986; Miyake & Shah, 1999), Ellis and Yuan (2004, 2005; Yuan & Ellis, 2003; Ellis, 2005) theorize that, when learners have the opportunity to engage in careful on-line planning, they are better able to access the LTM via the central executive or supervisory attentional system. This will assist the language formulation stage of speech processing, in particular, the planning of grammatical features, which are typically accessed in the planning process later than lexical items (Skehan, 1998). That is, when speech production is pressured, learners make use of the limited processing time available to them to search mainly for lexical material, but, when it is unpressured, they are better able to search their LTM for grammatical information, especially at morphosyntactic levels. Moreover, as suggested by Skehan and Foster (2005), on-line planning opportunities (either unpressured or mid-task provision) might help speakers not only formulate the language but also reconceptualize the message, monitor the utterances and take any action to perceived pragmalinguistic problems.

### *2.5.2 Focus-on-Form Effects on Task Performance in On-Line Planning*

In line with positive effects of planning on IL development, on-line planning also seems to contribute to SLA, particularly in terms of accuracy enhancement. As ‘tasks’ are widely regarded as vehicles to provide learners with meaning-focused activities (2.2.1), it is one of the significant aims in TBLT to consider how learners’ focus on formal aspects can be elicited despite engaging in meaning-focused activities. For this purpose, the on-line planning proposal seems closely linked to the current trend of focus-on-form study, because of its conscious effort to engage in the formulation process and to help more accurate language use.

In line with cognitive processing concerns, Doughty (2001) proposes three

pedagogical recommendations to promote focus-on-form instruction, which can be assumed to reinforce a good connection between focus-on-form and on-line planning:

1. **The noticing issue:** Do learners have the cognitive resources to notice the gap between their IL utterances and TL utterances around them?
2. **The interruption issue:** Is a pedagogical intervention that does not interrupt the learner's own processing for language learning even possible?
3. **The timing issue:** If so, then precisely 'when', in cognitive terms, should the pedagogical intervention occur?

For the first issue, planning implementations are expected to give more frequent opportunities to notice the gap between IL and TL due to more expanded cognitive space than spontaneous speech by engaging in conceptual planning prior to the task or being free from on-line communication pressure. However, as learners are still under on-line processing pressure in strategic planning, it can be assumed that on-line planning is more likely to create cognitive space on-task leading to noticing. The second issue, interruption seems unproblematic in planning, because, once speech has commenced, every decision about the speech processes is, consciously or unconsciously, left to the speaker and every focus-on-form experience is initiated by him/herself. Although planning may cause conscious awareness of linguistic form to intrude in the speaker's mind, an occurrence of this interruption can also be regarded as natural processing, because the decision whether or not they spend some time to think about the language is entirely up to the speaker, not the teacher. This learner-initiation seems to guarantee the last issue, timing because identification of problematic linguistic features should be incidentally drawn from their own production or monitored during their production process.

## 2.6 Conclusion

Planning research tells us that L2 learners tend to use more varied and complex language in planned conditions and this enhanced performance might break through the stabilized state of IL, but the underlying mechanism for this rationale has been less explored. I have attempted to promote our understanding of planning and its effects on L2 performance. One of the problems raised in previous research is that strategic planning tends to lead to more fluent and complex language, but does not guarantee more accurate language. To explore this unsolved issue, this chapter particularly focused on the distinction between strategic and on-line planning and observed the mechanisms of their differential effects on task performance. More specifically, on-line planning seems to lead to increased consciousness-raising of form and, beyond a simple improvement of accuracy, meets the focus-on-form requirements: the noticing, interruption and timing issues.

The last decade has seen a growing body of research on various aspects of planning, but, as suggested so far, several issues remain unsettled. A clearly observed trend is an increasing interest in the process of planning. Started by a rather simple distinction of detailed and undetailed planning (Foster & Skehan, 1996), planning research has evolved into more specific descriptions of learners' strategic planning process through introspective/retrospective methods (e.g., Ortega, 1999, 2005; Sangarun, 2005). On the other hand, research inquiring into the on-line planning process is in the initial stage. The present study argues that more focus-on-form effects of on-line planning are conceivable, but the relatively smaller number of studies in on-line planning does not yet provide convincing support for this argument, compared to the number of strategic planning studies in various pedagogical contexts. Together with a growing demand for the process-orientated view, it is important to

embark on more on-line planning research using both qualitative as well as quantitative approaches, to examine the proposed effects of on-line planning.

# CHAPTER THREE: METHODOLOGY

## 3.1 Introduction

Drawing on the theoretical rationale for strategic and on-line planning, this study investigates how Japanese speakers of English in a study abroad context tackle story-telling tasks in different planning conditions. From a global perspective, the study aims to better understand the nature of the two different planning conditions: strategic planning and on-line planning. As only a few studies of on-line planning have been done so far, we do not have a clear idea of specifically what sorts of IL changes on-line planning actually brings about. Thus, it is imperative to grasp the gist of this particular implementation and to examine the applicability of the theory to pedagogic contexts. Another aspect of the study is concerned with more specific research points: i.e., an inquiry into L2 formulation processes, particularly the link between on-line planning and focus-on-form. In contrast to the exploratory nature of the first global view, the main purpose of this aspect is to test the specific question: “Is on-line planning conducive to focus-on-form opportunities?”

This chapter will describe the methodological issues of the study. Firstly, I propose the research objectives, and then more specific hypotheses and questions in the study. Secondly, I describe the issues concerning the research design involving the participants, tasks and planning instructions. Finally, I explain the measures of the task performance analysis and the method of the verbal report analysis.

## 3.2 Research Objectives

In order to probe into the issues which were considered in the previous chapters, two general research objectives were posed:

1. How do strategic and on-line planning influence L2 oral performance?
2. How do L2 speakers plan their speech *on-line* to deal with linguistic or conceptual problems when their performance is underway? Do they focus on form in on-line planning more frequently than in strategic and non-planning conditions?

Assuming that on-line planning induces the speaker into using his/her upper limits of IL (i.e., more complex and more accurate language), the study attempts to scrutinize learners' psycholinguistic processes. That is, the main interest is in knowing what is happening while engaging in on-line planning, and whether there is any difference from the speech conditioned by strategic and non-planning. By proposing the hypothesis that on-line planning contributes to more accurate formulation of L2, a particular focus is to examine the processes of L2 formulation from a psycholinguistic perspective, and how much on-line planning satisfies the focus-on-form principles in SLA perspective. In order to explore these invisible processes, the study includes the analysis of verbal reports obtained from the immediate post-task interviews.

## 3.3 Research Hypotheses and Questions

The two research objectives are explored with reference to the task performance analysis and the verbal report analysis respectively. The first objective addresses five research hypotheses concerning the task performance analysis:

1. Strategic planning (SP) will have a more positive influence on fluency, complexity and accuracy than non-planning (NP).
2. On-line planning (OP) will have a more positive influence on complexity and accuracy, but a more negative influence on fluency than NP.
3. OP will have a more positive influence on complexity and accuracy, but a more negative influence on fluency than SP.
4. The effects of SP will interact with the proficiency levels.
5. The effects of OP will interact with the proficiency levels.

The first three hypotheses concern the three planning conditions to test the effects of strategic and on-line planning on performance dimensions. As mentioned, a number of strategic planning studies have tested the first hypothesis and reached the conclusion of positive effects on fluency and complexity, but have not confirmed the effects on accuracy. For the second hypothesis concerning on-line planning, this has only been tested by Yuan and Ellis (2003) and Ellis and Yuan (2005), finding the positive effects of on-line planning over non-planning in complexity and accuracy at the expense of fluency. Thus, it is reasonable to replicate the research to examine the effects of on-line planning in order to understand these expected effects on performance. As for the third hypothesis, the result of Yuan and Ellis (2003) did not show the relative difference of the two planning conditions in complexity and accuracy, although on-line planning theoretically seems to be more concerned with form-focused effects.

The last two hypotheses concern the relationship between planning and proficiency. As only a limited number of studies have investigated the relationship between strategic planning and proficiency and no study has investigated that between on-line planning and proficiency, the hypotheses are established as two-tailed (i.e., to leave open the possibility of the relationship being either direction) rather than one-tailed.

The nature of the verbal report analysis is exploratory rather than



hypothesis-testing and this stage was conducted based on the results that emerged from the performance analysis. For this reason, the specific questions will be presented in Chapter Five. However, on the whole, the two central research questions of the verbal report analysis are indicated here:

1. To identify the psycholinguistic operations in the three planning conditions.
2. To explore different psycholinguistic operations among high and low proficiency speakers

In the task performance analysis, the points of inquiry are to examine the effects of (a) the three planning conditions and (b) two proficiency levels. In contrast to the quantitative analysis of task performance, the verbal report analysis adopts a qualitative perspective. Traditional task performance analysis using statistical models is a powerful instrument to clarify the effects of planning and proficiency levels, but this sort of black-white procedure seems to sometimes miss significant points, considering that language is a complex working of human beings influenced by various factors. It is probably true to suggest that the task performance analysis can only look at the linguistic *products* of planning, and by nature it cannot reveal what sort of cognitive and affective *processes* occur in speech production. In this context, the use of qualitative analysis is beneficial in compensating for the drawbacks of statistical analysis to identify important clues to support the effects of planning. For example, if on-line planning leads to more accurate production than the other conditions, there must be different cognitive operations functioning in order to increase the level of accuracy in on-line planning. Therefore, the second part of my research aims to elicit any evidence to support the different effects of planning. For this, verbal reporting about on-task performance in the form of a retrospective interview was combined with stimulated recall methodology. In this way, this study

adopts a ‘process-product approach’ (Ortega, 2005) to utilize the advantages from two different research traditions and to shed new light on planning research in SLA.

### **3.4 Participants**

To research the effects of planning and its connection to learner proficiency, twenty-seven Japanese speakers of English were recruited. All of them were staying in the UK, either to complete a master’s or doctoral degree, or to attend academic seminars as visiting students in universities. Their participation was voluntary, and we confirmed that all details of the experiment would be confidential (see Appendix 3-A for Participant’s Consent Form, and Appendix 3-B for Pre-Task Questionnaire). The participants were not told about the specific goals of the research during the process of data collection.

A practical reason for the selection of Japanese speakers is that they share the same L1 with the researcher who can thus elicit reports on their complex mental activities during L2 performance by means of their L1. In the following sections, I describe the participants in more detail: their learning background, length of residence and different proficiency levels.

#### *3.4.1 Learning Background*

All the participants can be defined as EFL (English as a Foreign Language) learners, because they learned English as a foreign language in the Japanese formal education system. All of them started learning English in junior high school (aged 12-13) and their use of English outside class was very limited. Although supplementary speaking lessons by NSs of English were regularly given (usually one hour per week), the main

part of English learning focused on gaining and storing knowledge about the language (i.e., vocabulary and grammar) through reading textbooks. The instruction can be characterized as the Grammar-Translation Method. No participants had previously lived in English speaking countries in their childhood and thus such a traditional style of English teaching had been, for many of them, the only opportunity to learn and use English until they entered universities.

### *3.4.2 Length of Residence*

In addition to the three different planning conditions, the present study focuses on different speaking proficiencies, to inquire how strategic planning and on-line planning influence speakers of varying proficiency. As only a small number of studies investigated the relation of proficiency and strategic planning effects (see 2.3.4) and the results are not conclusive, it is worthwhile to include the proficiency factor in the present study. In order to collect different proficiency levels of speech data, first I paid attention to the length of residence (LOR). Although LOR does not guarantee the level of L2 development (for example, learners may not have sufficient opportunities to interact in the TL [Lapkin, Hart & Swain, 1995]), it can be one of the factors to decide proficiency levels. As the present research aims to grasp the features and tendencies of English oral performance in Japanese speakers, the necessary condition for the selection of participants is to have experienced formal English education in Japan (not those who had formal education outside Japan). So, all the participants have similar characteristics, but the LOR in adulthood is wide ranging from one week to more than three years (mean = 11.86, SD = 13.74, range = 0.25 – 41.00 months). Thus, the participants can all be regarded as EFL learners with varying LOR.

To balance the numbers between higher and lower proficiency speakers,

participants were recruited according to one of the three categories of LOR: several weeks, several months, and several years:

- **LOR A** (n = 9): the shortest LOR; the participants who had recently arrived in the UK (up to two weeks) and had no previous experience of living in any English speaking countries for a long period.
- **LOR B** (n = 9): the intermediate LOR; the participants who had lived, uninterrupted, in the UK for several months in succession (up to six months).
- **LOR C** (n = 9): the longest LOR; the participants who had lived, uninterrupted, in the UK for more than one year.

The first group is the participants of the shortest period of residence; they had just arrived in the UK and had no previous experience of living in any English speaking country. They are supposed to represent the typical state of ordinary Japanese speakers. The second group consists of people who had spent several months (up to six months) in the UK, implying that they are expected to be more proficient in speaking than the first group through being exposed to larger amounts of natural interaction. The last group includes people who had substantial experience of living in the UK (more than one year in succession), and are expected to have the highest proficiency in oral skills among the three groups.

### *3.4.3 Making Different Proficiency Groups*

LOR is a useful criterion to collect participants with a wide range of speaking proficiency levels, but it should be stressed that LOR does not necessarily predict levels of proficiency (e.g., Cummins, 1991, 2000). That is, we could say that the collected data include both high and low proficiency speakers, but it is necessary to consider how we could classify them into different proficiency groups.

Like the present research, a number of studies including the proficiency factor have debated how to assess the proficiency levels in a logically and practically satisfying way. Examining published articles, Thomas (1994) identified four major conventions for the assessment of proficiency: (1) use of institutional status as a proxy for proficiency level, (2) use of research-internal or in-house measures of proficiency, (3) standardized test scores, and (4) impressionistic judgment. Firstly, 'institutional status' could not be applied because none of the participants were language school students with a particular proficiency level (e.g., 'intermediate' course). 'Research-internal' or 'in-house measures' were not also used because testing time could not be secured in addition to the main research; also, assessing the validity of tests of speaking proficiency and tools such as cloze tests is problematic. Moreover, I did not resort to 'standardized test scores' (e.g., TOEFL), which many other SLA researchers have adopted; because, considering that most of them took the test before coming to the UK, the scores seem not to reflect their present proficiency, particularly those with long LOR. Also, it may not be appropriate to judge speaking proficiency by referring to TOEFL, which does not include a speaking component (cf. Yuan & Ellis, 2003). For such practical reasons, the present study adopted the last, so-called 'impressionistic judgment'.

To make the judgment as objective as possible, three raters assessed the 27 non-planning talks, which had been recorded in a digital voice recorder in random order in terms of 'fluency' and 'intelligibility' on a six-point global rating scale. The first and the second raters are applied linguists who have substantial experience of ELT (English Language Teaching), language assessment, and applied linguistics research, and the third rater is the researcher. The first rater is a native English speaker and the second and the third raters are non-native English speakers.

The results of the global rating are given in Appendix 3-C. The results of the

three rating judgments, were collected and compared; then we came together to discuss inconsistent cases (more than two points differences among the three raters) and agreed the judgment of proficiency for each instance. In order to make two (high vs. low) groups, an arbitrary cut-off score had to be established. The medium (i.e., 14th score) was 4.00, but, due to the three scores of 4.00 in total, the next lowest score (3.67) was categorized as the lower group, making the high proficiency group (n = 14) and the low proficiency group (n = 13). It shows that all LOR A participants were categorized in the low proficiency group, all LOR C participants were in the high proficiency group, and the LOR C participants were divided between the high and low groups. To illustrate the levels of each group, two examples of speech from the high and low proficiency groups are given in Appendix 3-D.

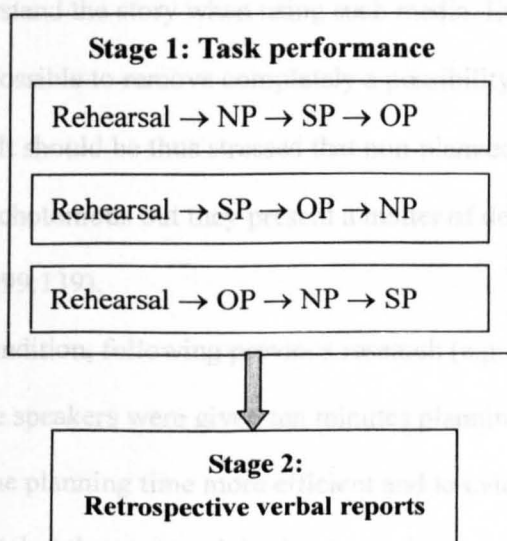
### **3.5 Research Design**

The present study investigates the performances of the same participants across the three planning conditions; also, it looks at the differences between the two proficiency levels. Therefore, I adopt both within and between-participants, or a mixed 3 x 2 research design (i.e., three planning conditions x two proficiency levels). In the following part, I describe the research stages, procedures and the task order of the main study. A pilot study (Nitta, 2004) was conducted along similar lines. It did lead to refinements in the design, instruments and confidence of the researcher. Specific influences on the main study are noted below (e.g., 3.7.1, 3.7.2). The summary of the pilot study is not included here because of a lack of space, but see Appendix 3-E.

### 3.5.1 Research Stages

In order to pursue the research objectives, a sequence of two research stages was designed: the task performance stage and retrospective verbal reporting stage (see Figure 3.1). Each participant was randomly assigned one of the three planning orders (i.e., (1) NP → SP → OP, (2) SP → OP → NP, or (3) OP → NP → SP; see 3.5.3 for the sequence of task planning). After the general instruction about the stages and rehearsal opportunity of a story-telling task, they completed the three tasks. In the second stage they were asked to report their psycholinguistic operations while being stimulated by listening to the audio recorded performance. Following the general procedure of retrospective interviews, the verbal report happened immediately after the task manipulation without any intervention (see 3.9 for more details of verbal report procedures).

**Figure 3.1:** Research Design



### *3.5.2 Research Procedures*

The process of the first stage took about 20-30 minutes, and the second stage took about 15-30 minutes; thus 35-60 minutes were spent to complete the whole process for each participant. The researcher sat facing the participant as a listener in order to create a natural situation. Before starting each task, the instructions were given in Japanese, and the speaker had the opportunity to ask about the meaning of the instructions, if they were deemed unclear.

In the NP condition, speakers were given thirty seconds for understanding and remembering the content from a picture story. So, they were not allowed to look at the picture while story-telling. In this situation, there may be a possibility of planning while looking at the picture; that is, the speaker might rehearse his/her speech during this period. Using a video story (as in Bygate, 1996; Wendel, 1997) enables the speaker to start telling a story immediately after watching the story and to decrease this kind of rehearsing possibility. However, it is unavoidable to give some sort of time to understand the story when using such media. Even in film-viewing procedures, it is impossible to remove completely a possibility of planning while watching the video. It should be thus stressed that non-planned and planned conditions are not dichotomous but they present a matter of degree and type of planning (Ortega, 1999:139).

In the SP condition, following previous research (e.g., Crookes, 1990; Foster & Skehan, 1996), the speakers were given ten minutes planning time. A blank sheet was given to make the planning time more efficient and to evidence participants' planning engagement, but they were advised not to write complete sentences either in L1 or L2. The sheet was removed during the speech production, but it was used to help them recall the performance in the interview. Following the method used in Yuan



and Ellis (2003), the speaker was required to finish his speech within a certain amount of time in both non-planning and strategic planning in order to create on-line processing pressure. The production time was set up for two minutes from the result of the pilot study (see 3.7.2).

In OP, the speakers were given thirty seconds (as in NP) to understand and remember the story and asked to start telling the story, but they were instructed to engage in planning their speech *on-line* as much as they wanted, when they found it difficult to describe their ideas or formulate the language. They were also advised to correct their utterances, when they found mistakes or were unsatisfied with their utterances. To eliminate on-line processing pressure, the time limitation was not set up in this condition.

### *3.5.3 Planning Order and a Rehearsal Task*

To minimize the possible practice effects, planning order was counterbalanced. In any order, the first performance is likely to be more difficult, because the speaker should not have any idea of what kind of task will be provided, and the subsequent task performance seems to be improved due to the task-type practice effect. Bygate (2001) shows that this effect was very limited, but there is still a possibility that it could influence performance. To combat this practice effect, I gave a rehearsal task with the same format as the following three tasks (see 3.6 for the tasks). In addition to this rehearsal opportunity for each LOR group of nine participants, there were three different task orders (i.e., NP – SP – OP, SP – OP – NP, OP – NP – SP) on three different stories, as illustrated in Table 3.1.

**Table 3.1:** Planning Order x Stories

		Stories		
		A	B	C
<b>Planning Order</b>	<b>1</b>	NP - Waiting for a bus	NP - The chase	NP - A surprise
		SP - A surprise	SP - Waiting for a bus	SP - The chase
		OP - The chase	OP - A surprise	OP - Waiting for a bus
	<b>2</b>	SP - A surprise	SP - Waiting for a bus	SP - The chase
		OP - The chase	OP - A surprise	OP - Waiting for a bus
		NP - Waiting for a bus	NP - The chase	NP - A surprise
	<b>3</b>	OP - The chase	OP - A surprise	OP - Waiting for a bus
		NP - Waiting for a bus	NP - The chase	NP - A surprise
		SP - A surprise	SP - Waiting for a bus	SP - The chase

### 3.6 Tasks

As briefly mentioned, I used three story-telling tasks, each of which was used for one of the planning conditions. Like many other task planning studies, I chose monologic rather than dialogic tasks. First I will clarify the reason for this particular type of task. In addition, as I used three different picture stories, it is necessary to standardize the difficulty of each story. Thus, I will also mention the difficulty of the three different stories.

#### 3.6.1 Monologic Tasks

To make the results more comparable with previous planning studies, I chose the most frequently used, *story-telling tasks* (Brown & Yule, 1983). This type of task requires the speaker to tell the storyline shown in pictures or films; ‘the stimuli given are purely visual and their verbal representations depend on the storyteller to a great extent’ (Albert & Kormos, 2004:286). Also, Ejzenberg (2000) suggests that a narrative

task is more efficient in assessing speakers' oral ability than a dialogue task. One of the main foci of the study being on exploring the nature of on-line planning, it is considered to be desirable to exclude complex interactional features caused by dialogic tasks.

As a number of other studies include interactional tasks, it is worth considering particular effects of monologic type of task on the performance outcome. Certainly, a decision of employing either monologic or interactive as a generic task type seems to affect learners' quality (and possibly quantity) of performance. A reason for the selection of the monologic task in the present study is practical rather than theoretical, as it is considered desirable to exclude complex interactional features, particularly for operationalizing the on-line planning implementation, because participants could be engaged in on-line planning while the interlocutor has the floor in the case of interactional tasks (Skehan & Foster, 2005).

One possible influence of the monologic type of task is an absence of an authentic listener. The effects of listener presence may have been potentially debilitating and facilitative consequences for learners' attention to the linguistic code (Ortega, 2005:101), but they seem to be largely dependent on individual learner types. Although the researcher participated as a listener in the present task session, the unique status of the researcher (clearly knowledgeable about the task materials already) may have a profound effect on task performance (Yule, 1997:76-7). The presence of the researcher may contribute to creating a formal situation (like a speaking test), because it was obvious for participants that the researcher knew the content of the story given. Iwashita, McNamara and Elder (2001:431) mention that a focus on accuracy may be paramount in the testing situation regardless of the conditions under which the task is performed, and this in turn may affect fluency and complexity. Thus, it should be borne in mind that the results obtained from monologic

tasks might not be equivalent of those garnered from dialogic tasks.

### *3.6.2 Task Difficulty*

To differentiate the effect of task repetition and to minimize practice effects, I prepared three different tasks (i.e., “Waiting for a bus,” “A surprise” and “The chase”; see Appendix 3-F for the three tasks and the rehearsal task). In the pilot study, the stories included a limited amount of text (e.g., a list of dishes on a menu). The influence of this seems very little but I cannot reject the possibility that, especially in the case of strategic planning it might multiply the amount of speech because the participants had sufficient time to be able to memorize the text. Thus with-minimal amount of text (e.g., a single work on a sign) cartoons were selected in the main research from an elementary picture book for EFL learners (Heaton, 1975).

One of the problems is the necessity to standardize the difficulty of the three tasks in order only to be able to examine the planning effects, not the task effects. Summarizing the previous studies (e.g., Nunan, 1989; Skehan, 1998; Robinson, 2001a), Ellis (2003:221-229) identifies input-related factors for grading task difficulty, including five sub-categories: medium (i.e., pictorial, written or oral), code complexity (i.e., high or low frequency vocabulary; simple or complex sentences), cognitive complexity (i.e., static, dynamic or abstract information; amount of information; degree of structure; context dependency) and familiarity of information. Applying these criteria to the three tasks, there were no clear differences identified, and most criteria could be judged as relatively ‘easy’ (see Table 3.2).

All the tasks include the same task type (i.e., story-telling) and format (i.e., 6 pictures). Lexical items needed for each story were listed in the resource book (Heaton, 1975), showing the high frequency vocabulary in all three tasks. Also, the

possible compositions to describe the stories were given, showing only short and simple sentences required (see Appendix 3-G for vocabulary and structures needed for descriptions of the three tasks). As for the information type, Ellis (2003:222-3) suggests three possibilities: ‘static’ (i.e., information remains the same throughout the performance), ‘dynamic’ (i.e., the information contains chaining events and activities), and ‘abstract’ (i.e., the information is used to form opinion). When no information is added *during* the task in any of the three tasks – as is often done during simulations which try to replicate real time sequences – they are identified as ‘static’. The obligatory use of the past tense was required in all the three tasks (see 3.7.1), so the context dependency was ‘there-and-then’ (Robinson, 1995). Also, as all the stories were not previously shown to participants, they did not have a prior knowledge about the contents (i.e., ‘unfamiliar’ information).

**Table 3.2:** Grading the Difficulty of the Three Tasks (Adapted from Ellis, 2003:228)

<b>Criterion</b>	<b>Three tasks</b>	<b>Grading</b>
<b>1. Medium</b>	pictorial	easy
<b>2. Code complexity</b>	high frequency vocabulary; short & simple sentences	easy
<b>3. Cognitive complexity</b>		
<b>a information type</b>	static	easy
<b>b amount of information</b>	6 pictures	?
<b>c degree of structure</b>	well-defined structure	easy
<b>d. context dependency</b>	there-and-then	difficult
<b>4. Familiarity of information</b>	unfamiliar	difficult

Focusing on the degree of structure, Tavakoli and Skehan (2005) show its effects

on L2 performance. Following previous studies (e.g., Foster & Skehan, 1996; Kobayashi, 2002a; Skehan & Foster, 1997), they identify several features contributing to macrostructure which impact on performance:

- a clear time line
- a script
- a story with a conventional beginning, middle and end
- an appeal to what is familiar and organized in the speaker’s mind
- a problem-solution

Looking at the three story-telling tasks (Appendix 3F), all of them can be regarded as having ‘well-defined structure’ according to these features, involving a complication – resolution pattern of narratives. Each task includes the following components.

**Table 3.3:** Complication – Resolution in the Three Tasks

	<b>Task 1 (Waiting for a bus)</b>	<b>Task 2 (A surprise)</b>	<b>Task 3 (The chase)</b>
<b>Complication</b>	Three boys could not get on the bus because senior boys jumped the queue.	Two thieves stole a big basket from an Indian man.	A boy didn’t notice that he dropped a small parcel, and was scared that a strange man was chasing him.
<b>Resolution</b>	The next bus they got on finally passed the first bus which broke down on the way.	When they opened the basket, they were surprised at seeing a big snake in it.	The boy realized that the man was following him to return the dropped parcel.

As Table 3.3 suggests, each task includes a clear logical pattern of beginning with a complication or a problem and ending with a resolution or solution. Such

well-structuredness is expected not to place much emphasis on understanding the concept and, as a result, to promote more focus on form. In the post-task interviews, every participant confirmed that they did not find it difficult to interpret the storyline in the three tasks. It is necessary to add here that several participants reported relative difficulty in translating their concepts into L2 for particular tasks, but this judgement is not consistent. Therefore, the three story-telling tasks were considered of similar difficulty level without complex content structures or the need for specific background knowledge required.

### **3.7 Instructions**

Having looked at the research design and the tasks, this section addresses the planning instructions. Firstly, I mention the provision of obligatory past tense form. Then, I describe the production time limitation set for strategic planning and non-planning. Finally, I explain the on-line planning specific instructions.

#### *3.7.1 Obligatory Occasion of the Past Tense Form*

In the pilot study (see Appendix 3-E for the summary), the use of the past tense form seemed to devour considerable attention, and consequently place a greater burden on the L2 speaker's processing than using the present tense. In addition, there are several studies to suggest that Japanese story-telling by Japanese NSs tends to engage in more frequent tense shifting than that by English NSs (Nakajima, 2005), and this might influence Japanese speakers' inconsistent tense usage in English. For either cognitive or linguistic reasons, it is probably true to say that low proficiency speakers could not deal with the verb formation problem efficiently. Because correct use of the past tense

can be one of the indicators to judge the levels of proficiency, this study provides obligatory occasions of the past tense form (i.e., ‘supplied in obligatory context’; Brown, 1973; Ellis & Barkhuizen, 2005; Larsen-Freeman & Long, 1991). Tense usage was inconsistent in the pilot study, and so in the main study it was decided to further encourage the participants to construct the main storyline in the past by providing the first sentence with the past tense embedded at the beginning of each condition (e.g., “one day three boys were waiting for a bus on a street”). To make them pay attention to the past tense, the researcher read out the sentence and asked them to write it down on a given sheet and to use it to start their story. In doing so, the participants were expected to be more or less conscious of the past tense usage, although the level of consciousness and the judgement of how much they maintained correct past tense forms were left to each one’s decision. As all the participants correctly wrote down the sentences I read out, it can be assumed that they were all conscious (perhaps at different levels) of the use of the past tense at the beginning of the story-telling.

### *3.7.2 Limitations of Production Time*

The time limitations for non-planning and strategic planning are also important, because this differentiates these from the on-line planning condition by increasing on-line processing pressure. The three minutes time limitation in the pilot study placed pressures on some participants but very little on other participants, probably because their preferred styles of speaking (fluency or accuracy-oriented) were different. It is not very difficult to request participants to follow this time limitation strictly, but this may go against the rules of naturalistic situations. To what extent they need to strictly follow the given instruction is open to debate, but, considering the focus of the present study is on-line planning, it is important to operationalize the



on-line planning instruction as clearly distinct from the other two conditions. Because most of the stories told in the pilot study were completed within two minutes and the difficulty of the cartoons used for the main research were similar to those in the pilot study, I decided to set two minutes time limitations for non-planning and strategic planning in the main research.

### *3.7.3 On-Line Planning Instruction*

Drawing on Levelt's speech production model, on-line planning is an operation where speakers pay special attention to the three monitoring phrases (i.e., preverbal message monitoring, internal speech monitoring, and external speech monitoring; see 1.5.4). A speaker pays particular attention to the first two, pre-production and post-production monitoring, which is very difficult to conduct in usual speech, as suggested by Krashen, due to processing constraints. In this sense, some unexpected results in the pilot research seem to come from the failure to operationalize the on-line planning rationale (see Appendix 3-E). To put the point more concretely, the on-line planning instruction, which was replicated from Yuan and Ellis (2003), does not include any clear suggestion of a particular focus on 'pre-production' monitoring, and as a result, it became difficult to differentiate this from the non-planning condition. Taking a lead from Yuan and Ellis (2003), the instruction (for the pilot study) was:

You can take as much time as you want when telling the story. If you think you say something not correct or not to your satisfaction, you can correct it as many times as you want. [the original instruction was given in Japanese.]

The problem with this instruction is that it does not seem to fully reflect the on-line planning definition given by Yuan and Ellis themselves. According to their definition,

it has two significant functions: (1) pre-production monitoring and (2) post-production monitoring. However, the task instruction seems only to promote post-production monitoring (i.e., “If you think you say ...”), whereas pre-production monitoring may have given direct impact on accuracy. The first sentence (i.e., “You can take...”) implies that the on-line planners are free from the time pressure in contrast to the other pressured conditions. As a result, two of the three participants in the pilot were careful about their language, but, due to still existing communication pressure, hesitated to engage in pre-production monitoring (i.e., taking sufficient amount of planning online) for accurate language performance. In other words, speakers may not necessarily have engaged in on-line planning following the original instruction.

In order to encourage speakers to engage more fully in on-line planning (i.e., not only for raising consciousness about form but also attaining accurate language), it was necessary to revise the on-line planning instruction as follows:

When telling the story, (1) if you find it difficult to say what you want, you can stop and think by taking as much time as you wish. (2) If you think you said something that was not correct or was not to your satisfaction, you can correct it as many times as you want. In these cases, don't worry about using less fluent English than you usually do. What is more important this time is correctness. Also, you have no time limitation to finish telling the story. [the original instruction was given in Japanese.]

As shown, it clearly states both (1) pre-production and (2) post-production monitoring as the significant on-line planning conditioning. It can be expected that, according to their own proficiency levels, each participant will employ their optimal level of on-line planning on getting this instruction. That is, it is expected that the less proficient speakers will tend to engage in (unnaturally) longer pauses, and the more proficient speakers will tend to engage in relatively slower but more natural length of

pauses, approaching to NS's level (see Appendix 3-H for the instructions of the three planning conditions).

To investigate the effects of planning on L2 oral performance, this chapter has described the issues relevant to the research design features involving the participants, tasks and instructions. Having explained the structure of this research, next I will explain the two research instruments of analysis in detail.

### **3.8 Task Performance Analysis**

The following sections consider the methodological issues in terms of two different approaches: the analysis of task performance (3.8) and verbal report (3.9). The two methods have different theoretical and research traditions, but their functions seem complementary; the former aims to grasp the linguistic and performance features of L2 speech production, and the latter aims to explore the underlying cognitive processes of these features. As a tool to test the effects of different planning conditions, first I consider the task performance analysis by reviewing various IL variables in the task-based and SLA literature, mainly focusing on the fluency, complexity and accuracy distinction (Skehan, 1996a, 1998). In contrast to this hypothesis-testing approach, the next section looks at the verbal report analysis combined with stimulated recall, which aims to explore learners' cognitive operations in different planning conditions.

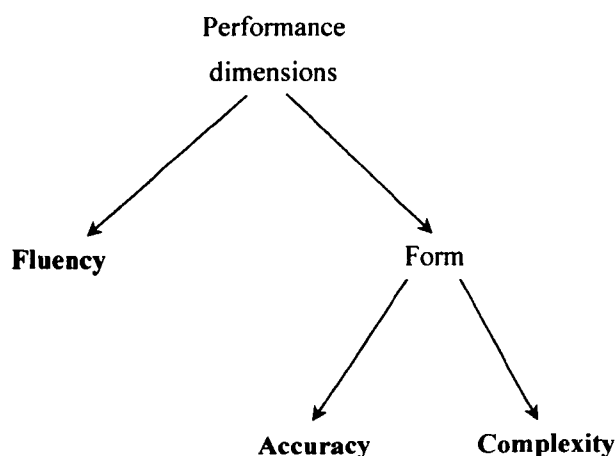
#### *3.8.1 Three Goals of Task-Based Research: Interaction of fluency, complexity and accuracy*

Taking account of processing theories, Skehan (e.g., 1996a, 1996b, 1998) proposes three distinctive areas of *fluency*, *complexity* and *accuracy*, and insists that 'achieving

an effective compromise between these three goals is more likely to lead not only to the capacity to be an effective communicative problem solver but also to longer-term linguistic development' (Skehan, 1996b:22). In a theoretical sense, the sequence of the three dimensions implies the three stages of change in the underlying system (i.e., complexity), in acquisition of greater control over the emerging system (i.e., accuracy) and in development of performance control, as elements are routinized and lexicalized (i.e., fluency; Skehan, 2003:8).

Figure 3.2 suggests that this three-way distinction is largely concerned with the tension between focus on meaning (i.e., fluency) and form, which was chiefly mentioned in 2.4.

**Figure 3.2:** Theorizing Dimensions of Performance (Skehan & Foster, 2001:190)



In task completion, focusing on fluency, for example, tends to make the learners take a 'getting the task done' strategy without paying sufficient attention to form, while focusing on form encourages them to commit conscious engagement with accurate and complex language use (Skehan & Foster, 2001). Assuming that a limited capacity of human cognition precludes simultaneous attention to both form and meaning aspects (VanPatten, 1990), the primary concern should be given to how we can get the

right balance between these conflicting areas, rather than exclusive focus on one at the expense of the others. In the following part, I explain the working definitions of each component and various IL measures used for the analysis of task performance.

### 3.8.2 Defining Fluency

The term, 'fluency', is difficult to define, because there is no unified understanding of the term and researchers have attempted to examine 'fluency variables' in their own definitions. Integrating manifold meanings of fluency, Lennon (2000:26) has generated a working definition of fluency as '*the rapid, smooth, accurate, lucid, and efficient translation of thought or communicative intention into language under the temporal constraints of on-line processing*'. As this definition suggests, understanding one's fluency concerns a very wide area, ranging from surface linguistic features to a pragmatic dimension, and thus only looking at one aspect may fail to grasp a real state of fluency.

Another significant point implied by Lennon's definition is that the fluency factor is largely related to on-line processing capacity. In other words, how fluent a speaker is is highly dependent on how much a speaker allocates his/her own on-line processing space to the fluency function. For example, a speaker would be able to correct grammatical errors if allowed to hear a recorded tape; given "enough time" (e.g., another couple of seconds or several days), the speaker would encounter the appropriate word which he/she could not find in spontaneous speech (Lennon, 2000: 27). This matter is not unrelated to the point that the primacy placed on fluency in speaking is in contrast to writing performance, which puts more emphasis on accuracy (Lennon, 1990: 391). Also, because fluency largely concerns temporal features in one sense, fluent speech literally gives pressures on one's on-line processing. It is

reasonable to consider that the more fluent a speaker is, the less accurate his/her language becomes, because in attempting to produce more fluent language, their attention seems to be more diverted (Brumfit, 1984). In a simple manner, such fluency-accuracy opposition can be a surrogate of the typical conflicts between CLT and traditional grammar-based teaching, but our current understanding of fluency, informed by processing theories, is more than such a simple sum. As a plausible strategy to deal with more communicatively pressured situations, speakers primarily rely more on lexicalized items and/or automatized processing ability, and when their resource is in deficit, access the rule-based system by strategically creating time through use of hesitation markers or lexicalized sequences and so on. It is thus assumed that to use fluent language more directly concerns the implicit knowledge system rather than the rule-based explicit system in one sense (Skehan, 1996a: 48), but, in another sense, occurrences of dysfluent features seem to underscore the speaker's attempt to solve the ongoing problem through using the explicit knowledge system. In capturing the whole picture of fluency dimension, it is not sufficient to only look at a particular feature, but it is essential to carefully examine a wide range of performance features with both quantitative and qualitative approaches.

There have been a number of attempts to explore particular quantifiable variables and qualitative linguistic features which particularly contribute to the distinction of fluent and nonfluent speech (e.g., Derwing, Rossiter, Munro & Thomson, 2004; Ejzenberg, 2000; Freed, 2000; Kormos & Dénes, 2004; Lennon, 1990; Riggenbach, 1991; Towell, Hawkins & Bazergui, 1996). Thus, findings from these fluency studies seem very informative for the task-based research tradition, which have tended to rely on a limited range of fluency variables (although recent years have seen several attempts to examine various variables [e.g., Skehan & Foster, 2005; Tavakoli & Skehan, 2005]). Despite numerous candidates, the fluency measure

can be largely categorized into two types: temporal phenomena and dysfluent markers. For the first of these temporal phenomena, I will apply two specific measures: the speech rate and the mean length of run.

### 3.8.2.1 FLUENCY MEASURES (1): TEMPORAL PHENOMENA

**Speech rate.** Many fluency studies employ the variable of speech rate (SR), and many of them concluded that SR is a reliable measure of L2 oral fluency. SR is usually calculated by dividing syllables (or words) by the number of seconds used to complete the task and multiplying by 60. In this variable, ‘unpruned speech’ includes all the words produced by the speaker, while ‘pruned speech’ excludes redundancies, such as repeated, reformulated and self-corrected words. There is a danger that, in the former measure, the level of fluency might rise by using less meaningful words such as retelling the message and fillers, thereby not reflecting the real fluency. In order to avoid such pseudo-fluency, I use ‘pruned’ SR in the present analysis.

**Mean length of run.** Another popular temporal variable frequently used as a reliable fluency measure is the mean length of run (MLR). This is a simple index of grammatical development because almost every new kind of knowledge increases length (Brown, 1973:53). The study by Towell *et al.* (1996) shows that an increase in the length and complexity of the linguistic units which are uttered between pauses, rather than the increase of the speed and the amount of pauses, has the strongest impact on the qualitative evaluation of fluency by trained raters. This is calculated as an average number of syllables produced in utterances between pauses and dysfluent markers (see 3.8.4). This measurement seems more useful for investigating monologues rather than dialogues, as speech running is often interrupted by the

interlocutor in the latter even if the speaker is able and willing to continue his/her utterance.

One important discussion concerning MLR has been on where to draw the cut-off point of pauses, which ranges from 0.2 seconds (Lennon, 1990) to 0.4 seconds (Derwing *et al.*, 2004). Towell *et al.* (1996:91) explain that the lower cut-off point may lead the analyst to be confused by displays in which an apparent pause is the stop phase of geminated plosives or other normal phenomena, while the higher point may lead to ignorance of significant amounts of pause time. With regards to this, Riggensbach (1991:426) gives an indication of the cut-off point: 0.2 seconds or less for micropauses; 0.3 to 0.4 seconds for hesitations; and 0.5 seconds to 3 seconds for unfilled pauses. Following this classification, pauses of 0.3 to 0.4 seconds may suggest learners' lack of resource in the middle of speech. Thus, I adopt 0.3 seconds as a cut-off point of MLR in my analysis.

### 3.8.2.2 FLUENCY MEASURES (2): DYSFLUENT MARKERS

Wendel (1997:70) argues that whereas SR captures the overall workings of performance, pause measures are concerned with the nonproceduralized or intentional aspects of the speech production system. Because variations in SR between speakers might reflect differences in pausing time rather than speed of articulation (Lennon, 1990), it is true to suggest that the temporal dimension is closely tied to dysfluent markers. Following this principle, Foster and Skehan (1996:305) adopted pausing and dysfluent measures as appropriate fluency indices. Also 'pausing' in L2 is regarded as one of the unique features (Bygate, 1998) different from L1. Thus this part touches upon another category widely used in the research on fluency, 'dysfluent markers'. This can be subdivided into 'pauses' and 'self-repairs'.



**Pauses.** Drawing on Griffiths (1991) and Beattie (1980), Bygate (1998:25) claims that ‘the distribution of pauses can be taken as a clue to some of the underlying processes involved in speech production’. In addition, Foster (2001) argues that the greatest time-gaining strategy for L2 speakers is pausing (see 1.5.6). Thus, analyzing pauses is important to understand L2 oral performance. In order to make the different length of speech comparable, the number of pauses is usually calculated by adding all pausing time and dividing the total amount of time to complete the assigned task.

According to Aitchison (1989:238), there are two main types of pause: breathing pauses and hesitation pauses. Whereas the former tends to come at grammatical boundaries and accounts for only 5 percent of the gap in speech, the latter does not have any obvious physical purpose and accounts for one-third to one-half of the time taken up in talking (ibid.: 238-9). As in linguistic and psycholinguistic studies, the primary concern of the present study is in the latter type of pause, or speech errors (in addition, a pause as intending rhetorical effects such as allowing a point to sink in or to stimulate laughter is possible but not included here). Van Gelderen (1994:303) calls this category ‘non-functional’ or planning pauses as these features are probably symptoms of problems a speaker encounters in generating and organizing content and in lexical access. The pause-related measures are significant in the present study, because it is conceivable that an indicator of ‘unnatural’ pauses, characterized by a relatively longer length and unusual positioning, may be indicative of an occurrence of on-line planning in speech processing. The parts of speech followed by such unnatural pauses can be regarded as unproceduralized elements of IL, and the speaker may engage in on-line planning.

For more precise analysis, the pausing factor needs to be categorized into two types: i.e., unfilled and filled pauses (or often called *fillers*). The former concerns a

silent pause, while the latter can be either nonlexical words such as ‘er’, ‘erm’, and ‘mm’ (Lennon, 1990) or lexical words such as ‘well’, ‘kind of’ and ‘you know’. In addition to these filled pauses, Riggenbach (1991:426) includes sound stretches – vowel elongation of 0.3 seconds or greater to this category. Lennon (1990:407) reports that the length of almost all filled pauses was 0.2-0.3 seconds (as his study categorized only nonlexical forms as filled pauses). It is conceivable that both filled and unfilled pauses function in a similar way (Lennon, 1990), but high proficiency speakers tend to use filled pauses as one of the useful speaking strategies to create more time to formulate what they plan to say next (Bygate, 1987); conversely low proficiency speakers tend to engage in more unnatural unfilled pauses. Thus the frequency of unfilled and filled pauses might become an important discriminator between proficient and less proficient speech. Presumably, fluent NNSs tend to use more lexical fillers, while less fluent NNSs rely more on nonlexical fillers and unfilled pauses (Riggenbach, 1991:431).

However, as NSs also engage in both filled and unfilled pauses, a simple sum of pausing occurrences may not indicate the level of fluency, against the received perception. Kormos and Dénes (2004:149) contend that the studies with a small number of participants (e.g., Lennon, 1990; Riggenbach, 1991; Freed, 2000) found a significant difference of the frequency of pauses between fluent and non-fluent speakers, while the studies with a higher number of participants did not find the correlation between them (e.g., van Gelderen, 1994), probably because the results in a small scale research tend to be more influenced even by a few idiosyncratic participants taking extreme amounts of pauses. It should thus be borne in mind that pausing composes a part of the fluency construct, but the relationship between them is not so straightforward.

Another important distinction of pausing is whether a pause can be

recognized as 'natural' or 'unnatural'. This recognition might originally come from rather intuitive and subjective impression, but an objective judging criterion is also necessary. A growing number of studies (e.g., Hawkins, 1971; Dechert, 1980; Deschamps, 1980; Lennon 1990; Riggenbach, 1991) suggest that NSs and highly fluent L2 speakers tend to pause at clause junctures or between non-integral components of clauses and clauses themselves, and pausing at other points gives the impression of dysfluency (Woods, 2001). Chamber (1997:538-9) demonstrates that a natural pause usually occurs at some clause junctures or after groups of words forming a semantic unit, allowing breathing space; these pauses may be either simply a silent gap or be marked by filled pauses, sound stretches (or drawls on words) or lexical fillers with no semantic information (such as 'you know' and 'I mean'); on the other hand, unnatural pauses are regarded as silence, which may be perceived as 'signaling poor functioning of mental processes.'

Similarly, Freed (2000:248) identifies 'dysfluent-sounding' as (1) silences occurred at places other than the predictable juncture boundaries (2) and tended to be of a certain time (0.4 seconds or longer in her study) in duration, arguing that 'dysfluent-sounding' pauses reflect 'a subject's struggle to find appropriate means of expression at either the lexical or the syntactic level' (p. 256).

More recently, Skehan (2004) observes that pauses at AS-unit boundaries (see 3.8.6) seem to represent 'normal' fluency, reflecting normal operations within WM constraints, i.e., mini-plans reasonably successfully achieved, while pauses in the midst of an utterance indicate various kinds of searches (e.g., an idea, word or grammatical form), which occurs far more frequently where learners are concerned, depending on their overall linguistic proficiency. In the case of NSs, monitoring morphological accuracy is unlikely to cause many hesitations (Chamber, 1997:538).

Synthesizing these arguments, the speakers tend to take appropriate length of

pausing at a clause (or AS-unit) boundaries as long as they could smoothly translate the thought into language; conversely, an unnatural pause deviating from these rules might illustrate a speaker's struggle with speech processing, occurring somewhere between the Conceptualizer, the Formulator and the Articulator (Levelt, 1989).

Fulcher (1996: 216-7) argues that such surface phenomena are easily coded but what is more important is to provide an explanation of the phenomena in terms of language usage. The occurrences of self-repairing subsequently after pauses might give us clues to identification of the explanations for pauses and to understanding processing hardships they encounter in speaking. Thus, the present analysis includes the number of end-clause pauses, mid-clause pauses, filled pauses and the total length of pauses.

**Self-repairs.** Not only temporal and pausing aspects, but also frequent occurrences of other dysfluent features, 'self-repairs', can be regarded as prominent characteristics of what Van Hest (1996:1) defines as those made by speakers on their own initiative, without intervention from the interlocutor(s). In this discussion, the category of self-repairs comprises the following elements; 'false-starts', 'repetitions', 'reformulations' and 'self-corrections'.

Firstly, a 'false-start' is regarded as an utterance which is begun and then either abandoned altogether, reformulated in some way (Foster *et al.*, 2000:368), or repaired in the next tone group (van Gelderen, 1994:301). An occurrence of false-start and further reformulation seems to provide evidence for the process in which the speaker notices a mismatch between what he/she intends to say and what is actually said and attempts to replace it through accessing the rule-based system. Riggenbach (1991: 427) observes that how much of the original utterance was rejected (only one word or more) is important, because more proficient speakers may be able to restart more quickly. while less proficient speakers can restart with only a small part of the

original utterance.

Secondly, 'repetition' means that the speaker repeats his/her previous production without any changes (Foster *et al.* 2000:368), which is necessary in order to distinguish between dysfluent repetitions and rhetorical repetition. Self-repetition seems to concern multifunctional nature, because, similar to filled pauses, it could reflect a way to buy time that actually gives an impression of fluency, or it could be perceived as a marker of dysfluent speech (Derwing *et al.*, 2004:672).

Thirdly, 'reformulation' is defined as either phrases or clauses that are repeated with some modification to syntax, morphology, or word order (Foster & Skehan, 1996:310). The speaker is considered to engage in more effort to reproduce the intended message than a simple repetition, possibly through the access to the rule-based system.

Finally, 'self-correction' is the speaker's identification of an error and formulation of a correct form immediately after the error, which may also underscore the speaker's access to the rule-based system.

It can be assumed that, as the speaker becomes more proficient, the number of these dysfluencies will decrease, but some studies (e.g., van Gelderen, 1994; Fulcher, 1996; Kormos & Dénes, 2004) argue that hesitation and dysfluent variables do not simply affect perceptions of fluency by listeners, because NSs also frequently engage in these features. Rather, it is probably more accurate to say that there is some critical zone below which all hesitation markers are accepted as perfectly normal (van Gelderen, 1994:314). In Lennon (1990), three of the four learners of English had in fact increased the number of self-corrections after six-months in Britain, suggesting that 'part of fluency development in the advanced learner may involve increased ability to reformulate, monitor, and self-correct production on-line' (p. 413). That is, cleared cognitive space caused by an increased automatization (together with a growth

of grammatical competence and the number of lexicalized items) might make the speakers more sensitive to the form aspect of their production. At the same time, such features might be evidence for the speakers' on-line planning involvement, because such dysfluent features make the listener aware of the production process under strain (Lennon, 1990:391).

Assuming that speakers have access to the rule-based system, self-repairs reveal their effort to process the language, i.e., translating thoughts into language formulation. Skehan (2004) argues that these features, particularly filled pauses, reformulation and false-starts, seem to reflect 'on-line' planning or engagement, as problems are perceived, but the speaker addresses these problems *resourcefully*. These can be seen as an on-line planning index (ibid.), showing speakers' conscious involvement in the process of perceiving the lack of lexical resources and attempting to solve the problem through language.

I have observed various fluency measures in terms of two macro-categories: temporal phenomena and dysfluent markers. As any single measure cannot detect learners' level of fluency, the present study applies various measures to examine it.

Understanding fluency is very important to capture one's performance but this is only one aspect of performance dimensions. Thus, I will look at the second dimension of speech performance, complexity, in the next part.

### *3.8.3 Defining Complexity*

The second goal, 'complexity', concerns 'the elaboration or ambition of the language which is produced' (Skehan, 1996b:22). Skehan (1996a:47) identifies complexity as a counter-product to accurate language use because learners who do not want to take

risks (i.e., complex language use) promote a conservative strategy, or more accurate performance.

In line with the planning perspective, one significant aspect is that this distinction seems to elucidate the proficiency levels of learners. For example, cognitively less demanding tasks may not clarify the different proficiency levels in terms of accuracy, if both advanced learners and beginners could pay sufficient attention to correct formation by taking ‘the safety first strategy’ to avoid attention-demanding structures in favour of already automatized language (Skehan & Foster, 2001:189). Therefore, the accuracy measurement does not always guarantee the proficiency levels. On the other hand, the competition model of accuracy and complexity implies that more advanced speakers adroitly attain well-balanced performance, because automatized morphosyntactic items and lexicalized complex structures seem to free up more cognitive space for other processing.

Looking at my own L2 learning experience, the accuracy-complexity characterization can be summarized as the differences between everyday conversation and academic discussion, which is reminiscent of Cummins’ (e.g., 1980) distinction of ‘basic interpersonal communicative skills’ (BICS) and ‘cognitive/academic language proficiency’ (CALP). Because what is needed in everyday life (e.g., buying a bus ticket, and ordering food in a fast food restaurant) is simple, it does not require intricate language, and thus we are, relatively soon, able to be communicative with accurate language use in such contexts. On the other hand, more formal and cognitively demanding contexts such as discussion in a seminar, due to its more complicated contents, usually requires more complex structures to express our intentions as fully as possible. In this example, the context (i.e., academic seminar) is likely to force speakers to realize the incompleteness of their IL (i.e., noticing a hole) to describe their opinions and take risks to try out their working hypotheses and

receive feedback from interlocutors verbally or non-verbally. It is thus very important to distinguish complexity from accuracy, because this can be one of the main reasons for stabilized L2 performance, which is accurate and fluent but still not target-like.

Compared to the fluency measures, there have not been many complexity measures used in the previous research. Among limited number of candidates, I apply 'syntactic complexity' and 'use of discourse organization devices' as complexity measures in the present study.

### 3.8.3.1 COMPLEXITY MEASURES (1): SYNTACTIC COMPLEXITY

More proficient speakers are expected to produce syntactically more complex language. One of the most frequently applied principles used to measure this feature is whether language consists of one or more dependent clauses attached to an independent clause. A key to measuring 'syntactic complexity' is concerned with a way of dividing transcribed data into units in order to identify a ratio of 'clauses' to particular 'supra-clausal segmentation'. Although most studies adopt clauses as the former unit, identification of the latter supra-clausal unit is largely inconsistent. As far as the task planning research tradition is concerned, the types of segmentation of speech unit can be generally classified into three types: 'utterance' (defined as a stream of speech produced under a single intonation contour by pauses [Sato, 1988: 375], and constituting a single semantic unit [Crookes & Rulon, 1985, cited in Crookes, 1990:187]), 't-unit' (defined as a main clause plus any other clauses which are dependent upon it [Hunt, 1965]) and 'c-unit' (including elliptical nature of spoken language in addition to the t-unit features [Loban, 1966]). However, surveying recent papers involving speech and written data analysis, Foster *et al.* (2000:357) argue that there has been 'a plethora of definitions of units of analysis,' and even these



frequently applied units do not provide clear definitions and sufficient applications. Among various types of speech unit, the present study adopts the AS-unit (analysis of speech unit) proposed by Foster *et al.* (2000) as a unit for analysis. For the analysis of syntactic complexity, I explain the working definition of ‘clause’ and ‘AS-unit’ applied in this research by referring to Foster *et al.* (2000) and other relevant studies.

**Clauses.** For the analysis of language, identification of clauses is linguistically significant, because ‘[t]he clause is the central processing unit in the lexico-grammar – in the specific sense that it is in the clause that meanings of different kinds are mapped into an integrated grammatical structure’ (Halliday & Matthiessen, 2004:10). With regard to the definition of clause (some use the term ‘s-node’ for the same meaning), however, there is some discrepancy observed in terms of non-finite clauses. For example, Wigglesworth (1997) clearly defines a clause as a unit containing a finite verb (thus excluding non-finite verbs), but Foster and Skehan (1996) and Ellis, Tanaka and Yamazaki (1994) include non-finite verbs as clauses. Standing on both views, counting the clauses in Example 3.1 is unproblematic by regarding the case as two clauses (i.e., an independent clause + a dependent clause; all examples are taken from my own data collected in the pilot study; errors are not corrected).

**Example 3.1:**

it seemed that \*// it ... this kind of situation often happen to the visitors  
\*// *represents a clause boundary*

However, it is critical to rely on one of these definitions in Example 3.2, which includes two possibilities of identifying clauses in terms of finiteness:

**Example 3.2:**

finally they couldn't have any meal and (.5) no choice // but to (1.1) walk around to such for some food restaurant for food something like that

Looking at these two examples, many of us would intuitively judge that the second example has a more complex structure than the first example. However, by excluding non-finite clauses for the clause category, Example 3.1 (2 clauses) would be more complex than Example 3.2 (1 clause), which suggest this exclusion of non-finite clauses seems not to reflect the psycholinguistic reality of speech production; regardless of finiteness, some sort of processing load seems to be imposed on the speaker's brain in engaging in verb formulation. Following the simple principle, 'the more clauses, the more complex and productive a language is', we should be very careful to identify clauses. Thus it would be more appropriate that the present study includes both finite and non-finite as a clause. To establish the criteria for analysis more clearly, I follow this clause definition by Leech and Svartvik (1994:246-251);

- (1) A clause can be analyzed into five different types of clause elements: **Subject (S)**, **Verb (V)**, **Object (O)**, **Complement (C)** and **Adverbial (A)**;
- (2) **Finite clauses** are clauses whose V element is a finite V phrase;  
e.g., She evidently *works* terribly hard.
- (3) **Non-finite clauses** are clauses whose V element is a non-finite V phrase;  
e.g., I used to lie awake at night, *worrying about the next election*.  
*Covered with confusion*, she hurriedly left the room.  
The best thing would be *for us to leave straight away*.  
All I did was *hit him on the head*.
- (4) **Verbless clauses** contain no V and often no S, but function like finite and non-finite clauses, and they can be analyzed in terms of one or more clause elements. Usually, a form of the verb has been omitted;  
e.g., Dozens of tourists were stranded, *many of them children*.

**AS-unit.** In addition to clause identification, the measurement of complexity in L2 performance requires a *supra-clausal* unit. Determining the appropriate analysis unit, one of the significant criteria is the psychological relation to the selected speech segmentation. Crookes (1990:191) argues that structural investigations of L2 are generally concerned with the results of the psycholinguistic processes of language production and analyses of such processes, and their results such as degree of complexity of speech would be based on a unit reflecting such processes. This is the underlying motivation of the present study, based on the current understanding of psycholinguistics, particularly information-processing theories, which is also in accordance with the orientation of much planning research (e.g., Foster & Skehan, 1996; Yuan & Ellis, 2003).

In the present study, the 'AS-unit' (Foster *et al.*, 2000) was chosen as the appropriate unit of analysis because, examining the advantages and disadvantages of other units, the unit is clearly defined for purposes of analysing a wide range of speech data. The AS-unit has a mainly *syntactic* basis, because syntactic units are genuine units of speech planning; also, it is greater than a single clause, because 'the ability to plan at the multi-clause level is important for establishing a speaker's level of proficiency' (Foster *et al.* 2000:365). Foster *et al.* (*ibid.*) define it as follows:

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.

This definition sounds very similar to the other units such as t-unit, but, importantly, the AS-unit includes independent sub-clausal units (*ibid.*:355-6). To make it clearer, the definition is further elaborated following Foster *et al.* (2001): (1) an independent clause minimally consists of a clause including a finite verb (e.g., 'That's right'); (2)

an independent sub-clausal unit consists of either one or more phrases which can be elaborated to a full clause by means of recovery of ellipted elements from the context (e.g., 'How long you stay here' and 'three months') or a minor utterance, which will be defined as one of the class of irregular sentences or nonsentences (e.g., 'Oh poor woman', 'Thank you very much' and 'Yes'); (3) a subordinate clause minimally consists of a finite, non-finite verb or verbless element plus at least one other clause element (i.e., Subject, Object, Complement or Adverbial).

Judgement of coordination function in speech is often problematic, because the speaker does not always use the coordinated verb clause in the same way as in writing. Surveying the recent research using unit segmentations, Foster *et al.* (2000) identify three broad categories, i.e. semantic, intonational, and syntactic. Relying on one of the categories frequently raises difficult cases to decide segmentations; so, it is useful or even imperative to consider all three elements to find appropriate unit boundaries. Taking intonational factor into account, Example 3.3 is counted as one unit, while the Example 3.4 is counted as two units:

**Example 3.3**

[and she was very kind // and she answer his question very detail (1.8)]\* (2 clauses, 1 AS-unit)

\*[...] represents an AS-unit boundary.

**Example 3.4**

[and they seem to a very starving and (.5) almost dying and...] [(8) um they seem to very (.8) they seem to be in (.8) so desperate] (2 clauses, 2 AS-units)

Both examples seemingly include a very similar syntactic structure. However, following the defined application that 'the coordinated phrases will normally be considered to belong to the same AS-unit, unless the first phrase is marked by falling

or rising intonation and is followed by a pause of at least 0.5 seconds' (Foster *et al.*, 2000:367), it is clear that the second clause in Example 3.4 is produced as a new start, due to its long unfilled pause (0.8 seconds) and the subsequent higher onset. Here the intonational feature is considered to be important, because the speaker might engage in the rapid on-line planning during this pausing.

However, the third clause in Example 3.5 below is, in spite of seamless intonational continuation, recognized as a separate segment, because this is comprised of the different subject ('he').

### **Example 3.5**

[and in order to get into the tower // the people have to pay fifty pounds] [and first of all he could paid fifty pounds for entrance] (3 clauses, 2 AS-units)

A discourse marker, 'first of all', complements this judgment. From this semantic difference, these clauses are counted as two AS-units.

Since a study by Crookes (1989) and further elaboration by Foster and Skehan (1996), the quantitative analysis of syntactic complexity has been widely employed in much task-based and planning research. However, whichever speech unit we choose, it seems that such numerical counts of subordinate clauses do not necessarily represent the reality of one's syntactic complexity level. For example, the following example, despite a rather simple syntactic structure, includes high level of complexity (five clauses in one AS-unit) by repeating similar phrases (i.e., 'call a taxi,' 'to get a taxi,' 'ask the taxi driver').

### **Example 3.6**

[then (2.9) all he did (.7) // after that was (.8) // call a taxi] [and (2.1) to get a taxi (.8) // and ask the taxi driver // to (1.0) get to the place (1.0)] (5 clauses, 1 AS-unit)

Thus, in addition to this statistical analysis of syntactic complexity, looking at these linguistic structures qualitatively is also important to grasp a real picture of language complexity.

### 3.8.3.2 COMPLEXITY MEASURES (2): DISCOURSE ORGANIZATION DEVICES

Pertaining to the fluency dimension, Pawley and Syder (1983) argue that larger macro-features, i.e., inter-unit cohesive devices, contribute to impressions of fluency. More specifically, speakers attempt to formulate spontaneous connected speech by taking markedly different syntactic strategies, that is, ‘clause-chaining’ strategy and ‘clause-integrating’ strategy. From their data, Pawley and Syder (1983:203) speculate that clause-chaining is more effective than the integrating style in terms of fluency, because the former makes speakers maintain grammatical and semantic continuity by ‘juxtaposing relatively independent clauses’ (Ejzenberg, 2000:294), while the latter involves ‘the speaker referring back to a previous structure while articulating a new phrase or clause’ (ibid.:295); as a result, it often leads to mid-clause pauses and structural breakdowns. To put it another way, clause-chaining seems to reflect a simple connection of lexicalized language by slotting ready-made phrases into appropriate places and connecting them by using conjunctures. On the other hand, clause-integrating concerns a speaker’s attempts to make a more complex structure, which is more likely to lead to dysfluent speech. In this case, more skilled speakers tend to pause or slow down at or near clause boundaries in lengthy connected discourses (Pawley & Syder, 1983:200). By doing so, they tactfully create micro-planning time within ongoing speech and avoid giving impressions of

dysfluency. Thus, it is conceivable that there are different psychological processes and different cognitive burdens imposed by the two organization devices, and the problems of the exclusive use of quantitative analysis of syntactic complexity lies in forcing compounds of both devices within one category.

Building on this distinction, Ejzenberg (2000) elaborates the following classification;

(a) Chaining devices

- Co-ordinating conjunctions (e.g., 'and', 'but', 'or' and 'so');
- Adverbial links (e.g., 'yet', 'moreover', 'furthermore', 'also', 'then', 'nevertheless', 'meanwhile', 'however', 'therefore', 'rather', 'in other words' and 'besides');

(b) Grammatical integration devices

- Subordinate conjunctions (e.g., 'when', 'after', 'unless', 'if', 'since' and 'because');
- Relative linking devices (e.g., 'that', 'which', 'who', 'whose' and 'whom');
- Restrictive linking devices (e.g., 'not only ... but also' and 'neither ... nor').

A noticeable aspect of these measures is that a simple count of chaining and grammatical integration devices does not necessarily bring about significant findings.

In her study, Ejzenberg (2000) did not find any quantitative difference between high-fluency and low-fluency speakers, but emphasizes that there is qualitative psycholinguistic evidence of a relationship between chaining and higher fluency, and grammatical integration and dysfluent episodes. Ejzenberg (2000) argues that more proficient speakers are better able to provide continuity and to successfully switch between chaining and grammatical integration, suggesting that 'the use of integration devices may be a conscious process whereby speakers' strategic competence is exhibited' (p. 299). It is noticeable that the main discussion lies in the relationship

between these linking devices and the level of fluency in the line of argument by Pawley and Syder, and Ejzenberg. Their speech data clearly illustrate a close relation of the use of these integrating devices to specific fluency features, such as the rate of speech and occurrences of dysfluent markers, which brings into consideration interaction between complexity and fluency variables.

### *3.8.4 Defining Accuracy*

The final goal, accuracy, has been very familiar with many foreign language teachers and researchers. It is probably true to say that exclusive accuracy orientation was a part of the major criticism of traditional foreign language teaching, and this was the primary motivation of the following development of CLT. However, as pointed out by much recent research (e.g., focus-on-form and French immersion studies among others), the issue of accuracy is back in the central position in language teaching, but not in the same sense as traditional classrooms were aiming at (see 2.4.2).

For the causes of inaccurate language use appearing, Skehan (1996a:46-7) observes two possibilities: one is that the underlying IL system is inaccurate; another is the result of the competence-performance relationship. It is conceivable that the first aspect can be frequently observed in L1 and possibly L2 development in ESL (English as a Second Language) contexts, because they engage in hypothesis-testing processes through a great deal of natural input, and developing their own working rules of a language. In such naturalistic learning, because learners are involved in an incessant revising process on their own, a particular moment of their IL system may be at an incomplete stage, so their production is likely to be erroneous. On the other hand, the second cause seems to be more often seen in the participants of this study, who learned L2 in EFL contexts, based on more traditional teaching methods. In such



formal classrooms, the learners are *deductively* given ‘correct’ rules of a language, rather than engaging in the process of exploring and discovering them through natural interaction. It is thus likely that their explicit knowledge of a language may be more or less accurate. Nevertheless, attempting actual use of a rule (e.g., subject-verb agreement), it is not easy to realize their ‘competence’ to the fullest extent due to high demands of processing burden under time pressure (i.e., competence ≠ performance). What happens here is that the learner is equipped with *declarative* knowledge, but has not fully incorporated it into his/her IL system in the form of *procedural* knowledge.

This distinctive two-way cause of inaccurate language is important, because accuracy tends to be regarded as purely a ‘competence’ problem, but for many L2 users, the real problem lies in their ‘performance’, suggesting the difficulty of effective translating processes from their generated message into a form of language. Only looking at an inaccurate product of L2 speakers, we would not be able to identify whether an erroneous form comes from inadequate competence or the processing problem. But this may be clarified if we also look at other dimensions of the speaking process; for example, a speaker’s attempt to use more complex language beyond their present level controlling L2 would bring more stress to the Formulator, and slowing down the speech rate implies the speaker’s lack of lexical resources and struggle to produce new language. In this sense, the issue of accuracy is inextricably linked to the other dimensions of IL variables. As the first step towards the interconnected relation of accuracy to other variables, the following section looks at accuracy measures widely employed in task-based research.

#### 3.8.4.1 ACCURACY MEASURES (1): ERROR-FREE CLAUSES

The accuracy variables employed in the preceding task-based studies vary widely, but

the underlying criteria for them are two-fold: either general (i.e., to count any syntactic errors as an inaccurate indication) or specific measures (i.e., to determine a particular grammatical entry as an index of grammatical or lexical accuracy). Another noticeable feature is that many accuracy measures did not reach a significant level of accuracy, as noted by Bygate (2001:43), either because the applied measures were too conservative, or because the nature of accuracy is a more complex phenomenon than other variables such as fluency and complexity. The main purpose of this section is to identify methodologically appropriate measures to examine the accuracy levels of L2 speakers.

The first important question for accuracy is whether to use general or specific measures. For example, Foster and Skehan (1996) reported that planning opportunity produced more accurate language (but not straightforwardly) by using general measures, while Crookes' (1989) study, emphasizing the results of specific measures, concluded that there was not any significant accuracy improvement as a result of planning opportunity. Specific measures should be appropriate if the research is designed to generate specific items (e.g., use of the past tense in Ellis 1987), but the problem with this is that it may be more subjective to differences in experimental conditions (Foster & Skehan, 1996:304). However, Mehnert (1998:86) observes that a frequently used global measure, the percentage of error-free clause proposed by Foster and Skehan (1996), may not be a particularly good measure in some cases, such as with relatively low proficiency learners or highly inflectional languages, because this does not identify the distinction between single and multiple errors within one clause. Taking these various factors into consideration, it seems sensible to employ global accuracy measures, a percentage of error-free clauses for a wide scope of analysis, together with specific measures to compensate each other (Crookes, 1988, 1989, as cited in Ortega, 1999:118).

The measure of error-free clauses is the percentage of clauses that did not contain any error in terms of syntax, morphology and lexical choice, but not including phonological errors in this analysis. When the speaker produces incorrectly and then the clause is self-corrected, the part is calculated as an error-free clause.

#### 3.8.4.2 ACCURACY MEASURES (2) CORRECT VERB FORMS AND ARTICLE

What grammatical items are to be chosen as specific measures of accuracy is a very sensitive question, because it may happen to be the case that a speaker makes a lot of errors but only with particular items. If the specific items are seldom used by the speaker, the validity of the measurement would be dubious. Therefore, the validity of a general measure of accuracy rests on the extent to which learners' ability to use specific forms correctly correlates with their overall grammatical competence (Ellis & Barkhuizen, 2005:151). In previous studies, word order was used by Mehnert (1998), and noun-modifier by Ortega (1999), but the languages investigated were German (Mehnert) and Spanish (Ortega). Thus, we should be careful about applying these items directly to the present research (i.e., English).

Taking account of psycholinguistic orientation, adopting an inflectional part of English grammar would be desirable, because transformation of correct form as well as searching for appropriate words seem to impose more cognitive burden on speech processing than only conducting lexical searches, and more proficient speakers are expected to be more capable of dealing with this difficult task. Among several candidates, the percentage of 'correct verb forms' (employed in Yuan & Ellis, 2003) was chosen to be suitable for the specific accuracy measure. The measure is the percentage of accurately used verbs in terms of tense, aspect, modality, and subject-verb agreement. That is, speculating that the verb formulation process

particularly devours much attention, and triggering more frequent occurrences of errors in L2 speakers, the measure shows the proportion of errors related to verb among all errors. Another advantage to this choice is that speakers are unable to avoid the use of verbs in each and every clause.

In addition to the measure of correct verb forms, I include that of ‘article’. Measuring the use of articles seems favourable, because Japanese speakers often find it difficult to use them correctly due to a lack of the equivalent system in their L1 (Kobayashi, 2002b:577). Since even advanced Japanese learners of English often fail to use articles correctly even in unpressured, writing tasks, the cause of this failure may be rooted in competence rather than performance problems.

### *3.8.5 Criteria for Error Identification*

In both general and specific measures, further problems arise as a result of the difficulties of determining exactly what constitutes an error (Ellis & Barkhuizen, 2005:151). It is necessary to decide the criteria for identifying errors, in other words, to what extent linguistic categories should be covered for the analysis of accuracy. Polio (1997) developed an error classification system for assessing accuracy in 38 written essays by ESL learners in an American university. Although Polio’s criteria provide us with useful guidelines for identifying errors, it is worth mentioning that the original system developed by Polio aimed to identify any types of error made by learners in great detail. The same amount of effort is not easy to achieve in spoken language, because, as mentioned above, speech is not necessarily composed of sentences, involving incorrect features such as ellipsis in writing. Polio (1997) also acknowledged that there were still disagreements of error identification between the raters, particularly in determining the native-like usage. This point should be

particularly problematic in the choice of vocabulary. Following Skehan and Foster (1997:195), therefore, only errors where a word used is 'nonexistent in English or indisputably inappropriate' will be regarded as inaccurate lexical choice. Similar attitude will also be applied to grammatical correctness in the present research. In the analysis, the researcher first identified errors in all coded data. Then, a second researcher (who is the second rater of the proficiency judgement; see 3.4.3) analyzed one-third of all data (nine participants x three planning conditions). The result of inter-rater reliability for error-identification indicated 95.1 percent agreement.

### *3.8.6 Summary of Interlanguage Measures*

As has been outlined, various kinds of performance features contribute to the constructs of fluency, complexity and accuracy. Much research has focused on examination of particular measures, but importantly they are not determined only by specific types of data. It is also noticeable that the distinction of fluency, complexity and accuracy is based only on methodological convenience; and these three areas of speech performance are highly interrelated and the borderline between them is often fuzzy. A complete understanding of the interaction between each measure is clearly beyond the level of the present study, but it should be borne in mind that the choice of measures has a direct influence on our conclusions concerning language performance. That is, failure to choose appropriate measures inevitably leads to erroneous conclusions. The only practical way to avoid this is by the inclusion of a wide range of measures. A precise examination by such extensive applications would be certainly useful not only to grasp one's performance more accurately but also to give clues to better understanding of the 'black box' of cognitive processing in L2 speaking (see Table 3.4 for the summary of the IL measures for the task performance analysis).

**Table 3.4:** Summary of the Interlanguage Measure for the Task Performance Analysis

---

<b>Fluency</b>
<ul style="list-style-type: none"><li>• Pruned speech rate</li><li>• Mean length of run</li><li>• Number of end-clause pauses per minute</li><li>• Number of mid-clause pauses per minute</li><li>• Total length of pauses</li><li>• Number of filled pauses</li><li>• Number of self-repairs (including false-starts, repetitions, reformulations and self-corrections)</li></ul>
<hr/>
<b>Complexity</b>
<ul style="list-style-type: none"><li>• Syntactic complexity (clauses per AS-unit)</li><li>• Number of chaining integration devices</li><li>• Number of grammatical integration devices</li></ul>
<hr/>
<b>Accuracy</b>
<ul style="list-style-type: none"><li>• Percentage of error-free clauses</li><li>• Percentage of correct verb forms</li><li>• Percentage of correct articles</li></ul>

---

### 3.9 Verbal Report Analysis

This section concerns the analysis of verbal protocols. After a brief statement of reasons for applying this analysis, I explain types of verbalization, criticisms of this method and several specific procedures.

#### *3.9.1 Performance and Competence-Oriented Research*

It is probably true to suggest that the mainstream of SLA and L2 learning research lies in approaches to comparing the current states of IL performance, which has been largely in accordance with the studies reviewed so far. As suggested by Færch and

Kasper (1987:6), this attitude is based on the assumption that the states of performance reflect essential characteristics of the underlying states of competence.

However, this assumption rather oversimplifies the sensitive and complex relation between the two distinctive domains. For example, Kohn (1982, cited in Lennon, 1989:380) suggests that, rather than such a simple congruence, L2 learners differ individually in their demands for making their performance accurately reflect their competence, and most learners shift position on this spectrum according to the situation. Regardless, all L2 planning research with a few exceptions (e.g., Ortega, 1999, 2005; Sangarun, 2005) have been exclusively based on this performance-analysis tradition. Although admittedly this trend has greatly contributed to our understanding of the usefulness of pre-task planning and the revelation of L2 processing, it should be stressed that such observations still remain 'descriptive' rather than 'explanatory' in terms of understanding IL competence underlying performance (Færch & Kasper, 1987:6). For example, in connection with the trade-off effects between fluency and accuracy and those between accuracy and complexity, the numerical results would offer us a great deal of clues to anticipating the states of competence but by no means make us aware of what is actually happening underneath all of these superficial linguistic (e.g., the rate of accuracy) or paralinguistic features (e.g., the total number of pauses).

Another important implication relevant to the present study is that such observations are generally limited to the moments that learners speak and do not tell us anything about the moments that they keep silent while engaging in speech planning. In tandem with L2 planning research, an investigation into this silent moment and speakers' processing would be worth focusing on, because erroneous speech production may not only occur at the moment of speaking but also come from the time of preparing for the utterance. We do not have an instrument to directly

observe one's mental operations, but it is worth considering another research paradigm, verbal protocol analysis, in order to enable us to examine the underlying states of mind in a more discernible manner. It is probably reasonable to suggest that the research tradition of SLA has tended to neglect the protocol analysis as a tool to observe the cognitive processes within task planning research. This statement does not intend to underestimate the values of performance-based studies, but it should also be stressed that there is a way to explore a number of important implications from a different angle. The rest of this chapter chiefly looks at the verbal reporting methodology, used to supplement and enforce the findings drawn from the performance analysis.

### *3.9.2 Types of Verbalization*

The above-mentioned characteristics of introspective methods are important to grasp the whole picture and to give theoretical justification to the method, but these descriptions cover any form of verbal reporting. Because the execution of verbal reports of thought is rigidly required to take specific steps to meet the expected validity of collected data, only looking at the general aspect tends to blur the important differences among various forms of introspection.

Verbal report protocols have been methodologically classified into several types, according to various criteria. Synthesizing Shavelson, Webb and Burstein (1986) and Joudenais (2001), there are four types of process tracing proposed: (a) think-aloud reports (b) introspective reports, (c) retrospective reports and (d) delayed interview (Table 3.5).



**Table 3.5:** Types of Protocol Reports (Based on Jourdenais, 2001:355)

	Concurrent with task	Subsequent to task	Delayed after task	Verbalizing process	Report process	Use of prompt
Think-aloud reports	*			*		
Introspective reports	*			*	*	
Retrospective reports		*			*	*
Delayed interview			*		*	*

With respect to the information-processing perspective, introspective reports and think-aloud reports are designed to tap directly into WM, while retrospective reports may require learners to retrieve information from LTM. Presumably, concurrent verbal reporting such as think-aloud seems more desirable for tracing more accurate mental processing in terms of immediacy of time frame, but the necessity to investigate the IL variables made us choose the retrospective option in the present study, because simultaneous attempts to conduct tasks and vocalize their thoughts would significantly harm their oral performance. On the other hand, retrospective reports are collected immediately after a task performance stage, while a delayed interview is conducted after a substantial time has past since the task completion (e.g., the next day or several months later). Thus, the former is supposed to allow participants to retain much clearer memory than the latter condition, and therefore be more reliable in reporting back on specific states of mind, while the latter tends to be useful to gain information on more global strategies. Additionally, in both methods the memory- stimulating procedure is expected to work effectively to revive partial or global memory. Therefore, the retrospective reports will be a particular focus in the present study.

### *3.9.3 Criticisms of Retrospection*

As mentioned above, the major distinction between introspective reports and retrospective reports is the 'freshness' of accessed information, and this temporal difference could lead to the different quality and quantity of information obtained from the reports. The nature of different time spans could cause some retrospective-specific problems. One of the most frequently raised objections is that, although the sequence of heeded information remains intact in the introspective levels of verbalization, retrospective reports require attention to additional information to retrieve particular memory structures (Ericsson & Simon, 1993:18-9). Thus, 'the retrieval operation is fallible, in that other similar memory structures may be accessed instead of those created by the just-finished cognitive process' (Ericsson & Simon, 1987:41). As this concern suggests, there is always a danger that the participant accesses untargeted information and misunderstands it as the correct information. That is, it is possible to retrieve similar information acquired previously, which is associated with the information the participant needs to retrieve (Ericsson & Simon, 1993), leading to 'inference' or 'fabricated intrusions' rather than reporting the mental states (i.e., 'errors of commission' argued by Ruso, Johnson & Stephens [1989:760]).

In an often-cited paper against the validity of verbal report, Nisbett and Wilson (1977:233) argue by illustrating invalid report examples:

People often cannot report accurately on the effects of particular stimuli on higher order, inference-based responses. Indeed, sometimes they cannot report on the existence of critical stimuli, sometimes cannot report on the existence of their responses, and sometimes cannot even report that an inferential process of any kind has occurred.

In response to this criticism, Ericsson and Simon (1993:27) demonstrate that the accuracy of verbal reports depends on the procedures used to elicit them and the relation between the requested information and the actual sequence of heeded information. From this ground, they suspect that invalid reports reviewed in Nisbett and Wilson (1977) may be due to a lack of these factors, and conclude that better methods for probing for that awareness would yield considerable insight into the cognitive processes (Ericsson & Simon, 1993:29).

This discussion illuminates the importance of careful formation of verbal report procedures. To preclude the contamination of inference, it is recommended to take into consideration: (1) the instruction to guide participants to report what they can remember and (2) careful analysis of the protocols, not simply trusting the participants (Ericsson & Simon, 1987:41). The following parts will mention specific procedures applied in the present research in order to minimize concerns for the protocol analysis.

#### *3.9.4 Stimulated Recall Methodology*

Recall support is crucial in retrospection in order to fill the temporal gap between the action and verbalization. In doing so, the participants are confronted once again with the performance situation (Færch & Kasper, 1987:17). It has been argued that such confrontations provide reactive traces in WM and counteract participants' tendency to conflate different events or confound them in retrospect. This danger should be more seriously taken in the case of retrospection. The only possible way to handle this drawback is, as mentioned above, to instruct the participants to report only on what they experience, not what they infer, and also to carefully analyze the protocols, not simply trusting every word by participants.

In addition to these cautions, to compensate for a partial memory loss caused by the time lag between task completion and the verbal reporting stage, the stimulated recall methodology was employed, which has been widely known as a useful tool to uncover cognitive processes in L2 research (Gass & Mackey, 2000). Gass and Mackey (2000) define stimulated recall as 'one subset of a range of introspective methods that represent a means of eliciting data about thought processes involved in carrying out a task or activity' (p. 1) by using a reminder of an event (audio recording and a written note in the present study) to stimulate recall of the mental processes in operation during the event. It is worth pointing out that the theoretical foundation for this elicitation technique is one focused form of introspection, based on an information-processing approach (ibid.), which should be useful in examining the planning processes informed by the same theoretical framework.

In a methodological sense, Nunan (1992:96) observes that this technique is particularly useful as an initial step in the research process, acting as a stimulus for the framing of questions for more formal investigation. This remark may be important for the present research aiming at accessing detailed steps of L2 learner's thought processes, to explore what mental activities are engaged in strategic and on-line planning, and more specifically to investigate whether any form of focus-on-form during on-line planning can be identified.

The form of the retrospection plans to frame a semi-structured interview, that is, the report was initiated by a set of pre-determined questions, but the following verbal reporting would be entirely left to the speaker. It was expected that, although they may not be very conscious of their cognitive processes before listening to their own recorded speech, their memory would be immediately activated through the listening and their own reporting, and was expected to eventually uncover various aspects of mental operations.

It is also argued that an important assumption of stimulated recall is 'recall accuracy' (Bloom, 1954, cited in Gass & MacKay, 2000:18). In researching classroom events, Bloom (1954) found that if the recalls were prompted a short period of time after the event (generally 48 hours), recall was 95 % accurate, and accuracy declined as a function of the intervening time between the event and the recall. In addition, stimulated recall provides contextual information, assisting the participants' memory retrieval. One of the major criticisms of mixture of report with inference can be overcome by segmenting the whole into specific listening parts; when encountering a focused point, the listening is stopped and the participant is requested to report about his/her thought at this point. In doing so, stimulated recall could create a similar condition to think-aloud without hindering their speech processing and might any participant training. In such conditions, the participant relives an original situation with great vividness and accuracy and 'this type of investigation can be carried on in such a way as to have only minimal effect on the nature of the original situation (Bloom, 1953, cited in Shavelson *et al.*, 1986:80).

### *3.9.5 Use of the First Language*

Use of either L1 or L2 in verbal reporting seems crucial in conducting the retrospective method in SLA because of the participant's incomplete L2 competence. Kormos (1998:354) suggests that ideally the participants should produce verbal reports in their L1, because L2 reports may not be a precise representation of their thought sequences, and describing mental states requires more precise and sensitive language than describing everyday events. Following this recommendation, the whole process of retrospection in this research was done in the speaker's first language (i.e., Japanese) in order to enable participants to describe their cognitive processes more

deeply without being thwarted by any linguistic difficulties.

### *3.9.6 Degree of Verbalization Structure*

‘The degree of verbalization structure’ refers to the extent to which it predetermines the content of the verbalization, ranging from the most highly structured instruments such as rating scales and multiple choice questionnaires to the least structured instruments such as leaving it to the participants to decide what, how much, when and how they verbalize (Færch & Kasper, 1987). The intermediate position between the two extremes is a way to elicit reports on specific cognitive processes, which is restricted by the researcher’s instructions or questions, while the specific content and form of the report is left to the participant’s decision (ibid.; 17). Given that the present study aims to explore L2 speaker’s cognitive and on-line planning processes during engaging in speaking, the attitude to give the participants freedom to verbalize their mental conditions seems more appropriate. Because there is a danger that the participants might not willingly take initiatives to report for affective (e.g., a fear of reporting silly matters) or social reasons (e.g., Japanese people’s passive personality), it is probably safer and more reasonable to take an intermediate position; that is, a set of basic questions is prepared by the researcher but the responsibility for the content of reports is given entirely to the participants.

## **3.10 Conclusion**

This chapter has described the various methodological issues of the study. Bearing in mind the theoretical accounts in the previous chapters, two general research objectives were proposed; do strategic and on-line planning differently influence L2 oral

performance?; and do participants focus on form in on-line planning more frequently than in strategic and non-planning? In order to explore these objectives, I set up a two-way research design: task performance and verbal reporting.

The latter half of this chapter considered the issues of the analysis of task performance and verbal protocols. Previous task planning studies used a number of IL measures, but many of them tended to focus on a limited range of measures. As our understanding of 'fluency', 'complexity' and 'accuracy' is still immature, failure to choose appropriate measures can lead to wrong conclusions. To avoid this risk and to promote our understanding of the construct, the present study applies a wide range of measures for the task performance analysis.

This chapter also considered the analysis of verbal report protocols, especially by focusing on retrospective and stimulated recall methodology. This particular research tradition is often criticized not only because the difficulty in accessing implicit knowledge can result in incomplete data, but also because its heavy reliance on participants' reports and researchers' interpretations could easily result in a lack of objectivity. There is a possibility that verbal report protocols include methodological problems, and it is thus necessary to carefully design the specific procedures. However, it is also true to say that it is the only established way to access thought processes in the field. Despite these recognized problems, it is believed that retrospective verbal reports combined with stimulated recall have the potential to go beyond our current understanding drawn from the performance-based research.

Assuming that verbal reporting can help reveal the cognitive and psycholinguistic processes underlying performance, it is conceivable that a combination of performance analysis with this methodology will allow us to go beyond the common practice of analyzing L2 speakers' competence solely on the basis of performance data (Kormos, 1998:354). Rather than choosing either

competence or performance-based traditions, it would be probably better to move towards the attitude that verbal reports be validated in line with the research on the performance dimension. In designing the present research, neither task performance analysis or verbal reports alone can answer all of the proposed questions; I believe that only a combination of both methods would reach more reliable and meaningful conclusions.



## **CHAPTER FOUR: TASK PERFORMANCE ANALYSIS**

Having looked at a number of issues involved in conducting the present study, I shall now describe the findings of the task performance analysis and the verbal report analysis. First of all, this chapter examines the task performance by means of several statistical models using SPSS (Statistical Package for the Social Studies); and the next chapter will explore speakers' cognitive processing in more qualitative ways by analyzing the post-task verbal reporting. Integrating these two approaches, Chapter Six will discuss the issues which are emergent from these analyses, exploring the nature of different types of planning and their effects on learners' psycholinguistic processes.

### **4.1 Stages of Statistical Analysis**

As the present study deals with a number of dependent variables, two methods of statistical analysis can be considered: a separate ANOVA (analysis of variance) used multiple times for each dependent variable and MANOVA (multivariate analysis of variance). The former method is often considered problematic because, if conducting ANOVAs multiple times, it is more likely to make a Type I error (i.e., incorrect rejection of a null hypothesis). That is, 'the more tests we conduct on the same data, the more we inflate the familywise error rate' (Field, 2005:572), which may finally result in a mistaken belief that there is a genuine effect, when there is actually not. In order to avoid this risk, the study performed MANOVA first to see the general planning effects.

However, some other concern with MANOVA should be taken into

consideration. Tabachnick and Fidell (2001) suggest that the dependent variables should not be correlated with each other, and each one should measure different aspects of the construct in MANOVA. Therefore, following Tavakoli and Skehan (2005), factor analysis was first performed in order to explore the measures representing independent construct, ‘to reduce a data set to a more manageable size while retaining as much of the original information as possible’ (Field, 2005:619). This data reduction method contributed to our understanding of the overall planning effects on task performance in the following MANOVA, but it inevitably ignore subtle differences brought by individual IL variables, which were not included in the MANOVA analysis. What is more, the output of MANOVA only tells us about the overall effect of planning on performance but does not give specific information on individual planning and IL variables. For this reason, despite a risk of Type I error, separate ANOVAs were also performed to grasp more local effects on task performance and also to see the different responses to each planning implementation between different proficiency levels. Table 4.1 summarizes the statistical models applied in the present analysis and the main aims of their application.

**Table 4.1:** Summary of Statistical Methods Applied

<b>Statistical Methods</b>	<b>Aims</b>
<b>1. Factor analysis</b>	To investigate the correlation of individual IL measures, and to extract the independent variables for MANOVA
<b>2. MANOVA</b>	To investigate the overall effects of planning on performance
<b>3. ANOVAs</b>	To investigate the effects of planning on (1) individual measures between (2) different proficiency levels

## 4.2 Results of Statistical Analysis

### 4.2.1 Basic Descriptive Statistics

Before launching into specific statistical analyses, the results of descriptive statistics for all measures are shown in order to overview varying tendencies between the three planning conditions (Table 4.2). Looking at the fluency measures, there is a tendency for on-line planning to be more dysfluent than non-planning and strategic planning, while there is no clear difference between non-planning and strategic planning (e.g., ‘pruned speech rate’ and ‘total length of pauses per minute’). As for complexity measures, syntactic complexity does not show any clear difference, but the ‘number of grammatical integration devices’ presents a relative difference between the three conditions. Finally, for accuracy measures, there seems a relative superiority in strategic planning and on-line planning to non-planning.

**Table 4.2:** Descriptive Statistics

		<b>Planning</b>	<b>Mean (SD)</b>
<b>Fluency</b>	Pruned speech rate	NP	98.13 (34.13)
		SP	101.65 (24.68)
		OP	73.95 (28.45)
	Mean length of run	NP	4.67 (1.94)
		SP	4.73 (1.67)
		OP	3.87 (1.14)
	Mid-clause pauses per 100 words	NP	12.61 (6.20)
		SP	11.06 (4.84)
		OP	15.56 (6.35)
	Total length of pauses per minute	NP	27.48 (5.93)
		SP	27.23 (5.43)
		OP	32.59 (6.50)
Total number of self-repairs	NP	7.63 (4.30)	
	SP	9.19 (5.64)	

		OP	13.30 (7.23)
<b>Complexity</b>	Syntactic Complexity	NP	1.38 (0.23)
		SP	1.43 (0.24)
		OP	1.45 (0.17)
	Number of chaining integration devices	NP	8.15 (2.58)
		SP	11.63 (5.15)
		OP	10.59 (4.55)
	Number of grammatical integration devices	NP	1.78 (1.45)
		SP	3.59 (2.45)
		OP	5.11 (5.14)
<b>Accuracy</b>	Percentage of error-free clauses	NP	57.38 (23.33)
		SP	62.41 (21.95)
		OP	69.16 (21.82)
	Percentage of correct verb forms	NP	70.17 (22.77)
		SP	74.94 (19.68)
		OP	79.24 (20.32)
	Percentage of correct article	NP	73.56 (20.88)
		SP	78.66 (18.39)
		OP	77.63 (23.26)

Note: NP = non-planning, SP = strategic planning, OP = on-line planning

#### 4.2.2 Factor Analysis

Factor analysis is primarily concerned with describing the variation or variance which is shared by the scores on three or more variables (Bryman & Cramer, 2001). That is, this method attempts to find a third variable that explains the relationship between two variables (Brace, Kemp, & Snelgar, 2003:278), and a factor is 'a construct operationally defined by its factor loadings' (Royce, 1963, cited in Kline, 1994:5). For practical reasons, the present study used this statistical model to find the measures which are not correlated, or more simply, to reduce the number of measures, in order to perform the subsequent MANOVA. However, it is also meaningful to use this model to identify which IL variables represent particular constructs. As explained in

Chapter Three, there have been a considerable number of IL variables proposed to pertain to one of the three performance dimensions – fluency, complexity and accuracy. Although these associations seem theoretically justified, it is unclear whether these measures in practice correspond to one of the three components in real performance. Therefore, factor analysis was performed to find the independent measures for MANOVA, but the results were also used to examine to what extent each measure is related to the theoretical rationale of performance variables.

#### 4.2.2.1 SAMPLE SIZE

Because correlation coefficients fluctuate from sample to sample, the reliability of factor analysis is also dependent on sample size (Field, 2005:638). There are a lot of indications of how sizeable many samples collected for factor analysis should be, although there is no consensus on how large the size should be (Bryman & Cramer, 2001:263). For example, Nunnally (1987) recommends 10 times as many participants as variables. Tabachnick and Fidell (2001) suggest at least 300 cases for the analysis. On the other hand, a Monte Carlo study (Guadagnoli & Velicer, 1988, cited in Stevens, 2002:395) indicates that the most important factors are components saturation (i.e., the absolute magnitude of the loadings) and absolute sample size, arguing that if a factor has four or more loadings greater than .6, then it is reliable regardless of sample size. Following this, Stevens (2002:395) recommends that any component with at least three loadings above .80 is reliable. Added to this, MacCallum, Widamanm Zhang and Hong (1999, cited in Field, 2005:640) indicate that with all communalities (i.e., the proportion of common variance within a variable) above .6, relatively small samples (less than 100) may be perfectly adequate. As the sample size in the present study (i.e., 27 cases) is far from the ideal numbers (i.e., 300) recommended by

Tabachnick and Fidell (2001), the reliability of factor analysis was judged by the values of loading and communalities, showing several reliable components (see 4.2.2.3). Following Stevens' (2002:388) recommendation, Bartlett's sphericity test was also applied with a small sample size to test the null hypothesis that the original correlation matrix is an identity matrix (Field, 2005:652), showing a significant level ( $p < .001$ ) in all the three conditions. Thus, it can be assumed that there are some relationships between the variables despite a small sample size.

#### 4.2.2.2 PRELIMINARY ANALYSIS

As a preliminary analysis, R-matrices (correlation matrices) in the three planning conditions were examined to check the pattern of relationship and whether any data set needed to be reduced to achieve parsimony by factor analysis. It is recommended that any variables that do not correlate with any other variables or that correlate very highly with other variables ( $R < .9$ ) should be eliminated because multicollinearity (i.e., variables that are very highly correlated) and singularity (i.e., variables that are perfectly correlated) are not appropriate for the statistics (Field, 2005:641).

As factor analysis aims to retain a small number of underlying constructs and to discard the rest of the information accounting for only 20 or 30 percent of the variance (Stevens, 2002:390), criteria for deciding the number of factors to retain for interpretation is necessary. Among various methods, Kaiser (1960) recommends retaining all factors with eigenvalues greater than 1, and research shows that Kaiser's criterion is accurate when the number of variables is less than 30 and the resulting communalities are all greater than .7 (Field, 2005:633; Steven, 2002:390). Thus, the present study consisting of thirteen dependent variables, adopting Kaiser's criterion (i.e., eigenvalues  $\geq 1$ ) while watching the values of communalities.

In order to increase the interpretability of factors, rotation of factors is needed while keeping the number of factors and communalities of each variable fixed. This is not only because factors with many large loadings are hard to identify and the loadings reflect the algebra by which they were computed, but also because there is no one perfect or ideal solution in factor analysis (Kline, 1994:56). There are two different types of analytic rotation methods that are free of subjective judgement – the method of orthogonal and oblique rotation (Kim & Muller, 1978:30). In the present analysis, orthogonal rotation was adopted because there are theoretical reasons to assume that the underlying factors (i.e., fluency, complexity and accuracy) could be independent (e.g., Foster & Skehan, 1996; Skehan & Foster, 2001). Among the three methods of orthogonal rotation (i.e., varimax, quartimax and equamax) in SPSS, varimax was selected because it attempts to maximize the dispersion of loadings within factors and simplifies the interpretation of factors (Field, 2005:636-7). A summary of the proportion of residuals (i.e., how many residuals have an absolute value greater than .05) was checked in reproduced rotations in each planning condition, and confirmed that none of them were more than 50 percent (Field, 2005:656).

#### 4.2.2.3 RESULTS OF FACTOR ANALYSIS

Table 4.3, 4.4, and 4.5 show the results of rotated component matrices in each planning condition. Factor loadings less than .4 have not been reported, following the recommendation that this cut-off point is appropriate for interpreting purposes (Stevens, 2002).

**Table 4.3: Results of Factor Analysis (NP)**

Measures	1	2	3	4	Communality
Total length of pauses	.950				.937
Mean length of run	-.894				.891
Pruned speech rate	-.871				.890
Mid-clause pauses	.870				.758
End-clause pauses	.783				.627
Error free clauses		.947			.936
Correct verb forms		.932			.888
Grammatical integration			.891		.879
Self-repairs	.477		.646		.792
Syntactic complexity		.403	.585	-.482	.876
Chaining integration				.749	.730
Correct article				.733	.564
Filled pauses				.519	.620

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

**Table 4.4: Results of Factor Analysis (SP)**

	1	2	3	Communality
Total length of pauses	.935			.877
Pruned speech rate	-.915			.872
Mean length of run	-.913			.842
Mid-clause pauses	.815			.678
End-clause pauses	.770			.642
Grammatical integration		.883		.791
Self-repairs		.696		.629
Chaining integration		.683		.514
Syntactic complexity		.670		.540
Filled pauses		.665		.468
Error free clauses			.977	.968
Correct verb forms			.872	.776
Correct article			.699	.534

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.



**Table 4.5: Results of Factor Analysis (OP)**

	1	2	3	4	Communality
Mean length of run	.947				.908
Mid-clause pauses	-.900				.880
Pruned speech rate	.849				.877
End-clause pauses	-.759				.683
Error free clauses		.964			.963
Correct verb forms		.919			.894
Correct article		.527			.517
Syntactic complexity			.825		.805
Self-repairs	-.469		.755		.835
Grammatical integration			.720		.637
Chaining integration				.821	.725
Filled pauses				.813	.705

Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization.

Looking at the results of the three factor analyses, several noticeable differences can be observed. Firstly, there were four factors extracted in NP and OP, while there were only three factors in SP. However, focusing on the first three factors in each, the results in three planning conditions similarly represented three independent dimensions, fluency, complexity and accuracy. That is, the first component includes ‘total length of pauses’, ‘mean length of run’, ‘pruned speech rate’, ‘mid-clause pauses’, ‘end-clause pauses’ and ‘self-repairs’ (i.e., fluency); the second (or third) includes ‘the number of grammatical integration devices’, ‘syntactic complexity’ and ‘self-repairs’ (i.e., complexity); and the third (or second) includes ‘error-free clauses’ and ‘correct verb forms’ (i.e., accuracy). On the other hand, the other measures (i.e., ‘the number of chaining integration devices’, ‘percentage of correct article’ and ‘the number of filled pauses’) were in the fourth, miscellaneous component in NP and OP without any logical reason for this grouping. As mentioned in 4.2.2.1, it is important to check the values to decide whether certain measures are included for interpretation

purposes because of a small sample size. In the fourth component of NP, there are four measures but the values in two of them are not very high (i.e., 'syntactic complexity' and 'filled pauses'). In addition, although two measures with high loading are in NP and OP, there may not be a definite reason to indicate the validity of this component. Kim and Mueller (1978:42-45) suggest that, given the complexity as well as uncertainties inherent in the method, the final judgment has to rest on the reasonableness of the solution on the basis of current standards of scholarship in one's own field. Taking these points together, these uncertain variables were excluded for the following MANOVA analysis.

With respect to the selection of the most representative measure in each component, Stevens (2002:394) shows the critical values for a simple correlation at  $\alpha = .01$  (two-tailed) for sample size, suggesting that  $.722 = >2 (.361)$  in absolute value would be declared statistically significant with 50 subjects. As the sample size of the present research is smaller than 50 and Stevens does not show any value for this, the measures at least more than .722 of a loading should be selected as a representing measure to be used in MANOVA. Following Stevens' criteria, 'total length of pauses', 'speech rate' and 'mean length of run' show a high loading and high communality consistently on one of the factors in the three planning conditions, the fluency component. Among them, 'mean length of run' was selected for MANOVA analysis, because this measure is credited as a reliable index in many fluency studies (e.g., Kormos & Dénes, 2004; Lennon, 1990; Riggenbach, 1991). For the complexity component, 'syntactic complexity' has been used in many planning studies, but the factor analysis did not show a high loading in the present research. Instead, 'the number of grammatical integration devices' involving a high loading in every type of planning was chosen for the complexity measure for MANOVA. For accuracy, 'the percentage of error-free clauses' was selected for MANOVA, because it consistently

showed a higher loading than the percentage of correct verb forms.

In addition to the point of data reduction, the results of factor analysis suggest other interpretations of several IL measures. The most striking one is the three factor analyses clearly represent the three-way distinction of task performance. The independence of fluency, complexity and accuracy has been theoretically justified (e.g., Skehan, 1998), but it seems to be lacking in empirical data. Thus, the results of factor analysis may support the independent nature of the three components and justify the present method of analysis in learners' language. However, the results also show that all of the measures do not correspond to the original assumptions. For example, the number of self-repairs was originally categorized in the fluency dimension, but the factor analysis shows that it is correlated to both fluency and complexity in NP and OP, and correlated only to complexity in SP, possibly because speakers showed hesitations before producing complex language. Therefore, this measure seems to concern complexity as well as fluency rather than an independent fluency measure.

### *4.2.3 Repeated Measures MANOVA*

#### 4.2.3.1 REPEATED MEASURES VS. INDEPENDENT MEASURES

Because each participant contributed to all the conditions, a repeated-measures (i.e., within participants) design was employed. One advantage of using this rather than an independent design is that the type of experiment controls individual differences and gives greater power to detect effects (Bray & Maxwell, 1985:69), by reducing the 'unsystematic variability' (i.e., differences created by unknown factors) in the design (Field, 2005:428). According to Field (2005:272), there are different reasons for the

different results between a repeated measures and independent measures design. In the former, differences can be caused by (1) the manipulation that was carried out on the participants, or (2) any other factor that might affect the way in which a person performs from one time to the next, but (2) is fairly minor compared to the influence of (1). On the other hand, in an independent design, differences can be caused by (1) the manipulation that was carried out on the participants, or (2) differences between the characteristics of the people allocated to each of the group, and the second factor may create considerable random variation within each condition. Planning studies seem to have traditionally favoured a between-participants design rather than a within-participants design, but there will remain the question of how to standardize the level of participants in each condition. That is, regardless of the different conditioning, there is a danger that a greater number of higher (or lower) proficiency participants gathered in one of the groups by chance. To solve this problem, random sampling has been usually applied, but, considering the relatively small number of participants in the present analysis, there is still the possibility of committing unsystematic variation. Therefore, the study used a repeated-measures design rather than an independent-measures design as a more appropriate method.

#### 4.2.3.2 RESULTS OF MANOVA

MANOVA is different from ANOVA in that it is designed to look at several dependent variables simultaneously. As the present study concerns thirteen dependent measures, using MANOVA is more appropriate than ANOVA, reducing the chance of making a Type I error. Another advantage of MANOVA is to provide information of relationships between independent variables. This information is important in the

present study, which attempts to see any different responses between different proficiency groups.

The output of MANOVA shows four different values (i.e., Pillai's Trace, Wilks' Lambda, Hotelling's Trace and Roy's Largest Root). To select the appropriate value is dependent on several factors such as the results of Box's Test of Equality of Covariance and the sample size; but, because the results of four tests showed the same significance level, only the result of Wilks' Lambda is shown in Table 4.6.

**Table 4.6:** Results of Repeated measures MANOVA

<b>Effects</b>	<b>Value</b>	<b>F</b>	<b>Hypothesis df</b>	<b>Error df</b>	<b>Sig.</b>	<b>Noncent. Parameter</b>	<b>Observed Power</b>
<i>Between-</i> <i>participants</i>							
<b>proficiency</b>	.487	8.08	3	23	.001*	24.231	.977
<i>Within</i> <i>-participants</i>							
<b>planning</b>	.283	8.45	6	20	.001*	50.697	.999
planning * proficiency	.524	3.03	6	20	.028*	18.184	.794

The result shows the overall effects of the three planning conditions on oral performance. Together with the strong observed power, it could be said that the strategic and on-line planning conditions had some sort of impact on oral performance, and the proficiency factor made a difference in the performance. It also shows the interaction between planning and proficiency, suggesting that different proficiency groups responded differently to the three planning conditions from the result of interaction between planning and proficiency. However, because the MANOVA test does not tell us in which relations the significant effects occur, it is necessary to consult the following univariate tests to identify the locations of significance, which

will be shown in the next section.

#### 4.2.4 Factorial ANOVAs

##### 4.2.4.1 UNIVARIATE TEST OF WITHIN-PARTICIPANT EFFECT

The Mauchly's sphericity test shows the significant value in the grammatical integration devices measure, suggesting the violation of the sphericity assumption and need for F-value correction. Thus, the values of the more conservative Greenhouse-Geisser were reported for grammatical integration devices in Table 4.7. On the other hand, the F values for the other measures (i.e., mean length of run and error-free clauses) met the value of the sphericity test ( $p > .05$ ).

**Table 4.7:** Results of Univariate Test

<b>Source</b>	<b>Measure</b>	<b>Type III Sum of Squares</b>	<b>df</b>	<b>Mean Square</b>	<b>F</b>	<b>Sig.</b>
planning	Mean length of run	12.21	2	6.10	9.75	.001*
	Grammatical integration	144.65	1.28	113.39	8.84	.003*
	Error-free clauses	2041.17	2	1020.59	4.36	.018*
planning * proficiency	Mean length of run	2.99	2	1.49	2.38	.103
	Grammatical integration	39.91	1.28	31.29	2.44	.122
	Error-free clauses	2276.64	2	1138.32	4.86	.012*

As suggested by Table 4.7, the univariate test shows significant effects in all three measures, implying that different planning conditions influence oral performance in terms of fluency, complexity and accuracy. The result also shows a significant effect of interaction between planning and proficiency in 'error-free clauses', suggesting that the high and low proficiency groups responded differently to the three planning conditions in terms of accuracy.

#### 4.2.4.2 OVERALL RESULTS OF ANOVA

The results of the MANOVA test and the following univariate tests showed the overall tendency of planning and proficiency effects on task performance, suggesting that the different planning conditions differently influenced the oral performance of the high and low proficiency groups in terms of fluency, complexity and accuracy. However, these tests do not provide specific information of how the three planning conditions influenced different proficiency speakers. Therefore, a series of ANOVAs and a *post hoc* Bonferoni were performed to find individual relationships. This part will look at the effects of individual measures on all participants (Table 4.8) and the results of different proficiency groups (Table 4.10 & 4.11).

**Table 4.8:** Results of ANOVAs in All Participants

Measure	F	Sig.	Mean (SD)			Location of Sig.
			NP	SP	OP	
Pruned speech rate	37.55	.001*	98.13 (34.13)	101.65 (24.68)	73.95 (28.45)	NP>OP, SP>OP
Mean length of run	9.75	.001*	4.67 (1.94)	4.73 (1.67)	3.87 (1.14)	NP>OP, SP>OP
End-clause pauses	5.84	.005*	11.63 (3.56)	12.99 (3.82)	14.00 (2.82)	OP>NP
Mid-clause pauses	11.37	.001*	12.61 (6.2)	11.06 (4.84)	15.56 (6.35)	NP>OP, SP>OP
Total length of pauses	22.29	.001*	27.48 (5.93)	27.23 (5.43)	32.59 (6.5)	NP>OP, SP>OP
Filled pauses	4.56	.022*	8.04 (6.18)	8.52 (7.29)	11.48 (10.2)	OP>NP
Self-repairs	11.78	.001*	7.63 (4.3)	9.19 (5.64)	13.3 (7.23)	OP>NP, OP>SP

Syntactic complexity	.93	.400	1.38 (.23)	1.43 (.24)	1.45 (.17)	
Chaining integration devices	8.79	.001*	8.15 (2.58)	11.63 (5.15)	10.59 (4.55)	SP>NP, OP>NP
Grammatical integration devices	8.84	.003*	1.78 (1.45)	3.59 (2.45)	5.11 (5.14)	SP>NP, OP>NP
Error-free clauses	4.36	.018*	57.38 (23.33)	62.41 (21.95)	69.16 (21.82)	OP>NP
Correct verb forms	2.84	.068	70.17 (22.77)	74.94 (19.68)	79.24 (20.32)	
Correct articles	0.56	.511	73.56 (20.88)	78.66 (18.39)	77.63 (23.26)	

One of the noticeable points in the results for all participants is the clear effects of planning on fluency. In many fluency measures, both NP and SP were significantly different from OP, although there were no clear differences between NP and SP. The significant difference between NP and OP suggests that the on-line planning instruction successfully differentiated this from the non-planning instruction, rejecting a possibility that the participants might ignore the instruction and react to the two conditions in completely the same manner. It is worth noting that no significant difference was identified between the SP and NP conditions in all fluency measures. That is, in the present study strategic planning did not lead to fluency improvement, which is contrary to the planning literature which shows its positive effects on fluency.

For complexity, there is no difference between the three conditions in ‘syntactic complexity’, which has been most frequently used as a complexity index in previous planning research; instead, both SP and OP outperformed NP in the number of chaining and grammatical integration devices. Moreover, the mean scores of grammatical integration devices suggest relatively more positive effects of OP over SP, although they did not reach a significant level. Thus, it can be said that there were



positive effects on complexity in strategic planning and on-line planning.

In the accuracy measures, there was a significant effect of OP over NP in 'error-free clauses'. In the mean scores, SP seems to have more positive influence on accuracy than NP, but this did not reach significance. Also, there were relatively more positive effects of OP over SP in this measure, although there was no significant difference between them. A similar tendency lies in the result of correct verb forms, which failed to reach significance. In the result of article use, there seems to be an improvement of SP and OP, but this also did not reach the level of significance.

To summarize, strategic planning did not have more positive fluency effects than non-planning, but on-line planning had negative fluency effects; strategic and on-line planning had positive effects on complexity; strategic planning had limited effects and on-line planning had positive effects on accuracy.

Referring back to the research hypotheses (3.3), **Hypothesis 1** ('SP will have more positive influence on fluency, complexity and accuracy than NP') is rejected, because strategic planning did not lead to a significant level of improvement in fluency and accuracy. **Hypothesis 2** ('OP will have more positive influence on complexity and accuracy, but more negative influence on fluency than NP') is accepted, because on-line planning led to less fluent but more complex and accurate production than NP. **Hypothesis 3** ('OP will have more positive influence on complexity and accuracy, but more negative influence on fluency than SP') is only partly accepted because the differences between on-line planning and strategic planning did not reach the level of significance, but the mean scores in many individual measures of on-line planning outperformed those of strategic planning.

#### 4.2.4.3 RESULTS OF ANNOVAS IN TWO PROFICIENCY GROUPS

This section shows results of multiple ANOVAs in the high and low proficiency groups. Table 4.9 shows the results of the individual measures in the high proficiency group. The results of this group were similar to the overall results, except for the results of the accuracy measures, showing no difference between the three planning conditions in any of the measures. Thus, it could be said that neither strategic nor on-line planning conditions influenced the high proficiency participants in terms of accuracy.

**Table 4.9:** Results of ANOVAs in the High Proficiency Group

Measures	F	Sig.	Mean (SD)			Location of Sig.
			NP	SP	OP	
Pruned speech rate	20.14	.001*	119.04 (30.74)	111.62 (27.45)	90.41 (26.27)	NP>OP, SP>OP
Mean length of run	7.04	.004*	5.70 (2.11)	5.40 (1.77)	4.48 (1.07)	NP>OP, SP>OP
End-clause pauses	7.83	.002*	10.54 (3.31)	11.95 (2.88)	13.38 (2.84)	OP>NP
Mid-clause pauses	1.79	.188	9.37 (5.19)	9.94 (5.43)	11.62 (3.64)	
Length of pauses	10.44	.001*	24.54 (5.6)	25.31 (5.8)	29.41 (5.57)	OP>NP, OP>SP
Filled pauses	2.88	.096	6.43 (6.48)	9.93 (8.31)	11.14 (12.98)	
Self-repairs	5.31	.012*	5.71 (4.43)	10.00 (6.84)	11.14 (8.48)	SP>NP, OP>NP
Syntactic complexity	.37	.698	1.41 (.25)	1.47 (.19)	1.46 (.21)	
Chaining integration devices	5.72	.009*	9.14 (2.8)	13.71 (5.94)	12.79 (5.28)	SP>NP

Grammatical integration devices	6.18	.020*	1.79 (1.67)	4.79 (2.64)	6.71 (6.66)	SP>NP, OP>NP
Error-free clauses	.04	.957	71.26 (18.88)	69.68 (19.98)	70.53 (21.63)	
Correct verb forms	.3	.744	82.55 (14.5)	79.11 (16.33)	79.31 (22.32)	
Correct articles	2.21	.154	75.71 (23.64)	83.62 (16.5)	82.3 (16.7)	

On the other hand, Table 4.10 shows the results of individual measures in the low proficiency group. The main difference between this group and the overall results lies in the fluency measures. In several fluency measures (i.e., pruned speech rate and mid-clause pauses), strategic planning outperformed non-planning, going against the lack of fluency effects in the overall results (4.2.4.2).

Looking at the results of the fluency measures in both proficiency groups, strategic planning has no effect on the high proficiency group, but some positive effects on the low proficiency group, which is consistent with previous studies in showing the positive effects of strategic planning on fluency. In the same way as the overall results, on-line planning had negative effects on both the high and low proficiency groups.

The results of complexity measures of each proficiency group seem similar to the overall results (4.2.4.2), but close examination reveals relative differences between the two levels. In particular, strategic and on-line planning had positive effects on the use of grammatical integration devices in the high proficiency group but only on-line planning had a significant effect on the low proficiency group, although the mean scores of strategic planning outperformed non-planning in the group. Therefore, there was a clear effect of complexity by strategic and on-line planning on the high proficiency group, but a rather weak effect on the low proficiency group.

**Table 4.10:** Results of ANOVAs is the Low Proficiency Group

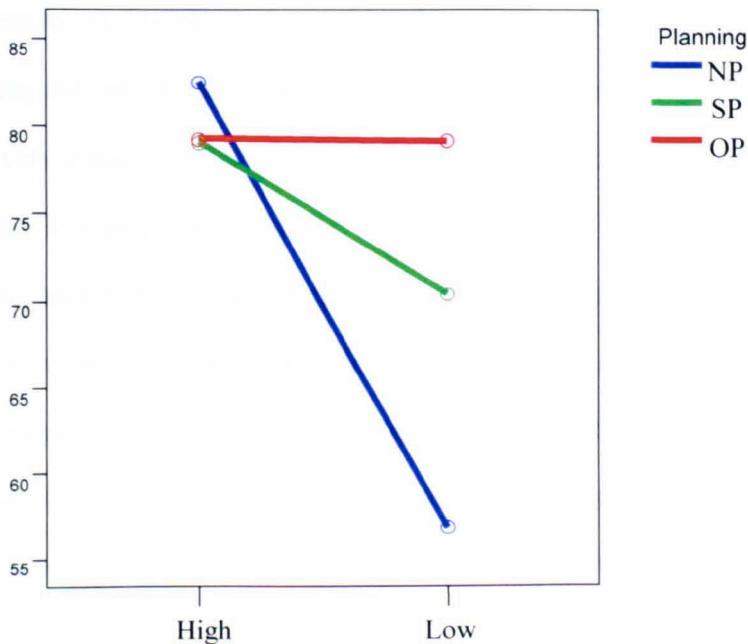
Measures	F	Sig.	Mean (SD)			Location of Sig.
			NP	SP	OP	
Pruned speech rate	22.31	.001*	75.61 (20.96)	90.92 (16.23)	56.23 (18.7)	SP>NP>OP
Mean length of run	4.78	.018*	3.56 (.87)	4.01 (1.24)	3.21 (.81)	
End-clause pauses	1.24	.307	12.81 (3.56)	14.11 (4.48)	14.67 (2.75)	
Mid-clause pauses	12	.001*	16.09 (5.36)	12.26 (3.97)	19.8 (5.95)	NP>SP, OP>SP
Length of pauses	12.27	.001*	30.65 (4.62)	29.29 (4.32)	36.02 (5.76)	
Filled pauses	6.82	.005*	9.77 (5.57)	7.00 (5.97)	11.85 (6.52)	OP>SP
Self-repairs	10.83	.001*	9.69 (3.15)	8.31 (4.07)	15.62 (4.93)	OP>NP, OP>SP
Syntactic complexity	.72	.496	1.34 (.2)	1.39 (.29)	1.44 (.13)	
Chaining integration devices	3.94	.033*	7.08 (1.89)	9.38 (2.96)	8.23 (1.74)	
Grammatical integration devices	4.61	.020*	1.77 (1.24)	2.31 (1.44)	3.38 (1.71)	OP>NP
Error-free clauses	7.6	.003*	42.44 (18.06)	54.57 (21.99)	67.69 (22.81)	OP>NP
Correct verb forms	6.28	.006*	56.83 (22.89)	70.46 (22.54)	79.16 (18.85)	OP>NP
Correct articles	.03	.926	71.24 (18.11)	73.32 (19.44)	72.6 (28.59)	

As suggested by the significant level of interaction between planning and proficiency in accuracy (4.2.4.1), the different results were identified in the three accuracy measures between the high and low proficiency groups. Neither strategic nor on-line planning had positive effects on any accuracy measures in the high

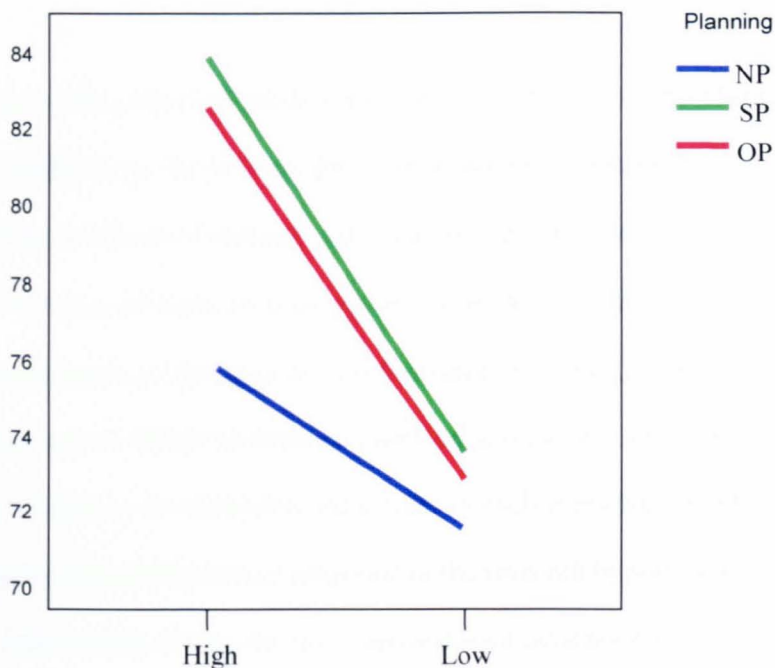
proficiency group, while on-line planning had positive effects on accuracy ('error-free clauses' and 'correct verb forms') in the low proficiency group.

A close examination clarifies the differential effects of planning in different linguistic items, and between the high and low proficiency groups. The results of correct verb forms are consistent with those of error-free clauses (Figure 4.1). That is, there was no significant difference between the three planning conditions in the high proficiency group, while there was a significant level of improvement in on-line planning and some relative improvement in strategic planning in the low proficiency group. In contrast to those, the result of correct articles shows some relative improvement in strategic planning and on-line planning in the high proficiency group, while there was no improvement in the low proficiency group, although none of them reached the level of significance (Figure 4.2).

**Figure 4.1:** Percentage of Correct Verb Forms



**Figure 4.2:** Percentage of Correct Articles



To summarize, in the high proficiency group, both strategic and on-line planning had the effects on complexity, but did not influence the accurate production. Also, similar to the overall results, on-line planning provided the negative effects on fluency, but strategic planning did not influence the fluency aspect. On the other hand, strategic planning positively influenced fluency and on-line planning led to the significant gain of accuracy in the low proficiency group.

Looking back to the research hypotheses concerning the relationship between planning and proficiency (3.3), both **Hypothesis 4** ('the effects of SP will interact with the different proficiency levels') and **Hypothesis 5** ('the effects of OP will interact with the different proficiency levels') are accepted, because a number of different responses between the two proficiency levels were identified. These hypotheses will be discussed in 4.3.1 below.

## 4.3 Implications of Task Performance Analysis

Following the results of the task performance analysis, the rest of this chapter will consider the implications drawn from the various statistical outcomes. The results suggested different effects of strategic planning and on-line planning on L2 oral performance between different proficiency levels. In this discussion, firstly I mention the interaction between proficiency and performance. Secondly, as the difference between form-focused effects in strategic planning and those in on-line planning is not very clear, I discuss the form-focused effects of each condition by referring to specific extracts. Beyond the issues proposed in the research hypotheses, finally I consider the relationship among the three performance dimensions.

### *4.3.1 Interaction between Proficiency and Performance*

The first matter to consider is the interaction between proficiency and performance, drawn from the results that multiple ANOVAs showed different effects of strategic and on-line planning on task performance between the high and low proficiency groups (4.2.4.3). The following section particularly focuses on the relationship between fluency/accuracy and proficiency.

#### 4.3.1.1 INTERACTION BETWEEN FLUENCY AND PROFICIENCY

Looking back to the results of fluency in the two groups, there were more positive effects of strategic planning on fluency in the low proficiency group (pruned speech rate, mid-clause pauses and self-repairs) than in the high proficiency group (self-repairs). This result is somewhat inconsistent with Wigglesworth's (1997) study

that showed that strategic planning was more effective in high proficiency speakers (2.3.4). It could be hypothesized that higher proficiency speakers are better able to use the strategic planning time efficiently.

One possible reason for this is that the monologic story-telling task type might reduce the effects of fluency improvement in high proficiency speakers. With reference to this, Ortega (2005) argues that the existence of an interlocutor has an effect on learners' production. In addition, Elder and Iwashita (2005:232) suggested that '[t]he absence of a live interlocutor, may have some bearing on the nature of performance reducing candidates motivation to 'lift their game' or to strive for enhanced performance even when opportunities to do so are provided'. It is possible that the participants did not attempt to increase the speech rate, being released from the communication pressure usually caused by interlocutors. This absence of interlocutor effect can be plausible for high proficiency speakers, but this does not fully explain why strategic planning was still beneficial to low proficiency speakers.

Another possible reason is that the high proficiency speakers might have shifted into a more careful, form-focused speech mode in strategic planning, while the low proficiency speakers did not do so for the absence of this ability. In other words, the former group of speakers was more likely to be equipped with a multiple speech mode than the latter group of speakers. Evaluating the given condition (i.e., planning time), the high proficiency speakers might have deliberately selected a 'careful' mode even in strategic planning, possibly because they interpreted that this condition was more 'formal'. Thus, it is necessary to examine how the high proficiency speakers responded to the strategic planning condition and any conscious mode shifting in the following post-task interviews. In contrast, it is conceivable that the low proficiency speakers are not developmentally ready to use such a multiple system. As a result, they might simply have engaged in improving fluency in strategic planning, ignoring



an aspect of linguistic form. This sort of dual speech mode has been widely discussed in the field of SLA, for example, in the interests of planned-unplanned discourse by Ochs (1979) and exemplar-ruled-based mode by Skehan, (1995, 1998; see 1.6). Because these are beyond the level of speculation obtained from the statistical outcomes, it is necessary to show the evidence from the verbal reporting to support this conjecture. This will be discussed in detail in Chapter Five.

#### 4.3.1.2 INTERACTION BETWEEN ACCURACY AND PROFICIENCY

The results of the task performance analysis also revealed the interaction between accuracy and proficiency. Different results of accuracy measures (i.e., correct verb forms and articles) may suggest different developmental stages in high and low proficiency learners, because English verb formation seems more learnable than the article system for Japanese learners of English.

The correctness of verb forms did not greatly vary in different planning conditions in the high proficiency group, while on-line planning led to a higher degree of correct verb forms in the low proficiency group. It can be assumed that high proficiency speakers have already acquired verb formations to a large extent and reached a ceiling of accuracy, being less influenced by planning. On the other hand, planning opportunities helped low proficiency speakers substantially, because they had not fully integrated much of this into their IL system. In contrast, strategic and on-line planning led to a higher degree of correct use of articles in the high proficiency group, while no difference over the three planning conditions was identified in the low proficiency group. Applying the same logic, high proficiency speakers might have approached the stage of article acquisition. On the other hand, planning did not have much influence on the low proficiency speakers, because they

had not reached the level to allow them to pay attention to the use of articles in speech production. However, although the mean scores of article use showed a clear distinction between the high and low proficiency groups (Figure 4.2), we should take these speculations with caution, because the difference did not reach significance.

#### *4.3.2 Different Form-Focused Effects of Strategic and On-Line Planning*

As mentioned before, strategic planning is, in theory, considered to push the IL to its limits, finally mobilizing stabilized language. The statistical results of task performance partially support this hypothesis, because it shows that speakers tended to upgrade their language in terms of complexity and accuracy in both strategic and on-line planning. Most importantly, the findings of the performance analysis indicate that strategic and on-line planning do not give equal, but markedly different, form-focused effects on the productions between the two proficiency levels. That is, strategic planning tended to lead to more complex language than non-planning (and also contributed to fluency improvement to some extent), while on-line planning brought on more accurate as well as complex language, which goes beyond the findings by Yuan and Ellis (2003) of the equivalent effects of complexity in strategic and on-line planning. Building on the statistical findings, this section looks at these different productions brought about by the three planning conditions by referring to the transcripts of the performance.

In order to illustrate the different form-focused effects, the extracts of production about the last two scenes (fifth and sixth scenes) of the ‘Waiting for a Bus’ task (Figure 4.3) is considered, because these include several complications, in which a lack of linguistic resources tends to force speakers to shift from the more smooth and comfortable lexical processing mode into syntactic processing. As a result of this,

speakers were more likely to suffer from the difficulty of maintaining the quality of performance and balancing fluency, complexity and accuracy. Despite an effort to describe the same pictures, this approach thus may illuminate the different qualities of performance made under different planning conditions by different proficiency levels.

**Figure 4.3:** ‘Waiting for a Bus’ Task



#### 4.3.2.1 PRIORITIZING FLUENCY IN NON-PLANNING

One of the clear tendencies in the non-planning condition is that several speakers (probably deliberately) made the story brief, which can be seen in both high and low proficiency groups:

**Extract 4.1:**

and (.4) on\* on the way to // (.2) their destination (.3) um they (1.1) they sh-  
watch (.7) they watched bus // um is (.8) clash (.4) clashing (.4) clashed so  
 (1.5) and er yeah that's all (LOW: 1NP8-10\*\*).

*\*Dysfluent markers and inaccurate utterances are underlined in the extracts.*

*\*\*All the extracts include a unique number at the end. This consists of proficiency group (High/Low), participants' ID number (1-27), planning condition (NP/SP/OP) and lines in the data (e.g., 2-3).*

**Extract 4.2:**

and then they noticed:: (.2) that um on the way (1.0) the:: (.8) bus (.7) the first bus (.6) like (.2) um (.3) they had some accident and (1.8) um the bus didn't work (.6) so (1.0) I think they are lucky [laugh] (1.4) (HIGH: 24NP10-14)

The number of dysfluent and hesitation markers in the first extract would imply that the speaker found it difficult to describe the scenes, finally abandoning the details. In particular, the speaker's search for an appropriate verb form in the last part ('bus um is clash... clashing... clashed') sounds unnatural, suggesting that the speaker seems to have lost control in making well-formed structures. This brevity and avoidance of telling the complex storyline can be seen even in the high proficiency speaker in Extract 4.2. Relatively frequent mid-clause pauses may suggest some hesitations in the second extract as well, but this speaker was certainly more able to cope with the predicament and thus avoided a serious breakdown by using fillers. Nevertheless, the production in Extract 4.2 also did not sufficiently describe the scenes, in particular the point of the story that the second bus finally passed by the first bus, and he finally finished up the explanation by saying 'I think they are lucky'. This final utterance may successfully function as a signal to conclude his story-telling, but did not provide the listener with sufficient information to understand the narrative satisfactorily.

In both examples, fluency seems prioritized over complexity and accuracy in non-planning. Confronted with the difficulty of describing in L2 and being unable to find appropriate language in the resources, there are usually two possibilities for coping with the difficulty: either to simplify the production, which is originally more

complex in the level of conceptualization; or to venture to formulate the language to match the conceptualized message as accurately as possible. The first choice is usually safer in terms of avoiding disrupting the flow of communication, but a significant drawback is that the learner stays in a comfortable state without being engaged in crossing the existing boundary of their current linguistic competence. An alternative way to challenge the uncertainty inevitably bears a risk of downgrading fluency and ultimately leading to communication breakdown but seems beneficial in allowing the chance to develop one's IL in the long-term. Particularly in learning situations, the second choice may be preferable, but in the two examples, the pressured situation made it more difficult for them to choose the more ambitious way. Interestingly, the different outcomes between the two speakers are that the speech of the high proficiency speaker was more successful by means of the 'safety-first' strategy in terms of maintaining the fluency, while that of the low proficiency speaker seems to have lost complete control of his performance. That is, the second speaker is more strategically as well as linguistically proficient in dealing with difficult situations and maintaining the quality of communication.

#### 4.3.2.2 ACHIEVING BALANCED PERFORMANCE IN STRATEGIC PLANNING

In the strategic planning condition, the improved performance of fluency and/or complexity was observed in both high and low proficiency speakers. In Extract 4.3, the low proficiency speaker made her production relatively smooth and accurate by relying on rather simple structures.

**Extract 4.3:**

on:: their way to:: the destination (.5) they found (.5) the bus (.4) the bus was (.6) um that was (.5) the very (.7) bus (1.5) he they (.2) couldn't (.6) get (.2)

on (.9) and they passed through (.3) the bus (.8) and (1.1) get to their destination [LOW:2SP17-21]

Frequent occurrences of unfilled pauses are suggestive of the speaker's caution in explaining the complicated scenes, but the speech completed the intended message without a serious breakdown. In contrast to many other speakers using relative clauses to describe the scene that the three boys' bus finally overtook the missed bus, this speech described the pictures well by chaining simple clauses. This simplified production within the comfortable state of fluency might allow the speaker to pay attention to accuracy. In this sense, this utterance seems to prioritize fluency and accuracy over complexity. It is interesting to note that even this low proficiency speaker attained the two performance dimensions (i.e., fluency and accuracy) at the expense of complexity in strategic planning, in contrast to the non-planned extract where none of the three dimensions were maintained (Extract 4.1).

Compared to the above productions, the strategically planned speech by a high proficiency speaker (Extract 4.4) includes more detailed information about the scenes such as the driver's fixing the flat tire and the boys' striped shirt, but possibly because of using only simple clause structures, the speaker did not describe the point that the second bus finally passed by the first bus, although this point may be speculated by the words that 'the stripe [sic] shirt guy looks very sad'.

**Extract 4.4:**

and (.5) the... um:: they saw (.8) the previous (.2) bus (.3) um:: (1.3) looked (.2) um (.5) er sorry has a (.8) problem (1.6) because um that driver (.3) um get outside (1.0) and (.3) fix the (.2) the (.2) tire (1.9) and also they saw the the:: the strip shirt guy (.3) looks very sad (.6) so (.3) er they um shouted (.5) to him (.7) "Good luck!" (.4) and then (.5) they got home (.5) safely [HIGH: 20SP21-29]

In addition, a number of fillers and unfilled pauses would suggest that the speaker

found it difficult to tell this part of the story, which also deprived her attention to accuracy of the forms. From her abandoning the initial utterance ('the previous bus um looked um er sorry...'), the speaker seems to have struggled in either recalling the scenes or formulating the intended message. Simply chained clauses and many inaccurate verb forms may imply that the speaker tended to prioritize meaning over form. That is, possibly because of an attempt to convey as much information as she remembered, little attention was left to the form aspect of her performance, finally downgrading the level of complexity and accuracy.

A rather successful case of describing the scenes can be found in Extract 4.5, in which the speaker completed the story with a limited number of dysfluent markers and inaccurate utterances:

**Extract 4.5:**

but during the journey (.3) bus number tw- si- twenty-six (.5) got a flat tire (.9) and finally this three boys are lucky because bus number thirty-three (.5) just (.5) could go forward (.5) than bus number six- twenty-six [laugh]  
[HIGH: 22SP10-12]

Although it does not give detailed information about the scenes, this utterance essentially includes the critical point that the three boys' bus overtook the broken-down bus. Because this is speech in strategic planning, this brevity of the conveyed production can be considered planned, rather than what the speaker improvised and produced on the spot. This planned brevity is different from the speech of Extract 4.2, which seems to have simplified the production on-line, meeting with the difficulty, in order to avoid disrupting the level of fluency. Although there are some dysfluent and inaccurate utterances and rather simple structures employed, this planned speech by the high proficiency speaker achieved a better balance of performance by maintaining the level of fluency, complexity and accuracy, than any

other performances including unplanned speech by the high proficiency speakers and planned and unplanned speech by the low proficiency speakers.

It is worth pointing out that, of the three performance dimensions, the production of Extract 4.5 does not seem to include very complex structures. The problem is that it is not evident, from the performance alone, whether the speaker deliberately chose these structures instead of more challenging, complex ones or, regardless of her resources involving more complex structures, whether her intended message only needed the structures that were actually produced. It may be necessary to take 'complexity' into consideration separately from the other dimensions, because, as the learners become more proficient, the occurrences of inaccuracies and dysfluencies would be on the decrease, while achieving the high syntactic and discursal complexity does not necessarily guarantee one's proficiency (Personal communication with Lourdes Ortega, September, 2005).

#### 4.3.2.3 CHALLENGING THE EXISTING LEVEL OF INTERLANGUAGE IN ON-LINE PLANNING

In contrast to the tendencies within non-planning and strategic planning, on-line planning production is characterized by speakers' clear attempts to describe the details of the scene and to use more challenging structures at the expense of fluency. In Extract 4.6, a number of fillers, unfilled pauses, repetitions and self-corrections clearly show that the speaker is struggling to formulate what he intends to say. A relatively small number of inaccurate verb forms may suggest that the speaker is attempting to pay attention to accuracy while trying out complex production:

**Extract 4.6:**

and then (.7) er (2.0) they found that (.8) the:: the bus (.4) er (1.3) the bus (3.3) the b- er (.4) the bus (1.2) boys (.4) who:: interrupt (.3) them (.8) ride



(.5) was (.2) there is because (.2) er (2.8) some er (5.3) so- er the bus (.4) had some trouble (.5) and (2.1) it was (1.5) it (1.6) er (1.2) it needed (.4) to repair (.8) so:: (1.0) the bus (.9) stopped (.4) and (.7) er three boy- (.2) er the bus (.3) whi- (.5) which (.4) the three boys (.7) rode (4.0) went (1.0) er faster than (1.0) that bus (.7) so (.8) they were sas- (.2) they they were (.2) satisfied with (.3) that (.4) resul- reselt (.2) result (1.8) (**LOW:3OP9-18**)

Similar tendencies can also be observed in a high proficiency speaker. Extract 4.7 includes a lot of unfilled pauses, repetitions and self-corrections as a result of an attempt to produce more complex language beyond the speaker's current linguistic competence.

**Extract 4.7:**

and (.8) from on (1.4) for a little (.5) whi- (1.0) after a little while (1.0) they (.5) from:: the bus (.5) the three boys (.3) saw the (1.6) another (.3) bus (.3) which departed (.8) before (.6) earlier (2.6) what (.3) they (.2) what they (.3) saw was (.2) the (.6) bus (1.0) had a problem the bus (.3) the (.9) they (1.0) try to get on (.6) had a problem with (.2) the (1.0) tire (.4) wheel (1.4) and (4.5) and the (.9) the (.2) boy (.2) the (.4) bus (1.4) with three (.2) boys (2.4) er (.2) went (.5) faster (.8) quicker than the bus (.5) the first bus (2.8) (**HIGH: 21OP23-31**)

This performance also includes only a small number of inaccurate verb forms, which may suggest the speaker's simultaneous attention to accuracy and complexity. Ample appearance of dysfluent features might be the by-product of this concurrent attention to two dimensions, so that few attentional resources were left to the maintenance of fluency.

In rather simple scenes which do not produce much cognitive burden, the three different conditions may not bring out these different qualities of production so sharply, but the need to relate complex situations is more likely to clarify the

differential effects of strategic and on-line planning. In the non-planning condition, the speakers tended to prioritize fluency over complexity by avoiding the descriptions of complex pictures. In particular, the high proficiency speaker coped with this predicament more strategically, probably by identifying the difficulty beforehand and simplifying the production (Extract 4.2). As a result of using simple structures, the speaker could also comfortably pay attention to accuracy. In contrast, the low proficiency speaker also took the safety-first strategy by avoiding the complex description but still experienced hardships in maintaining the level of fluency as well as accuracy (Extract 4.1). In terms of balanced performance, strategic planning proved effective for both proficiency levels. Although it was more successful in maintaining the balance of fluency, complexity and accuracy in the high proficiency speaker, it also seems to have been beneficial for low proficiency speakers in enhancing their performance.

Despite such clear improvements in performance, it is less clear whether strategic planning certainly challenged learners to go beyond their current level. As suggested by Skehan and Foster (1997:206), it is assumed that ‘the need to find language to express the complex relationships seemed to push learners to complex language which consumed a considerable proportion of available processing resources leaving little space capacity to attend to accuracy’, but the typical strategy taken in non-planning and strategic planning was to prioritize fluency, more or less abandoning the form aspects under the pressured conditions. That is, a significant degree of on-line computation (Skehan & Foster, 1997:206) seemed involved. It is on-line planning that pushed these speakers into using more complex structures while maintaining the level of accuracy. As a consequence of the on-line planning instruction, a marked number of dysfluent features emerged, but the speakers’ efforts to overcome the existing limitation were clearly different from ones observed in

strategic planning.

It seems obvious at first sight that an elimination of communication pressures made it possible to attempt to break through the current limitation; however, looking at this point from another angle, it may also imply that these on-line processing pressures tend to preclude L2 speakers from overcoming the present limitation in their everyday communication. Real world language use does not usually allow on-line planning engagement in the same way as the present research does, but the transcription analysis suggests that pressured situations noticeably reduced the possibility of venturing to upgrade the complexity even in high proficiency speakers. If the speaker stops at the current limited resources and only fosters the safety-first strategy by risk avoidance, their language is likely to fall into a stabilized condition (Selinker & Han, 2001).

The analysis of transcription as well as the statistical outcomes suggests a number of linguistic features brought about by strategic and on-line planning and possible mental conditions of the speakers to produce such performance. However, it should be stressed that these are only possibilities speculated from the linguistic/paralinguistic features. The different levels of accuracy and complexity improvement imply that different linguistic processing (i.e., the Formulator) operates between the different planning conditions, and between different proficiency levels. What sort of processing is brought about by strategic planning and on-line planning, and how these implementations specifically facilitate linguistic processing, will be discussed further when analyzing the verbal reports in Chapter Five.

### *4.3.3 Trading-off between Fluency, Complexity and Accuracy*

One of the general findings from a series of statistical models is the complex

interaction between the three planning conditions and the two proficiency levels, implying that the relationship of planning conditions and IL variables are not linear but that the variation seems systematic. Therefore, the chief objective in this section is to disentangle this complex relationship between various factors by paying attention to the three main points, and then to explore the issues further by considering the following verbal protocol analysis.

As mentioned in Chapter Two, previous studies have shown the trade-off effects between complexity and accuracy (e.g., Foster & Skehan, 1996, 1999) and fluency and accuracy (e.g., Wendel, 1997; Yuan & Ellis, 2003). It has been argued that the three performance dimensions are theoretically independent (e.g., Skehan, 1998; Skehan & Foster, 1997, 2001), but, as suggested by these trade-off effects, some aspects of the three variables seem closely interrelated rather than being completely independent of each other. The present study also shows several trade-offs in similar but slightly different ways from those of the previous studies. Thus, the following will touch upon the trade-offs between (1) fluency and accuracy, (2) fluency and complexity, and (3) complexity and accuracy.

#### 4.3.3.1 FLUENCY AND ACCURACY

Looking at the fluency-accuracy results of non-planning and on-line planning, the former was more fluent but inaccurate, while the latter was less fluent but more accurate in both proficiency groups (Table 4.7). Thus, it can be assumed that on-line planning gained greater accuracy at the expense of fluency, suggesting the trade-off between fluency and accuracy. This result is essential in supporting the on-line planning proposal, because the original conceptualization of this came from the limited strategic planning effects on accuracy, involving the accuracy-fluency trade

off (Wendel, 1997). From the result that showed a lack of accuracy in strategic planning in his research, Wendel (1997:146-7) conjectured that there is a competition between fluency and accuracy rather than between complexity and accuracy:

Having had time to plan, speakers come to the production part of the task with a complex representation of the narrative already in mind, but not with the details of the grammar worked out. Therefore, any attention that speakers give over to grammatical accuracy will have to be done on-line, attention to grammatical consideration on-line will cost speakers in terms of time. The result will be less fluent production.

The findings of the present study seem to support this speculation, because only on-line planning reached the significant level of accuracy by downgrading the level of fluency. Interestingly, strategic planning did not reach significance but the results of descriptive statistics (Table 4.1) show its relative effects on accuracy, accounting for the mixed results of rather weak strategic planning effects on accuracy in previous studies (2.3.2). Drawing on the limited WM capacity in the information-processing model, strategic planning could help speakers free up attentional resources especially in terms of planning contents of speech, but this function is limited in terms of accuracy. On the other hand, on-line planning could more greatly contribute to increasing the level of accuracy by more directly influencing the resources 'online'. As indicated by Yuan and Ellis (2003:21), this result may be particularly right in the case of EFL (especially East Asian) learners who have ample explicit knowledge through formal grammar instruction. That is, they are able to use the knowledge to formulate speech plans and to monitor their production to the fullest extent by comfortably accessing the explicit knowledge under the unpressured condition.

#### 4.3.3.2 FLUENCY AND COMPLEXITY

In addition to the relationship between fluency and accuracy, strategic and on-line planning made participants use a greater number of grammatical integration devices than non-planning. According to Pawley and Syder (1983), the use of such devices is detrimental to fluency, which would partly explain a lack of fluency in strategic and on-line planning in the present research (i.e., the more grammatical integration devices are used, the slower the speech becomes). That is, making a weak clause integration (i.e., 'chaining integration') is the most efficient in terms of speech planning strategy, whereas making a strong clause integration (i.e., 'grammatical integration') is the least efficient (Pawley & Syder, 2000:177). Recalling the results of factor analysis (4.2.2), there was some relative correlation between complexity and fluency measures. For example, the number of self-repairs was originally categorized in the fluency dimension, but the factor analysis shows its correlation both to fluency and complexity in non-planning and on-line planning, whereas it is correlated only to complexity in strategic planning. The fluency-complexity trade-off may be worth mentioning because this trade-off has not been well considered in the task planning literature. Ejzenberg's (2000) study did not show any quantitative difference of grammatical integration devices between fluent and nonfluent speech, but reached the conclusion, through the qualitative analysis, that fluent speakers were better able to successfully switch back and forth between chaining and grammatical integration. The trade-off effects of fluency and complexity being combined with those between fluency and accuracy (4.3.3.1), it is more appropriate to say that a trade-off between fluency and *form* was identified in the performance (see Figure 3.2).

#### 4.3.3.3 COMPLEXITY AND ACCURACY

Inconsistent with many other planning studies, no clear evidence of complexity and accuracy trade-off was identified in the results of task performance analysis. It is noticeable that the complexity factor in many other studies concerns *syntactic* complexity (3.8.6), not *discoursal* complexity (3.8.7), which was included in the present research. Syntactic complexity (i.e., clauses per AS-unit) may be more evidently competing with accuracy because, if speakers are challenged to use more complex structures, only a little attention is left to maintaining accuracy. However, it should be stressed that the present study includes use of grammatical integration devices, which involves an element of syntactic complexity, because both measures are identified by occurrences of subordinate conjunctions (e.g., ‘because’ and ‘when’) and relative linking devices (e.g., ‘which’ and ‘who’).

A possible explanation for a failure to show significant gains of syntactic complexity is its inclusion of relatively simple clause conjunctions. For example, the following utterance is identified as two clauses per AS-unit:

but they found that // in the box there was only snake (3NP7-8:LOW)

From an information-processing perspective, this sort of utterance seems not to place a heavy burden on the speaker’s WM because the integration of the two clauses can be regarded as a weak conjunction. Following the ‘one-clause-at-a-time constraint’ by Pawley and Syder (2000), grammatical integration tends to place much more burden on linguistic processing because speakers need to adjust language beyond clause-boundaries to make connections between clauses. All items concerning grammatical integration devices (i.e., ‘subordinate conjunctions’, ‘relative linking

devices’, and ‘restrictive linking devices’) seem psycholinguistically challenging because users manipulate the language beyond the clause boundaries by accessing grammatical resources. In other words, they need to go beyond the one-clause-at-a-time constraints in making production by using grammatical integration devices. Looking back to the overall results of planning effects (Table 4.7), strategic planning only influenced complexity while on-line planning gave impact on both complexity and accuracy. These results do not directly show the evidence of the trade-off effect of complexity and accuracy, but neither do they reject the possibility that participants prioritized complexity over accuracy in strategic planning. The results of the positive effects of on-line planning on both complexity and accuracy can be explained by the fact that the on-line planning condition might go beyond the level of competition between complexity and accuracy by giving ample attentional resources to on-line formulation and monitoring.

#### **4.4 Conclusion**

This chapter described the results of statistical analyses and also revealed the implications drawn from these findings and transcriptions. It is probably right to say that, despite a lot of minor differences, the findings of the task performance analysis are, on the whole, consistent with those of the previous studies on strategic planning and on-line planning; that is, the former had positive effects on fluency (only in the low proficiency group, though) and complexity, and the latter had positive effects on complexity and accuracy (but not in the high proficiency group).

A number of different effects were observed, but summarizing these individual results would tell us that different planning conditions cause different processing in speech production, and the extents of these differences vary between the



two proficiency groups. The analysis of task performance is useful for illuminating these speculations, but it does not tell us, more precisely, how such different mental operations occur when the performance is underway, nor how different they are between the two proficiency groups. In order to find more direct evidence of such distinctive speech processing, it would be helpful to examine the qualitative data collected from the verbal reporting.

The analysis of verbal reports also aims to explore different processing between high and low proficiency speakers. From the findings of the performance analysis, it seems that different proficiency speakers find different planning conditions more conducive to paying attention to form. That is, for high proficiency speakers, strategic planning may be sufficient to focus on form comfortably, while on-line planning may be more effective for low proficiency speakers to increase the level of complexity and accuracy. One of the main differences between the two types of planning seems to lie in the degree of attention required. Other factors being equal, the more proficient learners become, the less attention they need for language production. Therefore, on-line planning seems more effective for low proficiency speakers because it facilitates the monitoring function by releasing the on-line processing pressure on WM. Therefore, the main purpose of the next chapter is to seek evidence for such mental conditions from the qualitative data and triangulate this to the findings of the task performance analysis.

# CHAPTER FIVE: VERBAL REPORT ANALYSIS

## 5.1 Introduction

Following the task performance analysis in Chapter Four, I shall now present the findings obtained from the retrospective interviews, which were conducted immediately after task completion. The previous chapter revealed the general trends towards positive effects on complexity in strategic planning and those on complexity and accuracy in on-line planning, but the results were not straightforward between different proficiencies and different grammatical items (i.e., verb formation and use of article). There was also a trend observed towards relative focus-on-form effects in on-line planning over strategic planning in the mean scores, although the differences did not reach significance. For the further exploration of focus-on-form effects of planning and also the possible differences between strategic planning and on-line planning, this chapter considers the analysis of verbal reports with respect to learners' attention to form aspects and other cognitive processing underlying their oral performance.

This chapter is also important in terms of investigating the planning effects in task-based performance by using a qualitative research instrument, which has been used only in a limited number of studies. One example of this sort of inquiry in planning literature is given by Ortega (1999, 2005) investigating learners' planning strategies. However, her analysis was mainly of learners' reported strategies employed 'during planning time' rather than 'during production'. Ortega's study also showed evidence of learners' focus-on-form during planning time, but it is less clear whether such focus-on-form occurrences during planning would transfer to the following production. Therefore, the main focus of the present chapter is on learners' cognitive

processes on-line rather than those off-line, or strategic planning time (but see 5.4.2.1 for the summary of strategies during planning time). I attempt to explore the distinctive characteristics of their cognitive processes in speech production affected by different planning conditions, particularly in focus-on-form accounts.

In the following section, first I discuss the methodological issues relating to the research questions, interview procedure and data analysis method, which were briefly mentioned in Chapter Three. Having looked at these methodological issues, I present the findings of the verbal report analysis from two perspectives; summarizing the frequency and number of themes reported overall in my data; and characterizing each planning condition as fully as possible by referring to specific verbal reports.

## **5.2 Methodology**

### *5.2.1 Research Questions for the Analysis of Verbal Report*

Based on the second research objective (3.2), I conducted retrospective interviews in order to explore learners' mental processes in the different types of planning and to identify any distinctions between different proficiency levels. The analysis of task performance exhibited the general picture that strategic planning improved L2 in terms of complexity and fluency to a certain extent but not in terms of accuracy; on the other hand, on-line planning allowed speakers to produce more accurate and complex language at the expense of fluency. Understanding the different performance characteristics influenced by the three planning conditions, I raised a number of specific questions for underlying operations in the analysis of verbal report:

1. In NP, did participants pay more attention to meaning than form? What cognitive operations did they more specifically engage in?

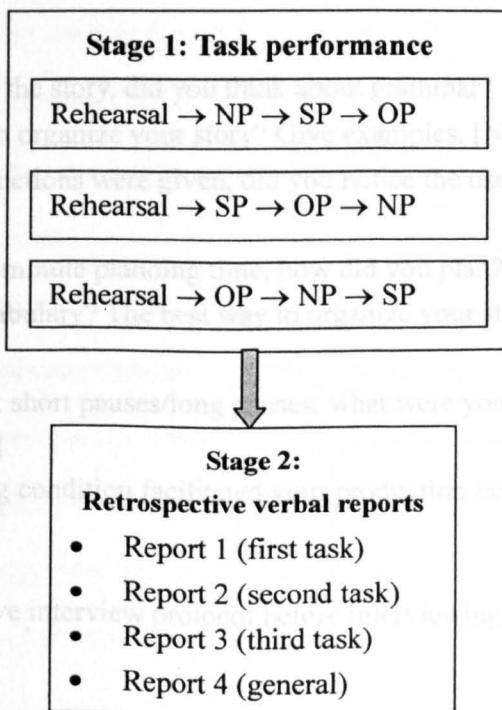
2. During ten-minute planning time in SP, did participants engage in focus-on-form operations prior to the production? What planning strategies did they employ (following Ortega [2005])?
3. In SP, did they pay attention to form more frequently than in NP production? What cognitive operations did they more specifically engage in?
4. In OP, did they pay attention to form more frequently than in NP and SP production? What cognitive operations did they more specifically engage in?
5. Are there any distinctive differences of cognitive processes between high and low proficiency speakers in any planning conditions?

This chapter aims to answer these questions.

### 5.2.2 Interview Procedure

For these purposes, verbal reporting about participants' thoughts during each planned production was elicited. The research design is reproduced here for the reader's convenience:

**Figure 5.1:** Research Design



The whole interview session was conducted in their L1 (i.e., Japanese) in order to minimize the danger that linguistic difficulty would block participants' verbalization of complex thought processes. Reporting of individual tasks was firstly elicited through stimulating their memory by listening to the recorded audio data, which was in the order of task (Report 1-3). Following each of their task performances, the opportunity for a general report (Report 4) was also provided. This elicited comments on issues not restricted to individual tasks but more general evaluations of their own performance and the conditions such as comparison of different planning conditions, task difficulty and learning and speaking styles.

### *5.2.3 Semi-Structured Interview Questions*

In the interviews, a number of questions were asked to investigate the cognitive processes in the different types of planning and also to explore the pedagogical benefits of planning implementation. Examples are as follows:

1. When you told the story, did you think about grammar? Vocabulary? Content? The best way to organize your story? Give examples. [NP/SP/OP]
2. When the instructions were given, did you notice the use of the past tense? [NP/SP/OP]
3. During the ten-minute planning time, how did you plan? Did you think about grammar? Vocabulary? The best way to organize your story? Give examples. [SP]
4. When you took short pauses/long pauses, what were you thinking? Give examples. [OP]
5. Which planning condition facilitated your production best?

I developed the above interview protocol before interviewing the first participant but

modified it slightly during the actual interview sessions; some questions (e.g., 1-3 & 5) were prepared from the beginning of the research while others (e.g., 4) were added in the middle of the data collection process.

#### *5.2.4 Data Analysis Method*

As there are only a limited number of studies investigating learners' reports on planned speech, I took up an inductive approach that allows theory to emerge from data (Richards, 2003). To analyze the verbal report data, I adopted one of the techniques of grounded theory (Glaser & Strauss, 1967), 'the constant comparative method'. This allows categories to emerge from the data, rather than how the researcher imposes preconceived categories on the data, and to compare each instance of a category with all other instances coded in the same category. That is, the discovery of relationship begins with the analysis of initial observations, which undergoes continuous refinement throughout the analysis process, and continuously feeds back into the process of category coding (LeCompte & Preissle, 1993). This comparison allows the researcher to refine the categories and to identify properties associated with the category (Lockhart & Ng, 1995:614). In doing so, the present study takes a qualitative method in which no categories were pre-established; rather they were data-dependent.

For the analysis, first I segmented individual episodes in the transcribed data and identified 503 episodes in total. In the present analysis, an 'episode' is defined as any segment of the protocol in which a participant reported about his/her cognitive thoughts concerning L2 production. I coded each episode and generated 25 themes, which were then categorized under the macro-themes of speech processing based on Levelt's (1989) speech production model (i.e., conceptualizing processes, lexical

operations, formulating processes and monitoring processes; see Chapter One). In these macro-categories, 'conceptualizing processes' are only relevant to meaning aspects, excluding any episodes related to form. 'Lexical operations' are limited to episodes reporting single word and lexical phrase levels, while 'formulating processes' are concerned with clause/sentence-level structuring.

After generating themes, I examined the reported data again and assigned one of the themes to each episode and repeated the same procedure several times to refine the analysis. To check the appropriateness of my analysis, five percent of all the episodes were analyzed by a second coder (who was also the first rater of the proficiency judgment; see 3.4.3), showing 74 percent agreement. In a similar analysis, Ortega (2005) used inter-coder reliability (following O'Malley & Chamot, 1990) and obtained 79 percent agreement. For inconsistent cases, she decided to use all her codings, because she considered them to be more trustworthy, given her intimate familiarity with the coding scheme and the data (Ortega, 2005:83). A possible reason for inconsistent analyses is that, as the original data were in Japanese and the second coder only looked at the translated episodes, different coding might have come from the accuracy of the translation. Following Ortega's decision, all results are based on my coding, not only because it is evident that this sort of coding is seldom very reliable, but also because I am confident from my frequent revision of the data and access to the original data in Japanese. For this difficulty of the reliable analysis of qualitative data, I used the sums of themes only for illustrating the general trend (5.3), and more weight will be on the analysis of specific episodes (5.4).

### *5.2.5 Limitations*

Before reporting the results of the verbal report analysis, it is necessary to mention

several limitations of the method. Firstly, as the present research takes a 'retrospective' (not 'concurrent') reporting stance, time lags between learners' thoughts and the reports were inevitable. I believe that stimulated recall compensated for this problem to some extent, but I should take great care in the interpretation of the learners' reports, in assessing whether or not they reported the accurate processes. Secondly, there may be a possibility of probing questions. I attempted to avoid guiding learners' answers in the whole interview sessions, but it was also very difficult to eliminate my own intentions completely. Thirdly, although the second coder checked parts of my analysis with 74 percent agreement, the reliability of the analysis might not be sufficient. That is, the analysis by other researchers might lead to more or less different results. Finally, as the interviews were conducted in Japanese, my translation might affect the interpretations of reported data. Although I carefully designed the retrospective interviews and analyzed the collected data as objectively as possible, all of these limitations may apply. The following section will report the number of themes in each planning condition, but this only aims to suggest a general tendency of the collected data.

### **5.3 General Overview of Verbal Reporting**

The aim of this analysis is to understand how L2 speakers' cognitive processes are influenced by the different planning conditions in a qualitative manner. The findings will provide an overview of the themes to see how common they are in the data. Categorization of themes is not mutually exclusive as there is some overlapping across themes. However, each episode was included in one theme only that which was most dominant in the present analysis. Table 5.1 shows the theme labels, their functions, the total number of occurrences of reporting and the percentages in the



three different types of planning. Examples of each theme are illustrated in Appendix 5-A. In the table, bracketed numbers are reported themes of high (left) and low (right) proficiency participants. The examples of each report are illustrated in Appendix 5-A.

Among the total numbers of themes, those in on-line planning occupy more than a half of all identified themes (54.9 %), probably because participants found it easier to reflect on their own production processes by speaking slowly and taking pauses more frequently and longer than in usual speech. This may suggest that the on-line planning implementation is useful not only for research purposes – to investigate on-line L2 processing, which is difficult to tap cognitive processes in the rapid flow of production – but also for pedagogical purposes to encourage learners to reflect on their own production and to encourage them to notice their own linguistic features.

A number of themes were found in all the planning conditions and many of them were distributed fairly similarly (e.g., recalling the picture and self-correction), suggesting that the different planning conditions did not greatly influence those processes; in other words, these processes occurred in every condition. Therefore, I will pay particular attention to the themes showing distinction between the planning types. The following sections summarize the tendencies in each macro-category (i.e., conceptualizing, lexical, formulating and monitoring).

**Table 5.1:** Summary of Themes Generated from the Retrospective Interviews

Themes: Functions	NP	SP	OP
<b>A. Conceptualizing processes</b>	30	25	67
A1. Task evaluation: To evaluate the difficulty of the given task	2 (1/1)		3 (2/1)
A2. Message generation: To generate the message	6 (4/2)	8 (6/2)	24 (9/15)
A3. Message regeneration: To regenerate the message without changing the meaning	3 (0/3)		11 (3/8)
A4. Recalling: 1. To recall the pictures	6 (5/1)	7 (6/1)	17 (9/8)
A4. Recalling: 2. To recall the planned ideas (SP)		5 (2/3)	
A5. Conceptual Simplification: To simplify the message	3 (0/3)	1 (0/1)	
A6. Conceptual Elaboration: To elaborate the message	1 (1/0)	2 (1/1)	1 (1/0)
A7. Avoidance of the intended message: To abandon the generated message	9 (1/8)	2 (1/1)	10 (4/6)
A8. L1 use: To generate ideas in L1			1 (0/1)
<b>B. Lexical operations</b>	52	26	95
B1. Lexical search: To search for an appropriate lexical items	34 (10/24)	12 (6/6)	47 (20/27)
B2. Lexical identification: 1. To identify the target lexical item	5 (4/1)	1 (1/0)	11 (4/7)
B2. Lexical identification: 2. Not to identify the target lexical item	4 (4/0)	1 (1/0)	13 (9/4)
B3. Generating lexical alternatives: To generate an alternative lexical item	8 (5/3)	4 (2/2)	11 (4/7)
B4. Recycling: To recycle the previously used lexical items		2 (0/2)	4 (1/3)
B5. Avoidance of lexical items: To avoid using the intended lexical item	1 (0/1)		3 (2/1)
B6. Lexical elaboration: To use elaborate lexical items		6 (5/1)	6 (5/1)
<b>C. Formulating processes</b>	18	28	56

C1. Structural formulation: To formulate the structure	12 (4/8)	16 (9/7)	40 (23/17)
C2. Structural reformulation: To reformulate the structure			3 (2/1)
C3. Self-correction: 1. To correct ill-formed production	3 (3/0)	4 (3/1)	7 (5/2)
C3. Self-correction: 2. Not to correct ill-formed production despite noticing errors	1 (1/0)	1 (0/1)	3 (3/0)
C4. Elaboration: To elaborate the structure		1 (0/1)	
C5. Avoidance of structural formulation: To abandon formulating the structure	1 (0/1)	5 (5/0)	3 (3/0)
C6. Testing a structural hypothesis: To apply an uncertain structure	1 (1/0)	1 (1/0)	
<b>D. Monitoring processes</b>	15	33	58
D1. Conceptual Monitoring	3 (1/2)	4 (3/1)	11 (7/4)
D2. Lexical monitoring	8 (5/3)	11 (3/8)	26 (14/12)
D3. Grammatical Monitoring	3 (3/0)	17 (11/6)	20 (13/7)
D4. Discourse monitoring		1 (0/1)	
D5. Phonological Monitoring	1 (1/0)		1 (1/0)
<b>Total number of Reported themes</b>	115	112	276

### 5.3.1 Conceptualizing Processes

In the conceptualizing processes, there is a greater number of themes identified in OP than NP and SP (NP – 30; SP – 25; OP – 67), suggesting that the on-line planning implementation might provide significant effects on changing learners' conceptualizing processes. A greater number of the 'message generation' theme (A2) were reported in OP, which may imply that participants more readily engaged in

extending and elaborating ideas when they were given extra time on-line, being released from processing pressure. As opposed to this attempt to elaborate in on-line planning, an observable tendency – though not very strongly suggested by the results – is that participants in NP might have simplified (A5; 3 against 1 in SP and 0 in OP) or even abandoned the originally generated message (A7; 9 against 2 in SP and 10 in OP) rather than elaborate the message (A6; 1 in NP, 2 in SP and 1 in OP), especially among low proficiency speakers. This may be because processing pressure without the pre-task planning opportunity precluded them from engaging in conceptualizing the intended message to its fullest extent, or pushed them into selecting an avoidance strategy after perceiving the linguistic difficulty in translating the original message into L2.

### *5.3.2 Lexical Operations*

A clear tendency is that a greater number of reports concerned lexical operations, 52 occurrences in NP and 95 occurrences in OP, in contrast to relatively small numbers in SP (26), suggesting that participants did not need to consider choice of lexical items on-task in strategic planning because they had been already engaged in lexical searches and scrutinized their own lexical resource during planning time. On the other hand, in NP they needed to primarily engage in lexical searches. This interpretation might be consistent with the relatively frequent number of lexical elaborations in SP (B6; 6 against 0 in NP; 6 in OP), in which participants might become ambitious to find more elaborated and sophisticated lexical items as a result of being equipped with a list of lexis prior to the speech. Also, the relatively frequent occurrences of ‘generating lexical alternatives’ (B3) in NP might be suggestive of participants’ efforts of lexical operations on-task; that is, realization of a lack of lexical resources might

have pushed them into an attempt to find alternatives.

### *5.3.3 Formulating Processes*

Looking at the total numbers of themes in the formulating process, OP produced a greater number (56) than the other conditions (18 in NP; 28 in SP). Taking specific themes into account, a greater number of 'structural formulation' (C1) was identified in OP (40).

An interesting result is that several reports about 'avoidance of structural formulation' (C5) were identified in SP (5 against 1 in NP; 3 in OP), and all of these identifications of abandoning original structures were collected from the high proficiency group. The fact that many participants did not engage in formulation during the ten-minute planning time and because this finding more readily applies to high proficiency speakers (see 5.4.2.1) suggests that they attempted to formulate the structure on-task, resulting in occasional difficulty in formulation and sometimes abandoned this process. On the other hand, because low proficiency speakers tended to engage more in formulation during planning time, the difficulty of some formulations was previewed and the risk of abandoning the formulation on-task was avoided.

### *5.3.4 Monitoring Processes*

In the monitoring processes, the greatest number of themes was identified in OP (58), followed by SP (33 against 15 in NP), suggesting that the greater processing pressures in NP made self-monitoring more difficult. Among various types of specific monitoring processes, grammatical monitoring (D3) tended to show a distinction

between the three planning conditions; OP led to the greatest amount of reporting on grammatical monitoring (20), followed by SP (17) against 3 in NP. Extra on-line processing space created by the strategic and on-line planning conditions is legitimate, but the stronger effects on monitoring processes in strategic planning than in on-line planning were not expected.

This section has reported the main themes in the three planning conditions. The identified themes were categorized into one of the four macro-themes; the conceptualizing, lexical, formulating and monitoring of L2 speech processes. These numerical results are useful for capturing the trend of learners' psycholinguistic operations, but they do not allow us to gain insights into more profound levels of their complex planning processes on-line. Therefore, the following section presents the content analysis of verbal reporting in detail.

#### **5.4 Content Analysis of Verbal Reporting**

The account so far has exhibited the general tendency of participants' thought processes, but it has only presented a partial picture of learners' responses. What is needed therefore in the rest of this chapter is to complement the above description of trends by looking at data episodes more closely. This is also important to support the findings of more explicit form-focused effects in on-line planning, which were observed in the task performance analysis in the previous chapter. In the following sections, I discuss the characteristic features of each type of planning via illustrative data episodes, arguing that different attentional allocation is achieved in strategic planning and on-line planning. This argument leads to better understanding of the processes hidden behind the three planning conditions.

### *5.4.1 Non-Planning Process*

As briefly mentioned in 5.3.1, there was a possibility that speakers engaged in simplifying (A5) and/or abandoning the originally generated message (A7) in non-planning due to the increased processing pressure, but the limited number of reports concerning this point did not yield any convincing conclusions. This part focuses in these episodes on the avoidance strategy in non-planning and explores the underlying reasons for this reaction. This avoidance strategy seems to often lead to learners' prioritization of lexis over grammar and that of meaning over form. Thus, I will consider these issues illustrated in the non-planning condition.

#### 5.4.1.1 AVOIDING PRODUCTION

One of the typical linguistic features identified in non-planning is learners' attitudes to avoiding the production of the intended message. As this condition represents a more natural response of communication in terms of the spontaneous nature of production and the processing burden, similar kinds of linguistic problems seem to frequently appear in actual communication. Focusing on this avoidance strategy, this part looks at the underlying reasons for this reaction, and further extends to the issues about general cognitive processing in pressured conditions.

Table 5.2 illustrates reasons for avoiding production in non-planning via examples of participants' reports, identifying three types of underlying process: (1) realizing a lack of lexical resources, (2) feeling processing pressure and (3) failing to access an appropriate linguistic item in the resources. In the extract, 'single-quotations' represent the words originally reported in English; and "double-quotations" are those originally reported in Japanese. All examples include a

unique number at the end. This consists of LOR group (A/B/C), participants' ID number (1-9), planning condition (NP/SP/OP), the number of episode (e.g., 16-17) and proficiency group (High/Low). The same convention applies to all the reported examples throughout the thesis.

**Table 5.2:** Reasons for Avoidance of Production in NP

Underlying processes	Example
1. Realizing a lack of lexical resources	<p>R: When you told the story, did you think about grammar? Vocabulary? Or content?</p> <p>A8: First I thought of the content here. I thought of the content but I couldn't find the words to describe it.</p> <p>R: How did you deal with this?</p> <p>A8: I think I skipped it. I mean, I couldn't tell the details... I couldn't describe the details of the story, yes.</p> <p>R: You had the picture in your mind then?</p> <p>A8: Yes.</p> <p>R: So, you thought how you could describe it into English...</p> <p>A8: Yes, I couldn't find the phrases to describe it, yes.</p> <p>R: Then, you skipped all these parts.</p> <p>A8: Yes. (A8NP5-6:LOW)</p>
2. Feeling processing pressure	<p>A3: I didn't tell this part [laughs]. I wanted to say "a police man did something" or things like that, but I went ahead because I felt I had no time.</p> <p>R: Did you skip this part consciously? Or you lost it completely while speaking?</p> <p>A3: Well, I remembered the scene, but I only had two minutes, so I avoided it to meet the time limit [laughs]. I thought I had to move ahead anyway. (A3NP3:LOW)</p>
3. Failing to access an appropriate linguistic item in the resources	<p>C8: In retrospect, I should have told "because the bus... because the bus had broken down on the way" or things like that, but I didn't notice it then...</p> <p>R: Saying it by using a past perfect</p> <p>C8: Yeah... um... yeah, it was fine with a positive sentence, but I said "bus being repaired" didn't I? Well... I don't think this is an economical expression, isn't it? Rather, I should have simply told</p>



---

“bus broken down”. Um... I should have told “had to be stopped down then repaired”. (C8NP5:HIGH)

---

It has been argued that learners tend to prioritize lexical operations, especially in speaking (e.g., Skehan, 1998), but insufficient lexical resources often preclude them from efficient speech production. Because of a gap of accumulated lexical items between NSs and NNSs (Nation, 2001; see also 1.5.6), the resource deficit problem is prevalent even in advanced L2 speakers. A significant decision is learners’ reaction in perceiving the resource deficit – that is, whether they shift their attention to attempting to formulate the language to realize the intended message, or whether they simply give up processing the original message. As illustrated in Table 5.2, realizing a lack of lexical resources tends to lead to the decision to abandon the whole production in non-planning rather than to challenge the construction of the language.

The second case also displays a typical reaction in non-planning, but the primary reason for this avoidance is perceiving the difficulty of processing the language under the pressured condition. This attitude might also be implicitly related to the resource deficit problem – learners’ perception of a lack of lexical items – but is more related to the fact that increased on-line processing pressure in non-planning pushed the speaker into taking the avoidance strategy.

In contrast to these relations to rather fundamental production processes, the third case concerns additional processing or a failure to attempt to formulate more elaborate and sophisticated production. Speaker [C8] reflected that she had produced rather redundant language to describe her intended message. What should be highlighted is her failure in this application despite her explicit knowledge of this particular structure (past perfect), which might be suggestive that the pressured condition made her metalinguistic knowledge inaccessible. The following sections

explore these reasons in more detail, particularly in terms of prioritization of lexical and meaning processing.

#### 5.4.1.2 LEXICAL PRIORITIZATION IN NON-PLANNING

Extending the identified cognitive processes underlying the avoidance strategy, I will take a further look at participants' reactions to prioritizing the lexical operations because the analysis in 5.3 also presented the non-planners' active engagement in lexis-related processes.

The extract below reveals a learner's struggle to search for the intended vocabulary:

##### **Episode 5.1**

R: Did you think about grammar or vocabulary while telling the story?

A6: Well... I didn't think about them in most of the time in this task. I couldn't find any word describing the intended idea. I really suffered from finding the target words.

R: So, what were you thinking in this task?

A6: In this task...I was thinking how I could tell what I wanted to say.... I mean I was thinking of vocabulary... I couldn't find the intended words, so I think I didn't have additional mental space to think about grammar, yes.

(A6NP7-9:LOW)

The learner reported his general reflection on his own NP production, implying his prioritization of lexical searches over grammatical processing. For this learner, identifying lexical items is prioritized, and only when the lexical identification is successful, grammatical processing will be taken into account.

A similar tendency could be found in high proficiency speakers. In the following episode, a speaker was conscious of the necessity to apply the past tense to the production, but perceiving the linguistic problem, she deliberately focused her

attention to lexical searches and, as a result, little attention was left for grammatical processing:

### **Episode 5.2**

R: Didn't you think of grammatical issues?

B4: In terms of grammar... well, because this story started in the past, I think I was conscious of maintaining the past tense. But, once I started telling the story, I was confused somehow, and I became more conscious that I would finish all the story before forgetting the content.

R: This means you wanted to use the past tense but you didn't have a mental space for this task?

B4: For example, I couldn't find the words for "walking following [the boy]", I couldn't pay attention to the past anymore, and while searching for vocabulary in mind, I couldn't think of such things as forming the past tense... I was saying 'catch up with' here [pointing to one of the pictures], and I should have said 'caught' or something like that, but the use of the tense became inconsistent and I managed to focus on telling the story anyway. (B4NP4-5:HIGH)

This report clearly represents her conscious selection of attention to lexical operations at the expense of grammatical operations under the pressured condition.

It has been mentioned that lexical operations tended to be prioritized in non-planning, but the underlying reason for avoidance of the accurate translation of a concept into the language and the prioritization of lexical operations is the increased on-line processing pressure under limited production time without any opportunity to consider the language in advance. A number of speakers in the low proficiency group reported their pressured psychological state and the resulting impact on the linguistic processing:

### **Episode 5.3**

R: Did you notice the past tense embedded when I gave the first sentence in the instruction?

A1: Yes. And I tried to take care of the tense usage, but when I listened to the

speech, I found my speech started in the present from the beginning [laughs]. In some very extreme cases, even choice of the subject was a mess... I was always thinking “I have to speak within two minutes!” in mind, and I felt pressured. (A1NP7:LOW)

#### **Episode 5.4**

*[one day two men were trying to steal a box at the airport (.8) they plan (1.0) plan to:: (3.7) they...(.3) a (.3) one boy (.6) er]\**

R: How about here?

A3: Here I was trying to say “setting up [a crime]” but I couldn’t find the words for this... So I changed the message into something like “one man is stealing while the other is talking”, but I was pressured under the “two minutes” and I was thinking “I have no time, I have no time!” [laughs]. I completely lost my head. (A3NP1:LOW)

*\*[italics bracketed] represent recorded speech;*

The above two examples illustrate the clear effects of this pressured state on the formulating process. In contrast to the simple application of identified lexical items, formulating the linguistic structure seems to demand more complex operations, requiring greater amount of attention on-line. The situation would be more problematic in the case of beginners, because being only equipped with a limited range of lexical items and the lower degree of automatization they would require more cognitive space for processing on-line than advanced speakers.

The next example also illustrates the pressured mental state, but this is suggestive of the tendency that the condition made the participant simplify the production against her intention:

#### **Episode 5.5**

A3: I was obsessed with a feeling “I must speak quickly, I must speak quickly”, but I think actually I only spent probably a minute or so. But I thought it went beyond the time later on in the speech. But I couldn’t return to the beginning [laughs], so I thought I should finish anyway, and I finished

it.

R: So, you felt pressured well?

A3: Yes, very much [laughs]. I was thinking “I only have two minutes!”

(A3NP9:LOW)

The speaker could complete the story-telling, but the pressured condition certainly deprived her of the opportunity to describe the pictures in more detail and elaborate the language, failing to satisfy her desires to challenge the production beyond her existing linguistic level.

Such reporting about pressures on processing is more typically found among low proficiency speakers, but similar mental condition and the effects on linguistic processing could also be identified among the high proficiency group:

### **Episode 5.6**

C5: Yes. For this task... how can I put it... I don't remember what I was thinking very well and also I don't think I paused frequently. And, because of two minutes, I was very conscious of the time in the speech. Because, in the first task [SP], I thought “I spoke too much beyond two minutes”, I was thinking, “finish the story immediately” and “finish the story immediately and briefly”, and because I understand I shouldn't stop, I shouldn't stop, I kept going on... I think. I was conscious of the time limitation.

(C5NP7:HIGH)

A possible distinction between low and high proficiency speakers is that, as exemplified above, the effect of the pressured condition may have been more on fundamental language processing in the former group, while this seems more limited in the latter group. Episode 5.6 shows that this speaker was urged to produce the language under pressure at a pace beyond her comfortable processing, but it also implies that she still maintained control over her language processing.

### 5.4.1.3 MEANING PRIORITIZATION IN NON-PLANNING

In addition to the lexical prioritization and processing pressures, the attitude to prioritize meaning over form was frequently reported in non-planning even in the high proficiency speakers. Consistent with the above reports by the low proficiency speakers, the pressured non-planning condition pushed even the high proficiency speaker into prioritizing the content:

#### **Episode 5.7**

R: When the instruction was given in this task... the first sentence was embedded with the past, wasn't it?

B1: Yes, certainly.

R: Did you notice it? I mean you noticed the use of the past tense?

B1: Yes, I noticed it.

R: So, did you try to apply it?

B1: Well, this time... yeah... I tried to apply it, but listening to the speech, [I noticed that] I didn't often apply it. I noticed I was mainly using the present tense.

R: You were using... this means you were thinking of something else?

B1: Yes, I think so. Yes. I was very conscious of "two minutes" and I had to tell the story within "two minutes", and I was in a rush. So, some of the parts became very obscure, but I thought telling the content was more important than grammar, and I tried to tell everything, everything I wanted.

(B1NP4:HIGH)

Among a number of focus-on-meaning accounts in non-planning, the reporting below clearly reveals the competition of attention to form and meaning in L2 processing:

#### **Episode 5.8**

B5: What I thought is... how can I say, I feel like, as the more I'm eager to focus on following the story, the less accurate the production would be. I mean, there is a trading off between something like "it is fine only if the content is transmitted" and "in addition to the content, is accuracy also necessary?"... In such a difficult condition like only thirty seconds for

preparation and two minutes given for the speech, I think, focusing on trying to tell only what happened is likely to lead to ignoring grammatical aspects.  
(B5NP5:LOW)

This statement underscores the processing trade-off between form and meaning proposed by the planning literature. Being aware of the two possibilities, to pay particular attention to either content or linguistic accuracy, the speaker seems inclined to prioritize conveying the content of the given story. What should be emphasized here is that this choice of focus of attention was not necessarily his own intention; rather, more naturally, his focal attention shifted to completing the story telling, resulting in only a limited amount of attention left for grammatical processing to monitor the accuracy of his production.

#### 5.4.1.4 SUMMARY

To summarize, the non-planning production is characterized by learners' limited access to linguistic repertoire and neglecting the engagement in formulation pressured by an absence of planning time and the limited production time. An increased processing pressure is likely to lead to focusing on the meaning aspects and prioritizing the lexical searches for minimizing the linguistic construction. These processes were supported by the episodes in this section. A particularly important implication drawn from these observations is that such a pressured condition is commonly seen in much speech production in the real world. It is wrong to claim that the NP condition created 'real world' conditions but it is at least true to suggest that learners tend to suffer from similar sorts of processing problem in communication. A possible consequence is that, being exposed only to this communicative condition, learners will only have a very limited chance to experience grammatical formulation,

especially in the case when appropriate lexical items are not stored in their Lexicon.

#### 5.4.2 Strategic Planning Process

The overview of verbal report analysis (5.3.1) suggested that participants tended to engage less in conceptual and lexical operations and engage more in the formulating and monitoring processes in the strategic planning production, particularly differing from processes associated with non-planning. Because the provision of ten-minute planning is the only difference between the two planning conditions, it is important to understand how the strategic planning process affected the production. In addition to the exploration of this aspect, more attention will be paid to participants' *on-task* production processes following the ten-minute planning period than their activities during the ten-minute planning. The following part briefly presents the findings of planning strategies which participants employed during the ten minutes. Then, the reports about processes during task performance are presented in detail.

##### 5.4.2.1 OVERVIEW OF PLANNING STRATEGIES

The generated themes of planning strategies were categorized under the three labels – ‘conceptual planning’, ‘linguistic planning’ and ‘metacognitive planning’ following Ortega (2005). From this categorization, it could be observed that learners generally conducted planning in a sequence of conceptual to linguistic planning, and might also engage in mental rehearsal, which is consistent with the general progression of speech processing, proposed by Levelt (1989), conceptualization – formulation – monitoring.

Looking at the percentages according to the three macro-categories (Table 5.3; see Appendix 5-B for the examples of learners' reports about each strategy), there



is a tendency for the participants to engage more in linguistic planning (68) than conceptual (25) and metacognitive planning (29).

**Table 5.3:** Planning Strategies during the ten minutes

Themes	Total	High	Low
<b>Conceptual planning</b>	25	19	6
Focus on understanding the story	12	11	1
Conceptualizing the message	4	1	3
Elaborating the message	5	4	1
Extending the given story	2	1	1
L1 use	2	2	0
<b>Linguistic planning</b>	68	29	39
Lexical search	24	11	13
Summarizing the picture by identifying the key words	11	5	6
Generating alternatives	1	0	1
Simplifying the structure	1	0	1
Formulating the structure	13	6	7
Accessing grammatical knowledge	14	6	8
Elaborating the structure	4	1	3
<b>Metacognitive planning</b>	29	14	15
Organizational planning	13	9	4
Mental rehearsal	9	3	6
Selective rehearsal	1	0	1
Selective planning	1	0	1
Brainstorming	1	1	0
Memorization	4	1	3
<b>Total</b>	122	62	60

Focusing on the two different proficiency levels reveals slightly different results between each group; high proficiency speakers were more involved in conceptual planning (19) than low proficiency speakers (6), while the lower group emphasized linguistic planning (39) more than the higher group (29). The total occasions of

metacognitive planning are more or less consistent between the two groups (15 and 14 respectively), where high proficiency speakers reportedly engaged in more organizational planning (9 against 4 in low) while low proficiency speakers reported their engagement in mental rehearsal more frequently (6 against 3 in high).

In addition to the overall trends, the ten-minute planning strategies are characterized by an active involvement in lexical searches, elaboration of language and rough planning. Taking into account that learners tend to prioritize lexical searching and consume much attention on these operations, prioritizing lexis during planning time seems a natural response. Planning time is also important for recognizing the limitation of their current linguistic repertoire.

Following the lexical operations, learners focus tended to be shifted to their attempts to elaborate the language. A good number of occasions concern 'structural formulation' (13) and 'access to grammatical knowledge' (14), being consistently distributed among both the high and low proficiency groups (6 and 7 in 'the structural formulation' and 6 and 8 in 'access to grammatical knowledge' respectively). Importantly, 'elaborating the structure' is not limited to fundamental form structuring and utilizing grammatical knowledge, but it can be extended to enhancing the quality of production by considering a range of linguistic items and selecting more appropriate language suitable for the given context beyond the level of simple message transmission.

Although a number of speakers were engaged in language elaboration during planning time, this does not simply mean their completion of planning – preparing every single structure to realize all of their intended messages. A more general approach is their conduct of 'rough planning' on the whole, while specifying the language only in particular parts. That is, they seem to have been engaged in selective planning – identifying the parts which could be further elaborated and also the parts

which could not be elaborated, and then primarily focusing on improving the former, 'more confident' part. Planning time would provide a previewing opportunity for language processing – considering the general structure of the story, searching for the necessary words and formulating the structures, but this experience would not often be sufficient for successful production.

As mentioned previously, strategic planning might be effective in enhancing the upcoming production by conducting lexical operations in advance, but structural formulation tended to be neglected even though there was time for considering the structure and grammar during the ten minutes. Therefore, I will consider how these planning strategies affected the actual production in the following sections.

#### 5.4.2.2 TRANSFERRING THE PLANNED PRODUCTION

As we observed above (5.4.2.1), participants selectively conducted detailed planning on some specific parts, while preparing other parts of the story only to a limited extent. Such unequal degrees of planning seem to have influenced their quality of language in the following performance. That is, the parts that were planned in detail could be smoothly delivered, or would ideally push learners in processing the language at more profound levels in production, while the parts that were not sufficiently planned needed to be considered again on task to find a solution. The content analysis of verbal reports reveals that a number of participants commented on their application of planned linguistic items:

##### **Episode 5.9**

*[...and then (2.0) some- (.7) a (.4) man (1.2) kept up with him (.7)]*

R: What about here?

A3: Probably I wanted to say "someone, unfamiliar person is following behind" [laughs]. Well, but only 'keep up with' for "tsuitekuru [follow]"

came up to my mind [laughs]. I said this phrase anyway for the moment.

R: Did you find 'keep up with' during planning time?

A3: I found it. I managed to find only this phrase somehow... I was spending too much time on thinking about this, because just only saying 'walking' doesn't convey a sense of "following behind". (A3SP3:LOW)

The speaker revealed that she identified the phrase, 'keep up with', and planned to apply it in the production. Interestingly, she only planned the use of the identified phrase but did not prepare more complete structure presenting her intended message. This is consistent with the result that a greater number of lexical search engagements (24) were reported than engagement with formulating the structure (13) during the planning time (Table 5.3), suggesting a tendency for learners to focus only on identifying the necessary lexical items during the planning time and to attempt to formulate the structure on-task by employing the prepared lexis.

#### 5.4.2.3 A GAP BETWEEN THE PLANNED AND PRODUCED LANGUAGE

To what extent learners could be engaged in formulating the language on-line by employing the identified lexical items during the planning time may be at least partially dependent on the speaking proficiency level, because this includes the degree of automatization of processing vocabulary and formulation in performance. The following report shows the possibility that less stabilized lexical items might not have been successfully applied, even though they were identified during the planning time:

##### **Episode 5.10**

B2: When I was actually speaking, somehow, I felt pressured for the time limitation and I didn't clearly understand the story, and also I don't have sufficient amount of vocabulary...

R: This means you didn't identify the intended words? Or you may have identified them but...

B2: I identified them, and I planned in what order I would use them, but

when I was producing them, I was unable to find them again... yeah.

(B2SP5-6:HIGH)

A low proficiency speaker described his preparation for vocabulary but did not plan the morphological changes of prepared words:

### **Episode 5.11**

B5: As for vocabulary... well... I don't think I'm able to find difficult words basically [even if planning time is given], but, I mean... I often mess up a singular and plural formation, and I attempted not to make a mistake for this, but when I listened to the actual speech, I found there were lots of inconsistencies [laughs]. That's it.

R: You didn't think about this during the planning time, did you?

B5: During the planning time, well... I wasn't thinking something like, I definitely use this in plural...

R: So, you played it by ear?

B5: Yes. (B5SP9-10:LOW)

It is conceivable from the instances reported that identifying and keeping lexical items in mind can be possible but maintaining morphological and syntactic information in the buffer seems to demand heavy pressure on WM. As a natural consequence of this limited processing capacity, learners tended to prepare only for vocabulary when planning time was given, and attempted to embark on formulating the production by accessing their metalinguistic knowledge in production. Therefore, a great part of the quality of produced language depends on learners' competence in on-line processing. Speaker [B5] attempted to apply morphological changes on-line, but the success of this application is dependent on how much attention can be left and allocated for this manipulation as well as the level of automatization required to process morphological alteration on-line.

Furthermore, the speaker went on to report about his concern about a gap between the planning and the actual production:

### **Episode 5.12**

R: Then, did you speak as you planned? Or do you feel a gap between the produced and planned language?

B5: Well... I was planning to produce in more dramatic and realistic ways, but, um... I think I could speak the development of the story as I planned, but well, I couldn't say what I was trying to produce during the planning time, I mean I couldn't speak such detailed parts. (B5SP11:LOW)

Speaker [B5] planned the details of the story as well as the general storyline and the necessary language to realize the planned concepts, expecting that he could attempt the production of more detailed descriptions by using on-line processing space created by planning. His evaluation of his own production is positive in terms of transmitting the general storyline but negative in terms of fleshing out the basic story. Despite his planning of the content and language prior to production, the speaker may have consumed so much attention realizing the planned items that there was no time left for the development of the planned content.

Grammatical processing tends to consume considerable attention even in more advanced speakers, but what should be highlighted to differentiate them from the lower proficiency speakers is their consciousness about the weakness of particular linguistic points and strategic attention allocation to such grammatical items. This is more prominent in a less pressured condition such as when planning time is provided:

### **Episode 5.13**

*[and:: (.5) and then the next- the next moment (2.0) um (.3) the man (.4) showed something (2.6) a a the man showed (1.3) um showed (.3) the boy something (.8) the man (1.2) was holding something in his hand (2.3)]*

C3: Here I was thinking of [laughs] grammar for using 'show'. I was thinking which 'show + someone' or 'show something to someone' is correct.

R: Is this something you planned?

C3: Yeah, I was thinking this, but I lost it in production.

R: So, you were thinking of how to use 'show'?

C3: Yes yes. I'm not good at using verbs like 'give' and 'send'.

R: You mean the verbs taking two objects?

C3: Yes yes yes. Because I'm not good at it, I forgot how to use it though I was planning. So I was thinking on-task while speaking. (C3SP7:HIGH)

This report illustrates the speaker's lack of confidence with regard to the use of verbs requiring two objects and having to process this on-line. Despite her recognition of the difficulty of this grammatical point, processing the language was not very efficient; but, most importantly, she monitored her production, accessed her grammatical system, and finally formulated the intended structure during the production.

#### 5.4.2.4 FOCUSING ON FORM AND MEANING IN THE STRATEGIC PLANNING PRODUCTION

It is probably true to suggest that the main theoretical underpinning of the benefit of strategic planning is that pre-engagement of L2 processing involving conceptualization and formulation would ease at least part of the essential linguistic processing and finally shift speakers' focal attention to other aspects of production, which are often neglected in more pressured situations. As observed in 5.4.2.1, learners may not have actively engaged in the formulating process during the planning time without being given specific instruction to focus on form aspects (Sangarun, 2005), but this point does not negate the proposal for positive focus-on-form effects in strategic planning. That is, being equipped with a clear understanding of the given story, conceptualizing messages and searching for lexical items beforehand is likely to release the on-line processing burden, contributing to increasing focus-on-form occurrences on-task.

The evidence of a learner's engaging in the Formulator pushed by the planning opportunity can be seen in a comment given by one of the high proficiency speakers. It is reported here that the speaker added detailed description about the characters by employing relative clauses despite a lack of consideration of this grammatical item during the planning time:

#### **Episode 5.14**

R: Weren't you thinking of relative clauses during the planning time?

C1: No, I didn't think about it. During the planning time, I only thought there were two people, and [in production] I added "who" to give more detail explanation. (C1SP8:HIGH)

It can be assumed that the proficiency level of the speaker has approached the acquisition of relative clauses in speech production, but there still seems to be a little more improvement needed for successful and automatic employment in the pressured condition. Speaker [C1] could apply the rule, if necessary, under the state in which other significant linguistic processes such as conceptualizing and formulating were cleared; such an optimal condition was achieved by having planning time.

In contrast to such incidental application of particular grammatical items, it is also possible that speaker [B8] was conscious of formulating a particular structure at the time of planning:

#### **Episode 5.15**

*[er (.2) sorry (.4) a small man open the box (.8) then (.5) there came (.2) a snake (.4) came out (.7) a snake seems to be a poisonous]*

B8: Well... um... I wanted to say "coming out, coming out" here. So, I said "snakes came out"... I was also thinking to say 'they found a snake coming out' or something like that... well... I found it difficult to include a sense of surprise. I found it difficult to describe the situation of their surprise.

R: Is this something you were thinking during the planning time? Or you found it at the time of production?



8: I was thinking of this during the planning time, but, well... planning time ran out, and I felt pressured to identify the words for “coming out”. But I was thinking of telling, at least, two things “they were surprised” and “this [snake] was coming out suddenly”. (B8SP17-18:HIGH)

In this excerpt, the speaker constructed only a basic structure while planning and intended to elaborate it on production. It is conceivable that she was uncertain how successfully she could develop the original structure. The produced language seems not to have met the quality she expected, but the key point here is that cognitive space created by planning time pushed her into formulating and made her attempt more ambitious language beyond the current linguistic limitation.

Consistent with the present line of discussion, a number of reports involving focus-on-form were identified in on-task production in the strategic planning condition. In the excerpt here, the speaker described that the strategic planning condition increased the accessibility to metalinguistic grammatical knowledge in production:

### **Episode 5.16**

R: Were you thinking of grammar while speaking? Were you thinking of vocabulary? Or, were you thinking of the content?

A1: In this time [SP], I was thinking of grammar, and content as well. Because of ten-minute for planning, I could understand the storyline. So, I tried to follow the story in the first place, and I was also conscious of the tense, or the use of the past tense.

R: Were you more conscious of the tense than the previous task [NP]?

A1: Yes, I was a little more conscious of this.

R: Did you find it easier?

A1: Yes. (A1SP5-6:LOW)

The following report clearly describes the processes that lexical operations in advance pushed the speaker into engaging in focus-on-form processing during the next production stage:

### **Episode 5.17**

B3: I was thinking of this [the past tense] [during the planning time] [laughs]. From the beginning, I was trying to recall things such as “this is not ‘hold’ here”. I mean I was recalling several sentences to mind... Probably, it was a little like reading out, but anyway I had a mental space to think of such things in mind.

R: What you are saying is during the ten minutes, isn't it?

B3: Yes, yes.

R: Then, were you monitoring this during the production?

B3: Yes, yes... Well, um... in the other conditions, when the production started, I had to do various things at the same time such as searching for vocabulary and adjusting the tense consistency, but in this time, I didn't need to conduct lexical searches, um... so I could focus on the next steps... Well, I had also already prepared for most structures, so I could pay attention to other lexical items and conjunctions... (B3SP9-10:LOW)

In this example, the speaker had already been engaged in formulating the structures as well as searching for the necessary words during the ten-minute planning period, so she could shift her attention to further elaboration of lexical choices and conjunctions, which seemed the next priority in her hierarchical preparation for speech production.

So far, we have seen form-focused occasions in the strategic planning production, following the theoretical rationale that planning opportunity prior to production is expected to process conceptual planning and lexical identification, encouraging learners to focus on form. However, the analysis of verbal reports also suggests that strategic planning does not necessarily guarantee learners' focus on form engagement. The episode below reveals learner's attitudes towards prioritizing the meaning aspects in addition to linguistic planning conducted during the ten minutes:

### **Episode 5.18**

B1: I was a little thinking of vocabulary during the planning time, but I think I couldn't use it during production.

R: I see. Though you were planning several words...

B1: Yes, yes.

R: How come?

B1: Well... I was focusing too much on content [in production]... Because I was trying to transmit the content, I couldn't pay attention to grammar and the planned words...I think.

R: Then, did you think of grammar? I mean, were you thinking of using this sort of sentence, or this sort of phrase?

B1: Yeah, I wrote down several key words, and I thought I would say by using this sort of sentence and this sort of word, but I couldn't use them well in actual speech...

R: You couldn't

B1: I think so... Yes, I think I couldn't use them. (B1SP1-2:HIGH)

It is also possible that, despite noticing the necessity of thinking of grammar, one learner deliberately paid focal attention to meaning, because of his recognition that focusing on form might impede task completion, or getting the message across:

### **Episode 5.19**

R: Then, did you think of grammar while telling the story? Were you thinking of vocabulary? Or, were you thinking of content...

A5: I was thinking of content while speaking.

R: So, you didn't think of grammar very well, did you?

A5: Yeah... I felt I would stumble very much, if...if I thought about grammar, so I was focusing on smooth transmission of the message. (A5SP7:LOW)

In another episode it is evident that there were focus-on-form effects of strategic planning but in a very limited way:

### **Episode 5.20**

B7: Probably, this time, I noticed the use of the past tense, but I couldn't pay attention to this while speaking...

R: You noticed it during planning...

B7: Yes, I noticed this probably. But, um... although I was thinking it would be better to tell this story in the past, I couldn't pay attention to this once the production started, I despaired... I shouldn't say 'despair' though...

R: You sometimes used the past tense, didn't you?

B7: I sometimes became a little conscious of this, but I lost it other times. Words were coming out before thinking of the changing of form, and I was saying “send” or something like that. (B7SP6-7:LOW)

Speaker [B7] noticed the necessity of producing the language in the past during the planning time, but once the speaker started, the difficulty to deal with on-line processing downgraded the focus on the grammatical item. What is interesting in this episode is that such meaning prioritization occupied a great part of his production but not every single utterance; when the linguistic processing burden was lessened, his attention automatically shifted back to the grammar.

#### 5.4.2.5 SHIFT OF ATTENTION TO FORM IN THE STRATEGIC PLANNING PRODUCTION

Having looked at focus-on-form and focus-on-meaning cases, this section looks at both scenarios together and considers how learners’ attention switched between the two states during production by referring to episodes by a single participant.

One detailed description about shift of attention to form in the strategic planning production can be seen in episodes reported by speaker [C5]. Firstly, she started reporting her thought process during the ten-minute planning time, in which she primarily focused on conceptualizing the messages and the organization of the given story; thus, focusing on particular grammar (i.e., past tense) and structural formulation were completely neglected:

##### **Episode 5.21.1**

C5: During the planning time, I didn’t think about grammar at all. Then... um, when taking this note [point to the written note], I wrote down the words for remembering the development of the story and what happened in each picture. Then, after that, I was thinking what kinds of story-telling could be

made if I connected each picture. At this time, I was not thinking like “I have to use the past tense here” or “I’ll use this sort of sentence”. I was only thinking “the story is going like this way”.

After focusing on meaning during the planning time, she naturally got engaged in focus-on-form processing in production. The following parts of the episode reflect her speech processing experiencing the focus-on-form:

### **Episode 5.21.2**

C5: As for grammar... when listening to the speech, I realized... I was thinking of grammar very much in speaking.

In this part, speaker [C5] realized her active engagement in grammatical processing, which was not planned at all prior to the production. It is suggestive that it is not easy for learners to recognize every single problem concerning the subsequent production in advance (as illustrated in 5.4.2.1), and it is not until the time of production that necessary grammatical items become clear. In the following episode, she uncovered her thought precisely when she paid attention to the use of the past tense:

### **Episode 5.21.3**

C5: For example, when I was using the past in the ‘until’ clause, ‘until...the man stopped him’... after using ‘until’, I was thinking “what tense I should use within the clause following “until”. While producing it, I was thinking “What is the tense? How should I change this?”... I didn’t think “how I can use the tense here” or things like that during planning. I was only thinking “I’ll remember the story”, “how should I connect the story?”... “how can I remember the story?” and “how can I develop the story later?” Yes, and, for the language, I was only thinking how I could use vocabulary... “should I use ‘reach’ or ‘catch up’ here?” and “should I use ‘chase’ or ‘following’ here?” I was only thinking of such kinds of things. Yes...yes. (C5SP29-34)

It seems that she could quickly find an appropriate structure to realize her intended message without effort, but some level of attention was required to construct the past

tense formation. A particularly noticeable point is that she started thinking of the tense formation after starting the production by monitoring the tense usage in the whole structure, accessing metalinguistic knowledge and applying the alteration. More importantly, this series of processing was conducted while maintaining a good level of fluency. Referring to the production she accounted for (*'...and a (.2) he didn't know (.3) a man a strange man (.5) chasing him until he (.3) the man stopped him'*), she neither took a lengthy pause nor made a serious breakdown.

#### 5.4.2.6 SUMMARY

To sum up, processing conceptual and lexical aspects during planning time may have pushed learners to engage in grammatical processing in production. The participants tended to spend the planning time on understanding the given story, organizing the conceptual structure and identifying the necessary lexical items, but not necessarily attempting structural formulation. The general trends of learner's strategic uses of planning time are consistent with those found by Ortega (2005) involving pervasiveness of retrieval operations, monitoring and rehearsal activities enabled by extra planning time. It should also be stressed that there are also consistent findings of different strategy applied between different levels of language expertise. Ortega (2005) reportedly identified that an advanced level of speakers showed a more balanced commitment to strategy, while low-intermediate speakers fostered retrieval strategies committed to solving lexical and verbal morphology problems. To be more precise, the present study also suggests high proficient speakers' more balanced attention to conceptual, linguistic and metacognitive planning, in contrast to an emphasis on lexical retrievals and grammatical processing in low proficiency speakers.

Cognitive space created by planning is likely to lead to occurrences of focus-on-form during the performance, despite limited focus-on-form effects in the task performance analysis. It can thus be argued that strategic planning induces L2 speakers to focus on form more frequently than non-planning, but this shift of attention is not always achieved by strategic planning.

Together with the episode shown here and all other evidence concerning focus-on-form, it could be assumed that strategic planning tends to push learners to focus on form not only during the planning time (Ortega, 1999, 2005) but also in the following production even in the 'unguided planning' condition, because planning opportunities facilitate conceptualization and lexical searches beforehand. This partial completion of linguistic planning tends to advance learners' linguistic processing naturally into formulation and focusing on specific grammatical items needed for the formulation on task. The study by Ortega (2005:101-104) showed evidence of explicit focus-on-form during planning, supporting her original hypothesis that strategic planning leads to focus-on-form, but the findings in the present study push this further, arguing that focus-on-form could also occur 'during' performance in addition to the ten-minute planning stage. A particularly noticeable point is that L2 speakers were engaged in focus-on-form episodes during production despite their neglecting it during the occasion of pre-production planning time.

However, it should also be stressed that this shift of attention to form is not always automatically triggered by providing planning time. As reported by several participants, speakers may have persisted in focusing on meaning, neglecting the focus-on-form, in production after planning, probably when producing the given story still pressured the speaker after the planning. A particularly significant indication is that the formulating process and associated grammatical processing cannot always be pre-planned due to a limited WM capacity and the difficulty of recognition of specific

linguistic problems prior to speech. Therefore, it could be said that strategic planning contributes to promoting focus-on-form but on-line processing pressure still interferes with linguistic processing.

### *5.4.3 On-Line Planning Process*

A number of themes concerning the on-line planning processing were generated from the episodes collected in the post-task interviews, which take up 55 % of all identified themes. Despite the various types of theme, most episodes were focused on a limited range, representing the general tendency of speech processing in the on-line planning condition. Looking back to the summary of reported themes (Table 5.1), a greater number of formulating (56) and monitoring processes (58) were found in on-line planning than in non-planning (18 and 15 respectively) and strategic planning (28 and 33 respectively). This tendency might have occurred as a result of an emphasis on grammatical accuracy and monitoring the on-going production encouraged by the on-line planning instruction and the resulting reduced on-line processing burden.

As in the two other planning conditions above, this section concerns a qualitative inquiry about the on-line planning processing. In the following parts, I particularly focus on the formulating process by referring to on-line planning episodes, through which I observe how language production was constructed and elaborated in on-line planning. Additionally, as longer and a greater number of pauses were observed in on-line planning, I then consider reasons for pausing as one of the characteristic features of its production.



### 5.4.3.1 ENGAGING IN THE FORMULATING PROCESS AND ACCESSING GRAMMATICAL KNOWLEDGE

As argued by Skehan (e.g., 1995, 1998), L2 speakers tend to firstly pay attention to the process of lexical searches, and when appropriate lexical items are not identified, then shift to the grammatical formulation. Emphasizing this proposal, a greater number of lexical operations (95) than formulating processes (56) were reported in OP, implying that the speakers may have been shifted into the formulation when lexical identification was not successful. The following report describes a learner's attempt to construct the structures in reaction to a failure of appropriate lexical items:

#### **Episode 5.22**

*[and a (.4) um there's a (2.1) um (.2) yesterday they had a rain (.3)]*

C7: Here, I couldn't find the word for "puddle", so I was thinking how I could let the listener know this situation... It was raining before but it is not raining this time, so I said this is the day before or things like this [laughs]. Then, assuming that it was raining, I started accounting for the situation of puddle – how water was sprinkled... yeah from this part I think I was saying the situation by using various sentences. [Then listening to the following production.]

*[and a (.8) um (1.6) there's a (.5) some waters beside the street (.7) and the truck just splashed (.4) er the water to the passengers er to the (.3) people (.3) waiting for the bus (.7)] (C7OP2:HIGH)*

Failing to identify the word "puddle", speaker [C7] embarked on describing the situation without using the target word. Although the on-line planning condition pushed the speaker into the grammatical formulation, it is also possible that an attempt at formulation was not successful. Another speaker describes the processing of changing the structure as a solution of failing to identify the necessary word:

### **Episode 5.23**

*[and:: (2.8) and:: (1.0) finally (2.0) he:: (1.5) the strange man (1.8) um? (.5) strange man (1.3) strange man reached (1.3) the boy]*

R: What did you do when you couldn't find the word?

A8: When I couldn't find the word...

R: Did you find it immediately when you were thinking the appropriate word?

A8: When I couldn't find the word, I was thinking of another way to describe the situation...um...well, probably in this occasion... I tried to make a sentence by using this [boy] as a subject at first, but then I changed to this [the strange man] as a subject, yes.

R: I see. You changed the sentence completely.

A8: Yes. (A8OP11:LOW)

After an unsuccessful lexical identification, the speaker resorted to the rule-based system for constructing the structure of the intended message.

As expected in the initial conceptualization of the on-line planning processing, reduced on-line processing pressure, which is made possible by an absence of time limitation, naturally pushed speakers into the processes concerning grammatical formulation. That is, as grammatical processing tends to consume considerable amount of attention, on-line planners are expected to be more readily engaged in this unpressured condition. Episode 5.24 below reveals that, although speaker [C7] was uncertain of the reason for her attention to the past tense formation in on-line planning, she speculated that the condition freed from the time limitation, as well as encouragement of self-correction, made it possible for her to access the grammatical system:

### **Episode 5.24**

R: Why did you correct it only in this task?

C7: I don't know, but this is... well... there was no time limitation in this task, so probably I had some mental space in mind to think about this... For the other tasks, the two minutes were required, and because I didn't know

how short it was, I was trying to keep talking and talking. But here, because I could correct, probably I thought “oh, ‘have to’ is not correct!”, I think. Yes, I think I was unpressured for time in this task. (C7OP15:HIGH)

Another speaker also reported her consciousness of grammatical processing but in a more certain way:

### **Episode 5.25**

B1: ... Because I wasn't limited by the time, um... because of the condition to allow the speech slowly, yes, I was trying to speak while thinking of grammar. (B1OP1:HIGH)

These episodes suggest that processing grammatical items consumes more attention and time, so that learners could be comfortably involved in this processing only when there was sufficient time.

In addition to exclusive focus on grammar, it is possible to observe participants' processing of both content and language, which means that further processing space could be needed for constructing the production:

### **Episode 5.26**

*[the... (.2) a man (.7) the man (.8) passed (2.2) the parcel (.3) which (.4) a... (.3) the boy dropped (.8) on the street (2.5)]*

R: There is a long pause here, isn't it?

B1: Yeah.

R: Were you making the following sentence in mind?

B1: Yeah... while thinking of content too...

R: Oh, both...

B1: There was a pause because I was thinking both content and form.

R: So, you need more time in doing so?

B1: Yes. (B1OP5HIGH)

In this excerpt, the speaker was engaged in the Conceptualizer as well as the Formulator in Levelt's term, affecting the fluency in a negative way.

So far, we have seen speakers' responses to general grammatical processing observed in on-line planning. In the rest of this section, I illustrate episodes on accessing the rule-based system of more specific linguistic items.

There are types of grammatical items which can be easily understood for their structural simplicity but not applied easily for processing difficulty. Such items tend to be taught in the early period of English teaching curriculum in Japan, but often continue to be a problem of accurate production even for advanced speakers. Making morphological changes in the past tense and plural forms are two typical examples.

Speaker [A1] reported his particular attention to such morphological changes:

**Episode 5.27**

A1: Well, well... the tense was changed several times... I was correcting it, but I was always saying the present tense... so I think I can't apply the past tense in an instant. Also, at the beginning, I firstly said 'men' for "someone". I thought "this is not plural here!". I was saying the plural for the case of singular. (A1OP11:LOW)

It is worth noting that speaker [A1] always identified the target item as an original root form and then changed it by accessing his metalinguistic knowledge. As it is conceivable that low proficiency speakers tend to suffer more seriously from processing these sort of changes on-line, it seems crucial to reduce the level of on-line pressure. Reflecting on the point that learners were not always engaged in grammatical processing in strategic planning production (5.4.2), the on-line planning implementation could be one method to encourage them to access metalinguistic knowledge in production, in order to develop their competence to make accurate production on-line.

Accessing the rule-based system could be conducted by making cognitive

space on-line, which can be attempted in various ways. The most obvious and probably easiest way is to allow pausing as long as the speaker wishes until he/she finds the lexical items and/or formulates the intended structure. Although the results of longer and more frequent pauses in on-line planning show that a number of participants chose this strategy on many occasions, this might not be a preferred approach for a particular group of speakers, especially in the high proficiency level, who may feel uncomfortable taking unnatural pauses. I will return to the issues of pausing in on-line planning later (5.4.3.2 & 5.4.3.3).

A more sophisticated approach to creating cognitive space on-line is the employment of time-creating devices such as fillers (Bygate, 1987). Speaker [C3] reported her engagement in grammatical monitoring by repeating and reformulating the same phrase:

### **Episode 5.28**

*[but um (.8) it was (.6) busy it's (.7) full of people and (1.2) it was full of people and um (.3) the three boys (.8) couldn't get on (.4) get on (1.8) so they (.3)]*

C3: I was thinking of grammar here.

R: You mean, you were thinking how you told the story here?

C3: Yes. For example, I was thinking "should I say 'bus', like 'get on the bus' or such kinds of expressions?". I was thinking how I could say more accurately.

R: You were thinking of accurate grammar...

C3: I was thinking whether I should say like this for accurate English, but it may not be necessarily important in speaking. This is not writing. While thinking such things, I repeated the utterance.

R: For buying the time?

C3: Um, I repeated it unconsciously and, while repeating, I was thinking if this sort of English would be acceptable. (C3OP3:HIGH)

It is noticeable here that this time-creating attempt is not a simple repetition of phrase.

After saying the first phrase 'it was busy', she reformulated the language by saying 'it's full of people'. And then, she repeated the same structure but she corrected the tense form in the second time ('it was full of people'). Another part of repetition in this episode is more simply recycling the phrase 'get on' while searching for more accurate production. In this part, she was tracing her stored knowledge on WM by repeating the phrase rather than accessing the explicit grammatical knowledge. It is interesting to see the example of 'juggling act of speech processing' (Ejzenberg, 2000), and such simultaneous attention to various parts of language processing probably could be achieved by advanced speakers who have solved the processing problem to a great extent, being equipped with a large amount of lexicalized items and automatized knowledge.

#### 5.4.3.2 SELF-MONITORING OF THE LANGUAGE PROCESSING: REASONS FOR PAUSING

A number of participants were clearly slowing down their speech and taking long pauses, which creates a different impression from the other planning conditions. As seen so far, such on-line planning incidences can occur not only at the time of pausing but also during utterances. However, it is probably true to suggest that speakers might be able to verbalize their thoughts while pausing more clearly at which point the level of processing seems to be lowered, releasing them from engagement in the Articulator. Such characteristic phenomena of pausing would lead to another question inquiring about the underlying reasons for these changes in on-line planning production. To explore this point, I asked a question about their reasons for pausing. Their answers include a wide range of themes, most of which could be categorized in one of the three processes; conceptualization, lexical operations and formulation. In the

following part, I will observe the three main reasons for pausing or self-monitoring by showing individual episodes, through which I search for the nature of on-line planning processing.

**Conceptual monitoring.** The on-line planning implementation primarily aims to encourage learners to process grammatical knowledge and to engage structural formulations, but it is certainly possible to facilitate the conceptualization as well. Because the present task takes a form of learners' telling the given story, not resorting to their own story, the content needs to be stored in WM, which increases the burden on other linguistic processing. Moreover, in on-line planning, as only thirty seconds was given for remembering the story, it is possible that speakers did not understand the given story fully.

An interesting tendency observed concerning pausing for conceptualization, or conceptual monitoring, is that this is particularly true of more advanced speakers. For example, an episode involving speaker [C9] represents her explanation of pausing for recalling the pictures:

**Episode 5.29**

C9: Um, yes, there were impressive pictures, weren't they? And, if I remember the picture very well, I could tell the story immediately, but for the less impressive picture, firstly I had to think what sort of picture it was. So, probably, I was taking a pause, while trying to recall it in such a case.

R: This is the reason for taking a long pause...

C9: Yeah, yes. I think I was thinking what sort of picture it was and what part I would start to explain.

R: And you remembered the story well when pausing only for a short time?

C9: Yes, that's right. (C9OP4:HIGH)

It should be noticeable in this episode that the speaker was engaged not only in

recalling but also in organizing the order of story telling. Her recalling process seems not to simply mean her failure of memorization in the first thirty seconds but that a memorable story includes a clear structure in it, making a distinction between telling as simply following the content and telling by organizing the story on her own (cf., Kobayashi [2002a] for the effects of information structure on learners' understanding). Therefore, the less memorable story is considered to require more attention in the production process, because speakers need to be engaged in the structural organization of the story.

Another example of conceptual monitoring is also reported by a high proficiency speaker. In this episode, the speaker did not find any problem translating the story into L2, but was concerned about the connections between each picture:

**Episode 5.30**

C8: Um... probably, I was thinking "um..." at the transition between pictures. I was not thinking of sentence structures, but I was thinking how I was able to explain this transition... and also... I think I didn't think about sentences and vocabulary... um... I think I didn't. (C8OP13:HIGH)

Speaker [B8] was not satisfied with a simple description of each picture but attempted to link or to fill the gap between the pictures by giving additional elements. She also mentioned her unconsciousness of lexical and grammatical processing during the production, suggesting the whole formulation processing is largely automatized.

The conceptualization process could occur in any planning condition, but the quality of this varied from occasion to occasion, and between the different proficiency levels. As shown in this part, the conceptualization in on-line planning seems more prominent in high proficiency speakers. The reason for this trend is not very clear here, but it can be assumed that, as implied by speaker [C8], the formulating process tends to be more automatized in advanced speakers, leaving greater attention to the



Conceptualizer similar to speech processing in NSs. As only a limited amount of attention is needed for linguistic formulation, speakers could focus more on developing conceptual monitoring on-line. It can be underscored by the fact that both episodes reported here concern the elements of enhancing the quality of production rather than generating simple messages.

**Lexical monitoring.** In contrast to the tendency that pausing for conceptualization could be seen mainly in high proficiency speakers, episodes on pausing for lexical searches were mostly seen among the low proficiency group. Episode 5.31 tells us that the success of lexical identification would determine the length of pauses:

**Episode 5.31**

A5: For both short and long pauses, I remembered the picture clearly in mind, but the difference lies in whether the vocabulary was coming up or not. I could relatively easily find frequently used words, but I needed to think what it was in case of less familiar words.

R: So, you could say it only for a short pause when the word came out immediately, but for the less familiar words...

A5: I was spending much time on thinking for the less familiar phrases, like "is this a right word?"

R: It took longer in such a case. (A5OP5-6:LOW)

What is interesting to note here is that the speaker's understanding is in accordance with the theoretical account in SLA that the accessibility to the target lexical items is determined by the frequency of the items (e.g., N. Ellis, 2002).

The next excerpt explains interference with lexical searches by a previously identified item. In the previous production, speaker [A2] identified a lexical item which could be recycled in the following production. In order to avoid the repetition, she attempted to find another appropriate word, but the previous word ('succeed') continued to impede her accessing the lexical resource:

### **Episode 5.32**

A2: When I was taking a long pause, I had only one word in mind... For example, I said 'succeed' before, and I couldn't find an alternative, and because the word remained in mind, I thought "other words are not coming out, not coming out". After all, I recycled the word 'succeed'... if I didn't find "this word" at first, I searched, found and said an alternative word, but if the word appeared in mind at first, it interfered with the search for alternatives. (A2OP18:LOW)

This report is suggestive of the failing case of searching for lexical items. The continuing lexical searches might successfully lead to finding the target lexical item, but they are also likely to lead to the identification of a lack of lexical repertoire in the resource. A likely response to the failure is shifting attention to grammatical processing and formulating the syntactic structure, as observed above (5.4.3.1), but it is also possible to go into less challenging production such as simplifying or abandoning the message. The following excerpt illustrates a participant's solution to a lack of lexical resources by selecting a simpler word, even though the selected words did not correctly represent his intended message:

### **Episode 5.33**

A6: Um... yeah, when I could say smoothly, I found out the word immediately at the point of production. And, when there was a pause, well... I didn't think about that, I mean, this was when I was confused how I could maintain the speech... So, I think I couldn't find the target words immediately, yeah.

R: How do you cope with when you can't find the target word?

A6: Um... thinking of the situation, and I think it would be good if I'm able to replace it by using simple words. (A6OP7-8:LOW)

More compelling evidence for continuing lexical searches by low proficiency speakers can be found in Episode 5.34. Speaker [A7] attempted to search for the target word, which was not in the other planning conditions involving time limitation

for the production:

#### **Episode 5.34**

A7: Um... probably, because I was told [in the OP instruction] that I was able to pause, I thought I tried to think until I found the word. Usually, I think I would think of other ways of telling, but, because I was allowed to pause in this task, I tried to find out the target word, but after all, I couldn't find it out.

R: So, in the case of long pauses, you were trying to recall the word...

A7: Yes, probably. (A7OP7:LOW)

A particularly noticeable comment in this episode is that his continuing attempt at lexical identification did not lead to a positive outcome. The reality is his failure of lexical identification despite further effort to continue to search for the lexis, implying that it is not necessarily fruitful to spend a considerable amount of time on lexical searches without accessing external assistance (e.g., teachers and dictionaries).

One of the most important findings drawn from the observation in this part is the tendency of the low proficiency speakers to persist in lexical searches, being released from on-line processing pressure. It may be demonstrating some evidence of lexical prioritization in L2 production, but it is interesting to see that, in the case where processing time is not a problem any more, learners were continuing to search for the target lexical items. Looking at this tendency from the other side, high proficiency speakers tended not to be engaged in lexical searches any more in on-line planning. There is no clear statement of this distinction in the reported data, but it can be speculated that high proficiency speakers more readily shifted into employing the rule-based system, possibly because many of them were aware of the limitations of their lexical resources and/or avoided serious breakdowns that might have arisen by continuing the lexical searches even in the on-line planning condition.

**Grammatical monitoring.** As repeatedly mentioned, speakers' attention is likely to shift into structural formulation and grammatical processing needed for the intended message, when the necessary lexical items are not stored in their resources. The previous section looked at the tendency of the low proficiency speakers to continue to prioritize lexical operations when they were released from on-line processing pressure. This observation leads to a possibility that the high proficiency speakers more readily engage in accessing the rule-based system. Underscoring this assumption, several episodes on prioritizing formulation in on-line planning were identified among the high proficiency group:

**Episode 5.35**

B1: Yes... I think, when I was pausing frequently, I was constructing the structure in mind.

R: Then, what about taking a short pause?

B1: When I was taking a short pause, I was producing simply what I was thinking.

R: I see... this is the same as usual speech.

B1: Yes... (B1OP2-3:HIGH)

**Episode 5.36**

C4: Well... when I'm able to explain spontaneously while looking at the pictures, I mean when I'm able to translate the picture into English, I can produce smoothly, but when I look at the picture and I'm unsure how I can say it in English, I'll try to construct the language in production.

(C4OP11-12:HIGH)

What these episodes reveal is that some sorts of concept are stored in their resources, which are accessible and ready to use immediately, while other sorts of concept are not stored, making the learners construct new language. For the former type of knowledge, the language appeared in their mind at the time of production or even at the time of looking at the picture. For the latter state of knowledge, on the other hand,

the learners needed to engage in structural formulation, which consumed a greater amount of attention, finally leading to frequent and lengthy pauses. What seems to be implied in these descriptions is the speakers' natural shift into grammatical formulation rather than continuing the lexical searches, as seen in the examples of low proficiency speakers (e.g., Episode 5.34). A possible explanation for this is that high proficiency speakers tend to understand the limitation of their lexical resource and strategically shift from the exemplar-based to the rule-based mode. Even though freedom of pausing was suggested in the on-line planning instruction, they might have felt uncomfortable with unnaturally long and frequent pauses, which would push them readily from staying in lexical operations into actively constructing the structure in order to transmit the intended message.

All the episodes so far concerned speakers' reasoning for relatively long pauses, but there is also a report concerning relocation of pausing. A pausing activity can be considered not only in quantitative terms of its length and frequency but also in the analysis of places where particular pauses occur. A distinction which has drawn attention in recent research is that between end-clause pauses and mid-clause pauses. Skehan and Foster (2005) argue that the number of pauses may not show clear distinction between NSs and NNSs of English, because NSs also pause in speech production. What may be showing a clearer distinction is the place of pausing; NSs usually tend to pause in the end-clause position, while NNSs may be more frequently pausing in the middle of a clause. In the following section, the episodes reporting places of pausing are presented.

### 5.4.3.3 PLACE OF PAUSING

One of the accounts of end-clause pauses was provided by a high proficiency speaker.

To the question about his thoughts during a pause, the participant reasoned:

#### **Episode 5.37**

*[and at that moment (.5) finally (.3) small (.2) boy (.2) notice that (1.0)]*

R: There is a pause after 'notice that'. Did you think about something here?

C1: Well... um... I think this is usual. I think I say like this even in Japanese.

I mean, I was pausing not particularly because of English, but even in English, how can I put it, I was pausing. (C1OP5:HIGH)

Speaker [C1] explained that pausing here is his usual habit, neither any linguistic problem nor structuring the language. For him, pausing at this sort of place is not unique to using English, but general transfer from his L1 production.

There is an episode about mid-clause pauses identified in a high proficiency speaker. Speaker [C5] reveals her engagement in accessing the grammatical knowledge in mid-clause pauses:

#### **Episode 5.38**

C5: ...then, as for pausing in the middle of speech, well, this is for the tense, I mean I was thinking how I could apply the tense... Such frequent pauses are for the tense. I was thinking of the tense. When I couldn't find the vocabulary, I paused [at the end of a clause] while thinking "how can I do?" (C5OP32-33:HIGH)

In addition to mid-clause pauses, she also commented on her long pausing at the end of the clause as a result of searching for lexical items. What is implied by this statement is that it is at the end of the clause that she conceptualized the message, searched for the necessary lexical items, and after processing all these, she started the

production while formulating the structure. As conceptualizing the message and identifying the lexical items would determine a great part of language production, these should be processed prior to production. On the other hand, processing the tense structure may consume much of her attention, but this tends to be marginalized in language production. This is not prioritized in production but is processed when cognitive space is created on-line and/or when the accurate production is paramount (cf., Krashen, 1982). Supporting this conjecture, speaker [C5] goes on to explain the importance of lexical identification in speech:

### **Episode 5.39**

C5: If this is a usual conversation, I would change the topic or try to say different things, but because I knew I had to account this story, I was lost and pausing, and I was thinking 'I cannot skip the part for sprinkling puddle.  
(C5OP34:HIGH)

It is also important to note the speaker's involvement in language processing in the case of lexical resource deficit. Speaker [C5] reveals her usual employment of an avoidance strategy when encountering a lack of lexical items, but the condition to describe the given story, together with the on-line planning occasion, made her attempt to describe the difficult contents.

#### 5.4.3.4 SUMMARY

Having looked at various features of the on-line planning production, what characteristics distinguish this from features of the non-planning and strategic planning productions? Although there was formulating and grammatical processing in every planning condition (Table 5.1), the content analysis of episodes suggests that, when the necessary lexical items were not stored in their resources, speakers clearly

engaged in formulation and accessing explicit knowledge in the unpressured on-line planning condition. This kind of focus-on-form processing may occur in the other planning conditions, but several reports suggest that on-line planning is more likely to encourage L2 speakers to employ the rule-based system.

The on-line planning production is also characterized by more frequent and longer pauses than the other conditions, implying learners' active engagement in self-monitoring. Although this tendency can be seen in both proficiency levels, the reasons for pausing were not consistent between them; that is, high proficiency speakers tended to pause longer in order to develop their conceptual planning, while many low proficiency speakers were engaged in lexical searches, suggesting that, the linguistic processing being largely automatized, the former group could pay attention to developing the content of the story, while the latter group needed to consume a large amount of attention on linguistic processing to describe the baseline of the given story. In the self-monitoring occasions, many high proficiency speakers also reported their conscious formulation processing. Although cognitive space was also created by the strategic planning condition, the existing on-line processing pressure seemed to block their full engagement in the Formulator. Because accessing the grammatical system consumes considerable amount of attention even in high proficiency speakers, learners were likely to be more readily engaged in switching to the rule-based mode in on-line planning.

## **5.5 Conclusion**

To summarize, the content analyses of the speakers' episodes collected in the post-task interviews offered a number of findings of distinctive mental operations among the different planning conditions and different proficiency levels. Firstly, in



non-planning, there was a tendency for speakers to neglect their engagement in structural formulation in the pressured condition caused by an absence of planning time and the limited production time, leading to focusing on the meaning aspects and prioritizing the lexical searches for minimizing the linguistic construction. Secondly, in strategic planning, processing conceptual and lexical aspects during the ten minutes planning time tended to push learners to engage in formulating and processing grammar in production. Cognitive space created by planning was likely to lead to occurrences of focus-on-form during the performance as well, despite limited focus-on-form evidence in the task performance analysis. This underscores the importance of retrospective interviews. Thirdly, in on-line planning, when the necessary lexical items were not stored in their resource, on-line planners were clearly engaged in formulation and accessing metalinguistic knowledge in the unpressured condition. This kind of focus-on-form processing can occur in the other planning conditions, but unpressured on-line planning is more likely to encourage L2 speakers to employ the rule-based system.

Following Ortega (2005) and integrating the findings of the psycholinguistic processes in the present analysis, it is conceivable that different linguistic proficiencies influenced the planning implementations. Ortega (2005) observed that planning time allowed a more balanced commitment among advanced level speakers, while it led mainly to linguistic operations among low-intermediate speakers. Consistent with and extending this trend, the present study also observes that high proficiency speakers tended to maintain balanced attention to various parts of speech production in on-line as well as off-line planning occasions, while low proficiency speakers engaged more in linguistic operations particularly in lexical retrievals in any planning opportunity. Having revealed the findings of the task performance analysis and verbal report analysis, I will attempt to integrate those from

information-processing perspectives and to draw a unified picture in the next chapter.

## **CHAPTER SIX: DISCUSSION**

### **6.1 Introduction**

Referring back to the theoretical rationale which was reviewed in Chapter One, the purpose of this chapter is to reconsider the original question, how syntactic-processing rich environments could be established by planning implementations through integrating the findings of the task performance analysis (Chapter Four) and those of the verbal report analysis (Chapter Five). Having looked at various findings, the following discussion is limited to the two main factors affecting L2 speech processing and learning; (1) task planning and (2) learner proficiency.

### **6.2 Integrating the Findings of the Task Performance and Verbal Report Analysis**

Following the proposal for learner-initiated focus-on-form induced by strategic planning (Ortega, 1999, 2005), the present study has explored the effects of on-line planning in comparison with those of strategic planning. This partially replicated the study by Yuan and Ellis (2003) in the first task performance stage, while being distinct from earlier studies in its use of post-task interviews to understand the underlying psycholinguistic operations behind the different planning effects. Integrating the findings from these two research stages, Table 6.1 summarizes identified performance features and psycholinguistic processes in the different types of planning.

**Table 6.1:** Summary of Performance Features and Underlying Psychological Processes

Planning	Performance features	Psycholinguistic operations
NP	<ul style="list-style-type: none"> <li>• Maintaining fluency</li> <li>• Downgrading complexity</li> <li>• Downgrading accuracy</li> </ul>	<ul style="list-style-type: none"> <li>• Prioritizing lexical operations</li> <li>• Avoiding the intended message</li> <li>• Pressures on L2 processing</li> </ul>
SP	<ul style="list-style-type: none"> <li>• Limited effect on fluency</li> <li>• Enhancing complexity</li> <li>• Some improvement in accuracy (but not a significant level)</li> </ul>	<ul style="list-style-type: none"> <li>• Transferring the planned language to the production</li> <li>• Shifting attention from meaning during planning time to form in production</li> <li>• Engaging in monitoring</li> <li>• Pressures on L2 processing</li> </ul>
OP	<ul style="list-style-type: none"> <li>• Downgrading fluency</li> <li>• Enhancing complexity</li> <li>• Enhancing accuracy</li> </ul>	<ul style="list-style-type: none"> <li>• Engaging in the syntactic formulation</li> <li>• Accessing explicit knowledge</li> <li>• Engaging in monitoring</li> </ul>

The results of the task performance analysis indicated not only expected features supporting the previous studies but also unexpected planning effects. For example, the fluency variables in strategic planning did not outperform those in non-planning, which is not consistent with the positive fluency effects in strategic planning in the planning literature. However, the content analysis of learners' reports displayed the different psycholinguistic operations under the three conditions, even though the statistical outcome did not capture any clear difference of their performance. This chapter approaches these planned performances in terms of learners' reduction and compensatory strategies because these two strategies seem to epitomize the distinctive difference between the three planning conditions. In the following sections, first I present the reduction strategies which are observed in non-planning. Then, I discuss the compensatory strategies in strategic planning and on-line planning.

### 6.2.1 The Reduction Strategies in Non-Planning

It has been widely assumed that NNSs differ from NSs in terms of the nature and use of strategies for dealing with communication problems, and this function can be regarded as one of the unique processing features of L2 (Bygate, 1998). In contrast to the lack of fluency improvement in strategic planning, learners were able to maintain the level of fluency in non-planning, but the underlying reason is their active engagement in reduction (or avoidance) strategies (Færch & Kasper, 1983). According to Færch and Kasper (1983), there are two types of reduction strategies, i.e., ‘formal reduction strategies’ and ‘functional reduction strategies’, both of which are further divided into several micro-strategies (Table 6.2). The former involves learner’s employment of a system that has been phonologically, morphologically, syntactically or lexically reduced, while the latter concerns the conceptual reduction adopted to avoid a communication breakdown.

**Table 6.2:** Reduction Strategies (Poulisse, 1993:161, based on Færch & Kasper, 1983:52-3)

<b>Formal Reduction Strategies</b>	Phonological
	Morphological
	Syntactic
	Lexical
<b>Functional Reduction Strategies</b>	Topic avoidance
	Message abandonment
	Meaning replacement

The findings of the task performance analysis are more concerned with formal reduction strategies, in which, for example, speakers tended to neglect the past tense formation in non-planning, even though many of them noticed the necessity of past

tense use at the time of the instruction given. On the other hand, the findings of the verbal report analysis give insights into their employment of functional reduction strategies.

Integrating the two research stages, Table 6.3 shows examples of reduction strategies by taking speaker [C7]’s (high proficiency level) NP production and her summarized comments on her thought processes in specific parts of the production. The production appears to be smoothly transmitted with a limited number of hesitations and inaccurate utterances, but speaker’s comments reveal her frequent employment of certain reduction strategies.

**Table 6.3:** NP Production and Speaker [C7]’s Commentary on Thought Processes

C7: NP	Speaker [C7]’s commentary on thought processes
1. [one day two men were trying to steal a box (.2) at the airport] (.6) um (.3) two men um (.3) they are talking to the (.4) er man from the Bombay from Bombay (.5) and he had a two boxes (.7) and (.5) and	Noticing incorrect production (“bags” not “boxes”) but not correcting it [ <b>formal reduction: lexical</b> ]
2. (.5) one of the man (.3) er a whi- while one of the man (.9) were speaking to the (.5) man (.3) er the other (.3) man (.5) um s-	Monitoring the choice of pronoun; noticing the ambiguities of the characters but not correcting it [ <b>functional reduction: message abandonment</b> ]
3. (.5) stole his- um basket (.4) and (.5) they ran away (.3) and	Noticing that the intended message was not produced but not adding it [ <b>functional reduction: message abandonment</b> ]
4. (.7) um the man from India (.3) found er they stole their luggage (.4) so they he call the police man (.4) but they (.4) drove off (.3) by his er (1.1) in hi- (.3) in their car	Noticing the incorrect use of preposition (“by”) and correcting it
5. (.5) and after that they (.3) a open the basket in the park (.3) and they found (.3)	Thinking which of “large” or “big” should be used and reformulating it

---

a large (.4) big snake (.6) er in the basket  
(.5) and [were] very shocked

---

In three out of the five cases, speaker [C7] used the reduction strategies by neglecting to take any action to solve the linguistic problems, despite noticing them during on-going production. This performance seems to show a typical response in many high proficiency speakers, i.e., maintaining a good level of fluency and accuracy as shown by the statistical analysis, but speakers' reports demonstrates that a number of problems actually occurred under the surface of a good level of delivery. In particular, it is conceivable that high proficiency speakers manage to deal with the linguistic problems to maintain the level of fluency and accuracy by employing various formal and functional reduction strategies. Referring back to the output hypothesis (1.4), these findings explain the likely conditions that lead to failure to induce learners to engage in syntactic processing, being not automatically triggered by the output opportunities. Another possible reason for reduction strategies in non-planning is that memory problems might also have played a role, since participants no longer had access to the picture story during production. Thus, there is a possibility that occurrences of reduction strategies would not have been fewer if they had been allowed to look at the picture. However, as will be mentioned, on-line planning participants were more likely to engage in compensatory strategies rather than reduction strategies (6.2.2). As participants in on-line planning were also not allowed to look at the picture, it is reasonable to say that the time pressure in non-planning had participants choose reduction strategies.

It is worth mentioning that frequent use of reduction strategies in NP might have come from memory problems since participants were not allowed to access the picture story during production. In addition, this possibility may apply to primary

focus on content in NP (5.4.1.3) and SP (5.4.2.5). It may be the case that retaining the story may have continued to place heavy burdens on WM, which may have finally prevented learners from focusing on form. Although we cannot eliminate this possibility, the effect of this condition is uncertain, because there was no statement of this in any participants' report. Thus, it may be meaningful to follow up the effect of with – and without picture stories during production as possible task implementation in order to investigate how this condition affects the degree of focus-on-form and choice of communication strategies.

### *6.2.2 Compensatory Strategies in Strategic Planning and On-Line Planning*

It is important to note that reduction strategies are not the only choice available when a speaker encounters linguistic problems. Taking Levelt's (1989) speech production model as a theoretical foundation, Poullisse (1993:179) argues that, when the intended lexical items cannot be accessed from the mental lexicon, a signal is sent to the Conceptualizer and to the Monitor, saying that part of the preverbal message cannot be encoded. In such a case, there are usually three options that can be taken; message abandonment, appeals for assistance and compensatory (or achievement) strategy use (ibid.). As mentioned earlier, the first, the message abandonment (or reduction) strategy is frequently applied in non-planning, but the second, appeals for assistance is not a useful tool in the present monologic type of task; the third, compensatory strategy is not preferred in many non-planning occasions, but is more common in strategic and on-line planning.

Among a variety of communication strategies (e.g., Bialystok, 1990; Færch & Kasper, 1983; Tarone, 1980), the present study particularly focuses on the



framework proposed by Poulisse (1993), because of its psycholinguistic orientation built on Levelt's speech production model as well as its inclusion of other previous studies. Poulisse (1993) further specifies three different types of compensatory strategies: substitution strategies, substitution plus strategies and reconceptualization strategies (Table 6.4).

**Table 6.4:** Three Types of Compensatory Strategies (Based on Poulisse, 1993: 180-183)

<b>Substitution strategies</b>	<ul style="list-style-type: none"> <li>• Approximations (or generalizations)</li> <li>• Borrowing (or code-switching)</li> </ul>
<b>Substitution plus strategies</b>	<ul style="list-style-type: none"> <li>• The out-of-the-ordinary application of L1 (e.g., foreignizing)</li> <li>• L2 morphological and/or phonological encoding (e.g., morphosyntactic creativity)</li> </ul>
<b>Reconceptualization strategies</b>	<ul style="list-style-type: none"> <li>• Encoding the conceptual features of the intended lexical item</li> <li>• Selecting two lexical items</li> <li>• Adding further background information</li> <li>• Gesturing</li> </ul>

'Substitution strategies' involve replacing the intended lexical item with another word, which can be either a related item (i.e., approximations; e.g., *animal* for 'rabbit') or the corresponding L1 item (i.e., borrowing; e.g., Japanese *matsuri* for 'festival'). The second, 'substitution plus strategies' are, as their name suggests, used to combine with substitution strategies, involving both L1 and L2 lexical items. The third type, 'reconceptualization strategies' can be defined as a change in the preverbal message involving more than a single chunk (Poulisse, 1993:181). As noted by Poulisse (1993:181-2), although all these approaches are similar in terms of changing the preverbal message in Levelt's model, most reconceptualization strategies (except for

'gesturing') must be considered more cognitively demanding than the other strategies from a processing point of view because the substitution strategies concern only one or two word alterations, while the reconceptualization strategies involve the substitution, addition or deletion of entire chunks, affecting the whole structure. Taking these features of compensatory strategies into account, the content analysis of verbal reports suggests the underlying reasons for adopting the reduction strategies in non-planning (and natural interaction): pressures continued to push speakers into prioritizing lexical operations and often hindered their engaging in syntactic formulation and monitoring the grammatical accuracy throughout the speech.

Certainly, the existence of processing pressure accounts for, at least partially, the neglect of the compensation, particularly reconceptualization strategies, in non-planning, but this does not capture the whole picture observed in the present research because of the lexical orientation of Poulisse's reconceptualization strategies. The first choice of her reconceptualization strategies partially pertains to speaker's engagement in syntactic formulation, but it seems to concern only simplified structures, which functions a role in compensating for a lack of lexical resources. It is conceivable that there is another choice when encountering the lexical deficit problem – constructing a more complex structure, which is believed to contribute to stretching the IL system.

Swain and Lapkin (1995) states that 'pushed output' can not only be externally induced by an interlocutor giving feedback about communicative problems (e.g., Pica *et al.*, 1989) but also internally induced by self-monitoring and self-correction. As the present research design does not include external feedback (i.e., it is monologic), the main focus has been on the other aspect, i.e., internal feedback: more specifically whether strategic and on-line planning implementations induce monitoring and self-correction.

In contrast to frequent reduction strategies in non-planning, because parts of the necessary language production (particularly generating the messages and searching for the necessary lexical items) are completed during the ten minutes planning time, learners in strategic planning tended to attempt more ambitious language beyond their current linguistic level, and as a result of this, the production showed more complex performance than in non-planning. Processing conceptual and lexical aspects often pushed their attention to formal aspects in production, showing a tendency towards accuracy improvement, but the level of improvement did not reach a significant level; presumably because, despite a pre-task planning time opportunity, the pressured condition continued to hinder speakers' full access to the grammatical system. It is also, as suggested by Robinson (2001a:37), possibly due in part to learner *avoidance* of problematic forms and to a narrowing of their productive repertoire to tried and trusted forms during the planning time.

The on-line planning production was characterized by more complex and accurate but less fluent language, implying that reduced on-line processing pressures contributed to the speakers' direct engagement in linguistic processing. Encouraged by the on-line planning instruction, learners were more readily engaged in on-going monitoring processes at the expense of fluency-related components, although the reported reasons for monitoring do not seem to be consistent between different proficiencies (see 5.4.3.2).

In order to illustrate the difference of performance and the underlying psycholinguistic processes between the two types of planning, I examine one participant's production and how her performance and psycholinguistic processes changed by taking strategic and on-line planning production. The performance and verbal data are taken from the same speaker (C7) as in the non-planning production (Table 6.3) in order to make the three occasions comparable.

Similar to the non-planning production, the speaker seems to maintain the level of fluency and accuracy in strategic planning, but her comments on thought processes tell a number of problems she perceived in speech (Table 6.5).

**Table 6.5:** SP Production and Speaker [C7]’s Commentary on Thought Processes

<b>C7: SP production</b>	<b>Speaker [C7]’s commentary on thought processes</b>
1, [a one day a boy dropped a small parcel on the way home a way to home] (.6) er (.3) it was already da- um:: (1.0) after dark	Noticing that the planned message was not produced but avoiding the formulation
2. and a (.3) he had to walk um [in] a (.7) dark (.3) lonely lane (.3) to his home (.5) so he walked	Forgetting to apply the planned word (“alone”)
3. very fast (.6) and (.3) after [a] while he found someone following him (.3) um with something in his hand (.5) and he was- um the boy was so shocked and scared and he	Correcting a personal pronoun (“he” → “the boy”) to avoid ambiguity for the listener
4. started running (.6) but (.4) the man also started running after him (.7)	Noticing that the planned content was not produced; monitoring the appropriateness of the word (“man”)
5. and (.2) finally um the man (.3) reached him (.2) and the boy was so scared but he turned around to see his face (.5) and (.8) um (1.8) but (.5) then	Thinking how to tell this part during planning time but not finding a good way; formulating the structure to describe the intended message
6. he found um he is a (.2) he was a very nice man and he <u>try</u> to:: (.3) um give (.3) the boy er give him (.5) the parcel he:: (.3) dropped	Monitoring the use of pronoun and changing from “the boy” to “him”

It is assumed that the speaker produced the language based on her planned elements, but there is also some evidence that planning did not necessarily push her into

syntactic formulation, as illustrated in the first and fourth comments. As suggested in Chapter Five, many speakers attempted to formulate the language on-task by planning the concepts and identifying the necessary words, but there is also a possibility that the formulation process is, nevertheless, avoided in production. An interesting example is in the fifth comment, in which the speaker is trying to explain the story by formulating the structure. She could not find the way to describe this part during the planning time, and attempted to formulate the language in production. A number of filled and unfilled pauses are also suggestive of her struggle to formulate the language.

The on-line planning production reveals a clear distinction from the other conditions in terms of the speaker's more active engagement in formulating the structures. In Table 6.6 below, the first and second comments illustrate that the failure to identify the target word led to reconceptualizing the message and formulating the language.

**Table 6.6:** On-Line Planning Production and Speaker [C7]'s Commentary on Thought Processes

<b>C7: OP production</b>	<b>Speaker [C7]'s commentary on thought processes</b>
1. [one day three boys were waiting for a bus on the street] (.7) and (.2) a while they are waiting (.3) a big truck (.3) pass by (.4) er (1.4) the pass by the sidewalk (.5)	Failing to identify the target word ("puddle"), so that reconceptualizing the message; identifying the necessary word
2. and a (.4) um there's a (2.1) um (.2) yesterday they had a rain (.3)	Reconceptualizing the message and formulating the necessary structures
3. and a (.8) um (1.6) there's a (.5) some waters beside the street (.7) and the truck just splashed (.4) er the water to the passengers er to the (.3) people (.3)	Noticing and correcting error ("passenger" → "people")

<p>waiting for the bus (.7)</p> <p>4. and (.2) they got um the boys' (.3) got the water (.7) and (1.1) after that (.3) er the bus (.2) they are (.3) they were waiting (.3) er (1.2) came (.5) and (.4) but (.8) the (.5) some boys (.5) um (1.4) who came after them (.5) um (.3) skipped their tur- (1.5)</p>	<p>Thinking how to formulate the structure by using the identified word ("turn"); reformulating the structure</p>
<p>5. um (.7) skip their:: turn? (2.9) and a (4.6) um (.7) er the four boys (.2) um (.7) just (5.0) um (1.5) just came (1.1) um (2.8) in front of the three boys and they just ignore (.3) um (3.2) um:: (.8) ignore (2.5) ignore the three boys (2.5) the- (.3) they came (.4) earlier and wait- waited for the bus before them (.8)</p>	<p>Monitoring the lexical items; thinking whether "boys" are clear for the listener</p>
<p>6. and (.7) when they (.4) er when the three boys try to:: get on the bus (.2) um (.5) the conductor told them (.3) er (.2) there's (.4) too many people on the bus already (.2) and they (.3) they have to wait another bus (1.5) s- th- so</p>	<p>Noticing the error ("have to") but not correcting it</p>
<p>7. they have- (.3) they had to wait (.4) for another bus for (.2) thirty minutes (.6) and (.5) finally they get (.3) er they got on the bus (.3) and (1.9) then (.2) er (.2) on (.2) on their way home (.3) they found (.4) the bus they:: (.2) couldn't (.3) um (.5) er which they couldn't (.5) er get on (.6) had a accident (.3) er had the trouble (.3) and a (.5) they</p>	<p>Monitoring the choice of lexis ("accident") and changing to an alternative word ("trouble"); conceptualizing the message and formulating the structure</p>
<p>8. (1.3) um the driver and (.3) conductor (.4) try to fix the bus (.5) so (1.0) after all (.3) er the three boys (2.5) um could er came (.2) come home (.2) earlier (1.0) than (.2) the other (.6) passengers who got on the (.3) first bus</p>	<p>Thinking how to describe the intended message but not finding a good way to produce it; reconceptualizing the message and formulating the structure</p>

The comments on her formulation are also observed in the fourth, seventh and eighth rows in the table. Regenerating the part of speech in the seventh comment, this includes her attempt to reconstruct the utterances in this unpressured condition:

and (1.9) then (.2) er (.2) on (.2) on their way home (.3) they found (.4) the bus they:: (.2) couldn't (.3) um (.5) er which they couldn't (.5) er get on (.6) had a accident (.3) er had the trouble

It can be assumed that this part of the story could not be easily verbalized by using her lexical resource, so that the speaker attempted to construct the language by employing relative clauses. A number of filled and unfilled pauses underscore her active engagement in formulating the structures to describe her intended message. As a result of this effort, the speaker produced the complex structure involving several dependent clauses. Despite such incompleteness of her production, it is obvious on this occasion that the unpressured on-line planning led her to formulate the message.

### *6.2.3 Shifting into Deeper Syntactic Processing*

It is now necessary to look back on the question, presented in Chapter Two, of the conditions eliciting learners' syntactic processing, more specifically of the differences between strategic planning and on-line planning effects on form; despite the theoretical similarities of released processing pressures, why did strategic planning and on-line planning bring about different qualities of performance? As argued by a number of researchers (e.g., Sinclair, 1991; Skehan, 1998; Widdowson, 1990; Wray 2002), it is conceivable that L2 speakers switch between the exemplar-based and rule-based systems, and that the former system is prioritized and the latter system is secondary in spontaneous speech. The analysis of verbal reports not only supports this

conjecture but also clarifies the problem of this natural tendency in spontaneous speech. Because the rule-based system is the second priority, it tends to be avoided even when access to the system needs to be gained. It is true to say that the pressured condition is not natural but experimental in the present study (due to the limited production time required), but a similar sort of processing pressure is expected to hinder speaker's shifting to the rule-based mode in natural settings as well. Due to an insufficient amount of lexical items stored in L2 lexicon (Poullisse, 1997), it is often the case that learners need to construct the structure to represent the intended message, but the pressured situation often forces them to select reduction strategies.

According to Poullisse (1993:184), speakers who are confronted with lexical problems will adhere to two general principles, the Least Effect Principle (i.e., the speaker should use communication strategies which require the least processing effort) and the Cooperative Principle (i.e., the speaker makes sure that his/her communication strategy is comprehensible to the interlocutor). Ideally, the speaker attempts to use a communication strategy that requires little effort and is also comprehensible, but there is a trade-off between the two principles. Following this assumption, the non-planning performance seems to lean to the least processing effort, as observed so far.

Consistent with Ortega's (1999, 2005) observation that strategic planning time enables a conscious shift to learner-initiated focus-on-form, the present study revealed focus-on-form effects in strategic planning but to a lesser extent than those in on-line planning. There is a tendency for pre-engaging in conceptual and lexical planning in strategic planning to induce learners to process syntactic aspects, but still existing processing pressures often block their full engagement in syntactic formulation, when the necessary lexical items are not present in their resources. A possible danger is that the pre-task planning opportunity sensitizes the learner to the



limitations of what he/she can say (Robinson, 2001a). That is, a realization of the lack of lexical resources may entice the learner to avert from the challenge of syntactic formulation to compensate for the resource deficit. On the other hand, the content analysis of on-line planning illustrates that the unpressured condition tended to push learners in not only increasing the grammatical accuracy (particularly in correct verb formation) but also the syntactic formulation affecting the linguistic complexity. In addition, both strategic and on-line planning led to a greater engagement in monitoring processes than non-planning in terms of the number of reports (Table 5.1), but the content analysis also demonstrates a more profound level of monitoring in on-line planning. As suggested by Izumi (2003:184), such engagement in grammatical encoding and monitoring can serve as an 'internal priming device' for grammatical consciousness-raising for the language learner, which is considered to trigger restructuring in learners' IL systems.

A probable distinction caused by strategic planning and on-line planning can be explained by 'depth (or levels) of processing' ( Craik & Lockhart, 1972) in cognitive psychology. According to Craik and Lockhart (1972:676), speed of analysis does not necessarily predict retention; rather, retention is a function of depth, and various factors, such as the amount of attention devoted to a stimulus, its compatibility with the analyzing structure, and the processing time available. That is, if learners are involved in more elaborate and deeper level of linguistic analysis, manipulated linguistic knowledge is more likely to be stored in LTM with a more durable and stronger state. Thus, deeper processing is likely to lead to more learning.

The original proposal of depth of processing concerns an aspect of input processing and is primarily applied to vocabulary learning research in SLA (e.g., Hulstijn, 2001; Laufer & Hulstijn, 2001). However, it can be assumed, taking it together with the output hypothesis, that deeper processing is expected in output more

than in comprehension. In a study investigating the noticing effects by input enhancement and an output production task and finding more positive effects in output, Izumi (2002:569-70) observes that the former may have caused mere recirculation or rehearsal at the same, relatively shallow processing level, which led learners to experience only a short-term retention of the attended form, while the latter triggered deeper and more elaborate processing of the form, which led them to establish a more durable memory trace.

Extending this finding, as argued by some researchers (e.g., Bygate, 1999; Izumi, 2003; Skehan, 1998), and considering also the analysis of non-planning production in the present study, there may be further distinctions of processing levels within 'deeper processing' of output production. As it has been argued so far, on-line planning more readily encourages learners to be engaged in syntactic processing (i.e., more attention and time is devoted to syntactic processing available on-line), and it may trigger deeper and more elaborate processing, which may lead to different learning effects from the other planning conditions.

As Craik and Lockhart's (1972) depth of processing was challenged for the ambiguities of its construct, the present study provides only partial evidence to support its application to the differential benefits of strategic planning and on-line planning. Thus, it does not go beyond the level of speculation, but it may be worth mentioning possible differences of depth of processing under different planning conditions in order to support the view concerning different degrees of focus-on-form effects brought by different planning types.

To summarize, strategic planning and on-line planning tend to bring shared as well as different benefits to L2 learning. As suggested by the present study as well as the previous planning literature, strategic planning may contribute to developing a learner's ability to produce language in a balanced way, but the on-going processing

pressure during performance mitigates learner's attention to formal aspects. This may provide important opportunities to conduct language processing under communication pressure. On the other hand, on-line planning may not be a preferred way for some speakers, because of unnaturally dysfluent production, implying that a lack of syntactic formulation often occurs, although this attempt arguably leads to processing the language at a deeper level. Although the ways of implementing on-line planning may need more consideration in the language classroom, some kind of 'on-line planning implementation' may be needed for developing learner's ability in syntactic formulation, as will be considered in Chapter Seven.

### **6.3 Learner Proficiency**

The effects of different types of planning have been discussed so far. In addition to the effects of planning, learner-related factors seem to affect the quality of planned performances. In this part, therefore, I will consider the effects of different proficiency levels of L2.

As has been widely discussed in SLA, there seem to be various psycholinguistic features to distinguishing proficiency levels. For example, Philip (2003) suggests two general factors modulating constraints on noticing in different proficiency levels: 'automaticity' and 'readiness'. With the former, more experienced learners may benefit from the increasing automaticity that comes with repeated practice, which allows attentional resources to be focused on higher order aspects of speech processing (1.5.5). In the present research, it is likely that high proficiency speakers are able to focus on conceptual rather than grammatical monitoring, as observed in the verbal report analysis (5.4.3.2). Another related area, readiness, is concerned with proficiency in terms of processing mechanisms and prior knowledge; i.e., it is not until the learner is ready in terms of psycholinguistic processing that

particular lexical and grammatical forms are retrieved at a high level of awareness and/or the forms need to be sufficiently encoded in LTM in order that the learner can recognize them (Philip, 2003:104-5). That is, high proficiency speakers tend to notice more developmentally advanced linguistic items than low proficiency speakers, which was illustrated as an improvement in the use of articles in the high proficiency group and that of verb-formations in the low proficiency group in the task performance analysis.

The subsequent section firstly considers the relationship between planning and proficiency by restating the previous relevant studies and the findings in the present research. Then, following on from the SLA discussion on learner proficiency factors, I focus on the two proficiency-related aspects; the level of psycholinguistic processes such as automatization and lexicalization, and psycholinguistic/developmental readiness. Synthesizing these issues affecting their planned performance, I shall propose the importance of attentional allocation as one of the distinctive differences between the different proficiency levels.

### *6.3.1 The Relationship between Planning and Proficiency*

As learner proficiency may affect in important ways how much of what is planned is actually reflected in the linguistic outcomes of planned performance (Ortega, 1999:136), recent planning studies are likely to incorporate the proficiency factor as an independent variable (e.g., Kawauchi, 2005; Ortega, 2005; Wigglesworth, 1997). While there are more positive planning effects in high proficiency speakers reported in Wigglesworth (1997), the planned performance of the advanced learners did not outperform that of high intermediate learners in Kawauchi (2005). We cannot draw any conclusion from the limited number of empirical studies investigating the

relationship between planning and proficiency as of now, but these inconsistent findings suggest the complexity of the learner proficiency factor.

Having looked at the findings of the task performance analysis, the statistical outcomes support the correlation of planning and proficiency advocated by Ortega (1999, 2005). However, there seem to be more complex relationships between different proficiency levels and different planning types than was originally assumed. The effects of strategic and on-line planning seem more limited in high proficiency speakers, as suggested by Kawauchi (2005), because both strategic and on-line planning had positive effects only on complexity (i.e., no positive effects on accuracy and fluency in both planning conditions). On the other hand, in the low proficiency group, there was an improvement in fluency in strategic planning, and in complexity and accuracy in on-line planning (4.2.4.3). A possible reason for the limited effects in the high proficiency speakers is that, as these advanced learners reached the ceiling of their performance to a certain extent even in non-planning, there was only a little room left for gains in strategic planning and on-line planning. However, it is premature to suggest a lack of improvement of strategic and on-line planning in high proficiency speakers because, despite a failure to reach significance, there is a tendency towards a relative improvement of article usage. Another noticeable point is that the effects on complexity are equivalent between strategic planning and on-line planning in high proficiency learners. Together with the result of negative fluency effects in on-line planning over strategic planning, it is conceivable that the effects of on-line planning on high proficiency speakers are very limited. On the other hand, there was some fluency improvement in strategic planning, and complexity and accuracy effects in on-line planning in low proficiency speakers. Integrating these analyses, it can be assumed that strategic planning tends to affect formal aspects in high proficiency speakers (though in a limited way) whereas on-line planning tends to

more readily contribute to enhancing complexity and accuracy in low proficiency speakers.

### *6.3.2 Some Psycholinguistic Evidence of Different Proficiencies: Automatization and lexicalization*

In addition to the issues directly linked to planning, some verbal reports provide evidence of qualitative psycholinguistic differences affecting learner proficiency. Although production practice is a target of criticism in some theories (e.g., input-processing by VanPatten [1996]) because of its ignorance of psycholinguistic constraints of learners' own development (Pienemann, 1985) and association of repetition of mechanical manipulation towards targeted grammar items, there is a growing body of research to support the idea that automatization is key to developing L2 fluency (e.g., DeKeyser, 1997, 1998, 2001; Johnson, 1996; Segalowitz, 1991, 2003; see 1.5.5). This section looks at some verbal evidence of learners' degree of automatization and associated effects of lexicalization, both of which are vital for efficient L2 transmission.

One clear distinction between the two proficiency groups is reports linked to automaticity. A trend towards higher degree of automatization of linguistic processing can be more commonly observed in high proficiency speakers. In the following excerpt, a comment by speaker [C3] implies her automatized state of syntactic processing:

#### **Episode 6.1**

C3: Yes, I was losing the picture in mind here. So, I was making it on my own. Yeah.

R: You mean you made your own story?

C3: Story...? I was making it while thinking 'probably it is like this'.

R: So, did you make it?

C3: Yes, I made it by myself [laughs].

R: Did you make only the content or the language as well?

C3: The language... I think I wasn't thinking of the language very much.

R: Did you think of the content mainly?

C3: Yes. But, um, of course I sometimes thought things like 'is this sentence right?', but I was not very conscious of making the language.

R: You feel like thinking of the content, do you?

C3: Yes, yes, yes. (C3OP10:HIGH)

As the speaker could not recall one of the pictures, she invented her own story, building on partially remembered information. An important point is that she was conscious of the message generation but less conscious of the linguistic construction, suggesting that her linguistic formulation is largely automatized. It is also noticeable that she commented on her engagement in monitoring the language and the generated story, which shows that the automatized linguistic processing seems to allow her to monitor her own production. This observation is consistent with the psycholinguistic model of distribution of attention that, when knowledge becomes more automatic and proceduralized, much less attention is spent on lower-level skills (e.g., the Formulator, the Articulator), and more attention goes to higher-level skills (e.g., the Conceptualizer; de Bot, 1996:546).

In addition to this indirect account of the automatized state of mind, there is more direct evidence of the strong connection between a high level of automaticity and proficiency. In the next excerpt, speaker [C8] reveals automatic emergence of pre-fabricated structures:

### **Episode 6.2**

C8: This is, well... at this time, I was thinking which phrase I should choose from "the boy realize that there... there was man who... there was man following him" or "the boy realize that the man was following him?" um...

then, why did I choose the latter one?

R: Did two sentences appear in your mind?

C8: Yes. When talking about this picture, I got these two sentences, but “there was the man following him...” um... “there was a man following him...” or “there was a man... who was follow... who has been”... If I choose “who has been following him”, I feel like this is boy’s point of view, but “there was ...” um if I say “he realize that a... the man the man was following him”, that would be our view point, wouldn’t it? This is the view of those who look at the picture, isn’t it? So, after all, I think I said “the boy realize that... the man was following him”... um, this is not right? [Then listening to the following speech]

*[he is (.5) he’s been followed or (.9) followed by this (.7)]*

C8: I was thinking a number of things at this time...and I was confused with the point of view...then, finally I abandoned both and said “he’s been following...”

R: See, you abandoned both...

C8: Yes, yes. I said “he’s been followed”. I wanted to include HIS point of view here. (C8OP5:HIGH)

Initially, the speaker spotted two different types of structure at the time of production, but she was not very comfortable with either candidate because the choice of the subject, or the viewpoint of story-telling, did not match what she thought was appropriate in this context; then, she abandoned both and finally applied a new structure. Looking at the production which she reported, relatively lengthy pauses and some hesitations can be identified, but this also shows that such various thoughts occurred within a very limited amount of time: *[and then the boy (1.9) realize that (1.5) he is (.5) he’s been followed or (.9) followed by this (.7) man (1.3)]*. A particularly noticeable point is that her formulation process attains a high level of automaticity, allowing her to prepare three different structures efficiently. Most importantly, these appeared in semi-structured, pre-fabricated forms rather than rapid, step-by-step construction of structures, illustrating that pre-fabricated, lexicalized structures ease on-line processing and enhance the level of automaticity (e.g., N. Ellis,



2001; Foster, 2001; Wray, 2002). As L2 speakers become more proficient, 'there is a shift from simple [morphosyntactic] error repairs to more complex discourse-level repairs' (Kormos, 1999:332), because as a result of automatizing these lower processes, more attention can be strategically allocated to such higher functions. The excerpt reported by speaker [C8] supports this conjecture that she could pay primary attention to the discourse element, the impact of thematic structure (Bloor & Bloor, 2004; Halliday & Matthiessen, 2004), without thinking of the linguistic processes.

As suggested by these examples of high proficiency speakers, automatizing and lexicalizing large parts of formulation are essential for efficient speech production in skill development perspective, because the rest (i.e., higher levels than morphosyntactic planning) largely remain nonautomatic. It is interesting to note that the semi-structured forms in Episode 6.2 are similar to what Pawley and Syder (1983) call lexicalized sentence stems (LSS) involving a nucleus of lexicalized elements while leaving some parts changeable; that is, like NSs, speaker [C8] seems equipped with a number of lexicalized, semi-structured items in order to achieve the degree of real-time fluency. With a limited number of LSS, low proficiency speakers, on the other hand, are considered to consume considerable attention in searching lexical items (more likely single-word level) and constructing the structure, building on the spotted words in syntacticization. Because this step-by-step construction tends to consume considerable amounts of time, speakers may neglect syntacticization to meet communicative effectiveness, even when it is needed. Significant effects of on-line planning on complexity and accuracy in low proficiency learners might suggest that it allows them to syntacticize the intended message; on the other hand, limited on-line planning effects in high proficiency speakers might be a consequence of accumulated lexicalized items, as typically illustrated in speaker [C8].

Such a heavy burden on linguistic formulation particularly in low proficiency

speakers has a serious impact on speech production processes. According to the ‘one-clause-at-a-time constraint’ (Pawley & Syder, 2000:163), it is not possible for a speaker to encode novel lexical combinations across independent clause boundaries in a single planning act. In this psycholinguistic constraint, it is not until the time of production that speakers are able to plan their actual language choice, implying that a great part of linguistic formulation needs to be conducted efficiently within a very limited amount of time. This mechanism may explain the limited effects of strategic planning in the present research; strategic planning time could contribute to the overall conceptual, and possibly some parts of syntactic, planning, but usually does not have a direct impact on accurate use of morphosyntactic aspects. The ‘one-clause-at-a-time constraint’ may be relatively unproblematic for choice of words not to be inflected but would be problematic for exploiting inflectional words because speakers, even if they engage in syntactic planning during strategic planning time, cannot conjugate the exact word formation until they actually start speaking and approach the point of morphological planning.

### *6.3.3 Psycholinguistic and Developmental Constraints on Focus-on-Form*

A clear distinction between high and low proficiency groups between accuracy scores in verb forms and those in articles (4.2.4.3) might indicate the influence of developmental stages among the participants. The morpheme studies in the 1970s provided strong evidence of common accuracy-acquisition order in ILs, but sufficiently consistent general findings for the order have been ignored (Larsen-Freeman & Long, 1991:92). Krashen’s (1977, 1982) Natural Order Hypothesis (based on Dulay & Burt, 1973, 1974; Fathman, 1975; Makino, 1979) proposed that regular and irregular past would be acquired later than articles,

irrespective of learners' L1 backgrounds, age and the medium (i.e., writing or speaking). However, this is not consistent with the findings of the present study as well as our experience of observing Japanese learners of English (Kobayashi, 2002b; also see Goldschneider & DeKeyser, 2005 for a recent review and a meta-analysis of a natural order of L2 morpheme acquisition research). In the following sections, I will discuss the findings of verb formation, especially in terms of tense consistency, and articles respectively.

**Tense consistency.** Learning regular past tense forms (e.g., 'walked', 'talked') is not linguistically complex, as suggested by the fact that this is usually taught at beginner's level in Japanese English education. To exemplify their understanding of the past tense formation, numerous participants reported their failure of the correct formation despite noticing of the necessity of its application. As suggested by Godfrey (1980), once a speaker selects a particular tense to narrate events at the beginning of a story, there is an obligation to continue in this tense sequence unless the sequence of events is interrupted in some way. The following example by a speaker illustrates his tense shifting violates Godfrey's temporal constraint (NB: verbs are underlined):

and then (.6) he:: (1.6) he open the box (.4) but (.8) all (.3) all they found (.7)  
in box (.2) is (.6) a big (.8) big snake (.8) and the (3.3) and he was so:: (.8)  
scared (.4) and the (1.5) his stealing (.6) is (1.4) er (3.3) er (2.3) and they  
(1.0) they are so disappointed (1.0) um (2.2) not to:: (3.0) no (.6) not to get  
(2.8) something (1.3) something rich (.9) er (1.1) er (1.0) (A6SP8-14:LOW)

The storyline of the example should continue the consistent past tense pattern as initiated by the first sentence embedded with the past, but the speaker busily moved back and forth between the present and the past without using 'a temporal or other adverbial to heighten the salience of the change in temporal reference' (Godfrey,

1980:94). With respect to such kinds of tense inconsistency, Godfrey (1980:95) elucidates that L2 speakers must not only control tense formation but also retain and attend to the tense continuities they establish. In terms of an information-processing perspective, both inadequate control over tense forms and inattention to tense continuities lead to tense errors.

Another interpretation of tense inconsistency is that tense continuity is a syntactic problem; that is, an established tense may be semantically maintained without continuing to produce correct tense forms. For example, if both the speaker and the interlocutor share the same information on a particular time of an event (e.g., yesterday), they would probably assume that all the events in the same storyline should have happened in the same time setting, even with ill-formed tense morphemes (e.g., present tense). It is probably right to say that, after some time, learners are likely to shift their attention from past tense formation to other aspects of production such as conceptualization and other morphosyntactic processing. In the present study, identifying this negligence of correct past tense formation seems very important because maintaining tense continuity devours considerable attention in WM in order to access LTM and engage in grammaticization.

It can thus be considered that such variations in verb formation result from amount of attention, as argued by Tarone's (1982, 1983) 'continuum paradigm' that new target language forms will first appear in the most careful style and gradually move to the least-attention-to-form, 'vernacular style'. It seems more susceptible to the amount of on-line attention made by different types of planning; i.e., the most accurate by on-line planning, followed by strategic planning in low proficiency speakers.

**Article.** In contrast to rather simple rules involving English verb formations, there are numerous studies to suggest the difficulty of English articles for L2 learners, particularly Japanese learners. Despite the abundant input, English articles are a remarkably complex system which encodes semantic notions of existence, reference, and attribution, discourse notions of anaphora and context as well as syntactic notions of countability and number (e.g., Beaumont & Gallaway, 1994; Master, 1990, 1997; Young, 1996). For example, taking the 'One to One Principle' by R. Anderson (1984) that acquisition of a new form is facilitated when there is a clear and unique correspondence between form and meaning, Young (1996:136) suggests that this principle clearly does not apply with the English article system. In addition to these linguistic difficulties, the influence of L1 might make the acquisition of English articles more difficult. The study by Oller and Redding (1971) suggests that those learners who had articles in their L1 (e.g., Spanish, French, German, Hebrew and Portuguese) significantly outperformed on the article test those who had no formal equivalent article system in their L1 (e.g., Japanese, Chinese, Persian, Korean and Thai). Similarly, researching morpheme acquisition orders by different L1 learners, Larsen-Freeman (1975) also reported that the rank of articles in morpheme accuracy orders was lower for Japanese learners than for other learners. In a study investigating the impact of instruction on different grammatical areas, Pica (1985:214) suggests that the impact of instruction may be little for highly complex grammatical morphemes such as the article, opposed to the positive effects of instruction on the third person singular 's'. In contrast to a number of reports concerning verb formation (partly because of the researcher's question of noticing the past tense), no comments on English articles were identified, implying that speakers were unaware of the use of articles at any proficiency levels. What is interesting is, nevertheless, the use of articles has shown a relative improvement in strategic and on-line planning in high

proficiency speakers but no difference in low proficiency speakers. It is premature to draw any conclusion from the findings because none of them reached significance; but there is a possibility that processing space created by planned conditions may contribute to improving the article system in high proficiency speakers, who are developmentally ready.

#### *6.3.4 Strategic Attentional Allocation in Different Proficiency levels*

It is not straightforward to draw any decisive reason for the different effects of planning in different proficiency levels, but, in addition to the factors which I mentioned so far, one of the possibilities is the different degree of attention required between strategic planning and on-line planning. With respect to the relationship between proficiency and communication strategies, Paribakht (1986:141) observes a directionality of transition in the learners' use of communication strategies towards that of NSs, as the learners become more advanced. However, because a high degree of automatization allows learners to pay attention to higher order production stages, the change of communication strategies is partly a consequence of available attentional resources. As has been repeatedly observed so far, particularly in the findings of the content analysis of verbal reporting, learners tend to suffer from processing problems more in strategic planning. Therefore, it is reasonable to suggest that, as in the original theory construction, on-line planning may release them more readily from the processing burden. In the present findings, strategic planning sufficiently provides cognitive space to process the language for high proficiency speakers, while low proficiency speakers need more processing space, which is more successfully given by on-line planning. Therefore, particular planning choice may differently affect learners' use of attention at different proficiency levels.

Together with the analysis of the number of generated themes in the post-task interviews (Table 5.1), the content analysis of learners' episodes suggests a tendency towards their active engagement in formulation and explicit focus-on-form in the on-line planning production. The ten-minute planning opportunity also tended to lead speakers into grammatical processing in the following production, but this did not necessarily guarantee their conscious awareness of formal aspects. Therefore, it could be argued that there are inconsistent tendencies between strategic planning and on-line planning despite the similarities caused by created cognitive space. An episode reported by one of the low proficiency speakers (in SP) might show a clue to understanding the differences between strategic planning and on-line planning:

### **Episode 6.3**

A2: While speaking, I had the content in mind, but once I started thinking of grammar, I couldn't maintain my speech production anymore. While thinking of grammar, I had no picture in mind, and when I thought 'I find out the grammar!', then the picture appeared in mind. And the next time I thought of grammar, I couldn't keep the picture in mind. When the picture appeared this time, I noticed 'I said the same thing twice'. (A2SP14:LOW)

This episode illustrates the competing state of meaning and form in the learner's mind, and her failure to pay simultaneous attention to both aspects. While thinking of grammar, speaker [A2] could not maintain her attention to the picture, and as a result, she described the same picture twice without noticing the repetition. It is true to suggest that pre-task planning opportunities lead to making cognitive space by processing some parts of language production prior to the speech, but on-line processing continues to put pressure on their linguistic processing, particularly grammatical processing. As argued by Ortega (2005), language expertise seems to have filtered the nature of benefits afforded by planning. In this respect, an interesting point in this episode is that focusing on grammar pushed out the meaning aspect from

her consciousness, suggesting that accessing explicit knowledge consumes a great amount of attention at this proficiency level. It can be thus assumed that cognitive space created by planning time would not always lead to focus-on-form for those who need much processing for grammatical accuracy.

It is also possible that some speakers might not be engaged in grammatical processing in on-line planning. Speaker [B3] explains the reason for her lack of focus-on-form despite the on-line planning instruction:

#### **Episode 6.4**

B3: Yeah, probably. I'm pressured [in production], so dealing with the past...or grammar is the least priority for me. Probably, the first is the information I want to tell, then search the phrases to present this, and finally grammar comes probably [laughs]. If I process the first two parts smoothly, I'm able to pay attention to the past tense, but if I stop one of these, I'll go to the next part of production, even if I notice the grammatical necessity.

R: You're going ahead...

B3: I'm going ahead. For this reason, I think I skipped [the grammatical processing]. (B3OP9-13:LOW)

This speaker clearly states that the grammatical aspect is regarded as the lowest priority; thus, focus-on-form can be achieved only when other processing such as conceptual information and lexical searches have been completed, which corresponds to the hierarchical stages of speech production proposed by Levelt's (1989) model. Despite the encouragement of focus on grammatical accuracy in the on-line planning instruction, some speakers might continue to prioritize meaning over form.

These two episodes shown above are suggestive of how focus-on-form is actually processed in L2 speech. Focusing on particular formal aspects in one's own speech production does not mean exclusive focus on form without paying attention to meaning. The speaker still keeps attention to meaning while focusing on form. To put it differently, focal attention is paid to the target form while maintaining the



conceptual information active on WM, or what Ortega (2005) calls ‘form-in-meaning’.

From her analysis of retrospective interviews, Ortega (2005:106) demonstrates:

...learners seemed to pay attention to the inextricable relationship between form and meaning simultaneously holding in long-term memory considerations regarding the message to be conveyed and the essential formal resources to convey it.

This statement may not be the case in mechanical grammar exercises, because overriding attention to form could be possible without understanding the meaning of the targeted language, in particular pedagogic contexts. However, in TBLT, some focus on the content should always be retained. Therefore, the principal question is not how focus-on-form can be achieved but how simultaneous attention to form *and* meaning can be achieved.

This theorization leads to the necessity of understanding of the importance of strategic attentional allocation in speech production in terms of one of the objectives in TBLT; ‘how can we achieve learners’ focus-on-form in meaning-centred tasks?’ The two different types of planning attempt to contribute at least partially to our understanding of this objective, especially in the case of on-line planning. However, taking into account that exclusive focus-on-form cannot usually be assumed in speech production, the more appropriate account of this objective is to foster learners’ strategic attentional allocation, balancing between attention to form and meaning. As probably the above episodes and discussion clarified in this chapter, one competence which high proficiency speakers seem equipped with is that of allocating their attention to various aspects of linguistic processing in on-line as well as off-line planning. As previously mentioned, a large part of formulating processes has already been automatized, and larger amount of lexicalized items are stored in high

proficiency speakers, so that the achieved high level of automaticity and lexicalization would make fluent and smooth communication possible. However, it is also true to say that their level of automaticity and lexicalization are not ‘fully’ achieved in many cases, and most speakers still have weaknesses in particular linguistic items, which would require higher levels of attention for processing on-line.

What differentiates such highly skilled speakers from low proficiency speakers seems to be their awareness about their own language characteristics, and the former type of learners tends to allocate their focal attention to such problematic areas. A revealing episode with a high proficiency speaker, for example, suggests her conscious engagement in strategic attention to problematic aspects of her production:

### **Episode 6.5**

R: Did you intend to use the past tense in this task too?

C4: Yes.

R: So, you paid attention to this aspect?

C4: Because I was aware of the past tense use from the beginning, I think I didn't always pay attention to the past tense formation. I paid more attention to other aspects, other grammatical and lexical aspects.

R: Was it hard for you to apply the past tense, compared to the use of the present tense?

C4: Well... because I have done similar kinds of task in a language school before, I got used to using the past tense as an experience. But, when I talked with my friend about what I did one week ago, I noticed I was using the present tense. Because I'm concentrating on this kind of task now, I'm able to rather naturally use the past tense without much attention. But I don't think it is right in everyday conversation. I think this also concern my experience. For example, if I'm asked to tell what I did one week ago, I would be confused with the use of the present and past tense.

R: Even if you are conscious of the past tense use?

C4: Even if I'm conscious of this, I would forget this while speaking. But, in the present case, because I was asked to tell the story according to the given pictures, the use of the past tense was switched on, and I kept using it, I think.  
(C4OP15-18:HIGH)

An interesting point to note in this episode is that the point of her attentional allocation varies depending on the situations and demands of communicative purposes. Speaker [C4] reveals her consciousness about past tense use in the story-telling task given in the present experiment, but she also suggests the possibility of lack of attention to the tense consistency in other situations. It may be true to say that the present task condition increased the learner's consciousness of formality and inclined her towards prioritizing accuracy. On the other hand, she speculates that she might not be able to keep attention to form in more casual communicative settings, probably because the low level of formality makes her decide to allocate her focal attention to other aspects rather than accuracy. She mentioned that she did not need a considerable amount of attention to maintain the appropriate tense usage in the present task, but her statement on the possibility of tense inconsistency implies that this grammatical item has not been fully acquired in her IL system, requiring some level of attention on-line for the production. Taking into account various factors in a communicative situation, the advanced speakers would judge the priorities from the communicative purposes and allocate their focal attention to the various problematic areas rather than always prioritizing any single aspect in a given communicative situation.

## **6.4 Conclusion**

Integrating the findings of the task performance and verbal report analysis, this chapter has further considered the effects of task planning on L2 oral performance in terms of not only types of planning but also learner proficiency, by referring to relevant theoretical accounts. Although the output hypothesis argues that output opportunity pushes learners into syntactic processing, the pressured condition

(particularly NP) tends to force speakers to use reduction strategies, exclusively focusing on meaning aspects at the expense of form focused processing. There is some focus-on-form evidence in strategic planning, but existing communication pressure continued to hinder learners' full engagement in constructing complex structures, while more active syntacticization was observed in on-line planning. In addition, learner proficiency might also affect the outcome of planned performance in terms of available linguistic resources and the degree of skill development. The analysis suggests that for high proficiency speakers, strategic planning is more facilitative in achieving most balanced processing, while on-line planning gives rise to better complexity and accuracy in low proficiency speakers; probably because a higher degree of automatization and a greater amount of lexicalized items allow high proficiency speakers to improve a good level of performance even in the pressured, strategic planning condition, whereas psycholinguistic immaturity requires more processing space in low proficiency speakers, as in on-line planning.

# **CHAPTER SEVEN: CONCLUSIONS**

## **7.1 Introduction**

This concluding chapter summarizes the whole thesis and proposes suggestions for teaching and future research as well as considering limitations of the study. Firstly I reiterate the significant issues which emerged in the previous chapters. Secondly, I consider pedagogic implications drawn from the present findings. Finally, I address the limitations of the study and issues for future research.

## **7.2 Summary of the study**

In order to investigate the underlying mechanisms of planned production and the effects of focus-on-form in different planning conditions, this thesis first looked at Levelt's (1989) L1 speech production model to understand speech processes. Despite similarities in fundamental speech production stages, there are a number of distinctive differences between L1 and L2. I also referred to Swain's (1985, 1995, 1998) output hypothesis. Output is likely to lead to more syntactic processing than comprehension but this does not mean automatic attention to form in any output condition. Searching for more promising pedagogic interventions triggering syntactic processing, Chapter One reviewed key issues in understanding L2 speech processing such as WM, attention, automatization and lexicalized language. These accounts led to an L2 speech model, proposed by Skehan (1998) among others, comprising the exemplar-based system and the rule-based system.

With these characteristics of L2 speech production in mind, Chapter Two considered types of pedagogic intervention for developing L2 processing capacity.

The study focused on the two SLA research traditions: task planning and focus-on-form. Following and extending Ortega's (1999, 2005) indication that strategic planning is likely to lead to learner-initiated focus-on-form and Yuan and Ellis's (2005) proposal for on-line planning as a tool to increase accuracy as well as complexity, the present study hypothesized that, with increased accuracy and complexity, on-line planning would lead to more focus-on-form occurrences.

Following the theoretical review, Chapter Three proposed the research objectives for the two research stages: do the strategic and on-line planning conditions differently influence L2 oral performance?; and, how do L2 speakers plan their speech on-line to deal with the linguistic or conceptual problems? The chapter described the research context, the target learners and various research procedures. I also explained various IL measures for the analysis of task performance and the issues in the verbal protocol analysis.

Chapters Four and Five showed the findings of each research stage in terms of planning conditions and learner proficiency. The task performance analysis suggested that there were form-focused effects in both strategic and on-line planning but the extents were different between different proficiency groups; that is, there were clearer form-focused effects of on-line planning than strategic planning in low proficiency learners, while the difference between the two conditions was less clear in high proficiency speakers.

To explore the cognitive processes underlying different performance variables, Chapter Five illustrated the findings of the verbal report analysis. There was some focus-on-form evidence in strategic planning, but the remaining communication pressures continued to hinder learners' full engagement in constructing complex structures, while more active syntacticization was observed in on-line planning. The analysis also observed that high proficiency speakers tended to maintain balanced

attention to various parts of speech production on on-line as well as off-line planning occasions, while low proficiency speakers focused more on linguistic operations, particularly lexical retrievals, in any planning opportunity.

Integrating the findings of the task performance and verbal report analysis, Chapter Six further considered the effects of types of planning and learner proficiency by referring to related theoretical accounts. Various task characteristics and the research design factors seemed to contribute to increasing learners' focus-on-form occurrences, together with less pressured mental states created by strategic and on-line planning. In addition, learner proficiency would also affect the outcome of planned performance in terms of available linguistic resources and the degree of skill development.

### **7.3 Implications for task-based pedagogy**

Having analyzed the findings of this experimental study, this part will touch upon teaching/learning issues. As the main focus has been on an examination of the nature of planning effects and L2 speech production processes, the scope for applying the present findings to pedagogy may be limited, but it is of importance that information about significant task-related variables acquired through research should assist teachers in deciding what tasks and what planning to use and when (Ellis, 2000). This section is also important for future studies exploring effective ways to incorporate on-line planning in the classroom.

As one of the reasons for the popularity of task planning in SLA research, Ellis (2005:33) addresses its pedagogical relevance; 'Planning, whether of the pre-task or within-task kind, is a variable that teachers can easily manipulate in their day-to-day teaching'. In order to consider the pedagogic implications of the present

study, it may be important to refer back to ‘task authenticity’ as one of the essential task characteristics proposed by Skehan (2.2.1). Stressing the connection of ‘task’ to the real world (e.g., Long, 1985), the application of strategic planning may be limited to practising oral presentations in formal settings, while that of on-line planning is uncertain. As indicated by Willis (1996:33), provision of planning time can be used to ‘offer learners opportunities to “upgrade” their task language to a version suitable for presenting in public and reflect on the changes that need to be made’. It is no doubt important not only to immerse learners in a situation to use spontaneous language but also to speak in a more formal style in front of large audiences. This pedagogic aim has clear connection between task planning and the real world language use, but adhering to the approximations of real-world needs not only blurs the application of on-line planning but also limits the scope of strategic planning for pedagogy.

On the other hand, understanding task authenticity as ‘some sort of relationship to comparable real-world activities’ (Skehan, 1998:95) would liberate us from limited use of planning in the classroom. Pedagogic tasks with planning implementation can create the condition where learners cannot avoid a description for lack of lexical resources. Beyond a simple preparation time prior to or within task, planning may function as an opportunity to scrutinize one’s own IL system and to control linguistic knowledge under communicative conditions. It is expected that planning integrates various L2 functions (e.g., attention, WM, monitoring, as reviewed in Chapter One), and through these it may contribute to developing the language processing capacity needed for real world language use. It is reiterated that the strategic and on-line planning conditions created opportunities for learners to focus on form on-task, which is not easily realized in the non-planning condition exclusively prioritizing meaning. In particular, the benefit of strategic planning would be to provide an opportunity to employ planned phrases under real communicative



pressure.

Understanding the benefits of strategic planning in the classroom, then, what are the specific pedagogic benefits of on-line planning? If these exist, in what ways can we implement on-line planning in language pedagogy? My answer to the first question is, as repeatedly mentioned, to regard on-line planning as a pedagogic tool to ease learners' cognitive engagement in the rule-based system, beyond the exemplar-based system. This may also be feasible with strategic planning, but the analysis suggests that, even in strategic planning production, learners often find it difficult to engage in syntactic processing under processing pressures, even when they notice the necessity to do so. This tendency is particularly the case in low proficiency speakers whose processing capacity is more limited, due to fewer amounts of lexicalized items and lesser extent of automatization. Taking this point into consideration, it can be said that on-line planning could contribute to one of the objectives in task-based pedagogy; how balance may be achieved among different performance areas (Skehan & Foster, 1999).

It may be possible to propose a planning sequence to minimize the processing burden, based on the present findings. If it is accepted that strategic planning tends to lead to upgrading fluency and complexity while on-line planning tends to increase the level of complexity and accuracy, a sequence of strategic planning to on-line planning, and probably to non-planning in the end, would direct learners' attention to meaning and then to form. Through this sequence, learners' attention may shift from meaning (SP) to form (OP) then back to meaning (NP) as their lexicalization and automatization proceeds. In doing so, there is a constant interplay of meaning and form connection. Also, this sequence may ease the task difficulty by reducing processing pressures before challenging the non-planned production. As the present study aimed to investigate the effects of planning

conditions (not any combination of different planning ways), it has little to say about the effects of particular planning sequences, but it may be of importance to search for the effects of this aspect in future research.

To the second question concerning specific on-line planning application, the immediate answer is to give the instruction which I used in the research (3.7.3), but it is premature to suggest just a simple transfer of the present study to teaching practices, because some learners' responses to on-line planning in the interviews tended to be rather negative. How they reacted to the on-line planning instruction was largely up to their own interpretation and the extent which they were engaged in 'on-line planning' varied between individual speakers. To generalize the tendency, low proficiency speakers seem to more clearly show on-line planning features such as pausing and dysfluencies; on the other hand, high proficiency speakers' engagement in on-line planning tends to be less clear, probably because, as suggested by several participants in the interviews, such learners felt less comfortable about more frequent and longer pausing than they do in their usual communication, finally blurring the difference between on-line planning and non-planning in terms of performance gains. Despite such apparent negligence of on-line planning engagement on the performance level, it is interesting to note that on-line planning provides many high proficiency speakers with opportunities for conceptual monitoring as an opportunity to elaborate their message, as suggested by the verbal report analysis.

For the pedagogic application of on-line planning, it is thus important to further consider what guiding principles for making the on-line planning instruction should be given, similar to attempts by Sangarun (2005) to construct SP instruction. For this aim, it may be useful, in future research, to differentiate 'true' on-line planners from others in the present participants and scrutinize what factors actually differentiate them (Peter Skehan, personal communication, April, 2006). In addition, it

may be important to design 'authentic' tasks that naturally include on-line planning. For example, a task where learners are applicants for a job and have to leave a voice message for the future employer saying why they want the job under conditions where the voice message can be edited before sending. This would allow as much time as available for on-line planning, with a natural pressure on accuracy. Together with the findings of on-line planning constructs (Skehan & Foster, 2005), this attempt may contribute to pedagogic applications of on-line planning.

Having considered these accounts, the study of pedagogic aspects of on-line planning is inconclusive. Although ways of implementing on-line planning may need more consideration in the language classroom, some kind of 'on-line planning implementation' may be useful for developing the learner's ability of syntactic formulation.

#### **7.4 Limitations of the study**

As this is a small scale study with limited perspectives, the findings presented have necessarily been modest and suggestive rather than conclusive. The rest of this chapter considers limitations of the study mainly in three aspects: situational, methodological and theoretical. These will illuminate issues for future research.

First of all, there are situational limitations in the study. As the number of Japanese learners who study abroad has increased in recent years, it is more and more important to investigate this particular type of learner, but it should be recognized that they are different from 'general' Japanese learners and other L2 learners (e.g., Chinese) in many respects. Although the present participants started learning in EFL contexts, their experience in an English speaking country distinguishes them from other learners who have received only limited amounts of L2 interaction in the

classroom. Therefore, it is less clear whether and to what extent the present findings can be applied to Japanese learners of English in other learning contexts. This also brings up the need for similar planning studies in Japanese EFL contexts, to supplement the relatively small number of studies in such research contexts (e.g., Wendel, 1997).

Secondly, there are several methodological limitations in both the task performance analysis and the verbal report analysis. The limited number of participants ( $n = 27$ ) might not have a strong explanatory power, particularly in the quantitative analysis of different proficiency groups (High = 14; Low = 13). The findings of task performance seem to follow the general trend of the previous planning studies, but considering the possibility of language expertise effects, there is certainly a need to conduct larger scale planning research with different proficiency groups.

Related to this respect, determining proficiency groups also needs to be reconsidered. The present study adopted the professional judgement on identification of learner proficiency (3.4.3). Together with the length of residence in English speaking countries, I believe that in this way I successfully created two relative proficiency groups. However, as suggested by Polio (1997), the reliabilities of holistic scale judgements may not be very high. This does not mean the lack of reliability in the present judgement; in this research, our agreement on inconsistent cases between examiners in the board meeting minimized the danger of inappropriate grouping, but, for more reliable proficiency judgement, it may be necessary to consider a combination of the professional judgement and other more objective judgements such as TOEFL and TSE scores.

It is also important to mention issues of experimental and classroom research. As this study is based on TBLT, it is hoped that the present findings have some

contribution to classroom practices, but the effects of experimental research design should be seriously considered in terms of pedagogic allocations of the study. From the study of negotiation for meaning, Foster (1998) argues that the performance under research conditions does not necessarily reflect performance in the classroom and ‘an approach that seeks to influence learners’ language through ever more tightly designed experimental tasks is moving itself further and further away from the classroom’ (p. 20). It is thus wrong to suggest that all the present findings have direct relevance to teaching practices, but it is probably the case that certain findings offer some clues to understanding the nature of planning and L2 speech production in pedagogic contexts. It may be that this is a common process from the laboratory into the classroom, as theorizing the concept develops. That is, in the situation that our understanding of on-line planning is preliminary, it may be a necessary and effective step to test the hypothesis and gather a wide range of data in a more controlled setting. Experimentation should obviously have its place in the researcher’s repertory, but it should also be stressed that it is no panacea (McLaughlin, 1987); in order to feed into teaching, ‘the research environment has to be willing to move out of the laboratory and into the classroom’ (Foster, 1998:21). For the next stage, together with our consideration of practical application of on-line planning, there is certainly a need to conduct on-line planning research in classroom settings.

There is also a need to address methodological limitations in the aspect of retrospective interviews. I took a great care to avoid guiding participants’ answers in particular directions in the interviews. Also, the categorization of episodes in the analysis was made as objectively as possible. However, it is very difficult to judge whether reports truly reflect participants’ mental conditions and whether it is possible to exclude the researcher’s subjectivity completely. It may be possible that some comments, even unintentionally, were distorted through the verbalizing process. There

is also a possibility of memory loss between task performance and the interview sessions, although stimulated recall helped participants recall memories accurately. Due to these methodological difficulties of verbal reports, the present study used these data essentially for exploratory purposes.

Despite several limitations, I believe that the study provides some avenues for exploring strategic planning and on-line planning in a narrower sense, and task-based research and SLA in a broader sense. In contrast to a number of studies on strategic planning, the study of on-line planning has been certainly limited. In this sense, this thesis not only supports the main argument from the previous on-line planning studies but also extends the planning effects into the cognitive processes underlying performance. In terms of the latter aspect, another significant contribution is the combination of process-product approach. Despite a growing body of research investigating strategic planning processes, studies investigating on-task speaking processes in any planning conditions have been largely neglected. There are methodological difficulties in tapping into learners' cognitive processes during on-going speech, but the retrospective interviews combined with stimulated recall can be a powerful instrument to uncover L2 speaking processes under different planning conditions.

## REFERENCES

- Aitchison, J. 1989. *The Articulate Mammal: An introduction to psycholinguistics (Fourth Edition)*. London: Rutledge.
- Albert, A. and Kormos, J. 2004. Creativity and narrative task performance: An exploratory study. *Language Learning*, 54/2: 277-310.
- Allen, P., Swain, M., Harley, B. and Cummins, J. 1990. Aspects of classroom treatment: Toward a more comprehensive view of second language education. In B. Harley, P. Allen, J. Cummins & M. Swain (Eds.), *The Development of Bilingual Proficiency*. Cambridge: Cambridge University Press.
- Anderson, J. R. 1983. *The Architecture of Cognition*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Anderson, J. R. 1993. *Rules of the Mind*. Hillsdale, NJ: Lawrence Erlbaum Associates, Publishers.
- Anderson, J. R. and Lebiere, C. 1998. *The Atomic Components of Thought*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Anderson, R. W. 1984. The one to one principle of interlanguage construction. *Language Learning*, 34: 77-95.
- Atkinson, R. C. and Shiffrin, R. M. 1968. Human memory: A proposed system and its control processes. In K. W. Spence & J. T. Spence (Eds.), *The Psychology of Learning and Motivation: Advances in research and theory* (Vol. 2). New York: Academic Press.
- Bachman, L. F. and Palmer, A. S. 1996. *Language Testing in Practice*. Oxford: Oxford University Press.
- Baddeley, A. D. 1986. *Working Memory*. Oxford: Clarendon Press.
- Beattie, G. W. 1980. The role of language production processes in the organization of behaviour in face-to-face interaction. In B. Butterworth (Ed.), *Language*

*Production Volume 1*. London: Academic Press.

Beaumont, M. and Gallaway, C. 1994. Articles of faith: The acquisition, learning and teaching of *a* and *the*. In M. Bygate, A. Tonkyn and E. Williams (Eds.), *Grammar and the Language Teacher*. New York: Prentice Hall.

Bialystok, E. 1978. A theoretical model of second language learning. *Language Learning*, 28: 69-84.

Bialystok, E. 1990. *Communication Strategies*. Oxford, MA: Blackwell.

Bloom, B. S. 1953. Thought processes in lectures and discussions. *Journal of General Education*, 7: 160-170.

Bloom, B. S. 1954. The thought processes of students in discussion. In S. J. French (Ed.), *Accent on Teaching: Experiments in general education*. New York: Harper.

Bloor, T. and Bloor, M. 2004. *The Functional Analysis of English: A Hallidayan Approach*. London: Arnold.

Brace, N., Kemp, R. and Snelgar, R. 2003. *SPSS for Psychologists (Second Edition)*. New York: Palgrave MacMillan.

Bray, J. H. and Maxwell, S. E. 1985. *Multivariate Analysis of Variance*. Newbury Park, California: Sage Publications

Breen, M. P. 1989. The evaluation cycle for language learning tasks. In R. K. Johnson (Ed.), *The Second Language Curriculum*. Cambridge: Cambridge University Press.

Brown, G., Anderson, A., Shillcock, R. and Yule, G. 1984. *Teaching Talk: Strategies for production and assessment*. Cambridge: Cambridge University Press.

Brown, G. and Yule, G. 1983. *Teaching the Spoken Language*. Cambridge: Cambridge University Press.

Brown, R. 1973. *A First Language: The early stages*. London: Allen and Unwin.

Brumfit, C. 1984. *Communicative Methodology in Language Teaching*. Cambridge:



Cambridge University Press.

Bryman, A. and Cramer, D. 2001. *Qualitative Data Analysis with SPSS Release 10 for Windows: A guide for social scientists*. Hove, East Sussex: Routledge.

Butler, C. 1995. Between lexis and grammar: repeated word sequences and collocational frameworks in Spanish. Paper presented at the 5th Dyffryn Conference on vocabulary and lexis, Cardiff.

Bygate, M. 1987. *Speaking*. Oxford: Oxford University Press.

Bygate, M. 1996. Effects of task repetition: appraising the developing language of learners. In J. Willis & D. Willis (Eds.), *Challenge and Change in Language Teaching*. Oxford: Macmillan Heinemann.

Bygate, M. 1998. Theoretical perspectives on speaking. *Annual Review of Applied Linguistics*. 18/ 20-42.

Bygate, M. 1999. Task as context for the framing, reframing and unframing of language. *System*, 27: 33-48.

Bygate, M. 2000. Introduction. *Language Teaching Research*, 4/3: 185-192.

Bygate, M. 2001. Effects of task repetition on the structure and control of oral language. In M. Bygate, P. Skehan & M. Swain (Eds.).

Bygate, M. 2002. Speaking. In R. B. Kaplan (Ed.), *The Oxford Handbook of Applied Linguistics*. Oxford: Oxford University Press.

Bygate, M. 2005. Researching tasks for language pedagogy: The need for a pedagogical agenda. Plenary presentation at BAAL SLLT Seminar, York.

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

Bygate, M. Skehan, P. and Swain, M. (Eds.), 2001. *Researching Pedagogic Tasks Second Language Learning, Teaching and Testing*. Harlow: Pearson Education.

- Candlin, C. 1987. Towards task-based language learning. In C. Candlin & D. Murphy (Eds.), *Language Learning Tasks*. Englewood Cliffs, N.J.: Prentice Hall.
- Carroll, D. W. 2004. *Psychology of Language (Fourth Edition)*. Belmont, CA: Wadsworth/Thomson Learning.
- Carter, R. and McCarthy, M. 1995. Grammar and spoken language. *Applied Linguistics*, 16: 141-158.
- Chamber, F. 1997. What do we mean by fluency? *System*, 25/4: 535-544.
- Cook, V. 1997. The consequences of bilingualism for cognitive processing. In A. M. D. de Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum.
- Corder, P. 1967. The significance of learners' errors. *IRAL*, 5: 161-70.
- Craik, F. I. M and Lockhart, R. S. 1972. Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11: 671-84.
- Crookes, G. 1988. *Planning, Monitoring and Second Language Development: A Review* (Tech. Rep. No. 6). Honolulu: University of Hawai'i, Centre for Second Language Classroom Research, Social Science Research Institute.
- Crookes, G. 1989. Planning and interlanguage variation. *Studies in Second Language Acquisition*, 11: 367-83.
- Crookes, G. 1990. The utterance and other basis units for second language discourse analysis. *Applied Linguistics*, 11: 183-99.
- Crookes, G. 1991. Second language speech production research: A methodologically oriented review. *Studies in Second Language Acquisition*, 13: 113-132.
- Crookes, G. and Rulon, K. 1985. Incorporation of corrective feedback in native speaker/non-native speaker conversation. Technical Report No. 3, The Center for Second Language Classroom Research, Social Science Research Institute, University of Hawai'i at Manoa.

- Cummins, J. 1980. The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly*, 14: 175-187.
- Cummins, J. 1991. Interdependence of first- and second-language proficiency in bilingual children. In E. Bialystok (Ed.), *Language Processing in Bilingual Children*. Cambridge: Cambridge University Press.
- Cummins, J. 2000. *Language, Power and Pedagogy: Bilingual children in the crossfire*. Clevedon: Multilingual Matters.
- De Bot, K. 1992. A bilingual production model: Levelt's 'speaking' model adapted. *Applied Linguistics*, 13/1: 1-24.
- De Bot, K. 1996. Review article: The psycholinguistics of the output hypothesis. *Language Learning*, 46/3: 529-555.
- De Bot, K. 2002. Cognitive processing in bilinguals. In R. B. Kaplan (Ed.), *The Oxford Handbook of Applied Linguistics*. Oxford: Oxford University Press.
- De Bot, K and Kroll, J. F. 2002. Psycholinguistics. In N. Schmitt (Ed.), *An Introduction to Applied Linguistics*. London: Arnold.
- De Bot, K. and Schreuder, R. 1993. Word recognition and the bilingual lexicon. In R. Schreuder & B. Weltens (Eds.), *The Bilingual Lexicon*. Amsterdam: John Benjamins.
- Dechert, H. W. 1980. Pauses and intonation as indicators of verbal planning in second language speech productions: Two examples from a case study. In H.-W. Dechert and M. Raupach (Eds.), *Temporal Variables in Speech: Studies in honour of Frieda Goldman-Eisler*. The Hague, The Netherlands: Mouton.
- DeKeyser, R. M. 1997. Beyond explicit rule learning: Automatizing second language morphosyntax. *Studies in Second Language Acquisition*, 19: 195-222.
- DeKeyser, R. M. 1998. Beyond focus on form: Cognitive perspectives on learning and practicing second language grammar. In C. Doughty & J. Williams (Eds.), *Focus on*

*Form in Classroom Second Language Acquisition*. New York: Cambridge University Press.

DeKeyser, R. M. 2001. Automaticity and automatization. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

DeKeyser, R. M. and Sokalski, K. 1996. The differential role of comprehension and production practice. *Language Learning*, 46: 613-642.

Dell, G. S. 1986. A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93: 283-321.

Derwing, T. M., Rossiter, M. J., Munro, M. J. and Thomson, R. I. 2004. Second language fluency: Judgements on different tasks. *Language Learning*, 54/4: 655-679.

Deschamps, A. 1980. The syntactic distribution of pauses in English spoken as a second language by French students. In H.-W. Dechert and M. Raupach (Eds.), *Temporal Variables in Speech: Studies in honour of Frieda Goldman-Eisler*. The Hague, The Netherlands: Mouton.

Doughty, C. 2001. Cognitive underpinnings of focus on form. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

Doughty, C. and Williams, J. 1998a. Pedagogical choices in focus on form. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom SLA*. New York: Cambridge University Press.

Doughty, C. and Williams, J. 1998b. Theoretical foundations of focus on form. In C. Doughty & J. Williams (Eds.), *Focus on Form in Classroom SLA*. New York: Cambridge University Press.

Dulay, H. and Burt, M. 1973. Should we teach children syntax? *Language Learning*, 23: 245-58.

Dulay, H. and Burt, M. 1974. Natural sequences in child second language acquisition.

*Language Learning*, 24: 37-53.

Dulay, H., Burt, M. and Krashen, S. 1982. *Language 2*. New York: Oxford University Press.

Edwards, C. and Willis, J. (Eds.) 2005. *Teachers Exploring Tasks in English Language Teaching*. Basingstoke, Hampshire: Palgrave Macmillan.

Elder, C. and Iwashita, N. 2005. Planning for test performance: Does it make a difference? In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.

Ellis, N. C. 1995. Consciousness in second language acquisition: A review of field studies and laboratory experiments. *Language Awareness*, 4/3: 123-146.

Ellis, N. C. 2001. Memory for language. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

Ellis, N. C. 2002. Frequency effects in language processing: A review with implications for theories of implicit and explicit language acquisition. *Studies in Second Language Acquisition*, 24: 143-188.

Ellis, R. 1987. Interlanguage variability in narrative discourse: Style in the use of the past tense. *Studies in Second Language Acquisition*, 9: 12-20.

Ellis, R. 1990. *Instructed Second Language Acquisition*. Oxford: Blackwell.

Ellis, R. 1991. *Second Language Acquisition and Language Pedagogy*. Clevedon: Multilingual Matters.

Ellis, R. 1993. 'Talking shop: Second language acquisition research: how does it help teachers? An interview with Rod Ellis'. *ELT Journal*, 47/1: 3-11.

Ellis, R. 1994. *The Study of Second Language Acquisition*. Oxford: Oxford University Press.

Ellis, R. 1997. *SLA Research and Language Teaching*. Oxford: Oxford University Press.

- Ellis, R. 1999. Input-based approaches to the teaching of grammar. *Annual Review of Applied Linguistics*, 19: 64-80.
- Ellis, R. 2000. Task-based research and language pedagogy. *Language Teaching Research*, 4/3: 193-220.
- Ellis, R. 2001. Introduction: Investigating form-focused instruction. *Language Learning*, 51/Supplement: 1-46.
- Ellis, R. 2002. Methodological options in grammar teaching materials. In E. Hinkel and S. Fotos (Eds.), *New Perspectives on Grammar Teaching in Second Language Classrooms*. Mahwah: Lawrence Erlbaum Associates.
- Ellis, R. 2003. *Task-Based Language Learning and Teaching*. Oxford: Oxford University Press.
- Ellis, R. 2004. The definition of measurement of L2 explicit knowledge. *Language Learning*, 54/2: 227-275.
- Ellis, R. 2005. Planning and task-based performance: theory and research. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.
- Ellis, R. and Barkhuizen, G. 2005. *Analyzing Learner Language*. Oxford: Oxford University Press.
- Ellis, R., Tanaka, Y. and Yamazaki, A. 1994. Classroom interaction, comprehension, and the acquisition of L2 word meanings. *Language Learning*, 44/3: 449-491.
- Ellis, R. and Yuan, F. 2004. The effects of planning on fluency, complexity, and accuracy in second language narrative writing. *Studies in Second Language Acquisition*, 26/1: 59-84.
- Ellis, R. and Yuan, F. 2005. The effects of careful within-task planning on oral and written task performance. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.

- Ejzenberg, R. 2000. The juggling act of oral fluency: A psycho-sociolinguistic metaphor. In H. Riggenbach (Ed.), *Perspectives on Fluency*. Ann Arbor: University of Michigan Press.
- Ericsson, K. A. and Simon, H. A. 1987. Verbal reports on thinking. In C. Færch and G. Kasper (Eds.), *Introspection in Second Language Research*. Clevedon: Multilingual Matters.
- Ericsson, K. A. and Simon, H. A. 1993. *Protocol Analysis: Verbal reports as data (Revised Edition)*. Cambridge, Massachusetts: MIT Press.
- Fathman, A. 1975. Language background, age and the order of acquisition of English structures. In M. K. Burt and H. C. Dulay (Eds.), *On TESOL '75: New directions in second language learning, teaching and bilingual education*. Washington, DC: TESOL.
- Færch, C. and G. Kasper. 1983. Plans and strategies in foreign language communication. In C. Færch and G. Kasper (Eds.), *Strategies in Interlanguage Communication*. London: Longman.
- Færch, C. and Kasper, G. 1986. The role of comprehension in second-language learning. *Applied Linguistics*, 7: 257-74.
- Færch, C. and Kasper, G. 1987. From product to process – introspective methods in second language research. In C. Færch and G. Kasper (Eds.), *Introspection in Second Language Research*. Clevedon: Multilingual Matters.
- Field, A. 2005. *Discovering Statistics Using SPSS (Second Edition)*. London: Sage Publications.
- Field, J. 2003: *Psycholinguistics: A resource book for students*. London: Rutledge.
- Fillmore, C. J. 1979. On fluency. In C. J. Fillmore, D. Kempler and W. S-Y. Wang (Eds.), *Individual Differences in Language Ability and Language Behavior*. New York: Academic Press.
- Foster, P. 1996. Doing the task better: How planning time influences students' performance. In J. Willis & D. Willis (Eds.), *Challenge and Change in Language*

- Teaching*. Oxford: Macmillan Education.
- Foster, P. 1998. A classroom perspective on the negotiation of meaning. *Applied Linguistics*, 19: 1-23.
- Foster, P. 2001. Rules and routines: a consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan & M. Swain (Eds.).
- Foster, P. and Skehan, P. 1996. The influence of planning and task type on second language performance. *Studies in Second Language Acquisition*, 18: 299-323.
- Foster, P. and Skehan, P. 1999. The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3: 215-247.
- Foster, P., Tonkyn, A. and Wigglesworth, G. 2000. Measuring spoken language: A unit for all reasons. *Applied Linguistics*, 21: 354-375.
- Fotos, S. and Ellis, R. 1990. Communicating about grammar: A task-based approach. *TESOL Quarterly*, 25: 605-28.
- Freed, B. F. 1995. Language learning and study abroad. In B. F. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context*. Amsterdam: John Benjamins.
- Freed, B. F. 2000. Is fluency, like beauty, in the eyes (and ears) of beholder? In H. Riggensbach (Ed.), *Perspectives on Fluency*. Ann Arbor: University of Michigan Press.
- Fromkin, V. A. 1971. The non-anomalous nature of anomalous utterances. *Language*, 47: 27-52.
- Fulcher, G. 1996. Does thick description lead to smart tests?: A data-based approach to raising scale construction. *Language Testing*, 13: 208-238.
- Garman, M. 1990. *Psycholinguistics*. Cambridge: Cambridge University Press.
- Garnham, A., Shillcock, R. S., Brown, G. D. A., Mill, A. I. D. and Cutler, A. 1982. Slips of the tongue in the London-Lund corpus of spontaneous conversation. In A.



- Cutler (Ed.), *Slips of the Tongue and Language Production*. Berlin: Mouton.
- Garrett, M. F. 1982. Production of speech: observations from normal and pathological language use. In A. W. Ellis (Ed.), *Normality and Pathology in Cognitive Functions*. London: Academic Press.
- Gass, S. 1988. Integrating research areas: A framework for second language studies. *Applied Linguistics*, 9: 198-217.
- Gass, S. M. and MacKay, A. 2000. *Stimulated Recall Methodology in Second Language Research*. Mahwah, NJ: Lawrence Erlbaum Associates Publishers.
- Gass, S. M., Mackey, A., Alvarez-Torres, M. J. and Fernandez-Garcia, M. 1999. The effects of task repetition on linguistic output. *Language Learning*, 49/4: 549-581.
- Glazer, B. G. and Straus, A. L. 1967. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. New York: Aldine.
- Godfrey, D. L. 1980. A discourse analysis of tense in adult ESL monologues. In D. Larsen-Freeman (Ed.), *Discourse Analysis in Second Language Research*. Rowley, MA: Newbury House Publishers, INC.
- Goldschneider, J. M. and DeKeyser, R. M. 2005. Explaining the “natural order of L2 morpheme acquisition” in English: A meta-analysis of multiple determinants. *Language Learning*, 55/Supplement 1: 27-77.
- Grainger, J. and Dijkstra, T. 1992. On the representation and use of language information in bilinguals. In R. J. Harris (Ed.), *Cognitive Processing in Bilinguals*. Amsterdam: North Holland.
- Green, D. W. 1986. Control, activation and resource: A framework and a model for the control of speech in bilinguals. *Brain and Language*, 27: 210-223.
- Green, D. W. 1993. Toward a model of L2 comprehension and production. In R. Schreuder and B. Weltens (Eds.), *The Bilingual Lexicon*. Amsterdam: John Benjamins.
- Griffiths, R. 1991. Pausological research in an L2 context. A rationale, and review of

- selected studies. *Applied Linguistics*, 12: 345-364.
- Guadagnoli, E. and Velicer, W. 1988. Relation of sample size to the stability of components patterns. *Psychological Bulletin*, 103: 265-275.
- Guariento, W. and Morley, J. 2001. Text and task authenticity in the EFL classroom. *ELT Journal*, 55/4: 347-353.
- Halliday, M. A. K. and Matthiessen, C. M. I. M. 2004. *An Introduction to Functional Grammar*. London: Arnold.
- Harley, T. 1984. A critique of top-down independent levels model of speech production: Evidence from non-plan internal speech errors. *Cognitive Science*, 8: 191-219.
- Harrington, M. and Sawyer, M. 1992. L2 working memory capacity and L2 reading skill. *Studies in Second Language Acquisition*, 14: 25-38.
- Hatch, E. 1992. *Discourse and Language Education*. Cambridge: Cambridge University Press.
- Hawkins, P. R. 1971. The syntactic location of hesitation pauses. *Language and Speech*, 14: 277-288.
- Heaton, J. B. 1975: *Beginning Composition through Pictures*. Harlow, Essex: Pearson Education.
- Holliday, A. 1994. *Appropriate Methodology and Social Context*. Cambridge: Cambridge University Press.
- Hopkins, D. and Nettle, M. 1994. Second language acquisition research: A response to Rod Ellis. *ELT Journal*, 48/2:157-161.
- Hulstijn, J. H. 2001. Intentional and incidental second language vocabulary learning: A reappraisal of elaboration, rehearsal and automaticity. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

- Hulstijn, J. H. 2002. Towards a unified account of the representation, processing and acquisition of second language knowledge. *Second Language Research*, 18/3: 193-223.
- Hulstijn, J. H. and Hulstijn, W. 1984. Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning*, 34/1: 23-43.
- Hunt, K. 1965. *Grammatical Structures Written at Three Grade Levels*. Champaign, Ill: National Council of Teachers of English.
- Iwashita, N., McNamara, T. and Elder, C. 2001. Can we predict task difficulty in an oral proficiency test?: Exploring the potential of an information-processing approach to task design. *Language Learning*, 51/3: 401-436.
- Izumi, S. 2002. Output, input enhancement and the noticing hypothesis: An experimental study on ESL relativization. *Studies in Second Language Acquisition*, 24: 541-577.
- Izumi, S. 2003. Comprehension and production processes in second language learning: In search of the psycholinguistic rationale of the output hypothesis. *Applied Linguistics*, 24: 168-196.
- Izumi, S. and Bigelow, M. 2000. Does output promote noticing and second language acquisition? *TESOL Quarterly*, 34: 239-78.
- Izumi, S., Bigelow, M., Fujiwara, M., and Fearnow, S. 1999. Testing the output hypothesis: Effects of output on noticing and second language acquisition. *Studies in Second Language Acquisition*, 21: 421-52.
- Johnson, K. 1979. Communicative approaches and communicative processes. In C. J. Brumfit & K. Johnson (Eds.), *The Communicative Approach to Language Teaching*. Oxford: Oxford University Press.
- Johnson, K. 1996. *Language Teaching and Skill Learning*. Oxford: Blackwell.
- Johnson, K. 2001. *An Introduction to Foreign Language Learning and Teaching*. Harlow: Longman.

- Jourdenais, R. 2001. Cognition, instruction and protocol analysis. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.
- Kaiser, H. F. 1960. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*, 20: 141-151.
- Kawauchi, C. 2005. The effects of strategic planning on the oral narratives of learners with low and high intermediate L2 proficiency. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.
- Kihlstrom, J. 1984. Conscious, subconscious, unconscious: A cognitive perspective. In K. Bowers & D. Meichenbaum (Eds.), *The Unconscious Reconsidered*. New York: Wiley.
- Kim, J. O. and Muller, C. W. 1978. *Factor Analysis: Statistical methods and practical issues*. London: Sage Publications.
- Klein, W. 1993. *Toward Second Language Acquisition: A study of null-prep*. Dordrecht: Kluwer Academic Publishers.
- Kline, P. 1994. *An Easy Guide to Factor Analysis*. London: Routledge.
- Kobayashi, M. 2002a. Method effects on reading comprehension test performance: Text organization and response format. *Language Testing*, 19/2: 193-220.
- Kobayashi, M. 2002b. Cloze tests revisited: Exploring item characteristics with special attention to scoring methods. *Modern Language Journal*, 86/4: 571-586.
- Kohn, K. 1982. Beyond output: The analysis of interlanguage development. *Studies in Second Language Acquisition*, 4: 137-152.
- Kormos, J. 1998. Verbal reports in L2 speech production research. *TESOL Quarterly*, 32/2: 353-358.
- Kormos, J. 1999. Monitoring and self-repair in L2: Review article. *Language Learning*, 49: 303-342.

- Kormos, J. and Dénes, M. 2004. Exploring measures and perceptions of fluency in the speech of second language learners. *System*, 32: 145-164.
- Kowal, M. and Swain, M. 1994. Using collaborative language production tasks to promote students' language awareness. *Language Awareness*, 3/2: 73-93.
- Krashen, S. 1977. Some issues relating to the Monitor Model. In H. D. Brawn, C. Yorio and R. Crymes (Eds.), *On TESOL '77: Teaching and Learning English as a Second Language: Trends in Research and Practice*. Washington: TESOL.
- Krashen, S. 1982. *Principles and Practice in Second language Acquisition*. Oxford: Pergamon Press.
- Krashen, S. 1985: *The Input Hypothesis*. London: Longman.
- Kroll, J. F. and de Groot, A. M. B. 1997. Lexical and conceptual memory in the bilingual: Mapping form to meaning in two languages. In A. M. D. de Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Kumaravadivelu, B. 1994. The postmethod condition: meaning strategies for second/foreign language teaching. *TESOL Quarterly*, 28: 27-48.
- Lapkin, S., Hart, D. and Swain, M. 1995. A Canadian interprovincial exchange: Evaluating the linguistic impact of a three-month stay in Quebec. In B. F. Freed (Ed.), *Second Language Acquisition in a Study Abroad Context*. Amsterdam: John Benjamins.
- Larsen-Freeman, D. 1975. The acquisition of grammatical morphemes by adult ESL students. *TESOL Quarterly*, 9: 409-419.
- Larsen-Freeman, D. 2003. *Teaching Language: From Grammar to Grammaring*. Boston, Mass: Heinle & Heinle
- Larsen-Freeman, D. and Long, M. H. 1991. *An Introduction to Second Language Acquisition Research*. Harlow: Pearson Education.
- Laufer, B. and Hulstijn, J. 2001. Incidental vocabulary acquisition in a second

- language: The construct of task-induced involvement. *Applied Linguistics*, 22/1: 1-26.
- LeCompte, M. D. and Preissle, J. 1993: *Ethnography and Qualitative Design in Educational Research (Second Edition)*. San Diego, California: Academic Press.
- Leech, G. and Svartvik, J. 1994. *A Communicative Grammar of English (Second Edition)*. Harlow: Pearson Education.
- Lennon, P. 1989. Introspection and intentionality in advanced second-language acquisition. *Language Learning*, 39/3: 375-396.
- Lennon, P. 1990. Investigating fluency in EFL: A quantitative approach. *Language Learning*, 40: 387-412.
- Lennon, P. 1998. Approaches to the teaching of idiomatic language. *International Review of Applied Linguistics*, 36: 11-30.
- Lennon, P. 2000. The lexical element in spoken second language fluency. In H. Riggensbach (Ed.), *Perspectives on Fluency*. Ann Arbor: University of Michigan Press.
- Levelt, W. J. M. 1989. *Speaking: From Intention to Articulation*. Cambridge, Mass: MIT Press.
- Levelt, W. J. M. 1992: Accessing words in speech production: Stages, processes and representations. *Cognition*, 42: 1-22.
- Levelt, W. J. M., Roelofs, A. and Meyer, A. S. 1999. A theory of lexical access in speech production. *Behavioural and Brain Sciences*, 22: 1-75.
- Libben, G. 1997. Psycholinguistics: the study of language processing. In W. O'Grady, M. Dobrovolsky and F. Katamba (Eds.), *Contemporary Linguistics: An introduction*. Harlow, Essex: Pearson Education.
- Lightbown, P. M. 1998. The importance of timing in focus on form. In C. Doughty and J. Willis (Eds.), *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press.

- Loban, W. 1966. *The Language of Elementary School Children*. (Research Report no. 1). Champaign, Ill: National Council of Teachers of English.
- Lockhart, C. and Ng, P. 1995. Analyzing talk in ESL peer response groups: Stances, functions, and content. *Language Learning*, 45: 605-655.
- Logan, G. D. 1988. Toward an instance theory of automatization. *Psychological Review*, 95: 492-527.
- Long, M. H. 1985. A role for instruction in second language acquisition: Task-based language teaching. In K. Hyltenstam & M. Pienemann (Eds.), *Modelling and Assessing Second Language Acquisition*. Clevedon, Avon: Multilingual Matters.
- Long, M. H. 1988. Instructed interlanguage development. In L. Beebe (Ed.), *Issues in Second Language Acquisition: Multiple perspectives*. Rowley, MA: Newbury House.
- Long, M. H. 1991. Focus on form: A design feature in language teaching methodology. In K. de Bot, D. Coste, R. Ginsberg and C. Kramsch (Eds.), *Foreign Language Research in Cross-Cultural Perspective*. Amsterdam: John Benjamins.
- Long, M. H. 1996: The role of the linguistic environment in second language acquisition. In W. Ritchie & T. Bhatia (Eds.), *Handbook of Second Language Acquisition*. San Diego: Academic Press.
- Long, M. H. 2000. Focus on form in task-based language teaching. In R. D. Lambert & E. Shohamy (Eds.), *Language Policy and Pedagogy: Essays in honour of A. Ronald Walton*. Amsterdam: John Benjamins.
- Long, M. H. 2001. Focus on form: A design feature in language teaching methodology. In N. Candlin & N. Mercer (Eds.), *English Language Teaching in its Social Context*. London: Routledge.
- Long, M. H. and Crookes, G. 1992. Three approaches to task-based syllabus design. *TESOL Quarterly*, 26/1: 27-47.
- Long, M., Inagaki, S. and Ortega, L. 1998. The role of implicit negative evidence in

- SLA: Models and recasts in Japanese and Spanish. *The Modern Language Journal*, 82/3: 357-371.
- Long, M. H. and Robinson, P. 1998. Focus on form: Theory, research, and practice. In C. Doughty and J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press.
- Loschky, L. and Bley-Vroman, R. 1991. Grammar and task-based methodology. In G. Crookes & S. M. Gass (Eds.), *Tasks and Language Learning: Integrating theory and practice*. Clevedon: Multilingual Matters.
- Lynch, T. and Maclean, J. 2000. Exploring the benefits of task repetition and recycling for classroom language learning. *Language Teaching Research*, 4/3: 221-250.
- Lynch, T. and Maclean, J. 2001. 'A case of exercising': Effects of immediate task repetition on learners' performance. In M. Bygate, P. Skehan, and M. Swain (Eds.).
- Lyster, R. 1998a. Recasts, repetition, and ambiguity in L2 classroom discourse. *Studies in Second Language Acquisition*, 20: 51-81.
- Lyster, R 1998b. Negotiation of form, recasts, and explicit correction in relation to error types and learner repair in immersion classrooms. *Language Learning*, 48: 183-218.
- Lyster, R. and Ranta, L. 1997. Corrective feedback and learner uptake: Negotiation of form in communicative classrooms. *Studies in Second Language Acquisition*, 19: 37-66.
- MacCallum, R., Widaman, K. F., Zhang, S. and Hong, S. 1999. Sample size in factor analysis. *Psychological Methods*, 4/1: 84-99.
- Makino, T. 1979. *English Morpheme Acquisition Order of Japanese Secondary School Students*. Unpublished doctoral dissertation, University of New Mexico, Albuquerque.
- Master, P. 1990. Teaching the English articles as a binary system. *TESOL Quarterly*, 24/3: 461-78.



- Master, P. 1997. The English article system: acquisition, function and pedagogy. *System*, 25/2: 215-232.
- McLaughlin, B. 1987. *Theories of Second Language Learning*. London: Arnold.
- McLaughlin, B. 1990. Restructuring. *Applied Linguistics*, 11/2: 113-128.
- McLaughlin, B., Rossman, T., and McLeod, B. 1983. Second language learning: An information-processing perspective. *Language Learning*, 33: 135-157.
- Mehnert, U. 1998. The effects of different length of time for planning on second language performance. *Studies in Second Language Acquisition*, 20: 52-83.
- Mitchell, R. and Myles, F. 1998. *Second Language Learning Theories*. London: Arnold.
- Miyake, A. and Shah, P. (Eds.) *Models of Working Memory: Mechanisms of active maintenance and executive control*. Cambridge: Cambridge University Press.
- Morrison, D. and Low, G. 1983. Monitoring and the second language learner. In J. C. Richards & R. W. Schmidt (Eds.), *Language and Communication*, London: Longman.
- Nairne, J. 1996. Short-term/working memory. In E. L. Bjork and R. Bjork (Eds.), *Memory*. San Diego: Academic Press.
- Nakajima, M. 2005. Tense switching in written narrative discourse by learners of Japanese. Paper presented at Second Language Research Forum, Columbia University, New York.
- Nation, P. 2001. *Learning Vocabulary in Another Language*. Cambridge: Cambridge University Press.
- Nattinger, J. R. and DeCarrico, J. S. 1992. *Lexical Phrases and Language Teaching*. Oxford: Oxford University Press.
- Nisbett, R. E. and Wilson, T. D. 1977. Telling more than we can know: verbal reports on mental processes. *Psychological Review*, 84/3: 231-259.

- Nitta, R. 2004. Online planning and its focus-on-form effects in L2 oral performance: A case of a single learner of English. Paper presented at 37<sup>th</sup> British Association for Applied Linguistics Annual Conference, Kings College, London.
- Nitta, R. and Gardner, S. 2005. Consciousness-raising and practice in ELT coursebooks. *ELT Journal*, 59/1: 3-13.
- Nobuyoshi, J. and Ellis, R. 1993. Focused communication task and second language acquisition. *ELT Journal*, 47/3: 203-210.
- Norris, J. and Ortega, L. 2000. Effectiveness of L2 instruction: A research synthesis and quantitative meta-analysis. *Language Learning*, 50/3: 417-528.
- Norris, J. and Ortega, L. 2001. Does type of instruction make a difference? Substantive findings from a meta-analytic review. *Language Learning*, 51/Supplement: 157-213.
- Nunan, D. 1989. *Designing Tasks for the Communicative Classroom*. Cambridge: Cambridge University Press.
- Nunan, D. 1992. *Research Methods in Language Learning*. Cambridge: Cambridge University Press.
- Nunan, D. 2005. *Task-Based Language Teaching*. Cambridge: Cambridge University Press.
- Nunnally, J. C. 1978. *Psychometric Theory*. New York: McGraw-Hill.
- Ochs, E. 1979. Planned and unplanned discourse. In T. Givón (Ed.), *Syntax and Semantics 12: Discourse and Syntax*. New York: Academic Press, 51-80.
- Oller, J. W. and Redding, E. Z. 1971. Article usage and other language skills. *Language Learning*, 21: 85-95.
- O'Malley, M. J. and Chamot, A. U. 1990. *Learning Strategies in Second Language Acquisition*. Cambridge: Cambridge University Press.

- Ortega, L. 1995. *The Effects of Planning in L2 Spanish Oral Narratives* (Research Note No. 15). Honolulu: University of Hawai'i, Second Language Teaching and Curriculum Centre.
- Ortega, L. 1999. Planning and focus on form in L2 oral performance. *Studies in Second Language Acquisition*, 21: 108-48.
- Ortega, L. 2005. What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.
- Paribakht, T. 1986. Strategic competence and language proficiency. *Applied Linguistics*, 6/2: 132-146.
- Pawley, A. and Syder, F. H. 1983. Two puzzles for linguistic theory: nativelike selection and nativelike fluency. In J. C. Richards & R. W. Schmidt (Eds.), *Language and Communication*, London: Longman.
- Pawley, A. and Syder, F. H. 2000. The one-clause-at-a-time hypothesis. In H. Riggenbach (Ed.), *Perspectives on Fluency*. Ann Arbor: University of Michigan Press.
- Philip, J. 2003. Constraints on "noticing the gap": Non-native speakers' noticing of recasts in NS-NNS interaction. *Studies in Second Language Acquisition*, 25: 99-126.
- Pica, T. 1985. The selective impact of classroom instruction on second-language acquisition. *Applied Linguistics*, 6/3: 214-22.
- Pica, T. 1988. Interlanguage Adjustments as an Outcome of NS-NNS Negotiated Interaction. *Language Learning*, 38/1: 45-74.
- Pica, T., Holliday, L., Lewis, N. and Morgenthaler, L. 1989. Comprehensible output as an outcome of linguistic demands on the learner. *Studies in Second Language Acquisition*, 11: 63-90.
- Pienemann, M. 1985. Learnability and syllabus construction. In K. Hyttenstan & M. Pienemann (Eds.), *Modelling and Assessing Second Language Acquisition*.

Clevedon: Multilingual Matters.

Pienemann, M. 1998. *Language Processing and Second Language Development*. Amsterdam: John Benjamins.

Pienemann, M. and Johnston, M. 1987. Factors influencing the development of language proficiency. In D. Nunan (Ed.), *Applying Second Language Acquisition Research*. Adelaide, South Australia: National Curriculum Resource Centre.

Pinter, A. 2005 Task repetition with 10-year-old children. In C. Edwards & J. Willis (Eds).

Pinter A. 2007. Benefits of peer-peer interaction: 10 year-old children practising with a communication task. *Language Teaching Research*, 11/2: 1-19.

Polio, C. 1997. Measures of linguistic accuracy in second language writing research. *Language Learning*, 47: 101-143.

Poulishse, N. 1993. A theoretical account of lexical communication strategies. In R. Schreuder & B. Weltens (Eds.), *The Bilingual Lexicon*. Amsterdam: John Benjamins.

Poulishse, N. 1997. Language production in bilinguals. In A. M. D. de Groot & J. F. Kroll (Eds.), *Tutorials in Bilingualism: Psycholinguistic perspectives*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.

Poulishse, N. and Bongaerts, T. 1994. First language use in second language production. *Applied Linguistics*, 15: 36-57.

Prahu, N. S. 1987. *Second Language Pedagogy*. Oxford: Oxford University Press.

Richards, K. 2003. *Qualitative Inquiry in TESOL*. Basingstoke, Hampshire: Palgrave Macmillan.

Riggenbach, H. 1991. Toward an understanding of fluency. *Discourse Processes*, 14: 423-441.

Robinson, P. 1995. Task Complexity and second language narrative discourse.

- Language Learning*, 45/1: 99-140.
- Robinson, P. 1997. Individual differences and the fundamental similarity of implicit and explicit adult second language learning. *Language Learning*, 47/1: 45-99.
- Robinson, P. 2001a. Task complexity, task difficulty, and task production: Exploring interaction in a componential framework. *Applied Linguistics*, 22/1: 27-57.
- Robinson, P. 2001b. Task complexity, cognitive resources, and syllabus design. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.
- Robinson, P. 2002. Learning conditions, aptitude complexes, and SLA: A framework for research and pedagogy. In P. Robinson (Ed.), *Individual Differences and Instructed Language Learning*. Amsterdam: John Benjamin.
- Robinson, P. 2003. Attention and memory during SLA. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition*. Oxford: Blackwell.
- Royce, J. R. 1963. Factors as theoretical constructs. In D. N. Jackson and S. Messick (Eds.), *Problems in Human Assessment*. New York: McGraw-Hill.
- Rubdy, R. 1998. Key concepts in ELT: Task. *ELT Journal*, 52/3: 264-265.
- Ruso, J. E., Johnson, E. J. and Stephens, D. L. 1989. The validity of verbal protocols. *Psychological Review*, 17/6: 759-769.
- Rutherford, W. E. 1987. *Second Language Grammar: Learning and Teaching*. London: Longman.
- Rutherford, W. E. and Sharwood-Smith, M. 1985. Consciousness raising and Universal Grammar. *Applied Linguistics*, 6/2: 274-282.
- Samuda, V. 2001. Guiding relationships between form and meaning during task performance: The role of the teacher. In M. Bygate, P. Skehan & M. Swain (Eds.).
- Sangarun, J. 2005. The effects of focusing on meaning and form in strategic planning. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*.

Amsterdam: John Benjamins.

- Sato, C. 1988. Origins of complex syntax in interlanguage development. *Studies in Second Language Acquisition*, 10/3: 371-95.
- Schmidt, R. 1990. The role of consciousness in second language learning. *Applied Linguistics*, 11/2: 129-158.
- Schmidt, R. 1992. Psychological mechanisms underlying second language fluency. *Studies in Second Language Acquisition*, 14: 357-385.
- Schmidt, R. 1993. Awareness and second language acquisition. *Annual Review of Applied Linguistics*, 13: 206-226.
- Schmidt, R. 2001. Attention. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.
- Schmidt, R. and Frota, S. 1986. Developing basic conversational ability in a second language: A case study of an adult learner of Portuguese. R. Day (Ed.), *Talking to Learn: Conversation in second language acquisition*. Rowley, MA: Newbury House.
- Scovel, T. 1998. *Psycholinguistics*. Oxford: Oxford University Press.
- Segalowitz, N. 1991. Does advanced skill in a second language reduce automaticity in the first language? *Language Learning*, 41/1: 59-84.
- Segalowitz, N. 2003. Automaticity and second languages. In C. J. Doughty & M. H. Long (Eds.), *The Handbook of Second Language Acquisition*. Oxford: Blackwell.
- Selinker, L. and Han, Z-H. 2001. Fossilisation: Moving the concept into empirical longitudinal study. In C. Elder, A. Brown, E. Grove, K. Hill, N. Iwashita, T. Lumley, T. McNamara & K. O'Loughlin (Eds.), *Experimenting with Uncertainty: Essays in honour of Alan Davies: Studies in Language Testing 11*. Cambridge: Cambridge University Press.
- Sharwood-Smith, M. 1981. Consciousness-raising and the second language learner. *Applied Linguistics*, 11/2: 159-168.

- Sharwood-Smith, M. 1993. Input enhancement in instructed SLA: Theoretical bases. *Studies in Second Language Acquisition*, 15/2: 165-179.
- Shavelson, R. Webb, N. and Burstein, L. 1986. Measurement of teaching. In M. Wittrock (Ed.), *Handbook of Research on Teaching*. New York: Macmillan.
- Shehadeh, A. 1999. Non-native speakers' production of modified comprehensible output and second language learning. *Language Learning*, 49: 627-75.
- Shehadeh, A. 2001. Self- and other-initiated modified output during task-based interaction. *TESOL Quarterly*, 35: 433-57.
- Sinclair, J. 1991. *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.
- Skehan, P. 1989. *Individual Differences in Second Language Learning*. London: Arnold.
- Skehan, P. 1995. Analysability, accessibility, and ability for use. In G. Cook & B. Seidlhofer (Eds.), *Principle and Practice in Applied Linguistics*. Oxford: Oxford University Press.
- Skehan, P. 1996a. A framework for the implementation of task-based instruction. *Applied Linguistics*, 17: 38-62.
- Skehan, P. 1996b. Second language acquisition research and task-based instruction. In J. Willis & D. Willis (Eds.), *Challenge and Change in Language Teaching*. Oxford: Macmillan Education.
- Skehan, P. 1998. *A Cognitive Approach to Language Learning*. Oxford: Oxford University Press.
- Skehan, P. 2003. Review article: Task-based instruction. *Language Teaching*, 36: 1-14.
- Skehan, P. 2004. Fluency and second language speech performance. Poster presented at 37<sup>th</sup> British Association for Applied Linguistics Annual Conference, Kings

College, London.

Skehan, P. and Foster, P. 1997. Task type and task processing conditions as influences on foreign language performance. *Language Teaching Research*, 1/3: 185-211.

Skehan, P. and Foster, P. 1999. The influence of task structure and processing conditions on narrative retellings. *Language Learning*, 49: 93-120.

Skehan, P. and Foster, P. 2001. Cognition and tasks. In P. Robinson (Ed.), *Cognition and Second Language Instruction*. Cambridge: Cambridge University Press.

Skehan, P. and Foster, P. 2005. Strategic and on-line planning: The influence of surprise information and task time on second language performance'. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.

Stevens, J. P. 2002. *Applied Multivariate Statistics for the Social Sciences (Fourth Edition)*. Mahwah, N.J.: Lawrence Erlbaum Associates, Publishers.

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*. Rowley, Mass.: Newbury House.

Swain, M. 1995. Three functions of output in second language learning. In G. Cook and B. Seidlhofer (Eds.), *Principles and Practice in Applied Linguistics: Studies in honour of H. G. Widdowson*. Oxford: Oxford University Press.

Swain, M. 1998. Focus on form through conscious reflection. In C. Doughty and J. Williams (Eds.), *Focus on Form in Classroom Second Language Acquisition*. New York: Cambridge University Press.

Swain, M. and Lapkin, S. 1995. Problems in output and the cognitive processes they generate: A step towards second language learning. *Applied Linguistics*, 16/3: 371-391.

Swain, M. and Lapkin, S. 2001. Focus on form through collaborative dialogue: Exploring task effects. In M. Bygate, P. Skehan & M. Swain (Eds.).



- Tabachnick, B. G. and Fidell, L. S. 2001. *Using Multivariate Statistics (Fourth Edition)*. Boston: Allyn & Bacon.
- Tarone, E. 1980. Communication strategies, foreigner talk, and repair in interlanguage. *Language Learning*, 30: 417-431.
- Tarone, E. 1982. Systematicity and attention in interlanguage. *Language Learning*, 32: 69-82.
- Tarone, E. 1983. On the variability of interlanguage systems. *Applied Linguistics*, 4: 142-63.
- Tarone, E. and Liu, G. Q. 1995. Situational context, variation, and second language acquisition theory. In G. Cook and B. Seidlhoffer (Eds.), *Principle and Practice in Applied Linguistics*. Oxford: Oxford University Press.
- Tavakoli, P. and Skehan, P. 2005. Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and Task Performance in a Second Language*. Amsterdam: John Benjamins.
- Thomas, M. 1994. Assessment of L2 proficiency in second language acquisition research. *Language Learning*, 44/2: 307-336.
- Thornbury, S. 1999. *How to Teach Grammar*. London: Longman.
- Tomlin, R. and Villa, V. 1994. Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, 16: 183-203.
- Towell, R. and Hawkins, R. 1994. *Approaches to Second Language Acquisition*. Clevedon: Multilingual Matters.
- Towell, R., Hawkins, R. and Bazergui, N. 1996. The development of fluency in advanced learners of French. *Applied Linguistics*, 17/1: 84-119.
- Young, R. 1996. Form-function relations in article in English interlanguage. In R. Bayley & D. R. Preston (Eds.), *Second Language Acquisition and Linguistic Variation*. Amsterdam: John Benjamins.

- Yuan, F. and Ellis, R. 2003. The effects of pre-task planning and on-line planning on fluency, complexity and accuracy in L2 monologic oral production. *Applied Linguistics*, 24: 1-27.
- Yule, G. 1997. *Referential Communication Tasks*. Mahwah, NJ: Lawrence Erlbaum Associates, Publishers.
- Van Gelderen, A. 1994. Prediction of global ratings of fluency and delivery in narrative discourse by linguistic and phonetic measures - oral performances of students aged 11-12 years. *Language Testing*, 11: 291-319.
- Van Hest, E. 1996. *Self-Repair in L1 and L2 Production*. Tilburg University Press.
- Van Lier, L. 1988. *The Classroom and the Language Learner*. London: Longman.
- VanPatten, B. 1990. Attending to content and form in the input: An experiment in consciousness. *Studies in Second Language Acquisition*, 12: 287-301.
- VanPatten, B. 1996. *Input Processing and Grammar Instruction*. New York: Ablex.
- Weinert, R. 1995. The role of formulaic language in second language acquisition: A review. *Applied Linguistics*, 16/2: 181-205.
- Wendel, J. N. 1997. *Planning and Second Language Narrative Production*. Unpublished doctoral dissertation, Temple University, Japan.
- White, P. 1982. Beliefs about conscious experience. In G. Underwood (Ed.), *Aspects of Consciousness. Vol. 3: Awareness and self-awareness*. London: Academic Press.
- Widdowson, H. 1989. Knowledge of language and ability for use. *Applied Linguistics*, 10: 128-37.
- Widdowson, H. 1998. Skills, abilities, and contexts of reality. *Annual Review of Applied Linguistics*, 13: 260-75.
- Wigglesworth, G. 1997. An investigation of planning time and proficiency level on oral test discourse. *Language Testing*, 14: 85-106.

Willis, J. 1996. *A Framework for Task-Based Learning*. Harlow: Pearson Education.

Woods, D. 2001. In search of fluency: What is it and how can we teach it? *Canadian Modern Language Review*, 57/4: 573-589.

Wray, A. 2000. Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21/4: 463-489.

Wray, A. 2002. *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.

# APPENDICES

## **Chapter 3**

Appendix 3-A: Participant's Consent Form

Appendix 3-B: Pre-Task Questionnaire

Appendix 3-C: Results of Global Ratings

Appendix 3-D: Examples of Speeches by High and Low Proficiency Speakers

Appendix 3-E: Summary of the Pilot Study

Appendix 3-F: Story-Telling Tasks

Appendix 3-G: Vocabulary and Structures for Stories

Appendix 3-H: Instructions for Planning Implementations

## **Chapter 5**

Appendix 5-A: Examples of Reports on Speech Production Processes

Appendix 5-B: Examples of Reports on Planning Strategies

## **Appendix: Chapter 3**

### **Appendix 3-A: Participant's Consent Form**

Experimental Research on Japanese Speakers of English in ESL Contexts

Term 1 2004/2005

Thank you for agreeing to participate in my research into Japanese speakers of English.

I plan to use the data (audio recordings of speech and interviews) essentially for research purposes – to analyse, identify features of oral performance. I will observe the usual anonymity practices. For example, names of people will be changed or removed as far as possible. Also, I would like to ask you not to tell anyone else about the content of the research, because it is likely that I ask other people to participate in the research too, and knowing what to be asked in the research beforehand may have significant influence on the results.

Many thanks for your generous co-operation. Please contact me if you have any questions or concerns about this.

Mr Ryo Nitta, MPhil/PhD Student, Centre for English Language Teacher Education,  
University of Warwick, Coventry, CV4 7AL,

Email: [R.Nitta@warwick.ac.uk](mailto:R.Nitta@warwick.ac.uk)

Phone: 01926 \*\*\*\*\*

## Appendix 3-B: Pre-Task Questionnaire

Date of Experiment \_\_\_\_\_

Task Type \_\_\_\_\_

Please answer the following questions

### 1. Gender

- Male
- Female

### 2. Age

- 18-21 years
- 22-26 years
- 27 years and up

### 3. Course

Major \_\_\_\_\_

- Undergraduate
- Master
- Doctor
- Others (Please specify \_\_\_\_\_)

4. When did you arrive in the UK? \_\_\_\_\_

5. Have you stayed in English-speaking countries more than half a year?

- Yes (Please specify \_\_\_\_\_ Years \_\_\_\_\_ Months)
- No

### Appendix 3-C: Results of Global Ratings

ID	LOR	Proficiency	Average	Rater 1	Rater 2	Rater 3
	Group	judgment	score			
1	A	Low	2.00	2	2	2
2	A	Low	3.00	3 (2)	3 (4)	3 (2)
3	A	Low	2.33	3	2	2
4	A	Low	2.00	2	2	2
5	A	Low	2.00	2 (4)	2	2
6	A	Low	2.33	3	2	2
7	A	Low	3.67	4 (6)	4 (3)	3 (2)
8	A	Low	2.34	3	2	2
9	A	Low	3.00	3 (2)	3 (4)	3 (2)
10	B	High	6.00	6	6	6
11	B	High	5.33	6	5 (4)	5
12	B	Low	3.00	3	3	3
13	B	High	4.00	4 (5)	4	4 (3)
14	B	Low	3.67	3	4	4
15	B	Low	2.00	2	2	2
16	B	Low	2.00	2 (5)	2	2
17	B	High	4.33	4	4	5
18	B	High	4.33	4	4	5
19	C	High	4.00	4	4	4
20	C	High	4.00	4 (5)	4 (3)	4
21	C	High	4.67	5	4 (3)	5
22	C	High	5.00	5	5	5
23	C	High	5.00	5	5 (3)	5
24	C	High	5.00	5 (6)	5 (4)	5
25	C	High	5.00	5	5	5
26	C	High	5.67	6	5	6
27	C	High	4.67	4	5	5
<b>Total</b>	-	-	3.72	4.07	3.44	3.59

Note: bracketed numbers are the first judgement.

## Appendix 3-D: Examples of Speeches by High and Low Proficiency Speakers

### [High proficient speaker]

when the bus arrived (.7) um (.5) there there was long queue so (.5) big boys (.9) er went first (.3) and eventually (.2) three boys couldn't get in (1.1) and:: also the bus driver (.2) er (.4) sh- sh- er shat them off (.2) saying 'it's full' (1.7) so (.4) it was just before (.2) three (.3) and they have to wait (.4) they had to wait another thirty minutes (.5) because the next bus arrived (.6) um (.3) thirty minutes later (1.4) so (.8) so this time they (1.0) successfully got on a bus (1.3) and on the way (.7) to:: (.6) their (.5) destination they (.4) saw (.5) the (1.7) previous bus (.5) was (.2) parked (.8) on (1.3) the road (.2) well (.5) side of the road (1.0) because it was being repaired (.8) er some kind of (.7) er (.3) break down or something (.8) so (.3) after all they got first because (.3) the (.2) the first bus (.2) had to be (1.0) repaired on the (.4) road where they could go on and on (.4) and they three boys happily waving (.8) well (.2) happily waving (.8) \* (.3) the (.6) big a (.6) boys who:: didn't (1.5) let them (.5) get on (.7) the first time

### [Low proficient speaker]

so um the truck splash the water (.3) and (1.6) they (.2) their their clothes dirty (.6) and (.2) um (.3) when (.5) when bus (.5) come (.8) bus come (.4) um their- (1.8) youn- youn- (.2) young people um (.3) push (.2) push them (.5) so they can't (.5) take a bus (.9) and (.2) er:: (1.3) three o'clock but um (.3) no no thirty minutes later (.2) she can (.4) get bus (.9) and (.4) on on the way to:: (.2) their destination (.3) um they (1.1) they sh- watch (.7) they watched bus um is (.8) clash (.4) clashing (.4) clashed (.6) so (1.5) and er yeah that's all



### **Appendix 3-E: Summary of the Pilot Study**

Before embarking on the main study, the pilot study (Nitta, 2004) was conducted with three Japanese speakers of English in a very similar way to the main study. All the participants were regarded as relatively advanced English users because they had stayed in the UK for a long time (more than one year) and they had had substantial experience of using English. The purpose of the pilot study was to establish whether the research questions of the study were researchable and whether any methodological problems could be identified prior to the main research.

In the pilot study, I examined the performance features, which are widely used in a number of task planning studies under the three distinctive areas of fluency, complexity and accuracy. One of the most important findings of the pilot study was that the on-line planning condition did not necessarily help speakers' performances to increase the level of accuracy. Table 3-A.1 shows that on-line planning was differentiated from non-planning in terms of the speech rate in all three participants; however, while participant 1 successfully improved his accuracy in on-line planning, the accuracy levels in the other speakers (participant 2 & 3) were more or less equivalent to or even lower than those in non-planning.

To put it another way, as a natural consequence of their effort to speak carefully and the resulting freed cognitive resources, these two participants were also expected to become more conscious of paying attention to linguistic aspects and to monitor their speech production more carefully and efficiently, but this conditioning could not actually lead to an improvement of their accuracy.

**Table 3-E.1:** Summary of the Task Performance of the Pilot Study (\*all the numbers reported here are mean scores)

	Participant 1			Participant 2			Participant 3		
	NP	SP	OP	NP	SP	OP	NP	SP	OP
<b>Unpruned SR</b>	130.00	151.70	102.38	141.86	143.56	132.38	110.00	117.60	108.80
<b>Pruned SR</b>	123.64	139.81	89.52	131.14	135.46	129.00	88.00	104.40	86.40
<b>Pauses</b>	14.00	27.00	39.00	33.00	25.00	38.00	12.00	7.00	11.00
<b>Self-repairs</b>	4.00	7.00	6.00	6.00	6.00	4.00	2.00	4.00	9.00
<b>Clauses per AS-unit</b>	1.90	2.00	2.06	1.95	2.24	1.54	1.50	1.33	2.50
<b>Error-free clauses</b>	57.89	60.71	81.82	58.97	65.96	47.50	40.00	87.50	55.00
<b>Correct verb forms</b>	63.16	78.57	81.82	58.97	65.96	47.50	40.00	87.50	55.00

Despite such results, it should be stressed that this failure to gain similar results to Yuan and Ellis' (2003) research does not entirely contradict the effects of on-line planning, because even those speakers who failed to improve accuracy in on-line planning, as reported in the following interview, still noticed differences of cognitive processes between on-line planning and the other planning conditions. Clearly, they became linguistically ambitious in on-line planning and tried to use more syntactically and pragmatically elaborated language. Although this attempt was not very successful in terms of quantitative results of accuracy variables, the retrospective interview evidently revealed their conscious involvement in syntactic processing in on-line planning, which was not realized in the other conditions.

Understanding the findings in the pilot study, on-line planning is likely to help L2 speakers involve syntactic processing, by releasing on-line time pressure to some extent, but, as the pilot study suggests, it is doubtful whether it also contributes to making language production accurate, because the speakers are still not entirely freed from time pressure even in on-line planning. On-line planning is a different condition compared to writing, which allows 'unlimited' time to formulate and modify the language. L2 speakers, being given the on-line planning implementation,

could take a significant step to maximize their upper limits of IL, i.e., pay conscious attention to linguistic form and access declarative knowledge, but this condition does not always guarantee focus-on-form and accuracy improvement because communication pressure still remains.

Looking back to the different results of the three speakers in the pilot study, the question, why only one of the speakers (i.e., participant 1), in accordance with the results in Yuan and Ellis (2003), could improve accuracy in on-line planning, remains unresolved. A key to unfolding this result is the extent to which speakers are involved in on-line planning. It is highly likely that the participants engaged in on-line planning (i.e., taking sufficient pausing) in different degrees, according to their interpretations of the instruction of the three planning conditions, and probably preference of their speaking styles (e.g., more fluency-oriented, or accuracy-oriented). Compared to the speech rates in on-line planning with those in non-planning (see Table 3-A.1), there is a significant gap between the accuracy-improved speaker (participant 1) and the unimproved speakers (participant 2 & 3). The results showed that the latter type of speaker ‘partially’ engaged in on-line planning, while the extent of on-line planning that participant 1 engaged in was significantly larger. It can be thus conceivable that such an unpressured state could allow him to comfortably engage in morphosyntactic processing, successfully leading to accuracy improvements. His verbal report on the condition clearly reveals that this condition pushed him into syntactic processing:

**Participant 1:** In the usual speech, because I’m more conscious of conveying a message to the interlocutor, I tend to speak more rapidly. But this time, I kept in mind that I was allowed to correct my utterance when I thought this was wrong.

**Researcher:** What did you think about while pausing?

**Participant 1:** Mainly grammar, I think... because I remembered the most content.

(Translated from Japanese)

Because the other participants also reported that they became ambitious to produce more complex language in on-line planning, it can be said that they were also consciously involved in morphosyntactic processing, but I could not elicit such explicit reporting on the formulation process, particularly accessing the rule-based system. The important aspect of participant 1's performance is that he not only consciously engaged in morphosyntactic processing but also successfully completed the whole formulation stage. In the retrospective interview, the other speakers also revealed their (more or less) consciousness of formulation process in on-line planning, but not leading to accuracy improvement in the end. To support this result, the impression of their on-line planning speech is not very different from natural speech rates. As shown in Table 3-A.1, the gap between the rates of non-planning and strategic planning and those of on-line planning in Participant 1 was larger than the other participants who could not greatly differentiate the rates in on-line planning from those in non-planning and strategic planning.

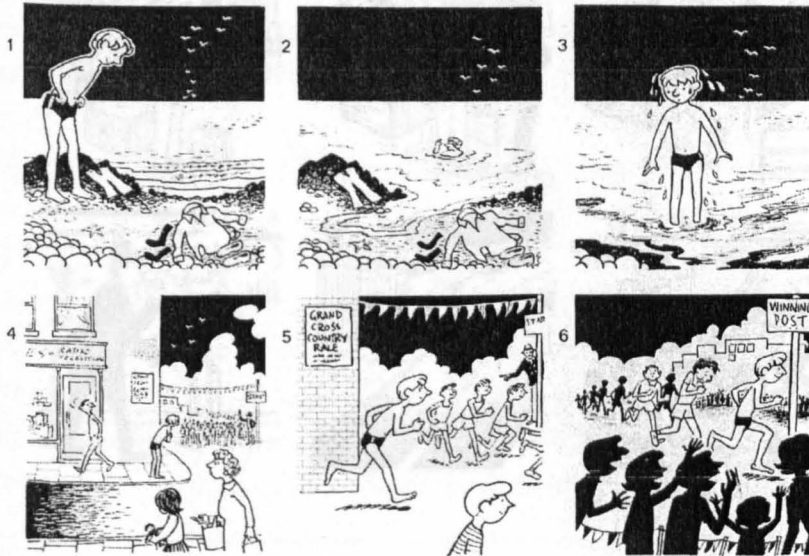
These observations drawn from the pilot study seem very important because they produced methodological concerns of how we differentiate on-line planning from strategic and non-planning. Following the argument that some sort of on-line planning always happens in speaking, and can also be described as 'careful speech' (see 2.2.3), on-line planning does not necessarily lead to better accuracy, in the same way as Yuan and Ellis (2003). Looking back to the performance of participant 1, his on-line planning performance was unnaturally slow, compared to those of the other two participants who spoke slowly and carefully but still within the natural rate of speech. It is necessary thus to emphasize speakers' attempts to speak carefully by monitoring their language before and after production through the instructions of planning.

## Appendix 3-F: Story-Telling Tasks

26 A surprise

### Rehearsal task

#### 22 The winner!



48

### Task 1

77 The choice

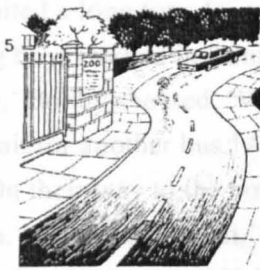
#### 25 Waiting for a bus



54

Appendix 3-G: Vocabulary Structures for Stories Task 2

26 A surprise



Task 2: 'A surprise'

Vocabulary: airport, Indian, porter, case, basket (dark) glasses, care, arm, policeman, whistle, corner, woods, surprise, fright, not large, huge, quietly, carefully, quickly, arrive, pick up, steal, follow, blow (a whistle), drive, open, appear

Task 3

27 The chase



Task 3: 'The chase'  
 Vocabulary: run, chase, catch up with, hold, attack, drop

Composition: One day last week I went shopping. It took me a long time to do all my shopping and I caught a bus home late in the evening. It was very dark when I got off the bus. I had to walk home through a lonely wood and I felt a little frightened.

## Appendix 3-G: Vocabulary and Structures for Stories

### Task 1: 'Waiting for a bus'

Vocabulary: queue, bus stop, lorry, bus conductor, arm, window, clock, country, driver, engine, passenger; full, steep, tired; (to) queue, splash, pass, come, get on, push (in front of), hold out, look out for, travel, stop, break down, try, mend, wave

Compositions: One day three small boys called Bobby, Charles and Peter went to visit their friend in a hospital. They waited a long time for a bus. When the bus arrived, it was nearly full. Suddenly four big boys pushed in front of them and got in the bus. 'We were in the queue before you,' Bobby shouted. 'We don't care,' one of the big boys answered. 'You'll have to wait for another bus.' After half an hour, the next bus came and the three boys got on. On their way to the hospital, they saw the first bus half-way up a hill in front of them. 'It's broken down,' Charles cried. 'Now those four boys will have to wait a long time for another bus,' Bobby laughed.

### Task 2: 'A surprise'

Vocabulary: airport, Indian, porter, case, basket, (dark) glasses, care, arm, policemen, whistle, corner, woods, surprise, fright, snake, lid; large, huge; quietly, carefully, quickly; arrive, pick up, steal, follow, blow (a whistle), drive, open, appear

Composition: One day an Indian with a basket arrived at an airport. It was too heavy to carry, so he put it down and began to look for a porter. Just then he noticed a small boy. The boy came up to him and began to talk to him. At the same time, a man with dark glasses picked up his basket. Then the man and the boy ran away. The Indian hurried to a policeman with a whistle and tried to stop the thieves. But it was no use. The boy and the man got into a car and drove off. They passed a zoo and then went towards a wood. When it reached the wood, it stopped and the man and the boy got out. They opened the basket and, to their great surprise, they saw a snake.

### Task 3: 'The chase'

Vocabulary: parcel, moon, starts, footpath, smoke, chimney, direction, bridge, river, something; parked, dark, windy, alone, distant, lonely, close(r), relieved; get off, hurry, blow, turn round, chase, catch up with, hold, attack, drop

Composition: One day last week I went shopping. It took me a long time to do all my shopping and I caught a bus home late in the evening. It was very dark when I got off the bus. I had to walk home through a lonely wood and I felt a little frightened.

Suddenly I heard a noise. There was a strange man behind me. I began to run home very quickly, too. Then I crossed a bridge over the river and saw my house a long way off. The next moment the man caught me and held my arm. 'I followed you to give you this parcel,' he said. I was very surprised and very glad, too. 'Thank you very much,' I laughed. 'You are very kind.'

(From Heaton, 1975)



## Appendix 3-H: Instructions for Planning Implementations

[Original instructions were given in Japanese.]

### **General instruction:**

I will give three story-telling tasks. Each task has the same format but different stories and no connection between them. What I would like you to do is that you tell the stories one by one, and after finishing the three tasks, answer several interview questions. I plan to finish all the process within 30 minutes at length.

### **Rehearsal task:**

Before starting, I would like to give a rehearsal opportunity to get to know what a story-telling task is like. The story I will give has the same format as tasks for the real tasks but different contents. Please remember the story within 30 seconds. You cannot look at the pictures while telling the story.

Please begin by the following sentence.

“One day a boy took his clothes off on the seashore.”

I will not record your speech this time.

### **Non-planning (NP)**

You will see a set of pictures. These pictures tell you a story. Now I would like you to retell this story in English. Imagine that somebody has never seen these pictures and this is his/her first time to learn about the story from you. So please tell the story as detailed as you can. I will not ask any question while you are telling the story.

You will be given 30 seconds to remember the story, and I will take the picture away while telling the story. In addition, you have only 2 minutes to tell the whole story.

Please write down the sentence I will read out, and begin by following the sentence:

### **Strategic planning (SP)**

You will see a set of pictures. These pictures tell you a story. In a short while, I would like you to retell this story in English. Imagine that somebody has never seen these pictures and this is his/her first time to learn about the story from you. So please tell

the story as detailed as you can. I will not ask any question while you are telling the story.

Before you retell the story, you have 10 minutes to plan what you are going to say. To assist you to prepare, you are given a sheet of paper and a pencil. You can use them to write some notes. But please don't write a complete sentences either your first language or in English. That is, I do not want you to remember and read out the whole sentences, but only make a note of the impression of the pictures and the outline. When you begin to tell the story, I will take the paper and the picture away. You have only 2 minutes to tell the whole story.

Please write down the sentence I will read out, and begin by following the sentence:

### **On-line planning (OP)**

You will see a set of pictures. These pictures tell you a story. Now I would like you to retell this story in English. Imagine that somebody has never seen these pictures and this is his/her first time to learn about the story from you. So please tell the story as detailed as you can. I will not ask any question while you are telling the story.

When telling the story, (1) if you find it difficult to say what you want, you can stop and think by taking as much time as you wish. (2) If you think you said something that was not correct or was not to your satisfaction, you can correct it as many times as you want. In these cases, don't worry about using less fluent English than you usually do. What is more important this time is correctness. Also, you have no time limitation to finish telling the story.

You will be given 30 seconds to remember the story, and I will take the picture away while telling the story.

Please write down the sentence I will read out, and begin by following the sentence:

## Appendix: Chapter 5

### Appendix 5-A: Examples of Reports on Speech Production Processes

Themes: Functions	Examples
<b>A. Conceptualizing processes</b>	
A1. Task evaluation: To evaluate the difficulty of the given task	A5: Here [when I was looking at the story] I couldn't understand the first picture very well [laughs]. I was pressured very much because only thirty seconds were given. And I looked through all the pictures quickly... (A5OP1:LOW)
A2. Message generation: To generate the message	B9: Well, in this part... um... without explaining the action during being followed, he would be caught up very quickly [laughs]. But how can I describe... I mean, there is a development of the story, and there is temporal development, but I was thinking whether I would explain this and how... yeah. (B9NP2:HIGH)
A3. Message regeneration: To regenerate the message without changing the meaning	R: How did you cope with, when you couldn't find the language? What sorts of way... B7: I felt like skipping the part... I felt like that. Skipping the part, or I might try to change the sentence...simplify it. I returned to the beginning and thought 'are there any good way?' (C7OP16:HIGH)
A4. Recalling: 1. To recall the pictures	C3: I was thinking how I could tell. I was recalling what pictures were like. R: Did you think about pictures, or the content, not the language? C3: Yes, I was thinking of the pictures. (C3OP1:HIGH)
A4. Recalling: 2. To recall the planned ideas (SP)	A4: In this case, because I made a plan, I was trying to recall the story, I mean, the plan itself. So... well... I could speak what I planned smoothly... I could tell by recalling the content... (A4SP8:LOW)
A5. Conceptual Simplification: To simplify the message	A3: I wanted to say 'succeed' here, I mean I wanted to describe 'did well', but I couldn't find these words, so I only tried to say 'did' anyway. (A3NP2:LOW)
A6. Conceptual Elaboration: To	C8: Well... I was thinking how detail I should tell the story. Because the story is very simple, I would finish it for about

elaborate the message	five minutes, wouldn't I? So, I was thinking how I could extend the given story. (C8OP9:HIGH)
A7. Avoidance of the intended message: To abandon the generated message	B9: I was thinking "isn't it a large snake called 'snake'?" R: You mean you were looking for a more appropriate word? B9: Yeah, I felt there is another words to describe this animal, but I couldn't find it [laughs]. Yes, this is [what I thought]. (B9OP9:HIGH)
A8. L1 use: To generate ideas in L1	B5: I mean, in my mind, well... "the bus which the boys took caught up with the earlier bus which they attempted to take", I knew this in Japanese, but I didn't know how to say it [in English] ...um ... I couldn't find it immediately, I mean, how to construct the structure by using 'the bus' as a subject. (B5OP4:LOW)
<b>B. Lexical operations</b>	
B1. Lexical search: To search for an appropriate lexical items	C1:sudden... I was thinking how to say here. I mean, how to describe it in English. R: You mean you remembered the content but... C1: Yes. But any appropriate language did not come up in mind, so I was thinking of this. I mean, vocabulary. How can I say this black shadow man. (C1OP2:HIGH)
B2. Lexical identification: 1. To identify the target lexical item	A5: Yeah, I was bearing in mind that I wanted to tell the story smoothly. R: You mean to tell the content? A5: Yes. R: Did you find the words efficiently? A5: Yes, I could find the words efficiently. (A5NP3:HIGH)
B2. Lexical identification: 2. Not to identify the target lexical item	B7: [...] the most difficult thing was finding the vocabulary. R: Vocabulary B7: It didn't appear in mind... R: Did you find it after searching? B7: There were cases I couldn't find ... (B7OP12-15:LOW)
B3. Generating lexical alternatives: To generate an alternative lexical item	A6: Here... I wanted to say "they were splashed water by bus twice", but I didn't know how to say it, and I was thinking [how to say it]. Then, after all, I told... What did I tell? R: You said 'wet'. A6: Yes, I only found 'wet' for "soaked".

	R: You were searching for this? A6: Yes. (A6NP1:LOW)
B4. Recycling: To recycle the previously used lexical items	A3: I only said 'scared' [laughs]. I said it about three times. I should have said other phrases like 'afraid', but only 'scare' for "frighten" in mind. (A3SP4:LOW)
B5. Avoidance of lexical items: To avoid using the intended lexical item	A1: Also, I wanted to say "a bus has come to a bus stop", but words for "bus stop" didn't come to my mind at once. (A1NP3:LOW)
B6. Lexical elaboration: To use elaborate lexical items	B9: I mean, yeah... I was wondering how I could say "catch up with him... and passed it to him". I didn't find the appropriate words here. R: This is, rather than just saying, you attempted more appropriate... B9: Yes. I mean, yes... I felt like more [laughs] more appropriate words, but I only found very simple verbs... (B9NP3:HIGH)

### C. Formulating processes

C1. Structural formulation: To formulate the structure	C5: Probably, when I was telling, I thought because they are close to the house, at the side of the house... I mean... there is a streetlight here, that person's face appeared... I said 'since it was near his home' and 'it wasn't completely dark'. Then, here I remember the road was very dark, so I think I was trying to tell "his faced was lighten up". Yeah. (C5SP17:HIGH)
C2. Structural reformulation: To reformulate the structure	B4: Their... I thought "oh, no, I can't continue the sentence in this way" and "OK, I'll change the way to tell", and I tried to return to the previous part. R: This means... B4: I wanted to say "they are planning to steal it", but I couldn't find the following structure after saying 'their plan' as a subject, so I noticed "I should use 'they' as a subject". R: I see, you changed the grammatical structure, didn't you? B4: Yes. R: Without changing the content very much... B4: Yes. I didn't have any structure following 'their plan' in

---

mind, and I noticed that the sentence using “they” as a subject must be easier., so I decided to change. I was instructed to correct as many times as I wish. (B4OP1:LOW)

---

C3. Self-correction:  
1. To correct  
ill-formed  
production

A1: Here I said ‘men’, but I recalled “there was a single person”, then I corrected it like ‘someone’. (A1OP4:LOW)

---

C3. Self-correction:  
2. Not to correct  
ill-formed  
production despite  
noticing errors

C7: In this part, I planned to tell “this boy didn’t notice that he dropped it”, before telling ‘after dark’, but I forgot telling this. After telling ‘after dark’ I noticed I didn’t tell, but I went ahead. (C7SP4:HIGH)

---

C4. Elaboration: To  
elaborate the  
structure

C5: In this part, when I said ‘caught him’... both are male, and I called one of them ‘boy’ in the first sentence, and I thought this is the best way.... But when I had to tell both in one sentence, both became ‘he’, and I thought “it is not easy to understand [to the listener]” in mind... I mean... After saying ‘the man caught him’, I said ‘the boy’, while thinking “what I mean ‘he’ is the boy” in mind, thought I thought it is repetitious... (C5SP9:HIGH)

---

C5. Avoidance of  
structural  
formulation: To  
abandon formulating  
the structure

C4: I thought that I should tell he wears a turban because there is no time limitation, but I didn’t tell it after all.  
R: Is this because you didn’t find the language?  
C4: I remembered the details of the pictures... yeah I think so. I didn’t know how to construct the language. And I felt like making it complicated, so I didn’t say it. (C4OP1:HIGH)

---

C6. Testing a  
structural  
hypothesis: To apply  
an uncertain  
structure

A1: Here I wanted to say the expressions meaning ‘escaped’, but I was thinking ‘I used “escape” before, so I want to use the other way’ [laughs]. I said ‘ran over’ while thinking ‘is this a correct phrase?’  
(A1OP3:LOW)

---

#### **D. Monitoring processes**

D1. Conceptual  
Monitoring

C7: When I said “boxes”, I thought ‘they were not boxes but bags’... yeah I thought it. (C7NP1:HIGH)

---

D2. Lexical  
monitoring

C8: In this part, I was thinking whether I should use ‘previous bus’ or ‘earlier bus’... I mean I found the phrase ‘previous bus’ quickly, but I felt like ‘previous’ is too formal in this

---

---

situation, and I thought I should use a different word, but I couldn't find 'earlier' immediately... Yeah, then I used 'previous'. (C8NP3:HIGH)

---

D3. Grammatical  
Monitoring

A1: Well, I wanted to say '[he] hurried home', and I wondered if "way to his home" is grammatically correct? [laughs].

R: Um, you noticed it after saying it, didn't you?

A1: While telling this, I was thinking that 'it seems not very correct'. (A1OP1:HIGH)

---

D4. Discourse  
monitoring

C5: Um, here, when I said "he was totally scared", I thought 'can I use "totally" here?', but I think I said "scared" at the beginning. And 'he was more scared...' er? Yeah, I wanted to say 'he is more scared', so I remember I added the adverb. Because I wanted to say he was more surprised than the last time he was hailed and started running. I wanted to describe the boy's psychological state like 'what will happen to me if I'm caught by him?' But I couldn't do this well... (C5SP11:HIGH)

---

D5. Phonological  
Monitoring

C5: "two of them...ran... ran away..." um... 'ran away'... um what did I think? Yes, I am pausing here... "Ran away"... perhaps the pronunciation of "ran".

R: To change it in the past form?

C5: Yeah... um... I was thinking 'this is R...R', and thinking 'am I pronouncing this?' "Ran away"... yeah, I didn't search for the other words. I think I was thinking of 'R' pronunciation. (C5NP3:HIGH)

---

## Appendix 5-B: Examples of Reports on Planning Strategies

<b>Conceptual planning</b>	
Focus on understanding the story	B4: Because of ten minutes, I could have the time to look at the pictures more carefully, so I planned the content in more detail. For example, “this is a bus at what time” and “how many men are approaching”. (B4SP1:HIGH)
Conceptualizing the message	C7: Well... I mean, first I looked at the pictures for a while, then...yeah...first of all the main, I mean, well... to tell the baseline of the story, I was thinking what I had to tell, so I was writing down this first, I mean I was writing down the main point of each picture in order. (C7SP1:HIGH)
Elaborating the message	C6: But, he is Indian, isn't he? I mean he is wearing a turban. So, I think he is Indian. Moreover, what is inside the box is a king cobra, isn't it? I added these things from my imagination. R: You mean you arranged the story or made it more interesting? C6: Yeah. (C6SP4:HIGH)
Extending the given story	C8: There isn't any picture that he is approaching, isn't it? When I said 'came', I noticed that my story doesn't match the given pictures. When I first looked at this story, there are scenes I need to predict between each picture, aren't they? [...] I mean, the story is expected to fill the gap between the pictures by myself, aren't they? (C8SP10:HIGH)
L1 use	C9: Um, yeah, I wasn't thinking very well, but as there was a lot of [planning] time, I took note in Japanese like this, and then after making myself clear about the story, I was thinking how I could explain it in English, yeah. (C9SP6:HIGH)
<b>Linguistic planning</b>	
Lexical search	B2: I was thinking of vocabulary rather than grammar. Yeah... and expressions, like “how can I say these expressions?” For example, in the last, number six picture [in 'Waiting for a bus'], I was thinking for a long



	time “Um, here it has a flat tire... How can I say a “flat tire” in English? ...Yes. I spent a lot of time... yes. (B2SP3:HIGH)
Summarizing the picture by identifying the key words	C1: First I summarized each picture by simple words, and then made the overall structure. (C1SP1:HIGH)
Generating alternatives	A9: This time, I thought the vocabulary was difficult. For example, I couldn't find the word “splash the water”, and I couldn't also know how to say “push out”... At the last scene too... I was thinking that the vocabulary is difficult this time, and how I could replace them with other words. (A9SP2:LOW)
Simplifying the structure	A6: Yeah, I was just thinking [of grammar], but I didn't want to make a mistake by using difficult words [laughs], so I tried to tell it by only using easy words. (A6SP4:LOW)
Formulating the structure	B7: For the rest of the time, I was practising the story [in mind], while thinking how I could formulate the sentences. (B7SP1:LOW)
Accessing the grammatical Knowledge	A5: Grammar... yeah, I see... ‘perception verb’? I was thinking of ‘notice’ or something like that, because these words might be useful. R: You mean how to give Object after ‘notice’? A5: Yes. ‘Notice’... I mean, I can describe it briefly by saying “notice someone doing something’ instead of saying it by a long sentence. (A5SP2:LOW)
Elaborating the structure	B5: Yeah... I'm not sure whether this is grammar, but ‘as fast as possible’, is this idiom? I was trying to use this kind of expressions... R: Did you think of this while speaking, not during the planning time? B5: No, I was thinking it during the planning time... yeah, I mean... to make the scene alive, I thought this kind of phrases would be good... (B5SP7:LOW)
<b>Metacognitive planning</b>	
Organizational planning	B5: Well... first of all, to understand the whole story... I mean I tried to understand only the point of each picture, after finishing this, I was fleshing out the story.

---

(B5SP1:LOW)

---

Mental rehearsal

B4: While ten minutes, I could repeat to practise the story three times. While practising it, I noticed new information, and I tried to include it in the story, and the first organized story was gradually diverted. (B4SP6: HIGH)

---

Selective rehearsal

R: Did you rehearse the story?  
A3: Um... I mumbled it. But I didn't practice the whole part but only the main parts, like I was practicing it just a little. (A3SP1:LOW)

---

Selective planning

B6: I thought the parts I could describe and "how can I say it in English?" But I avoided it, I mean I avoided the parts I couldn't describe, I mean the parts I couldn't explain accurately... (B6SP6-7:LOW)

---

Brainstorming

C3: Well, first, I did brainstorming for all the pictures and wrote down the keywords, but because I couldn't speaking with this note, I attempted to formulate sentences in mind for the last five minutes. I told the story in mind twice. (C3SP1:HIGH)

---

Memorization

B2: Well... it is memorable if I'm given ten minutes like this and write down the words...(B2SP11:HIGH)

---