

**TOWARDS A CATEGORIZATION AND DESCRIPTION OF READING  
STRATEGIES THROUGH THE USE OF INTROSPECTION WITH  
HONG KONG UNIVERSITY STUDENTS**

**JONATHAN CHARLES WILLIAM HULL**

**Thesis submitted for the degree of Doctor of Philosophy**

**Department of English Language  
University of Glasgow**

**March 2000**

# **TOWARDS A CATEGORIZATION AND DESCRIPTION OF READING STRATEGIES THROUGH THE USE OF INTROSPECTION WITH HONG KONG UNIVERSITY STUDENTS**

**Jonathan Hull**

## **ABSTRACT**

The research reported in this thesis explores approaches towards categorizing and describing reading behaviour. In the literature, these behaviours are widely known as 'strategies'. The focus of this study is on adults reading in a second language. The subjects, whose first language was Chinese (Cantonese), were a group of 38 first-year students at a university in Hong Kong. In order to try and gain access to these readers' mental behaviour while reading short academic articles, the research used introspective methodology. Two studies were conducted, one using introspection in the written mode, the other in the spoken mode. There was also variation in task across the two studies, one minimizing researcher intervention with a free-reading task conducted individually and silently, the other using a cloze task conducted orally and in pairs.

The bulk of the thesis describes the development of a categorization scheme of reading strategies that was driven by both the mentalistic data from the two studies and influenced by reading theory. The scheme developed consists of three levels, those of strategy category, subcategory and descriptor; the lower the level, the greater the level of detail recorded. The two categories represent the divide between mental actions, which are unobservable, and physical actions, which are observable. At subcategory level, there were various manifestations of mental and physical behaviour reported by the subjects. For mental actions, these involved the use of both text and non-text sources of information; for physical actions, these were reading that was recursive, ongoing and involved skipping as well as annotation of the text. The descriptor level, which contained groups of strategies, represented an attempt to ascribe wording to every strategy recorded. These strategy descriptors included slots permitting variability for part of text and aspect of language reported as well as the capacity to record such factors as task variation. Accommodating this level of detail should facilitate comparison of the findings of this study with those of other studies.

Current theory suggests that fluent reading involves complex behaviour requiring a high level of word recognition skills along with an interaction of text and non-text information, and this interactive quality was reflected in the classification scheme. Reports of behaviour that involved the integration, or even fusion, of two or more processes were represented in the scheme as strategies with combined categories; these were called 'compound' strategies.

While the main focus of the research was the development of this categorization scheme, some findings, derived from the categorization scheme, are presented in the thesis. These findings compare the reading behaviour of the subjects across the two studies. There was evidence among some readers that the free-reading task precipitated relatively global reading whereas the cloze task encouraged local reading. However, among the subjects, there was a group of readers who appeared to approach both tasks in a very similar, local manner. Consonant with theory and its bottom-up interactive view of reading, there was ample evidence in both studies of the subjects decoding, or attempting to decode, words.

# CONTENTS

<b>INTRODUCTION.....</b>	<b>1</b>
<b>CHAPTER 1: Theoretical foundations.....</b>	<b>7</b>
1.0 Introduction.....	7
2.0 Research into reading: A historical overview.....	8
2.1 History of L1 reading research.....	8
2.2 History of L2 reading research.....	11
2.3 Section summary .....	15
3.0 Reading theory and models of reading.....	16
3.1 Development of theory and models in L1 reading.....	17
3.1.1 Top-down models.....	18
3.1.2 Bottom-up models.....	21
3.1.3 Interactive models.....	24
3.1.4 Towards bottom-up interactive models .....	27
3.1.5 Models that incorporate affective factors.....	30
3.2 Development of theory in L2 reading.....	31
3.3 Section summary.....	37
4.0 The reading process.....	39
4.1 The reading process in L1.....	39
4.1.1 Converging evidence.....	39
4.1.2 The roles of words and context.....	41
4.1.3 The role of phonology.....	45
4.1.4 The role of writing systems .....	46
4.1.5 Summary.....	48
4.2 The reading process in L2.....	48
4.2.1 L2 deficiency.....	48
4.2.2 Schemata.....	50
4.2.3 Summary.....	52
4.3 Section summary.....	53
5.0 Chapter summary.....	54



<b>CHAPTER 2: Strategies</b>	<b>56</b>
1.0 Introduction	56
2.0 A brief historical outline	57
2.1 Paradigm shift	57
2.2 Strategy training	58
2.3 Section summary	59
3.0 Learning strategies: The problem of definition	61
3.1 Elusive quality	61
3.2 Overlapping terminology	63
3.3 Section summary	66
4.0 Reading strategies: The problem of definition	67
4.1 Lack of consensus	67
4.2 Operational constraints	68
4.3 Section summary	70
5.0 Strategies: The problem of classification	71
5.1 The cognitive psychology perspective	71
5.2 The applied linguistics perspective	74
5.3 A blending of perspectives	76
5.4 Section summary	80
6.0 Chapter summary	82
<b>CHAPTER 3: Mentalistic approaches</b>	<b>84</b>
1.0 Introduction	84
2.0 Theoretical considerations	87
2.1 Impact of prevailing theories in psychology	87
2.2 Impact of prevailing theories in linguistics	91
2.3 Section summary	94
3.0 Practical considerations	95
3.1 Task design	95
3.2 Research design	100
3.3 Naturalness	104
3.4 Section summary	106
4.0 Introspective methodology and reading strategies	108
4.1 Introspection in recent studies	108
4.2 Data-gathering procedures	113
4.2.1 Some procedures for verbal report data	113
4.2.2 Some procedures for written report data	119
4.3 Section summary	123
5.0 Chapter summary	125



**CHAPTER 4: Rationale, context and method.....127**

1.0 Introduction.....127

2.0 Rationale.....129

2.1 Breakpoints and breakpoint cycles.....129

2.2 Defining reading strategies.....131

2.3 Towards a research design.....133

2.4 Research aims and objectives.....134

3.0 Context: Hong Kong.....139

3.1 Language and tertiary education .....140

3.2 Section summary.....146

4.0 Method.....147

4.1 Development of the research design.....147

4.2 Pilot-testing.....150

4.3 Study 1 and Study 2.....155

4.4 Section summary.....161

5.0 Chapter summary.....163

**CHAPTER 5: Study 1.....164**

1.0 Introduction.....164

2.0 Development of the task.....166

2.1 Selection of the text.....166

2.2 Pilot-testing.....168

2.3 The written introspection task.....169

3.0 Data analysis .....172

3.1 Establishing strategy categories.....177

3.2 From descriptors to initial strategy groups.....181

3.3 Initial strategy groups.....183

3.4 From initial groups to interim strategy categories.....187

3.5 Interim strategy categories.....192

3.6 From interim to final categories..... 196

3.7 Final categorization.....208

3.8 Section summary.....216

4.0 Chapter summary.....217

**CHAPTER 6: Study 2.....218**

1.0 Introduction.....218

2.0 Development of the task.....219

2.1 Selection of the text.....219

2.2 Pilot-testing.....223

2.3 The oral introspection task.....227

3.0 Data analysis.....230

3.1 Procedures.....235

3.2 Initial phase.....236

3.3 Describing the pair-work dimension.....240

3.4 Review phase.....245

3.5 Section summary.....249

4.0 Some tentative findings from Studies 1 and 2.....251

5.0 Chapter summary.....264

**CHAPTER 7: Conclusions.....265**

1.0 Research review.....265

1.1 Reading theory and the study of reading behaviour.....265

1.2 Review of research aims and objectives.....269

1.3 Limitations.....276

2.0 Future directions.....282

2.1 Taxonomy for categorizing reading strategies.....282

2.2 Future research.....288

3.0 Pedagogic implications.....294

4.0 Chapter summary ..... 299



**FIGURES**

4.1 Proposed processing sequences in breakpoint cycles.....130

5.1 Proposed processing sequences in breakpoint cycles.....165

6.1 Contrasting interactive dynamics across Studies 1 and 2.....231

**TABLES**

1.1 Hedge’s (1991) reading modes.....36

2.1 Kirby’s (1988) definitions of learner activities.....63

2.2 Interconnections between and among learner activities (Kirby 1988).....64

3.1 Cohen’s (1984 and 1987) categories of mentalistic data.....96

3.2 Ericsson and Simon’s (1980) reporting paradigm.....97

3.3 Three types of triangulation (Grotjahn 1987).....102

3.4 Considerations in task design and research design .....106

3.5 Studies of reading strategies using mentalistic approaches.....110

3.6 Pros and cons of oral reading.....114

3.7 Pros and cons of cloze.....115

3.8 Pros and cons of interviewing.....118

3.9 Pros and cons of questionnaires.....119

3.10 Pros and cons of introspections on breakpoints.....121

3.11 Pros and cons of journals.....122

4.1 Research aims and objectives.....134

4.2 Features of research design.....148

4.3 Subjects’ ‘Use of English’ scores.....156

4.4 Schedule of data-gathering sessions.....159

5.1 Initial strategy groups.....173

5.2 Final categorization of L2 reading strategies.....174

5.3 Coding principles for analysis of reading strategies (Study 1).....176

5.4 Meanings covered in strategy descriptors.....179

5.5 Initial strategy groups.....186

5.6 Interim strategy categories.....193

5.7 Final categorization of L2 reading strategies.....209

5.8 Vertical continuum of reading behaviours.....212

5.9 Descriptors of Reflective Text Meaning-Centred Strategies.....213

5.10 Descriptors of Reflective Text Language-Centred Strategies.....213

5.11 Descriptors of Reflective Strategies in compounds.....214

5.12 Descriptors of Sensory Strategies.....215

5.13 Descriptors of Sensory and Reflective Strategies in compounds.....216

6.1 Profiles of some cloze passages in recent studies.....221

6.2 Cloze deletions by word class in the pilot version of the task.....222

6.3 Average word scores (%) for Pilot Study 1.....225

6.4 Cloze deletions by word class in the final version of the task.....228

6.5 Design differences between Study 1 and Study 2.....230

6.6 Final categorization of L2 reading strategies.....232

6.7 Coding principles for analysis of reading strategies (Study 2).....233

6.8 Variation in strategy descriptors across the two studies.....239

6.9 Reflective Text Meaning-Centred Strategies (singles).....246

6.10 Reflective Text Language-Centred Strategies (singles).....246



6.11 Reflective Text Meaning-Centred/Text Language-Centred Strategies  
(compounds).....247

6.12 Reflective Extra-Text-Centred/Text Meaning-Centred Strategies  
(compounds).....248

6.13 Reflective Non-Specific Strategies (singles).....248

6.14 Use of Reflective Strategies compared across studies.....252

6.15 Text categories.....254

6.16 Part of text utilized with Reflective Strategies compared.....255

6.17 Use of Reflective Strategies among selected reader groups compared.....257

6.18 Part of text utilized among selected readers compared.....259

6.19 Comparison of reading and overall ‘Use of English’ grades.....262

7.1 Features of research design .....279

**APPENDICES .....301**

4.1 Training sequence .....301

5.1 Study 1 text .....303

5.2 Study 1 task .....315

6.1 Study 2 text and task .....318

**REFERENCES .....320**

## ACKNOWLEDGMENTS

It is impossible to thank everyone who helped me in the writing of this thesis by name, but I would particularly like to thank the following people:

- my father, who was terminally ill when the project began but encouraged me to go and do my field work in Hong Kong, and my mother, who consistently supported me throughout
- my brothers, Julian and Philip, Philip having done his own PhD a decade before me but who still had a very vivid memory of the process
- friends who gave their active support, particularly Nicki, who was a mainstay throughout the process and who, on numerous occasions, invited me to her home for the most wonderful meals, and Stella and David, who phoned me on the evening of my submission deadline to check that I had completed
- my advisor, Florence Davies, who was always both challenging and supportive, who provided an exceptional breadth of knowledge of, and insight into, the area under study, an extraordinary generosity both intellectually and in terms of time and who conveyed an infectious enthusiasm about the project, even at times when I was at a low ebb
- the staff in the Department of English at City University of Hong Kong, particularly Jack Richards and Martha Pennington, for encouraging me to begin the project, and Lindsay Miller, who gave generously of his class time with his students, my subjects
- the subjects who participated in the research in Hong Kong, particularly the thirty-eight BA TESL students, who, despite the constant pressure they were under with their course work, gave generously and graciously of their time

Glasgow, March 2000

## INTRODUCTION

Different areas of study require different types of challenge. For an activity such as reading, which is normally silent and private, the most obvious challenge is how to get hold of the data without destroying, or at least altering, the process under study. On the one hand, empirical studies of eye movements (e.g., Rayner and Pollatsek 1989) require the use of equipment, such as the tachistoscope, which seem likely to create an artificial setting for the research. This kind of setting may be stressful to the subject and so affect their reading behaviour. There also seems to be a limit to what the findings from eye movement studies can say about mental behaviour, which is invariably unobservable. On the other hand, studies using introspective methodology (e.g., Hosenfeld 1977, Sarig 1987a and 1987b, Li and Munby 1996) run the risk of gathering mentalistic data that are, at best, incomplete, and, at worst, inaccurate. Readers may only report part of their behaviour and they may even, though not maliciously, report behaviour that did not take place. So, findings from mentalistic studies also have to be treated with caution.

Of the two types of research, the latter appealed to me more for a number of reasons. One factor was that there appeared to be an intriguing challenge in developing tasks that would tease out the most reliable data possible. The principal conundrum appears to be that, in order to capture reports of reading behaviour, the reading process itself has to be interrupted, but this very interruption is likely to skew the very process under study. Most introspective studies have used tasks that force interruptions on their readers, but some have included reader-generated pauses (e.g., Cavalcanti 1987). Another factor that appealed to me about conducting research using introspection to gather data was that it seemed to me that the majority of subjects would enjoy the experience of introspecting on their reading more than being surrounded by strange equipment. There are also clear pedagogic uses for introspective techniques, but they appear to be under-utilized in the average reading class.

Having decided on the type of methodology to use, another issue to consider was how to categorize the data. Initially, this might seem to be an easier challenge than that concerning data type; however, a quick look at the literature on reading strategies would soon disabuse anyone of such an impression. There are very divergent views



on what constitute a ‘reading strategy’, and, without a definition of what one seeks to categorize, there can be no categories. To illustrate the problem of defining strategies, here are some instances, taken from studies of reading in both the first language (L1) and the second language (L2):

“a subject’s use of context in responding to a mutilated text” (Clarke and Burdell 1977)

“deliberate means of constructing meaning from a text when comprehension is interrupted” (Kletzien 1991)

“processes for enhancing comprehension and overcoming comprehension failures” (Palincsar et al. 1984)

“those mental actions which readers consciously choose to use in accomplishing reading tasks” (Cohen 1986)

“a physical or mental action used consciously or unconsciously with the intention of facilitating text comprehension or learning” (Davies 1995)

From these five examples of definitions of reading strategies, it can be seen that there are several issues that have divided researchers for the past quarter-century, during which the vast majority of mentalistic studies of reading, especially L2 reading, have taken place. These issues include whether strategies have operational constraints such as a problem orientation, whether they are conscious or both conscious and unconscious, whether they are mental or both mental and physical, and whether they are comprehension-oriented or both comprehension-oriented and learning-oriented. Additional issues, not illustrated here, include whether they involve a single action or two or more related actions and whether they might include personal reactions, such as affective responses, to the text.

Assuming the adoption of a definition of reading strategies, the way may seem to be clear to categorize them. However, another look at the literature reveals an extraordinary number of classification schemes; in fact, every study of reading strategies seems to have its own scheme. These schemes vary in several ways. Given that the roots of reading strategies research can be traced to work investigating learning strategies, some schemes have a clear foundation in cognitive psychology. These have typically been developed in studies of L1 reading and the categories display an awareness of the reader processing information but do not address language

issues (e.g., Wade et al. 1990). In contrast, early studies of L2 reading are characterized by a focus on linguistic behaviour, the categories typically describing text features rather than the reader's mental behaviour (e.g., Barnett 1988a).

Thus, while some categorizations of reading strategies focused on the reader, others focused on the text, and the problem with this state of affairs is that it ignores reading theory. As early as 1977, Rumelhart's model made a good case for the 'interactive' nature of reading, that is, that it simultaneously combines top-down and bottom-up processing. Although Rayner and Pollatsek's (1989) influential bottom-up interactive model later modified this view somewhat to emphasize that rapid and accurate decoding of words is paramount in successful reading, it is clear that a principled classification of reading strategies needs to take account of both the reader and the text. More recent work on reading strategies, such as that of Sarig (1987a and 1987b) and Pritchard (1990) appears to have captured the interaction between reader and text.

An additional problem with many studies of L2 reading strategies is the manner in which the data are assigned to categories. Although the development of categories is necessary to provide a framework for analysis, too often, once developed, categories seem to be treated as discrete elements when, as has been observed, current theory suggests that reading can comprise simultaneous processing from more than one source. Where subjects report such integrated or fused processing, as for example with the combined use of prior knowledge and text information, the categorization scheme should have the flexibility to accommodate it. One of the motivations for this study was to investigate ways of building flexibility into the scheme developed.

As mentioned above, the norm in the literature on reading strategies is for each study to have its own categorization scheme. This makes it hard to compare findings across studies. While it may be too early to talk in terms of a definitive scheme, there are ways in which researchers developing their own schemes can facilitate comparison across studies and thus create a climate where there can be a greater degree of agreement about what a scheme should minimally contain. One way to do this is to record important information about how a study was conducted. For instance, since qualitative variation in strategy use may be precipitated by external input, it seems important to record whether strategies were reported spontaneously by subjects or



whether they were generated by researcher probing or peer interaction. However, recording this level of detail runs the risk of the information in the scheme becoming too dense, and the challenge here is to devise ways to include it without the scheme becoming overwhelmed by it.

Beyond these concerns about designing and conducting the proposed research, there is a concern about the continuing mismatch between reading theory and orthodox L2 reading pedagogy, which all too often still clings to the discredited top-down model of reading. It is hoped that this research can help to substantiate the evidence that already exists that readers need to be able to engage in fast and efficient bottom-up processing in order to free cognitive space for processing other sources of information.

There are therefore several challenges motivating this study. Among them is the development of a task that minimally changes the normal reading process. Exploring the possibilities will involve considering the options offered by introspective techniques. There is also the need to establish a definition of 'reading strategy' which embraces as wide a range of reading behaviour as possible. The categorization scheme developed needs to be able to accommodate a range of interactive behaviours, some of which, if the theory is correct, will straddle discrete categories. In addition, there is the motivation to represent the data in the scheme in a manner that is easily accessible to other researchers.

In pursuit of these goals, Chapter 1 of this thesis sets out to trace the development of reading theory from its origins to its current state. This shows that the polarized situation in the early 1970s, when there were proponents of extreme top-down and bottom-up views of reading, shifted in the late 1970s to a situation where both levels of processing were seen to interact. Reflecting mounting evidence about the necessity for readers to have efficient word recognition skills, this view was modified in the late 1980s to one that might be characterized as 'bottom-up interactive'. If the decoding of words is essential in reading, then it can be assumed that varying levels of language deficiency among L2 readers is likely to have varying effects on the process; indeed, a developmental theory of L2 reading, yet to be superseded, was put forward in the



early 1990s. The chapter goes on to review both L1 and L2 reading studies in the literature for more detailed evidence of what reading entails.

With its relatively narrow scope, Chapter 2 reviews the literature on L1 and L2 reading strategies. In seeking to determine the range and variety of behaviour that various researchers have included under the term 'strategy', it also considers such terms as reading 'styles', 'tactics', 'moves', 'skills' and 'abilities'. The chapter shows how elusive it has proved both to describe and to categorize reading behaviour and how further work in the area remains to be done.

Chapter 3 focuses on introspective methodology. It traces the impact of prevailing theories in cognitive psychology and linguistics on mentalistic approaches to the collection of data. From this basis, practical considerations in designing research projects and tasks are considered. Among these, at least for a portion of the data collected and where 'normal' reading is under study, is the importance of seeking to minimize research and researcher impact on the subjects' reading. There is also a review of L1 and L2 reading research studies that have used introspective techniques. From this, there is an evaluation of various procedures that can be used to collect introspective data in both the spoken and written modes.

Chapter 4 begins with a rationale for the research and outlines its aims and objectives. As already mentioned, in part of her study, Cavalcanti (1987) used reader-generated pauses in ongoing reading to tap her subjects' mental behaviour. The chapter describes how these 'breakpoints', as Hedge (1991) called them, provide a unique opportunity to record subjects' introspections concurrently with reading but without researcher-generated distractions. Processing routes showing variation in the manner in which readers deal with these internal interruptions and then continue reading are proposed in 'breakpoint cycles'. The chapter then moves on to describe the research context, Hong Kong, with particular reference to tertiary-level education there. Following this, the research design is outlined. This comprises two parallel studies, Studies 1 and 2, each with some features held constant and other features that vary; among the latter is the use of both written and spoken introspective data. The chapter ends with a description of the pilot-testing of the tasks, the training of the subjects and the data-gathering procedures for the two studies.

Chapters 5 and 6 describe Studies 1 and 2, respectively. Both chapters begin by outlining the development of the tasks and the texts utilized with them. A substantial part of each of the two chapters is then devoted to describing the manner in which the data were analyzed and the categorization scheme, with its three levels, was developed. Chapter 6 includes some findings; these were derived from the scheme and compare the reading behaviour recorded in the two studies.

Chapter 7 serves to review and evaluate the entire project. It also provides recommendations for future directions in this area of research. Finally, pedagogic implications are considered. These are particularly important because of the disparity between reading theory and much research evidence, on the one hand, and orthodox classroom practice, on the other.

## **CHAPTER 1: THEORETICAL FOUNDATIONS**

### **1.0 INTRODUCTION**

This chapter presents a review of the theoretical foundations and the nature of the reading process. While reading models can provide a theoretical foundation for researchers to design studies investigating the reading process, the reverse can also occur; that is, research findings can influence the evolution of theory. The rise of cognitive psychology in the past three decades has triggered a renewed interest in the investigation of mental processes in general as well as, in particular, the mental processes involved in reading. Although a substantial amount of work has now been done in L1 reading and a consensus has emerged on many of the essential elements in the reading process, L2 reading is still a relatively recent area to attract attention, either in terms of theory development or of practical investigation.

The following research questions are addressed in this chapter:

- 1.1 What have been major influences on the development of the study of reading?
- 1.2 What model or models of reading best describe the reading process?
- 1.3 What differences are there, if any, between reading in a first language (L1) and reading in a second language (L2)?

The chapter begins with a historical overview of L1 and L2 reading research. It continues with a discussion of the terms ‘theory’ and ‘model’ and a description of the progression in the development of theory and models of L1 and L2 reading; details of the most influential theories and models are presented. The chapter ends with a review of the literature on reading processes in both L1 and L2.



## **2.0 RESEARCH INTO READING: A HISTORICAL OVERVIEW**

Historically, reading research has had a far from smooth progression. As will be shown in this section, a key reason for this uneven development is that most of the processes that occur during reading are unavailable to the direct observation of investigators. Consequently, the study of reading processes is, in some measure, subject to prevailing conventions and beliefs surrounding the collection of mentalistic data.

The first part of this section is a historical overview of research into L1 reading from its beginnings more than a century ago. The second part is an overview of L2 reading research, whose history is barely three decades old.

### **2.1 History of L1 reading research**

Venezky (1984) has observed that “the history of reading research is a thing of many colors. It is not a single, continuous stream of human endeavor but at least four and perhaps as many as six independent threads” (p. 3). The four main strands he nominates are: basic research into reading processes, research into reading instruction, research into testing, and the study of literacy. As will be shown, the current study can be placed firmly in the first; that is, it investigates reading processes, an area of research which is itself multi-faceted. The discussion that follows, therefore, focuses on reading processes, or, specifically, mental processes in reading.

In recognition of the fact that the history of research into reading processes closely reflects developments in the field of psychology, this section is divided into three phases: early research work, the fallow period and the recent revival.

#### **• Early research work**

The study of perceptual processes in reading began in two European universities in the late 1870s. The world's first experimental psychology laboratory was founded by

Wundt at Leipzig University, Germany, in 1879. However, Wundt was not the main figure in the early reading studies; this was Cattell, an American PhD student working under him. Cattell investigated letter and word recognition, legibility of letters and print types, and span of attention. Cattell (1886), for example, found that his subjects needed twice as much time to read aloud unconnected words as connected ones (i.e., sentences) or unconnected letters as connected letters (i.e., words). However, Venezky (1984:6) reports that this result was ignored until the 1950s.

At the same time, Javal, working at the University of Paris, was conducting experiments on eye movements in reading. Javal (1879), in collaboration with Lamare, noted that eye movements are not, as hitherto assumed, smooth across text but that they are jumps, or ‘saccades’, and between each jump, there are ‘fixations’, when a reader's eyes remain relatively still.

During the next thirty years, reading processes were a central focus in experimental psychology research in Europe. Venezky (1984:7) notes that most of the basic processing problems considered important today were studied during this period, including cues for word recognition, the eye-voice span, the role of peripheral vision, memory for connected text, and subvocalization. However, around the turn of the century, the main thrust of the research effort moved from Europe to the United States.

A milestone in the history of research into reading processes was the publication of Huey's book ‘The psychology and pedagogy of reading’ in the U.S. in 1908. Huey was interested in eye movements in reading, the nature of the perceptual span, word recognition processes, inner speech, reading comprehension, and reading rate. Interestingly, his book (Huey 1908/1968) was republished at a time when research interest in reading was experiencing a renaissance after a long period of dormancy. Indeed, in the introduction to the re-published book, Kolars wrote that “remarkably little empirical information has been added to what Huey knew, although some of the phenomena have now been measured more precisely” (p. xiv). However, before this renaissance, there was a period of standstill.



## **The fallow period**

In the interim between Huey's original work and its republication, that is, from the early 1910s till the late 1950s, the behaviourist revolution in psychology, ushered in an era in which there was a very narrow view of what research was considered acceptable. In particular, research into essentially hidden processes that necessitated the collection of mentalistic data was deemed to lack validity. Since such phenomena had to be reported by informants to the investigator, the resulting data were not directly observable. Since little is observable in the study of reading processes, the field became a casualty to the new movement in psychology. As Venezky (1984:5) notes, "when psychological interests changed, concern for reading usually changed also". Attention in reading research switched to applied areas, especially assessment.

In his extensive review of the history of reading research, Venezky (1984) notes that the study of cognition in reading comprehension "was never pursued as systematically as word recognition ... and research on comprehension processes was so sparse up to the 1950s that even the phrase 'reading comprehension' was seldom found" (p. 13). Among the reasons for this may be that the main concern of verbal learning studies was the serial learning of nonsense words and lists of real words; also, text comprehension was ignored in favour of the study of vocabulary in texts.

However, during this long fallow period, there were two notable exceptions: Henderson (1903) and Bartlett (1932). Henderson had his subjects read short narrative and expository texts and then recall in writing as much as they could. The protocols were scored with reference to the number of topics, subtopics and words recalled. Quantifying semantic content of texts in this way was unusual for the time, and it was not widely practised until the late 1950s. Nonetheless, Bartlett (1932) conducted a similar experiment, in which he added a longitudinal element. He asked his subjects to write recalls at varying intervals from two weeks to ten years after reading the texts. He found that the overall "scheme or plan of a prose passage" tended to be evident across all the recalls. This study laid the groundwork for the development of schema theory and story grammars (e.g., Rumelhart 1975, 1977, 1986, Mandler and Johnson 1977, van Dijk and Kintsch 1983, Mahoney et al. 1997).

## The recent revival

By the late 1950s, it was becoming clear that behaviourism was unable to account for language learning or language processing. Chomsky, the linguist, published ‘Syntactic Structures’ in 1957 in which he suggested that the study of language and the mind are closely related and presented a new theory of *phrase structure grammar*. This was, according to Rayner and Pollatsek (1989), “a reasonable metaphor to represent the basic information in a reader’s head when the syntax of the sentence is being comprehended” (p. 239). Chomsky’s ideas triggered a new interest in the study of cognitive processes. While attention was first directed at language processes in general, around 1970, this was extended to the reading process.

One manifestation of this renewed interest was the sudden interest in the construction of reading models. Samuels and Kamil (1984/1988) say that, up to this time, “there simply was not a strong tradition of attempting to conceptualize knowledge and theory about the reading process in the form of explicit reading models” (1988:22).

However, the research culture in the field of psychology changed dramatically, and, especially during the 1970s and 1980s, several models of the reading process were proposed. (Some of these models are discussed in Section 3.0 below.)

Chronologically, there was a surge of interest in *top-down*, or concept-driven, models (e.g., Goodman 1967, Smith 1971), and this was quickly followed by an opposing *bottom-up*, or data-driven, view (e.g., Gough 1972). Later, these two opposing perspectives were modified into an *interactive* view of the reading process (e.g., Rumelhart 1977). A large number of research studies theoretically founded on models proposing one or other of these perspectives have been conducted in the past three decades.

## 2.2 History of L2 reading research

In contrast to L1 reading research, L2 research has an extremely short history, the research effort only beginning in a concerted way in the 1970s, a century after its L1 counterpart. For the purposes of this discussion, L2 can be understood in the rather



broad sense used by Stern (1983), who refers to “the chronology of language learning, a second language being any language acquired after the native” (p. 12).

Behaviourism had a major impact on L2 reading research, just as it did on its L1 counterpart, seeing the reader as a passive recipient of textual information; in other words, it saw reading as essentially a one-way process. In addition, as will be shown below, behaviourist concepts long maintained a particularly strong stranglehold on L2 teaching practice. More recently, with the advent of reading models proposing a more active role for readers, investigators, using various methods, have pursued several avenues in L2 reading. At the same time, certain socio-political changes taking place globally have provided further pressure and incentive for L2 reading research, which, by the early 1990s, had become a small, but vibrant, field of inquiry.

### **The influence of behaviourism**

Research into L2 reading began rather later than that into speaking and writing. This may be due, in large measure, to the dominance of audiolingualism, the pedagogic manifestation of behaviourism. Audiolingualism de-emphasized the so-called ‘passive’ skill of reading in favour of the so-called ‘active’ skills, especially that of speaking. Silberstein (1987), among others, has noted that reading was seen as little more than a reinforcement of oral language instruction (e.g., Fries 1972). My own initial experience of language teaching in the mid-1970s confirms this observation.

Carrell (1988) reports that early work in L2 reading assumed that the process was “rather passive, bottom-up” (p. 1). It was seen as a matter of constructing meaning by decoding text from the smallest textual units (letters and words) to increasingly large units (phrases, clauses, sentences, etc.) (e.g., Rivers 1968, Plaister 1968, Yorio 1971). Rivers (1968) and Fries (1972), among others, attached great importance to phoneme-grapheme relationships in the teaching of L2 reading. While both Rivers and Fries acknowledged the importance for L2 readers of culture-specific knowledge in texts, Carrell (1988) maintains that “these concepts played no real role in early theories of second language reading, and the methodological and instructional focus remained on decoding, or bottom-up processing” (p. 2).

## **The influence of Goodman's model of reading**

The first shift from this early, but stagnant, phase in the history of L2 reading research came when the first models were developed for L1 reading. In particular, Goodman's concept-driven, or top-down, model of reading (e.g., 1967), one of whose main claims is that good readers do not use all the available textual cues, had a profound impact on the field of L2 reading. Goodman posited that good readers 'leap' through the text, making predictions that are confirmed or disconfirmed, partly through the use of cues from the graphophonic, syntactic and, in a broad sense, semantic systems of the language used, and, crucially, also with reference to their own life experience and knowledge of the language.

The impact of Goodman's model was to transform the view of an L2 reader from a passive figure to, in Carrell's (1988) words, "an active participant in the reading process" (p. 3). Thus, in the mid-1970s, there was a sudden surge of interest in researching L2 reading processes (e.g., Eskey 1973, Neville and Pugh 1975, Hosenfeld 1977). Apparently reflecting the fact that this field of study had a very sudden onset, Bernhardt (1991) chose the year 1973 as the starting point for her exhaustive database of L2 reading studies.

Furthermore, concurrent with readers being perceived as active participants in the reading process was the perception that they were also a useful, even essential, source of data. At first, data collection was done using relatively conventional methods, such as recalls of content (e.g., Mandler 1978). However, once there was general acceptance of mentalistic data in cognitive psychology, the way was clear to investigate readers' mental processes in studies using introspective data-gathering techniques (e.g., Hosenfeld 1977, Hastrup 1987).

## **Recent research strands**

In the past twenty years, several focal shifts, using various research methods, have taken place in the study of L2 reading. Initially, there was a top-down phase that focused on what the reader brought to the process. One strand of the research effort



investigated the role of L2 proficiency (e.g., Cziko 1978, 1980, Clarke 1979, 1980, Devine 1987, James 1987). Clarke's (1980) 'short circuit' hypothesis, for instance, posits that L2 deficiency would lead to relatively more bottom-up processing in L2 than in L1 reading. In some cases, this research strand made specific reference to the relative importance of lexical and grammatical proficiency (e.g., Devine 1987, Strother and Ulijn 1987).

A related research strand started from a schema-theoretic standpoint (i.e., schemata other than language). For instance, some researchers looked at the effects of reader knowledge of the rhetorical structure of the text (e.g., Mandler 1978, Carrell 1984, Mahoney et al. 1997) while others researched the effects of content knowledge (e.g., Hudson 1982, Carrell and Wallace 1983, Afflerbach 1990, Lee 1991). Lee's (1991:200) 'bi-orientation' hypothesis, for instance, posits that adult L2 readers who are already proficient L1 readers, particularly when reading content that is familiar, combine top-down and bottom-up processing, just as do proficient L1 readers. Other researchers looked specifically at the effect of readers' cultural schemata on reading processes. For example, Pritchard (1990) studied the effects of cultural content while Parry (1996) studied the effects of readers' cultural contexts.

Finally, there was a research strand that sought to investigate L2 reading processes by categorizing behaviours as strategies (e.g., Hosenfeld 1977, Block 1986a and 1986b, Sarig 1987a and 1987b, Casanave 1988, Li and Munby 1996, Jimenez et al. 1996, Jimenez 1997). (A full literature review of this strand is given in Chapter 2.)

### **Socio-political factors**

Concurrent with changes in L2 reading research, socio-political forces involving the increasingly global use of English appear to have added to the interest in investigating L2 reading processes and new approaches in its pedagogic outcomes (e.g., Grabe 1991, Bernhardt 1991).

In particular, there have been profound changes worldwide at the tertiary level of education in at least three distinct contexts. First, increasing numbers of non-native-

speaker university students are studying for degrees in English-speaking countries, and these students are required, as part of their studies, to read academic articles, both widely and in depth, in English. More recently, English-medium tertiary-level education around the world has proliferated in non-English speaking contexts such as Hong Kong, Singapore, Japan and Latin America. In addition, even where non-native speakers of English study in their own countries at universities officially using the local language as the medium of instruction, the academic texts these students have to read are often only available in English. Third, there have been changes at the pedagogic level. In advance of either a strong theoretical framework or a major research effort, English as a Second Language (ESL) instruction, particularly at relatively high proficiency levels, shifted in the 1970s from a focus on listening and speaking skills to one that emphasized reading and writing skills.

### **2.3 Section summary**

In answer to Research Question 1.1, which was posed at the beginning of the chapter and inquired into major influences on the study of reading, this section described some of the differences and similarities in the histories of L1 and L2 reading. The history of reading research as a field of study has been, even in the L1, relatively short and spasmodic. It began with a brief period of about three decades from the late 1870s in which studies investigated such matters as eye movements across text, letter and word recognition, the eye-voice span and the role of peripheral vision. This was followed by a longer period of about five decades, coinciding with the dominance of behaviourism, in which very little research into reading was undertaken. The past three decades, coinciding with the rise of cognitivism, have witnessed a resurgence of interest in L1 reading and the start of research into L2 reading, especially in young adults and adults. This new interest has extended to include both theoretical development and an increasing number of studies whose research questions and methods reflect the changes in theory. In pursuit of an answer to Research Question 1.2, which asks about models that best describe reading, the chapter now turns to a review of theoretical development in both L1 and L2 reading.



### 3.0 READING THEORY AND MODELS OF READING

Cohen and Manion (1994) note that the words ‘model’ and ‘theory’ are sometimes used interchangeably, both being “explanatory devices or schemes having a broadly a conceptual framework” (p. 16). However, models may be distinctive in giving more graphic or visual representation of a particular phenomenon or set of phenomena. Rayner and Pollatsek (1989), basing their definition of a model on that of Carr (1982), say that it is a description of the major working parts of a real-life process, such as reading. They also say that it incorporates “the most important characteristics of each part's operation, though it might leave out large amounts of detail” (p. 25).

Davies (1995) describes a model as “a systematic set of guesses or predictions about a hidden process, which are then subjected to ‘testing’ through experimental studies” (p. 57). Samuels and Kamil (1984/1988) say that “an absolutely crucial characteristic of a good model is that it be precise enough to lead to testable hypotheses” (1988:27). Interestingly, this definition would preclude one of the most influential ‘models’ of reading, that of Goodman (e.g., 1967) (described below).

Samuels and Kamil (1984/1988:25) consider two major problems in evaluating models. The first is that one needs to take account of the knowledge base as well as the intellectual conceptualizations prevailing at the time of the development of the model. Second, researchers are influenced by data gathered during experiments and may ignore four key factors that impede external generalizability: the age and skill of the subjects, the tasks the subjects undertake, the materials used, and the research context.

Writing in the L2 literature, Bernhardt (1991) says much the same thing as Samuels and Kamil; however, she takes as her starting point *readers* rather than *reading*. She says that “principled theory development requires fairly extensive knowledge of readers as multifaceted, multiproficient learners” (p. 225). She also argues that second language literacy research “should recognize itself as an entity distinct from first language literacy research” so that it can become more theoretically viable and principled (p. 226).

The ensuing discussion of the theoretical development of reading begins with a detailed discussion of several models of the reading process that have been proposed in the past three decades. Most assume L1 reading and fall into three basic views: that reading is primarily *top-down* (i.e., concept-driven), *bottom-up* (i.e., data-driven) or *interactive* (i.e., driven by a combination of both higher-level and lower-level processes). More recently, there has been a modification in model development towards the concept that reading is *bottom-up interactive*. In addition, as will be shown, there is a limited number of models incorporating affective factors into the reading process. Finally, theory development in L2 reading is discussed.

### **3.1 Development of theory and models in L1 reading**

It is important to trace the history of model development in part because it is a way of explaining the current theoretical position in the field and in part because some of the models have been very influential both in driving the research effort and in determining pedagogical outcomes. Divergent views of the reading process, as reflected in the various models, have led to what has been widely referred to as ‘reading wars’, or what Stanovich and Stanovich (1995) have called “overheated rhetoric” (p. 87). Specifically, the debate has been between those who adhere to a *top-down* view, and to the associated whole language approach to the teaching of reading, and those who take a wholly, or mainly, *bottom-up* view and call for a phonics dimension, or even an emphasis on phonics, in the teaching of reading. This shows that the motivation for model construction may often have been generated by pedagogical motivations.

The discussion of reading models that follows is largely chronological in character. It begins with the polar opposite positions of those who have top-down and bottom-up perspectives and ends with diverse interactive views.



### **3.1.1 Top-down models**

Top-down models view reading as an active concept-driven process in which readers draw extensively on their knowledge of the world. This information, stored in long-term memory, is used to formulate hypotheses about upcoming text; and, as reading proceeds, these predictions are either confirmed or rejected. Among others, Goodman (e.g., 1967), Hockberg (1970), Levin and Kaplan (1970), and Smith (1971, 1973, 1988) propound top-down views of reading.

#### **Some influential top-down models**

An early and influential top-down model is that of Goodman (1967, 1969, 1975/1988). Goodman claims that the reading process is characterized by universals; he says that his model “has been built through the study of English reading, but it must be applicable to reading in all languages and all orthographies” (1975/1988:20). Further, he sees reading as a ‘psycholinguistic guessing game’ in which readers ‘leap’ through the text to derive meaning. In contrast to its focus on the role of prior knowledge, the model pays little heed to the visual or orthographic decoding of text.

It is perhaps significant that Goodman's primary source of data for his model was miscue analysis, which was based on data yielded from the oral reading of beginner L1 readers. (Chapter 3 provides more information on miscue analysis.) While the use of performance data and miscue analysis to develop a model may be, as Davies (1995:62) has said, “theoretically robust”, it is perhaps not surprising that, with these child subjects, Goodman found little evidence of the use of text and much evidence of a ‘psycholinguistic guessing game’. In any case, as Eskey (1988:94) observes, since the mode of analysis used only deals with miscues, it tells us very little about what readers do when they decode accurately.

Further, as Davies (1995) points out, while there is evidence that the Goodman model may offer “a reasonably ‘truthful’ representation ... of the behaviour of beginning L1 readers” (p. 62), it does not offer an explanation of it. Moreover, despite Goodman's

claims to universality, it seems likely that the model offers neither a representation nor an explanation of the behaviour of proficient L1 readers.

A similar psycholinguistic view of the reading process is propounded in the work of Smith (e.g., 1971, 1973, 1988). Smith hypothesizes that, due to sensitivity to semantic and syntactic redundancy available in discourse, the good reader develops hypotheses about upcoming words and can then confirm the identity of a word by sampling only a few features in the visual display of a new word. Like Goodman, he sees reading as showing a procedural preference for reliance on prior knowledge, including language, rather than on graphic information.

Also like Goodman, Smith maintained his position over a long period in a developing field. Seventeen years after his original assertion that word decoding is unnecessary for comprehension, Smith repeated it word for word:

“Procedures permitting the identification of words without the prior identification of letters also permit comprehension without the prior identification of words.” (1971:120, 1988:116)

Further, he believes that the good reader relies less on graphic cues than the poor reader; and he refers to the poor reader as ‘a slave to print’. He states:

“The more difficulty a reader has with reading, the more he relies on the visual information; this statement applies to both the fluent reader and the beginner. In each case, the cause of the difficulty is inability to make full use of syntactic and semantic redundancy of nonvisual sources of information.” (1971:221)

Samuels and Kamil (1984/1988) believe that Smith's work is a description of the linguistic and cognitive processes that a model of reading should take into account rather than a model. They say:

“Perhaps the greatest contribution of Smith's work is to explain how the redundancy inherent at all levels of language (letter features, within letters, within words, within sentences, within discourses) provides the reader with enormous flexibility in marshaling resources to create a meaning for the text at hand.” (1988:24)



## **Some problems with top-down models**

Even by the mid-1970s, it was evident that there were several clear problems with a top-down view of reading. For instance, research investigating the effects of orthographic structure on readers provided findings that questioned the basis of the top-down models. Golinkoff (1975-76) cites Coomber and Hogie's finding that poor readers are more sensitive than good readers to spelling pattern violations and concluded that "slow decoding may not be due to a failure to utilize intraword information in the form of spelling patterns" (p. 654).

Commenting on this and related findings, Stanovich (1980) says that there is evidence that "the poor reader not only uses orthographic redundancy to facilitate word recognition, but ... in some circumstances ... may even rely more on this stored knowledge source than the good reader" (p. 41). As Stanovich points out, this presents problems for top-down models, which posit that good readers' fluency is due to greater use of redundancy at all levels.

Another line of inquiry into top-down models was conducted through the measurement of reading speed in order to test the notion of readers 'leaping' through text. If readers do indeed predict as they leap through text, one would expect that relatively predictable parts of a text would be read faster. Mitchell and Green (1978) conducted a study that, rather artificially, required their subjects to press a button connected to an on-line visual display, and each time they did, they were presented with the next three words of the passage. The researchers found no evidence of readers speeding up during 'predictable' portions of text and concluded that reading speed is more dependent on the speed at which a reader can recognize words and construct a representation than on the ability to use predictions to facilitate word recognition. In a related experiment, Mitchell (1984) found "no statistical support for the view that reading time was influenced by the predictability of the material" (p. 83).

Despite such findings, many prominent top-down proponents conduct themselves in isolation and continue to maintain their position. Indeed, a hallmark of Goodman's work is that he simply asserts his beliefs; for instance, in his 1975/1988 article

propounding the top-down model, there are only two citations, one of which was of an article written by himself.

Recently, Stanovich and Stanovich (1995) have made a similar point about Grundin's (1994) article, which claims that proponents of the importance of word recognition skills make the simplistic assertion that “word recognition -- or rather word pronunciation -- amounts to reading” (p. 9). Stanovich and Stanovich say that Grundin's article had:

“no citations – no quotes from actual researchers who had stated such things. In the absence of any evidence that a researcher actually has made such a statement, such charges should be taken for what they appear to be – strawmen assertions designed to trigger negative associations in the mind of the reader.” (p. 97)

Perhaps this embattled position is caused by mounting evidence from several different strands of L1 reading research over the past two decades indicating that the top-down view of reading has severe problems. These strands include not only research into orthographic structure and reading speed, as discussed above, but also research into context, individual differences, and visual information. Moreover, studies contributing to these doubts have subjects take part in a wide range of tasks, including oral reading, cloze, word boundary activities, and proof-reading.

### **3.1.2 Bottom-up models**

Unlike their top-down counterparts, bottom-up models of reading are data-driven. The flow of information is seen to be very fast and there is little recourse to prior knowledge. Much of the evidence for these models came from laboratory investigations in which proficient L1 adult readers took part in letter and word recognition tasks.

Rayner and Pollatsek (1989:465) report that, though models such as those of Mackworth (1972), LaBerge and Samuels (1974) and Massaro (1975), are often cited in the literature as bottom-up, they are either not ‘comprehensive’ models of the reading process or they have not received much attention.



Clearly the most comprehensive and influential bottom-up model is that of Gough (1972). It was originally presented in May 1971 at a research conference at the Smithsonian Institution on the relationships between speech and learning to read. Unlike the top-down models, Gough's model was very specific about the reading process and much of what he claimed was testable. It spawned much research and it is for this reason that it warrants description in some detail.

### **Gough's (1972) bottom-up model**

Very much against the prevailing views of the time, Gough stressed the need for the reader to decode words for comprehension to take place. For that to happen, visual information from a fixation is entered into iconic memory, where it is stored until another fixation takes place. This recognition takes place serially and letter by letter. At the same time, the 'scanner' consults pattern recognition routines stored in long-term memory. The 'decoder' then maps letter strings onto a string of systematic phonemes; it makes use of graphophonic rules, and the products of this process are stored as inner speech, or 'phonemic tape'. The phonemic representation, along with the 'lexicon', identifies the fixated words, which are held in short-term memory until the sentence can be parsed and placed in what Gough called TPWSGWT AU ("the place where sentences go when they are understood"). A comprehension device then makes use of syntactic and semantic rules to understand the sentence.

In essence, Gough (1972) argued that a reader "plods through the sentence, letter by letter, word by word" (p. 354) at about 400 words a minute; his use of the word 'plod' to describe reading is in sharp contrast to Goodman's 'leap'. Of top-down models, Gough states:

"A guess may be a good thing. But rather than being a sign of normal reading it indicates that the child did not decode the word ... rapidly enough to read normally. The good reader need not guess. The bad reader should not." (p. 354)

Gough argued for tuition in phonics because he believed in the importance of letter to sound correspondence. However, as Smith (1973:89) points out, Gough did not

account for the impracticability of teaching graphophonic rules in languages that have a considerable number of them (e.g., English, which has at least 166).

### **Some problems with Gough's (1972) model**

From the above, it is clear that there are a number of problems with Gough's (1972) model. First, Gough proposes serial or linear processing, which imposes a heavy burden on working memory. Brewer (1972), in a discussion paper following Gough's original article, says:

“An adequate model of reading will have to include some mechanism to allow an interaction between the higher linguistic processing and the control of eye fixations. It is not clear how Gough's linear inflow model can handle this problem.” (p. 360)

Brewer goes on to say that, without this interaction, there should be no difference in the rates of processing words and non-words. He cites several studies that show that letter search times are faster for words than non-words. Questioning Gough's claim that the mapping of letters onto systematic phonemes is the “exclusive” path from letter to meaning, Brewer cites evidence that deaf subjects can read without a stage of phonological transformation. In a response to this point, Gough (1972) says that the reading of deaf people may differ from that of people with hearing and that his model “does not pretend to explain it” (p. 367).

Samuels and Kamil (1984/1988) have outlined several deficiencies in linear models such as Gough's. They point out that each stage in a non-interactive model works independently and “passes its production to the next higher stage” (1988:27-28). Indeed, in a question and answer session after Gough presented his paper, Kolars questioned the serial nature of the model, and Gough maintained his position, saying that there was “no influence from the syntactic level” (p. 369). However, like Kolars, Rumelhart (1977) believes that linear models are deficient in that they fail to account for a number of processes known to take place in reading. In his groundbreaking interactive model, Rumelhart tried to account for these processes, including the use of syntactic and semantic constraints.



### 3.1.3 Interactive models

Interactive models of reading allow for an interface between top-down and bottom-up processing, that is, they posit that reading is both a concept-driven and a data-driven process. This section selects two of the most influential interactive models, those of Rumelhart (1977) and Stanovich (1980).

#### Rumelhart's (1977) interactive model

Rumelhart's (1977) model represented a major step forward in the development of models of reading. Rumelhart (1977) states that “reading is at once a ‘perceptual’ and ‘cognitive’ process” (pp. 573-574). His objective was to propose an alternative to serial bottom-up models by incorporating parallel processing, in other words, the simultaneous processing from more than one source of information.

What made Rumelhart's model distinctive at the time was that the reader was seen to draw simultaneously but selectively on a range of sources of both *sensory* and *non-sensory* information: orthographic, lexical, semantic, syntactic, schematic. While the LaBerge and Samuels (1974) model can account for the effects of orthographic structure on letter perception, Rumelhart (1977) says “it is somewhat more difficult to see how the effects of syntax and semantics can be mediated within such a model” (p. 581).

Rumelhart's model allows higher level processes to influence lower level processes, something not accommodated into the earlier linear models. He outlines how a reader's perception of words depends on both the syntactic and semantic environment in which they are encountered, and how a reader's perception of syntax depends on the semantic context in which the string appears.

These sensory and non-sensory knowledge sources converge on the ‘pattern synthesizer’, which uses all the information to produce what Rumelhart (1977) calls a “most probable interpretation” of the graphemic input (p. 588). He also discusses a

mechanism he calls a 'message centre' that holds and redirects the information as needed; it is, in effect, responsible for monitoring strategies.

### **Some strengths and weaknesses in Rumelhart's model**

As Davies (1995) has observed, Rumelhart's model has several strengths. It provides an alternative to top-down and bottom-up models, and it allows for learners to rely on more than one knowledge source for comprehension, not merely prior expectations as in the Goodman model, or letter-sound correspondences as in the Gough model. It does not prescribe a single direction or route to comprehension, allowing for various predictions during different stages of L1 or L2 reading development, thus leaving the way open for researchers. Eskey (1988) recognizes that the model can accommodate L2 readers who need to focus on simple decoding or "holding in the bottom" (p. 94).

There are, however, a number of problems with Rumelhart's (1977) model. As Rayner and Pollatsek (1989:467-468) point out, it is not a comprehensive model of the reading process for several reasons. It does not account for how hypotheses are generated, nor does it specify the relative importance among the knowledge sources. Rayner and Pollatsek are further concerned that it does not account for eye movements, the phonological route in word recognition, backup strategies used in word recognition or comprehension issues beyond the level of the sentence.

Rumelhart himself acknowledged many of these shortcomings, and continued to develop the 1977 model in several subsequent papers. For instance, Rumelhart (1984) presents a 'schema-theoretic' view of the comprehension process. Though lower level processes are still acknowledged, Rumelhart incorporates schema theory, with its greater emphasis on the semantic level of processing, into his model. He describes a schema as a unit of knowledge, "a data structure for representing the generic concepts stored in memory ... a kind of informal private unarticulated theory about the nature of the events, objects, or situations we face" (pp. 2-3). Schemata are frameworks for interpreting the world and what we read. They are fluid and dynamic by nature, not "structures" but coalitions of hypotheses or units that cohere or go together or confirm each other.



Later, Rumelhart et al. (1986) embellish this definition of schemata, saying they are “a kind of organic element which grows and fulfills itself within its environment” (p. 36); moreover, schematic knowledge includes knowledge of visual features, words, syntax as well as semantic, formal and content knowledge.

McClelland's (1986) model, though, again, not a comprehensive model of the reading process, builds on the Rumelhart model and partially accounts for eye movements and the way in which information is integrated across fixations. However, Rayner and Pollatsek's (1989) model, described below, accounts more fully for eye movements.

### **Stanovich's (1980) interactive-compensatory model**

Stanovich (1980) writes about a conflict taking place at that time between what models were predicting and what researchers were finding. He noted that both top-down and bottom-up models predict that higher level processes will be more implicated in the performance of better readers than in that of poor readers. However, his comprehensive review of the literature on the use of orthographic structure and the effect of contextual information on word recognition did not bear out this prediction. He therefore set out to account for individual differences among readers.

In his view, good and bad readers differ in that good readers have more highly automatized word recognition *context-free* to enable them to give more attentional capacity for integrative comprehension processes. Poor readers' greater tendency to use context to guess word meaning may be “purchased” (p. 64) at a price, since this conscious-expectancy process uses attentional capacity, leaving fewer resources for comprehension operations that work on integrating larger units of text. Poor learners focus on personal associations, which serves as a compensatory function (cf. Block's 1986b *reflexive* non-integrators).

Stanovich (1980) writes “toward” an interactive-compensatory model of reading. (The word is part of the title of his article.) He takes Rumelhart's (1977) assumption of simultaneous parallel processing and adds a compensatory assumption that “a

deficit in any knowledge source results in a heavier reliance on other knowledge sources, *regardless* [his italics] of their level in the processing hierarchy” (p. 63). Stanovich has thus accounted for the finding in many studies (e.g., Perfetti, Goldman and Hogaboam 1979, Perfetti and Roth 1981, Stanovich and West 1981) that a poor reader deficient in decoding skills might show a greater reliance on contextual factors.

Although Stanovich did not mention L2 readers when describing his model, it is nevertheless particularly relevant to L2 reading. Whatever their L1 reading ability, many L2 readers may need to ‘compensate’ for a low language proficiency level. As Davies (1995) notes, “when syntactic knowledge is poor, a greater reliance may be placed on orthographic or lexical information” (p. 65). She notes that this kind of prediction is testable and that the model provides a basis for investigations of performance and processing strategies of different groups of readers.

#### **3.1.4 Towards bottom-up interactive models**

It is becoming increasingly hard to draw a dividing line between bottom-up and interactive models. This partly reflects an emerging consensus, perhaps fed by increasing evidence in the literature that, while reading is an interactive process, rapid word decoding is a necessary part of fluent reading. It also, perhaps, reflects increasing rigour among theoreticians, cognizant of the fact that only models with a distinctive bottom-up element provide testable predictions.

This section describes some models that are hard to classify as either bottom-up or interactive and then moves on to describe Rayner and Pollatsek’s (1989) highly influential bottom-up interactive model.

#### **Some models that are hard to classify**

Rayner and Pollatsek (1989) caution that what are often cited as ‘bottom-up models’ may not really qualify as either bottom-up or as models. For instance, they believe that the Massaro (1975) model has “significant interactive aspects” (p. 465).



Likewise, LaBerge and Samuels' (1974) automaticity model seems hard to classify; it claims that fast and automatic word recognition is an important determinant of fluent reading and that this can free up capacity for higher-level processes. While Rayner and Pollatsek say that LaBerge and Samuels' model is not really a 'comprehensive' model of the reading process, Rumelhart (1977) refers to it as "a strictly bottom-up process" (p. 578). Later, Samuels (1977) revised the model to account for interactive-compensatory processing at the word level when word identification is problematic, thus moving the model away from a largely bottom-up perspective.

Just and Carpenter's (1980) model has also been hard to classify; specifically, there is controversy as to whether it is 'interactive'. The model accounts for comprehension processes based upon studies of eye movements, and it has an integration component designed to capture the need for the reader to realize the coherence of a text by relating clauses and sentences. It emphasizes bottom-up processes more than top-down, and lower-level information controls the flow, but top-down processes can influence bottom-up processes. Moreover, the model allows for parallel processing and interactions between processing stages and working memory, or short-term memory, and between short-term memory and long-term memory. With these features, Rayner and Pollatsek (1989:470-471) say they would classify the model as interactive.

At the same time, Rayner and Pollatsek make the vital point that, in contrast to bottom-up models, interactive models seem unable to provide a clear prediction of the outcome of an experiment (cf. Samuels and Kamil's 1984/1988:27, and Davies' 1995:57 definitions of 'models'). Nonetheless, they rightly concede:

"To a certain extent, this criticism is unfair because it may well be that many of the complex processes that occur during reading will be impenetrable to psychological experimentation." (p. 471)

And they conclude:

"While interactive models are very good at explaining behavior, it is yet an open question as to how well they will predict and how much research they will generate." (p. 471)

The most recent trend, then, is for models to be ‘bottom-up interactive’; that is, they stress that fast decoding is necessary, though not always sufficient, in fluent reading. Rayner and Pollatsek's (1989) model is the most influential bottom-up interactive model; it is described below.

### **Rayner and Pollatsek's (1989) bottom-up interactive model**

In the development of their model, Rayner and Pollatsek (1989) seek to demonstrate the relationship between eye movements and cognitive processing. They posit that the automatic identification of words partially influences the interaction of lexical knowledge in the lexicon with background knowledge in long-term memory, syntactic knowledge in the parser and semantic knowledge in the thematic processor. At the same time, they posit that automatic recognition of words frees processing space and time for interaction with higher-level sources of information.

Rayner and Pollatsek readily admit gaps in their model. They call it “a (temporary) working model or convenient way of summarizing the evidence that we place the most credence in” (p. 471). They admit that, beyond the recognition of words, the model becomes vague and speculative. Therefore, to explain construction of meaning, they discuss several possibilities: the linking of propositions, the creation of a causal ‘if-then’ structure, the construction of a mental model and fitting current material into a schema.

Despite Rayner and Pollatsek's modesty, theirs is the most recent serious model to be propounded in the literature; moreover, it has several strengths. It is derived through a careful synthesis of research findings and it is detailed on the interaction between the bottom-up and top-down processing involved in word recognition. In addition, it proposes parallel processing of visual features and letters, it allows for, but does not insist upon, a phonological route, and it takes into account constraints on the perceptual span. Finally, as Davies (1995) points out, it lends support to Eskey's (1988) calls for greater attention to lower-level decoding processes in L2 and for attention to monitoring processes at all levels of reading in L1 and L2.



### **3.1.5 Models that incorporate affective factors**

With hindsight, it is clear that, while behaviourism held sway, it acted to preclude many fruitful avenues of research. In the same way, it is possible that the current dominance of cognitivism may be forcing the research effort into blinkers with the consequence that other aspects of psychology, such as affection (i.e., involving emotion) and conation (i.e., involving volition or a striving for change), are being ignored. This section describes two models of reading, both developed in the mid-1980s, which incorporate an affective element: Mathewson (1985) and Ruddell and Spence (1985).

#### **Mathewson's (1985) model**

In explaining the theoretical background to the development of his model, Mathewson (1985) cites the “historical tripartite division” of the study of the mind into cognition, affection and conation (p. 841). He says that not only did behaviourism generally ignore affection and conation but that the subsequent information-processing and computer-based models did so, too. Hilgard (1980, cited in Mathewson) notes that “it is obvious immediately that cognitive psychology is ascendant at present, with a concurrent decline of emphasis upon the affective-conative dimensions”. This lack of interest in affective factors on the part of psychologists also pervades the study of reading, something Mathewson sought to amend.

Mathewson's (1985) model posits that attitude, motivation, affect, and physical feelings contribute to the initial decision to read and to on-going decisions to continue reading. This activates the ‘primary processes’: physical orientation and activity, attention, and comprehension. Mathewson adds the ‘secondary processes’, recall, reflection, and application, “because it appeared to be important that readers do more with materials than simply read them” (p. 847).

The model has a number of strengths and weaknesses. Its strengths include the fact that it adds a welcome additional dimension ignored in the mainstream cognitivist literature. It reflects the evidence of the effect of real-world and individual variables

on the reading process. It has potential for further research as it pays attention to affective and text-based factors. Further, it has the intuitively appealing hypothesis, and one increasingly supported by evidence (e.g., Devine 1988), that, outside the laboratory, the nature of success in reading is typically under the control of individual and personal attitudes and decisions. However, as Davies (1995) has noted, the model lacks detail about the ‘primary’ processes. In addition, it does not account for such aspects as eye movements; nor does it consider the interaction between affective factors and primary and secondary processes.

### **Ruddell and Spence’s (1985) model**

Ruddell and Spence's (1985) model incorporating affect has four “components of the interactive reading process” that function simultaneously and in no fixed sequence (pp. 751-753). The ‘knowledge utilization and control’ component controls the process, and it includes the reader's affective, cognitive, and metacognitive states as well as the text representation, or the reader's interpretation of the text meaning. The ‘reader environment’ includes textual, conversational, and instructional features the reader uses in constructing meaning from text. The ‘declarative and procedural knowledge’ component represents the reader's decoding, language, and world knowledge schemata. Finally, the ‘reader product’ is the outcome of the interaction of the other three components; eight products are specified, including comprehension, word recognition, and affective, cognitive and metacognitive state change.

Again, Ruddell and Spence add a non-mainstream element in their model, and it is backed by a substantial review of literature relevant to its components. Nonetheless, it can be criticized for similar reasons to critiques of Mathewson's model in that it does not account for such key aspects of the reading process as eye movements.

## **3.2 Development of theory in L2 reading**

So far in this discussion of reading theory and models of reading, the focus has been on L1 reading. I will now show that, in comparison, very little work has been done on



the development of theory in L2 reading; moreover, there appears to be surprisingly little interaction between investigators of L1 and L2 reading. Even though, as has been described above, the past three decades have seen the appearance of several models of L1 reading propounding views other than top-down, they do not seem to have had much impact either in the L2 literature or in L2 pedagogy. While the top-down models do not drive teaching methodology in the L1, their pedagogic outcomes are still very obvious in the L2, where emphasis tends to be given to guessing and prediction skills at the expense of word recognition skills (e.g., Grellet 1981, Nuttall 1982, Doff 1988, Wallace 1992).

Grabe (1991) has observed that “two efforts to translate this [top-down] theory into ESL contexts have been extremely influential on ESL reading theory and instruction from the late 1970s to the present” (p. 377). First, Clarke and Silberstein (1977) encouraged teachers to schematize their students to text content and organizational structure with pre-reading activities and to have students guess meaning from context. However, they also suggested that teachers provide students with strategies for dealing with difficult syntax and vocabulary. Second, Coady (1979) reinterpreted Goodman’s model for L2 reading, arguing that, while beginners might need to focus on word identification, more proficient readers should concentrate on the use of prior knowledge and minimize the use of information in the text for confirming hypotheses and predicting. While Coady’s work might be termed top-down-dominant, it nevertheless acknowledged a developmental element in L2 reading, a dimension that was to be developed more than a decade later.

In a groundbreaking book on L2 reading research, Bernhardt (1991) shows that top-down models are by far the most frequently cited of any models in the literature (pp. 22-25). Her comprehensive review of the literature on L2 reading research revealed that 66.40% of references to reading models were to those of Goodman and Smith (both described above). She posits two explanations for this. First, she believes there may be a prevailing acceptance of the top-down view of reading; second, there may be “a basic lack of awareness” of other models and how they might account for L2 reading phenomena (p. 22).

Also seeking an explanation for the dominance of top-down conceptualizations in the L2 reading literature, Paran (1996) speculates that there may be “a time lag between developments in psychology and in linguistics and developments in L2 teaching” (p. 33). While this may be a partial explanation, the fact is that, with the publication in the L1 research literature of Gough's bottom-up model in 1972 and Rumelhart's interactive model in 1977 (both described above), the weight of evidence has shifted against top-down models.

Another possible reason for the dominance of top-down models in the L2 literature is their potential for an instant, yet superficial, resonance for teachers of English as a Second/Foreign Language. When confronted with texts they find difficult, learners with a low L2 proficiency level can indeed be seen engaging in what Grellet (1981) calls “a constant process of guessing” (p. 7). This parallels what Goodman found with his L1 child subjects, the common factor being that both young child readers and low proficiency L2 readers of any age lack automatic word recognition skills and so resort to predicting and guessing. Thus, the top-down view of reading may seem to confirm what many teachers of English as a Second/Foreign Language observe in their own classrooms, and an assumption is made that this represents all L2 reading processes.

However, as Eskey (1988), among others, has pointed out, if L2 researchers focus exclusively on top-down models, it leads to insufficient attention to the decoding problems L2 learners face, that is, the ‘short circuits’ they experience when their language proficiency falls short of that required for a given text. Eskey's position has considerable support in the L1 reading research literature (e.g., Mitchell and Green 1978, Yuill and Oakhill 1991, Stanovich and Stanovich 1995, Perfetti 1995), where converging findings seem to indicate the importance of rapid word identification for successful and fluent reading. Deficiency in L2 may impede word decoding, and Bernhardt's (1991) developmental theory of L2 reading appears to be the most serious attempt so far to incorporate this element into theory.



### **Bernhardt's (1991) developmental theory**

Bernhardt's (1991:168-171) theory seeks to account for readers who are proficient in their L1 but who are at varying proficiency levels in an L2. Her theory is based on performance data from L2 (German and Spanish speakers) secondary school learners in the United States, involving analysis of errors in recall protocols.

The theory integrates *knowledge-driven* and *text-driven* operations in the reading process. The former materialize in the theory as 'background knowledge', 'intratextual perceptions', and 'metacognitive activity'; the latter materialize as 'word recognition', 'phonemic/graphemic feature recognition', and 'syntactic feature recognition'. With the exception of 'metacognitive activity', these are plotted on a line graph whose x-axis represents language proficiency level and whose y-axis is labelled 'error rates'. Bernhardt explains that she does not plot metacognitive activity on the line graph since, though present in all her subjects' data, it cannot be plotted against 'error rates'.

A key feature of the theory is that it is developmental, and it makes predictions about changes in behaviour with varying L2 proficiency. Error rates in the knowledge-driven operations are predicted to change very little as L2 proficiency level increases; they merely drop slightly. However, there are dramatic changes in the error rates of language-driven operations. Word recognition and phono-graphemic feature recognition are very high at low L2 proficiency levels and very low at high L2 proficiency levels, crossing the knowledge-driven components at about lower intermediate L2 proficiency level. Syntactic recognition, on the other hand, is represented by an almost normal bell-shaped curve, with error rates low at low and high L2 proficiency levels and rising at intermediate levels. On the basis of these findings, Bernhardt presupposes that L2 readers at low L2 proficiency levels would not attempt to read difficult texts but that they would take more 'risks' at intermediate levels, thus confronting texts with such challenges as complex syntax. However, it is not immediately clear why Bernhardt does not apply this presupposition to lexical recognition.

Bernhardt's theory is a continuum showing the accumulating importance in L2 reading of linguistic knowledge and the relative importance of knowledge-based inferencing at low levels of language proficiency. It accounts for the fact that adult L2 readers may be very frustrated at the slowness of their reading in the L2 compared with the L1 and so adopt certain strategies to compensate. These strategies might include a bottom-up focus on visual information or isolated lexical items, leading perhaps to context-dependent guessing; alternatively, they might include top-down strategies such as skipping-guessing or a reliance on prediction with minimal attention to visual information and lexical items.

Interestingly, in discussions of the differences between good and poor L1 readers, these are routes predicted by, among others, Rumelhart (1977), Stanovich (1980), and van Dijk and Kintsch (1983). For example, van Dijk and Kintsch (1983):

“What exactly distinguishes a good reader from a poor reader? ... It has been found over and over again that the best discriminator between good and poor readers is performance on simple letter and word identification tasks. What is really wrong with poor readers is that they recognize isolated words inaccurately and too slowly, and compensate for their lack in decoding skills with context-dependent guessing or hypothesis testing.” (pp. 23-24)

There are a number of strengths and weaknesses in Bernhardt's theory. Dealing first with the strengths, it is the first theory of L2 reading, and, as such, it is likely to foster interest in further L2 reading research and theory development; it may also encourage comparative studies of L1 and L2 reading. Second, as a consequence, it provides a basis for specific predictions about L2 reading performance at different stages of language proficiency; these are testable. Third, it is the first theory to pay attention to the acquisition of language proficiency; since children usually already have a substantial knowledge of their L1 when they start reading, this may be particularly relevant to L2 reading. Fourth, it is the first theory to take account of the text variable in terms of the difficulty of syntactic operations; however, this is done without any specificity, and, as has already been pointed out, it is not clear why text variability should not apply equally to word recognition.

Weaknesses in Bernhardt's theory include the fact that it does not seek to describe the perceptual and cognitive processes that reading entails; nor does it account for



affective factors. Furthermore, the concept of ‘error rates’ seems a negative basis for a theory, and it has been suggested to me that it might be worth considering the use of ‘success rates’, which would have the effect of flipping the theory over (Davies, personal communication 1995). While Bernhardt’s concept is based on her extensive work on errors in performance data, ‘error analysis’ elsewhere in the literature has largely been overtaken by ‘performance analysis’ (i.e., the attempt to include in analysis all aspects of the learner data recorded). (This issue of research methods is developed in Chapter 3.)

**Hedge's (1991) modes, anchors and drivers**

Before leaving the subject of reading theory, it is worth describing a study that, while not positing a theory or model, presents an interesting taxonomy of reading behaviours. This is based on analysis of introspective data from L2 subjects, many of whom were beginning courses of study at tertiary level in Britain. Hedge (1991:304) calls the reading behaviours she observed ‘modes’ of reading. These are presented in Table 1.1.

**Table 1.1: Hedge's (1991) reading modes**

Reading mode	Description of reading behaviour
Top-down, relative data exclusion	Uses predominantly conceptual knowledge to the relative or selective exclusion of text data
Top-down, deferred interactive	Uses all available knowledge sources from text to concepts but processes top-down before bottom-up before synthesizing to attain an interactive network of comprehension
Interactive	Uses all available knowledge sources from text to content, genre and world knowledge
Bottom-up, recursive, deferred interactive	Uses all available knowledge sources from text to concepts but processes bottom-up, recursive, before top-down before synthesizing to attain an interactive network of comprehension
Bottom-up, recursive	Uses predominantly text data to the relative exclusion of conceptual knowledge but does reread or consider previous text
Bottom-up, non-recursive	Uses predominantly text data to the relative exclusion of conceptual knowledge and does not reread or consider previous text

This taxonomy does not include utilization of actual visual input, which as Davies (1995) points out, is “an inevitable limitation of classroom-based, non-laboratory

studies” (p. 77). However, it allows for several possibilities known to take place in the reading process, including predominantly top-down or bottom-up, as well as automatic interactive and deferred interactive, processing.

From her data, Hedge also developed knowledge-based ‘anchors’ that “run alongside, or in parallel to, modes of reading” (p. 309), and hence are drawn upon selectively during reading. These are divided into linguistic sources and content sources, the former subdivided into lexico-semantic and discourse categories, the latter into general and specific. Taken together, the taxonomies of modes and anchors provide a coding frame for analyzing the shifting of subjects' behaviours as they read.

In follow-up but unpublished work, Hedge also noticed in her data that readers' purposes could have an effect on the modes and anchors utilized. She used the term ‘drivers’ to describe these purposes. (Further reference will be made to Hedge’s drivers in Studies 1 and 2 of the current research.) The four drivers she found were: reading for meaning, gist, language acquisition, and a combination of both meaning and language acquisition. Hedge's choice of the term ‘driver’ suggests that she saw purpose as having the capacity to drive the reading process. As Davies (1995:78) observes, this is consistent with the roles of the ‘message centre’ in Rumelhart's (1977) model and that of the ‘thematic processor’ in Rayner and Pollatsek's (1989) model. It is also evidence of the importance of monitoring (e.g., Sarig 1987a and 1987b).

### **3.3 Section summary**

This section has traced the development of theory in both L1 and L2 reading. Essentially, it has shown that, over the past three decades, there has been a progression from polarized top-down or bottom-up views of reading towards a converging view that fluent reading is a bottom-up or bottom-up interactive process. Thus, in answer to Research Question 1.2, which was posed at the beginning of this chapter and which sought the model or models that best describe the reading process, my own view is that Rayner and Pollatsek’s (1989) model comes closest.



However, this conclusion needs to be moderated by a pervasive doubt in the literature as to whether we are yet in a position to postulate a comprehensive model of reading. Indeed, Rayner and Pollatsek (1989) themselves take a very cautious view of all models, including their own: “we feel that most ... models are little more than general frameworks which provide some biases about which aspects of reading are really important” (p. 25). In a similar way, Oakhill and Garnham (1988) pose the fundamental question: “can we formulate an overall model of reading?” (pp. 32-33). Their own answer to this question is that “the task is a very difficult one” and go on to suggest that overall models of reading have become “less common in recent years” because the more we learn about reading, the more complex we realize it is.

As far as L2 reading is concerned, Stanovich’s (1980) interactive-compensatory model appears highly relevant. It might particularly apply to readers at relatively low levels of L2 proficiency who need to compensate for this deficiency. Moreover, Bernhardt’s (1991) developmental theory aimed to be at least a partial description of L2 reading and Hedge’s (1991) ‘modes’ provide a very clear view of the diversity in her own L2 subjects’ reading behaviour.

It is in some ways ironic that the increasing levels of consensus about what constitutes the reading process should actually impede the development of models to describe it. However, while model development seemed stalled in the 1990s, much recent work has deepened our knowledge of such key aspects of reading as the roles of words, context, phonology and writing systems. The following section presents a review of these aspects of the reading process.

## 4.0 THE READING PROCESS

Although reading is an immensely complex process and one that is very hard to research, certain clear and consistent trends have emerged in the literature, particularly over the past two decades. Indeed, as suggested in the section above, it is becoming clearer which of several processes need to take place for successful reading to occur. This raises the issue, posed in Research Question 1.3 at the beginning of this chapter, of differences between L1 and L2 reading. In an attempt to answer this question, the following section reviews the literature on the reading process, first in the L1 and then in the L2.

### 4.1 The reading process in L1

This part of the discussion begins by outlining an emerging consensus on L1 reading processes. It continues by discussing in detail each of the following factors, generally accepted as implicated in the reading process: the roles of words and context, phonology, and writing systems.

#### 4.1.1 Converging evidence

Carr (1981) noted the prevailing tendency at that time for researchers to attribute individual differences in reading to a 'single factor'. Apart from the fact that there was no consensus as to what that single factor might be, he cited extensive literature to show that "a very wide range of component skills and abilities has in fact been shown to correlate with reading success" (p. 73). Carr outlined four 'theories of reading ability' emerging from the literature in the early 1980s, three of which point to *bottom-up* processes as the source of reading disability and one of which points to *top-down* processes.

The most salient single factors for reading disability that Carr noted are deficits in:



- visual discrimination and visual code formation; i.e., the reader has difficulty dealing with closely related structural properties of the writing system, such as poor discrimination of letter location and order, and inability to use orthographic regularity as an aid to perceptual encoding;
- phonological and semantic recoding; i.e., the reader has difficulty translating from spelling patterns to pronunciation and to meaning;
- short-term memory: the reader has recoding difficulties, a by-product of a fundamental deficit in short-term memory;
- the utilization of linguistic knowledge and context to guide perceptual recognition and encoding; i.e., the reader has difficulty in spontaneous identification and use of syntactic relations, and in flexible use of semantic context to aid prediction.

Fourteen years later, in an article on how cognitive research can inform L1 reading education, Perfetti (1995) isolated four factors that emerged from his extensive review of the literature to date. He states that there are now “solid results from converging research” and that these results “span a range of school levels from skilled reading, through middle-grade to beginning reading” (p. 106). The four factors which Perfetti noted are:

- skilled readers read words to identify them;
- less skilled readers use context to identify words;
- skilled readers use phonology;
- successful readers learn how their writing system works.

What is striking, given the time span, is how similar Perfetti's four factors are to those of Carr, suggesting a consolidation of the findings available to Carr in the early eighties. Again, it is clear from these four statements that the bottom-up or bottom-up interactive models appear to have far greater support than their top-down counterparts. While Perfetti's factors provide a very useful framework for analyzing the key issues in the literature, they are interrelated and thus there is some overlap in the discussion that follows.

#### **4.1.2 The roles of words and context**

There follows a discussion of research findings on the roles of words and context in reading in which an important distinction is made between two types of contextual processing. Next, there is a review of converging evidence, emanating from studies utilizing diverse elicitation tasks and procedures, of the importance of word recognition in reading. After that, there is a discussion of how some top-down theorists have chosen to ignore the mounting weight of evidence about the importance of the bottom-up process of decoding text.

#### **Research evidence**

The primacy of words in reading has been a common finding since the earliest studies. Huey (1908/1968) observed that more letters can be apprehended in a given unit of time if they spell a real word than if the same letters are used in a non-word. Related to this finding, Miller, Bruner and Postman (1954) found that more letters can be apprehended in a nonsense letter string that conforms to English spelling rules than in a nonsense letter string that does not.

More recently, a common finding has been that skilled readers read words to identify them. For instance, while Biemiller's (1970) developmental study found increasing attention to graphic information over a period of eight months in a class of children aged seven, Weber (1970) found greater attention to graphic information among her better readers. Allington and Strange (1977), using a variant of an oral reading task on children aged about ten, also found that their good readers paid greater attention to graphic information than their poor readers.

Juel (1980), researching the reading of children aged seven and eight, found that her good readers were predominantly text-driven while her poor readers were relatively context-driven. Likewise, Perfetti and Roth (1981) found good readers were less reliant on contextual information than poor readers. They argue that good readers decode faster than poor readers and so word recognition takes place before reliance on context becomes necessary; good readers can therefore be at once more sensitive to



context and yet less reliant on it, due to the availability of information from other knowledge sources.

In a similar vein, Oakhill and Garnham (1988) claim that good readers may have greater awareness of context but that they do not need to use it while reading. West and Stanovich (1978) found that their less skilled child subjects relied more on context than their more skilled readers. They argue that the use of context supplements poor decoding skills. Parenthetically, this finding may well have led to Stanovich's (1980) interactive-compensatory model (described above), a possible example of a research finding informing the evolution of reading theory.

### **Two types of contextual processing**

Mitchell and Green (1978) distinguish between two types of contextual processing, that for comprehension (e.g., Bransford and Franks 1971, on semantic integration) and that for word recognition (e.g., Smith 1971, Goodman 1976, on contextual hypothesis testing). As Stanovich (1980:44) points out, this is a useful distinction in understanding the literature on individual differences in the development of reading fluency. While top-down theorists assume that fluent readers make considerable use of all higher level processes, Stanovich (1980), in an extensive literature review, shows that the weight of evidence indicates that poor readers rely more on context for word recognition than good readers. Later, Stanovich (1982) reviews twenty-two studies, none of which showed that good readers use context to aid word recognition more than poor readers. However, as Stanovich and Stanovich (1995:90) observe, there is considerable evidence that good readers are better able to use contextual information to facilitate their comprehension processes (e.g., Baker and Brown 1984, Stanovich and Cunningham 1991).

### **Evidence from studies using diverse tasks**

Adding considerable weight to the evidence of the importance of word recognition is the fact that it emanates from studies using diverse reading tasks. Examples of tasks

used are: lexical decision tasks; cloze tasks; various tasks, including reading disrupted text, whose purpose is to study eye movements; and, for the youngest child readers, oral reading.

Many studies have used priming paradigms where a context, ranging from a word to a paragraph, precedes a target word about which the subject is asked to do a naming or a lexical decision task (e.g., Perfetti 1985, Stanovich 1986). While oral reading errors may have diverse causes, the findings of these studies are consonant with those using other tasks (e.g., Biemiller 1970, Perfetti and Roth 1981). Furthermore, Stanovich and Stanovich (1995) report that the results of studies investigating the effects of text disruption and timed text reading also converge, showing that the effects of contextual information are reduced as word recognition increases.

A review of the literature on eye movements also fails to support the top-down hypothesis that fluent readers can recognize words fast as a result of extensive use of context. For example, McConkie and Zola (1981), investigating eye fixation patterns on a cloze task, found that the highly constrained words (chosen more than 85% of the time) were just as likely to be fixated as relatively unconstrained words (chosen less than 15% of the time). Further, they found that the fixation durations on constrained words were only 14 milliseconds shorter than on the unconstrained items. Bearing in mind that Rayner and Pollatsek (1989:118) report that normal fixation time is about 231 milliseconds (depending on the genre of text being read), this would seem to be a minute difference.

Rayner and Bertera (1979) conducted an experiment in which a computer masked one letter in *foveal* vision on each *fixation* and found that this apparently small loss reduced reading speed by fifty percent. (Foveal vision is defined by Rayner and Pollatsek 1989:9 as the area of vision that is within two degrees of a fixation and that has the greatest acuity.) This strongly suggests that foveal vision is crucial in normal and fluent reading.

Ten years later, Rayner and Pollatsek (1989) described several experiments on eye movements using a wide variety of genres, including light fiction, newspaper articles and biology texts. They report that readers fixate, on average, on every 1.1 words of a



text and that *saccade* length is approximately 7-9 characters. They also report that, if fewer than four character spaces to the left of a fixation and fifteen character spaces to the right (i.e., a total of 19-20 characters) are visible, reading does not proceed normally. Nonetheless, only about seven characters in the middle of a fixation are processed through the *fovea* (i.e., the part of the retina that produces the greatest visual acuity), the other characters being available only to *parafoveal* vision. (Parafoveal vision is defined by Rayner and Pollatsek 1989:9 as the area of vision that is between two and ten degrees of a fixation and that has less acuity than foveal vision.) Thus, it appears likely that, since saccades need to be 7-9 characters long, foveal vision is necessary for word identification.

### **Ignoring the evidence**

Despite the weight of this evidence from studies using a variety of tasks and data-gathering techniques, there is continued opposition from some die-hard top-down theorists. For example, as recently as the late eighties, Smith (1988), ignoring the fact that eye movements are one of the few ways in which observable and measurable evidence about reading can be obtained, dismissed studies looking at eye movements as focussing on “rather superficial characteristics of reading” (p. 249).

Even more recently, Goodman (1992) has continued to use rhetoric and a tone of personal rivalry to obscure the increasing weight of evidence against top-down models. He declared himself “amused that ... Tom Nicholson and others [Nicholson et al. 1988] have devoted their careers to proving I was wrong” (p. 192). He also appeared to suggest that he should have a monopoly in developing reading models when he continued by stating that “people like Gough (1972) and LaBerge and Samuels (1976) felt obligated to explicate their own models if only to refute mine” (p. 193).

### 4.1.3 The role of phonology

Research over the past twenty years has shown the important role of phonology in written word identification. The question at issue is whether phonology mediates all word identification or whether it only mediates low frequency words with regular spelling patterns. The former view has been supported by van Orden et al. (1990), and by empirical results from research conducted by Perfetti et al. (1988), Lukatela and Turvey (1990), and Perfetti and Zhang (1995a). The latter view, on the other hand, conforms to the ‘dual route theory’ (Coltheart 1978). This posits that much word identification occurs along the faster of two mechanisms, that is, either the conversion of letters to phonemes prior to word identification or simply using the spelling of a target word to ‘look it up’ in the lexicon. This is supported by, among others, Seidenberg et al. (1984) and Paap and Noel (1991).

As already noted, Rayner and Pollatsek's (1989:472-473) model is the only one that specifically accommodates the possibility of a dual route, thus acknowledging what remains an open question. The model posits a direct route to lexical access (from printed letters to the lexicon) and an indirect route (from printed letters via “the application of rules and/or analogies to create an auditory route to the lexicon”). Rayner and Pollatsek say that the indirect route would “obviously” be the primary means for identifying words that are relatively infrequent or unfamiliar in print (p. 474).

Perfetti, Zhang and Berent (1992) propose a ‘universal phonological principle’ in reading whose central principle is that ‘most’ printed words automatically lead to phonological activation. This is supported by recent evidence that retrieval of phonological information may be universal across writing systems. Perfetti and Zhang (1995b) conducted a study into reading in Chinese, and they found that what they termed the ‘word identification reflex’, or the retrieval of a phonological word form as part of word identification, occurs in Chinese as well as English.

The significance of this finding is that reading in Chinese has been commonly thought to be a strictly print-to-meaning process. This is based on the widespread belief that the Chinese writing system is entirely based on a logographic principle, although, as



Perfetti and Zhang point out, this is not entirely true. While it is true that the writing system is based on a morphological principle, DeFrancis (1989) has shown that there is a considerable amount of syllable-based phonology as well. Moreover, and anecdotally, Perfetti (1995) notes that Chinese education “acknowledges the value of alphabetic reading” (pp. 112-113). It does so by initiating the teaching of reading to schoolchildren with eight weeks of instruction in letter-sound correspondences using an alphabetic script known as ‘pi yin’ that has been in use in the People's Republic of China since 1957 (Norman 1988).

Recently, Bernstein and Carr (1996) conducted a study whose results raise the possibility that “normal mature readers may differ in their reading system architectures” (p. 110). They recommend follow-up studies on individual differences to impose order on the vexed literature on visual word recognition. Their results make them more confident of the separate status of the assembly route and hence of the viability of the dual-route or parallel-coding system approach for mature readers. But they cannot determine how the assembly route operates except that it requires more attentional capacity than does retrieval.

The weight of evidence seems to suggest, then, that some phonological coding takes place in reading, probably in all languages, but that its level of use may vary among individual readers and, in all readers, in response to high and low frequency words in a text.

#### **4.1.4 The role of writing systems**

Perfetti (1995) says that “the central fact of learning to read is that a child must learn a writing system, specifically how the writing system encodes his or her language ... whatever else learning to read is, it is a kind of learning” (p. 109). As Perfetti puts it:

“The alphabetic principle is difficult, as indicated by the lateness and uniqueness of its discovery by the Phoenicians and Greeks. Its discovery (or invention) lagged far behind the writing systems that merely provided symbols for objects and meanings (Gelb, 1952).” (p. 110)

Perfetti also quotes Gleitman and Rozin (1977), who note, somewhat sardonically, that learners might have some trouble “replicating” this discovery.

Children learning an alphabetic writing system need phonemic awareness; that is, they must be able to associate letters with phonemes. Several studies have found a relationship between phonemic awareness and learning to read (e.g., Fox and Routh 1976, Lundberg, Olofsson and Wall 1980, Stanovich, Cunningham and Cramer 1984). Perfetti (1995) says that the research showing the details of this relationship has become “substantial” and that “the consensus from the research is that the relationship between phonemic awareness and learning to read is not one-directional but reciprocal” (p. 111). He points out that this reciprocity between literacy and phonemic awareness reflects the fact that specific phonemes are not a salient part of the perceptual experience of hearing words. He argues convincingly that this lack of salience “can be compounded in the reading situation, which requires the child to coordinate temporally represented phonemes with spatially represented print” (p. 112).

Pertinent to this discussion is the debate in the literature as to whether the process of reading acquisition is ‘natural’ or ‘unnatural’. Top-down theorists, such as Goodman (1986), claim the former, equating the learning of written language with that of the spoken language; in contrast, bottom-up theorists, such as Gough and Hillinger (1980) with their two-stage model, claim the latter. Citing several cognitive psychologists, Stanovich and Stanovich (1995) observe:

“... the use of the speech/reading analogy ignores the obvious facts that all communities of human beings have developed spoken languages but only a minority of these exist in written form, that speech is almost as old as the human species but that written language is a recent cultural invention of only the last three or four thousand years, and that virtually all children in normal environments develop speech easily, whereas most children require explicit tuition to learn to read and substantial numbers of children have difficulty even after intensive efforts on the part of teachers and parents.” (pp. 92-93)



#### **4.1.5 Summary**

On the basis of current evidence, it seems that the acquisition of reading cannot be regarded as a 'natural' process in the same way as that of the spoken language. The process of developing phonemic awareness, necessary for reading, particularly in English, is highly complex as it requires the coordination of phonemes and print. In addition and crucially, fluent reading requires the rapid and accurate decoding of words in text. This leads to the question of the extent to which reading is a thinking process and/or a language process, a question that has profound import for L2 readers, particularly those at relatively low proficiency levels. The next part of this section investigates the reading process in L2.

### **4.2 The reading process in L2**

In the discussion on the history of L2 reading research above, brief reference was made to various strands or avenues of inquiry in the literature. The first involved the question of L2 proficiency or, more accurately, the role of L2 deficiency in reading. The second related to various routes of inquiry based on a schema-theoretic perspective. The third line of inquiry mentioned was that into reading strategies. The first two of these strands are discussed below while Chapter 2 is devoted to a full literature review of work relating to reading strategies.

#### **4.2.1 L2 deficiency**

By the time most children start learning to read in their L1, they have a fairly sound grasp of the language. They have learned between 5,000 and 7,000 words (Singer 1981) and, according to Grabe (1991), they have a good intuitive sense of the grammar of the language. Most beginners and elementary students of an L2 have neither of these, though it is safe to say that, depending on the L1, there would be variation in their passive knowledge of both vocabulary and grammar, resulting in varying levels of difficulty in language processing. Some languages have a lexical base close to that of English while others have no such relationship (cf. French and

Chinese); syntactic variation (e.g., word order) across languages may also cause varying levels of difficulty for L2 readers.

Setting this inevitable variation to one side, there are some general points that can be made about the possible consequences of language deficiency in L2 reading. Eskey (1988) has observed that, while top-down models of reading “make the perfectly valid point that fluent reading is primarily a cognitive process, they tend to *deemphasize* [his italics] the perceptual and decoding dimensions of that process” (p. 93). He goes on to say that he believes that “simple language decoding has a major role to play in the process – that good reading is a more language-structured affair than the guessing-game metaphor [e.g., Goodman 1967, 1992] seems to imply” (p. 94).

One argument for this position is that, when decoding has reached a level of automaticity that requires no conscious cognitive effort, fluent readers have free cognitive space to employ higher-level top-down interpretative processes such as the use of prior knowledge (e.g., Neely 1977, McLaughlin 1990). This is reflected in Bernhardt’s (1991) theory of L2 reading, which predicts higher ‘error rates’ for such bottom-up features as word recognition and phonemic/graphemic features at lower levels of L2 proficiency. However, possibly counter to this argument, Bernhardt predicts fairly stable ‘error rates’ for such top-down features as prior knowledge and intratextual perceptions (i.e., knowledge gained during the reading of a text) across all L2 proficiency levels.

Several studies have provided evidence of the importance of language proficiency level in successful L2 reading. Analyzing oral reading errors in L1 and L2 readers of French, Cziko (1980) found that the less proficient L2 speakers of French in his study relied more on graphic information than either the more proficient speakers or native speakers of French. Using oral reading errors and cloze test results, Clarke (1980/1988) found that “limited control over the language ‘short circuits’ the good reader’s system causing him/her to revert to poor reader strategies when confronted with a difficult or confusing task in the second language” (1988:120). (Clarke’s ‘good’ readers were selected on the basis of cloze test scores in Spanish, their L1; by ‘poor reader strategies’, Clarke was referring to a marked focus on syntactic cues.)



In a more recent study, Davis and Bistodeau (1993) found support for the ‘short circuit’ hypothesis. Using both L1 and L2 subjects introspecting concurrently with reading texts that had each sentence typed on a separate card, they found “strong” evidence to suggest that there was more bottom-up processing, especially with regard to vocabulary items, among the L2 readers than the L1 readers (p. 468). However, they also noted that the L1 readers tended to express irritation with the writer when they could not understand something whereas the L2 readers tended to blame themselves for lack of comprehension. This variation in affective response may mean that L2 readers make more effort to decode text for comprehension than their L1 counterparts.

Devine (1987) saw reading as such a language-centred process that she subscribed to a language threshold for L2 reading abilities; however, following analysis of data in two case studies of two middle-aged “beginner/low intermediate” subjects, Devine (1988:130) modified her position. From her 1988 data, which was derived from an oral reading interview, a sample of oral reading and a retelling of the oral reading, she identified three ‘internalized models of reading’: ‘sound-centred’, ‘word-centred’ and ‘meaning-centred’. The first of these is the most bottom-up while the last represents “a successful balance between text-based and knowledge-based processing” (p. 137).

#### **4.2.2 Schemata**

While L2 readers may suffer the handicap of language deficiency, those who have reached adulthood actually have some advantages over child L1 readers. They have a developed sense of the world both factually and conceptually and can make inferences about it from texts. They may also be fluent L1 readers and may even have metacognitive awareness of the strategies they use when encountering reading difficulties. A central question, then, is to what extent prior knowledge balances lack of knowledge about the L2. I will review two areas of knowledge covered in the literature: content and rhetorical.

## **Content schemata**

Hudson (1982/1988) challenged the ‘short circuit’ hypothesis. He attempted to override or compensate for L2 deficit by inducing content schemata and providing vocabulary deemed essential to understand the text. Using adult students of varying L2 proficiency levels attending an intensive language institute in the United States, he found that, especially among the subjects at lower L2 proficiency levels, the induced schemata apparently allowed access to language decoding that was otherwise not available. He concluded that “the linguistic ceiling is only one determinant of reading comprehension” (1988:198). Although this finding acknowledges a role for L2 proficiency in reading comprehension, it begs the question of the practicality of teachers and lecturers having to induce schemata for lower-level L2 students whenever a new text has to be read. It might be more generative for teachers to concentrate on their students’ L2 proficiency. Nonetheless, several researchers recommend the presentation of at least minimal background knowledge to students who have low levels of L2 proficiency and who lack content knowledge (e.g., Barnett 1989, Dubin and Bycina 1991).

Pritchard (1990) looked at a particular aspect of content schemata, that of the reader’s cultural knowledge of text content. Using eleventh grade high school students (i.e., aged in their mid to late teens) in the United States and in Palau, he found that both groups of students read more locally (i.e., used more bottom-up strategies) when reading culturally unfamiliar content than when reading content that was culturally familiar. In retellings, he found that students remembered more, and remembered more accurately, culturally familiar material. In an ethnographic study looking at culture in a wider sense than text content, Parry (1996) compared readers’ strategies in Nigeria and China. She found that her Nigerian subjects were more likely to use top-down strategies and her Chinese subjects were more likely to use bottom-up strategies and concluded that setting could play an important part in the development of literacy.



## Rhetorical schemata

Several studies have investigated the effects of reader knowledge of the rhetorical structure of texts. Carrell's (1984) study adapted an earlier L1 study conducted by Mandler (1978) investigating narrative recall. Working with intermediate L2 students from several countries enrolled in an intensive pre-session course, Carrell gave half her subjects a 'standard' version (i.e., with story nodes following a normal narrative sequence) of three short narratives while she gave the other half 'interleaved' versions (i.e., with nodes following a distorted sequence). She found that the readers of the 'standard' versions of the stories recalled significantly more nodes (at  $p < 0.05$ ) than the readers of the 'interleaved' versions.

Working with university students in an English-medium university in Hong Kong, Mahoney et al. (1997) replicated Carrell's study. Higher recall rates were found for both groups of readers than in the Carrell study, but, again, the readers of the 'standard' versions recalled more than those of the 'interleaved' versions, though not at the same level of significance ( $p < 0.15$ , compared to  $p < 0.05$ ). The researchers concluded that one reason for the higher recall rates in their study was probably the higher L2 proficiency level of the Hong Kong students than those in the Carrell study.

### 4.2.3 Summary

The principal argument about the necessary components for successful L2 reading has been between those who believe that the main determinant to successful L2 reading is L2 proficiency and those who see it as being due to various schemata, whether content, rhetorical or cultural (i.e., excluding L2 schemata). However, it would be wrong to suggest that the field is currently polarized. While many researchers suggest a balance between these two views, there are also those, including myself, who, while acknowledging the evidence of the importance of various schemata, nevertheless see language development as a *prerequisite* to rapid and accurate L2 reading.

### 4.3 Section summary

Converging evidence from L1 studies using a variety of tasks and data-gathering techniques indicates that skilled reading involves reading words, not skipping them; it also involves reading words context-free. Context is typically used for the higher-order processes of meaning interpretation rather than for word identification. While it is not yet clear whether phonology is used for the identification of literally every word, it is clearly strongly implicated in word identification. However, it is safe to say that reading is not a ‘natural’ activity and that it requires the learning of a writing system, which entails fostering the child's phonemic awareness and builds on his or her knowledge of the syntax and vocabulary of the spoken language.

In terms of L2 reading, studies have provided evidence for the importance of both L2 proficiency and various schemata. However, as Eskey (1988) has observed,

“knowledge of the language of a text must be an integral part of whatever background knowledge is required for the full comprehension of that text. Language is a kind of schema, too, albeit one that for fluent native users may be activated automatically, and there is no easy top-down route around the fact that fluent reading entails bottom-up perceptual and linguistic skills as well as higher-order cognitive processes.” (p. 96)

Thus, in answer to Research Question 1.3 in Section 1.0 above, which asks about differences between L1 and L2 reading, the principal and most obvious answer is that, except at extremely advanced proficiency levels, L2 reading is characterized by varying levels of language deficiency. Since rapid and automatic decoding of words is a prerequisite to fluent and accurate reading, the inability to do so in an unfamiliar language represents the defining difference between L1 and L2 reading. Many other factors, including the reader’s knowledge of the text content and rhetorical structure, certainly play a part in making L2 reading an experience that is different from, and usually more difficult than, L1 reading. In addition, L2 readers from different cultural backgrounds may approach texts in varying ways.



## 5.0 CHAPTER SUMMARY

This chapter posed and sought to answer the following research questions:

- 1.1 What have been major influences on the development of the study of reading?
- 1.2 What model or models of reading best describe the reading process?
- 1.3 What differences are there, if any, between reading in a first language (L1) and reading in a second language (L2)?

The chapter began by reviewing the history of reading. While research began in the late 1870s, it was effectively aborted with the rise of behaviourism in the 1910s. However, with the resurgence of interest in the study of cognition came new interest in the mental processes involved in reading.

During the 1970s and 1980s, several models of reading were proposed. At first, there was a stark division between those who believed that reading involves mainly top-down processing and those who held that it requires mainly bottom-up processing. Later models, developed with the benefit of an increasing number of research findings, posited that reading is essentially an interactive process between reader and text. The most recent influential model (Rayner and Pollatsek 1989), following much research seeking to establish links between eye movements and mental processes in reading, refines the notion of interactive processing into one that is bottom-up interactive. Interestingly, as models have become increasingly comprehensive, their proponents have become increasingly cautious about their positions (cf. Goodman 1967 and Rayner and Pollatsek 1989).

Turning to L2 reading, the chapter reviewed theory development in this very young field of study. The only serious theory of L2 reading to date, and one founded on a considerable body of research evidence, that of Bernhardt (1991), is based on the tenet that reading processes in L2 adults vary considerably with language proficiency level. It seems clear that automatic decoding of text is a necessary but not sufficient skill in reading comprehension. There is also evidence that L2 readers' schemata (e.g., content, rhetorical, cultural) also have an impact on reading comprehension.

Much work remains to be done to discover what L2 learners in various contexts actually do when reading texts. Some researchers have attempted to increase our knowledge of this area through the study of readers' mental processes and their representation as reading strategies. This avenue of the research effort is reviewed in the following chapter.



## **CHAPTER 2: STRATEGIES**

### **1.0 INTRODUCTION**

One of the themes running through Chapter 1 was that investigating the reading process is no easy matter. In order to extend our knowledge of reading behaviour, the researcher needs to be able systematically to describe readers' reports of their actions, both mental and physical. One way that has been used to do this is to study what have widely been termed 'strategies' reported by readers, though, as will be shown, there is little consensus as to a precise definition of what strategies are. With this in mind, this chapter aims to address the following research questions:

- 2.1 What is the context behind research into reading strategies?
- 2.2 What are reading strategies and how have they been defined?
- 2.3 How have they been classified?
- 2.4 Can they provide at least a partial description of reading behaviour?

In an attempt to find answers to these questions, this chapter begins by outlining the history of research into strategies and the motivations of investigators. Since research into reading strategies has evolved from work on learning strategies in cognitive psychology, the problems of defining learning strategies are first addressed. This is then followed by a review of corresponding definitional problems in the literature on reading strategies. The chapter ends with a discussion of the difficulties that have been experienced in classifying strategies.

## **2.0 A BRIEF HISTORICAL OUTLINE**

This section, which seeks to provide at least a partial answer to Research Question 2.1 above, briefly describes the historical development in the study of learning strategies. The section is divided into two parts, each presented chronologically. The first part outlines research investigating successful learners' strategies and strategies that lead to successful outcomes; the second part reviews work conducted in the area of strategy training.

### **2.1 Paradigm shift**

This section describes a paradigm shift that has taken place in this area of research in the past quarter century. From a focus on strategies used by successful learners, there has been a move towards an interest in strategies that lead to successful outcomes, irrespective of the overall success of the informant reporting the use of particular strategies.

Initially, research into learning strategies was motivated by the notion that successful learners might use certain strategies or combinations of strategies that less successful learners do not (e.g., Rubin 1975, Stern 1975). It was hoped that discovering which learning strategies best facilitate learning, including language learning, might yield an understanding of how to assist slow learners (Larsen-Freeman 1991). This line of research was based on several assumptions, one of which was that consciousness-raising of learning processes could lead to automaticity in those processes (e.g., McLaughlin 1978, Smith 1981). (See Rubin 1987:15-19 for a full review of the assumptions underpinning this work.)

Many early studies focussed on identifying and classifying learning strategies that were regarded as important for successful language learning (e.g., Stern 1975, Naiman et al. 1978, Rubin 1981). Hosenfeld (1977) was the first to report on L2 reading strategies, and, following the trend of the time, she focussed on 'successful' and 'unsuccessful' learners. Wesche (1979), having analyzed her data, taken from 'successful' L1 adult learners, hypothesized that "it may be complexes of them



[learning behaviours] rather than specific ones which characterize different kinds of learners” (p. 419).

Several years later, Wesche’s interesting insight was pursued in terms of L2 reading strategies when Sarig (1987a) investigated what she called ‘comprehension-promoting’ and ‘comprehension-detering’ strategies in her data. This idea added to this line of inquiry the important ingredient of whether or not a particular strategy was successful *in a particular context* rather than *per se*.

## 2.2 Strategy training

Hand in hand with the notion that what ‘successful’ learners do can be used to help ‘unsuccessful’ learners is that of strategy training. With the drive to find ways to encourage learners to become more autonomous and, at the same time, more responsible for their own learning, the spotlight has shifted in recent years towards the effects of strategy training. Much research has been conducted in this area (e.g., Weinstein and Underwood 1985, Padron 1985, Carrell et al. 1989, O’Malley and Chamot 1990, Oxford 1990).

Ground-breaking research in the L1 literature (e.g., Weinstein 1978, Brown and Palincsar 1982, Weinstein and Underwood 1985) has shown that relatively poor readers can improve their skills through training. Casanave (1988), among other L2 researchers, has advocated the implementation of strategy training studies similar to those in the first language. She stresses the importance not only of explicit instruction in formal and content schemata for successful reading, but also of what she calls ‘strategy schemata’. This taps the metacognitive domain and refers to a reader’s underlying knowledge of monitoring and repair, or compensatory, strategies.

The following year, Carrell et al. (1989) conducted a study that focussed on metacognitive strategy training. They did so because of what they regarded as the importance, in successful reading, of metacognitive control, that is, where the reader “consciously directs the reading process” (p. 650). Using semantic mapping and the ‘experience-text-relationship’ method (see pp. 650-654 for their explanation of these

methods), their study found that metacognitive strategy training can be effective. They concluded, of poor readers, that successful training in metacognitive control might be a key way to facilitate learner autonomy in reading.

Several issues related to the efficacy of strategy training among poor readers have emerged from L1 research. These include choice of strategies or combinations of strategies to be taught as well as the degree and duration of explicitness required for strategy use to be initiated, maintained over time and transferred to other reading tasks (e.g., Palincsar et al. 1984).

There are numerous advocates for L2 replications of this and similar studies as well as of closely related research (e.g., Oxford and Nyikos 1989, Skehan 1989, O'Malley and Chamot 1990, Oxford 1990). In terms of strategy training research design, there appears to be a need for the use of both experimental and qualitative designs; moreover, longitudinal studies are advised wherever possible to determine change over time. Some studies (e.g., Palincsar et al. 1984) have used both quantitative and qualitative data for separate analyses as well as a combination of the two in a case study approach (pp. 153-157).

In a comprehensive review of developments in L2 reading research over the past quarter century, Grabe (1991), while warning of the complexities of strategy training, writes that “this line of research is particularly important because of the promise it holds for reading instruction” (p. 393). It follows that it may also be of potential importance in teacher education. Certainly, in order to encourage L2 reader independence, fundamental changes are needed in the typical reading curriculum and in materials development as well as in teacher education.

### **2.3 Section summary**

In partial answer to Research Question 2.1, which was posed at the start of the chapter and inquires about the context behind research into reading strategies, this section has shown that early work investigating the learning strategies of successful learners developed into research into strategies that lead to successful outcomes. Some of the



work on strategies focused more specifically on language learning and on L2 reading. In the 1980s, investigators looked into the possibility that poor learners and poor readers might benefit from strategy training, and there is some evidence that they can. However, perhaps the most enduring, and intractable, issues in the strategies literature remain those of defining and classifying strategies. These issues are now addressed.

### **3.0 LEARNING STRATEGIES: THE PROBLEM OF DEFINITION**

This section begins by showing that a definitive definition of learning strategies has proved elusive. Issues where researchers are at variance include whether strategies are neutral descriptors or carry an evaluative and/or affective element, whether they are conscious and/or automatic processes, and whether they are always problem-oriented. Difficulties in identifying a widely accepted definition of strategies are exacerbated by the fact that the term is often confused with related terms such as 'styles', 'skills' and 'tactics'. This section seeks to add substance to the response above to Research Question 2.1. (The research questions can be found at the beginning of the chapter.)

#### **3.1 Elusive quality**

Defining the term 'strategy' has been very problematic. It is sometimes narrowed down to learning strategy or narrowed still further to, say, reading strategy or comprehension strategy. Weinstein and Underwood (1985) say the term 'learning strategy' is "used in a very broad sense to identify a number of different competencies that researchers and practioners have postulated as necessary, or helpful, for effective learning and retention of information for later use" (p. 241).

However, while this definition suggests that there is a positive evaluative element to the term, it has also been used in a neutral sense. This view acknowledges, at least tacitly, that it is, to a large extent, an open question as to which strategies are 'good' or 'successful' and which are 'bad' or 'unsuccessful'. It is worth remembering that even strategies that can be shown to be valuable in one context may not necessarily be equally effective in other contexts (e.g., Sarig 1987b).

Referring to L2 learners, Wenden (1987) gives a neutral definition of learning, or learner, strategies: "learning behaviors learners actually engage in to learn and regulate the learning of a second language" (pp. 6-7). This is a non-evaluative statement about the behaviours involved in the learning process and its management. Wenden posits six criteria that characterize strategies:



- They are “specific actions or techniques”, not a general approach.
- They may be observable or non-observable.
- They are problem-oriented (i.e., they respond to a learning need).
- They contribute directly or indirectly to learning.
- They may be consciously deployed or have become automatized.
- They are amenable to change.

The characterization of strategies as ‘problem-oriented’ is interesting in that it seems to preclude on-going or uninterrupted learning. Feldman and Stemmer (1987) have a similar view when they say that a strategy is “a potentially conscious plan for solving what, to the individual, presents itself as a problem in reaching a particular goal” (p. 258). (Problem-orientation in reading strategies is pursued later in this thesis.)

However, Dansereau (1985) does not incorporate problem-orientation into his definition, preferring to opt for a more neutral view, saying that a strategy is “a set of processes or steps that can facilitate the acquisition, storage, and/or utilization of information” (p. 210). Further, he characterizes strategies as being either algorithmic (i.e., consisting of sequences of processes that remain fixed over tasks) or heuristic (i.e., consisting of a sequence of processes that may be modified, “depending on task conditions and the needs and skills of the individual learner”) (p. 210).

Dansereau's definition adds an important element to the quest for a robust definition of strategies by specifying that they are “a set of processes or steps”. Garner (1987, quoted in Kletzien 1991:69) echoes the idea of the possibility of a multiplicity of steps: “a reading strategy is an action (or a series of actions) that is employed in order to construct meaning”. (The notion that learning strategies have the potential to encapsulate a series of related actions will be developed in this thesis.)

Oxford (1990), in her book on language learning strategies, does not pursue this possibility. She feels that Dansereau's definition is “technical ... [and], while helpful, does not fully convey the excitement or richness of learning strategies” (p. 8). She adds further elements to what might qualify as a strategy with her rather unusual definition: “specific actions taken by the learner to make learning easier, faster, more

enjoyable, more self-directed, more effective, and more transferable to new situations” (p. 8). Oxford's definition assumes that strategies are ‘good’ or ‘successful’ and adds an affective element to the concept. Further, she sees the teaching of learning strategies as a route to greater learner autonomy. Also, she seems to assume that this will somehow be transferable, though evidence for this would need a substantial research effort in diverse educational settings.

While there has been difficulty in pinpointing the essential quality or qualities of strategies, there has been a corresponding confusion over terminology. This question is now addressed.

**3.2     Overlapping terminology**

One of the problems in trying to define learning strategies is that they are often confused with several other related terms, such as learning styles, skills and tactics.

There has, in particular, been confusion and overlap between ‘styles’ and ‘skills’. This is not surprising since there are considerable similarities among them. Indeed, according to Kirby (1988), strategies and skills are inextricably linked and constantly interact; and, by his own definition, styles are also closely connected. Kirby’s definitions of these and other related terms are shown in Table 2.1 below (p. 230).

**Table 2.1: Kirby’s (1988) definitions of learner activities**

<b>Styles</b>	The habitual use of a class of similar strategies
<b>Strategies</b>	Means of selecting, combining or redesigning cognitive routines, choices among tactics
<b>Tactics</b>	Decisions to employ certain skills
<b>Skills</b>	Existing cognitive routines for performing specified tasks
<b>Abilities</b>	Set limits to the development of skills

Table 2.2 illustrates how Kirby envisages the connections between the learner activities within what he terms the ‘skills domain’ and the ‘strategies domain’ (p. 230).



**Table 2.2: Interconnections between and among learner activities (Kirby 1988)**

<<<<<<< Skills domain >>>>>>>			<<<<<<< Strategies domain >>>>>>>		
Abilities .....	Skills .....	Tactics .....	Strategies .....	Styles	

While there is greater consensus on the notion that a learning style has the characteristics of a habit and represents an individual's prevailing use of particular strategies, there is no such agreement over Kirby's definitions of strategies and skills. He seems to confine strategies to a managerial role while his definition of skills appears close to some definitions of strategies (e.g., Wade et al.'s 1990 definition, quoted below).

Williams and Moran (1989) define skills in a similar way to Kirby but highlight their subconscious nature by saying that they are “an acquired ability, which has been automatized and operates largely subconsciously” (p. 223). In contrast, they define strategies as “a conscious procedure carried out in order to solve a problem”. Armbruster et al. (1982) also add a metacognitive component, arguing that “a technique becomes a strategy only if students have the metacognitive knowledge of when, where, and how to use it” (p. 18). Likewise, Paris, Lipson and Wixson (1983) believe that “strategic behavior connotes intentionality and purpose on the part of the learner” (p. 294).

This suggests that the difference between a *skill* and a *strategy* can be seen as one of awareness. Thus, while note-taking may be viewed as a learned *skill*, it can also be viewed as a *strategy* chosen by a reader to fulfil an individual reading purpose. In other words, note-taking might be regarded as a skill if conducted subconsciously but as a strategy if selected consciously from among other possibilities. A difficulty with this interpretation, however, is that there is no consensus in the literature as to whether a strategy is conscious or subconscious (though they can be researched more easily if they are conscious). This is one of many definitional problems that is addressed in greater detail later in this chapter.

There is a further reason for the difficulty of defining learning strategies. If they are to be regarded as a set of related processes (e.g., Dansereau 1985, Levin 1986, Garner 1987, Wade et al. 1990), then there may need to be umbrella terms for the overall

process and for any subprocesses that make up the whole. Several researchers have attempted to do this by differentiating between learning *strategies* and learning *tactics* (e.g., Armbruster et al. 1982, Derry and Murphy 1986, Wade et al. 1990). Wade et al.'s definition of strategies uses the related term *tactics*: “a strategy is defined here as a configuration of study tactics used together in a purposeful way to accomplish a particular learning task” (p. 147). In contrast, “a learning tactic is an individual study technique such as underlining, note-taking, outlining, summarizing, visualizing, or using mnemonic devices” (p. 149).

The term ‘tactic’ is not used as extensively in the literature as ‘strategy’, and Wade et al.'s examples of tactics appear as strategies in many studies (e.g., ‘visualizing’ is a tactic for Wade et al. 1990 but a strategy for Kletzien 1991). Moreover, the overlapping use of definitional terminology is evident here, where Wade et al. use the term ‘technique’ to define tactics whereas Wenden (1987, see above) uses it to define strategies.

Cohen (1987:92) notes that some researchers prefer to use the term ‘strategy’ to describe “more general categories of behaviour” such as ‘monitoring’ and terms such as ‘technique’ or ‘tactic’ to describe what he calls “lower-level activities” such as ‘checking that nouns and adjectives agree in gender and number’. Cohen chooses to use the term ‘strategy’ for both the levels of behaviour he describes.

Setting terminology aside, Wade et al.'s notion of a strategy being the sum of many parts is an interesting one, and they are explicit about this notion:

“A strategy consists of more than simply the use of a tactic. Rather, a strategy is a configuration of different tactics, deliberately selected for a particular purpose, and carefully monitored for effectiveness. Thus, readers’ strategies may vary in the diversity of tactics used, in their purposefulness, and in the consistency or flexibility of their use.” (p. 150)

(The notion that strategies may be multi-faceted, little explored in the literature on reading, is taken up in Chapters 5 and 6.)



### **3.3 Section summary**

This section has sought to add substance to the partial answer, given in the preceding section, to Research Question 2.1. The research question asks about the context surrounding research into strategies and this section has focussed on the considerable definitional and conceptual issues surrounding learning strategies, including such concerns as whether they describe behaviours that are necessarily conscious and problem-oriented. There was also a discussion of additional terms such as ‘styles’, ‘skills’ and ‘tactics’, which have been used to describe similar behaviours. One of the main areas where there is a lack of consensus is the precise scope of each term and where lines can be drawn between one level of behaviour and the next. Comparison of findings in different studies is therefore very difficult. It will be interesting in the discussion that follows to see the extent to which the literature on reading strategies, which largely evolved from that on learning strategies, reveals similar problems to those just outlined. In the following two sections, the discussion will cover difficulties with defining reading strategies and categorizing them.

## **4.0 READING STRATEGIES: THE PROBLEM OF DEFINITION**

This section seeks to answer Research Question 2.2, which inquires about the nature of reading strategies. While the term ‘reading strategies’ may represent a narrower range of behaviour than that of ‘learning strategies’, the definitional difficulties seem to proliferate. Those already mentioned remain, and there are additional areas of disagreement. The first part of the section discusses issues surrounding this lack of consensus. This is followed by a discussion of how some of the differences among the definitions may be the consequence of operational constraints.

### **4.1 Lack of consensus**

Barnett (1989) refers to a “lack of consensus” about what constitutes a reading strategy (p. 70). Casanave (1988) feels that “no research yet exists that has definitively identified what reading strategies are” (p. 285). Davies (1995) says that, across the different studies she reviewed, “the kinds of behaviour classified as strategies, at least initially, appear to be almost unlimited” (p. 49).

In a similar way, Paris et al. (1991) note that there is “no consensus” on a definition of reading strategies, and refer to three persisting problems. First, they feel that it is not clear how to differentiate reading strategies from “other processes that might be called thinking, reasoning, perceptual, study or motivational strategies” (p. 610). Second, they note the absence of agreement on the scope of strategies in the literature. They cite Levin (1986), who argues that strategies include multiple components, and Derry and Murphy (1986), who argue that they are general learning plans implemented through tactics. Paris et al. justifiably conclude that “strategies are difficult to demarcate when they are embedded in complex sequences of behavior or hierarchies of decisions” (p. 610).

Paris et al.'s (1991) third reason for the identity problem of reading strategies is the issue of consciousness, and they cite contrasting definitions taken from the L1 literature, definitions that are, incidentally, echoed in the L2 literature. For instance, Cohen (1986) writes, “reading strategies refers to those mental processes which



readers consciously choose to use in accomplishing reading tasks” (p. 133). Pritchard (1990) agrees, “a strategy is defined as a deliberate action that readers take voluntarily to develop an understanding of what they read” (p. 275).

Several definitions suggest consciousness but stop short of a full commitment to it. For example, Olshavsky (1976-77) defines reading strategies as “purposeful means of comprehending the author's message” (p. 656). Likewise, in a conference paper cited in Pritchard (1990:275), Rowe and Rayford (1984) define them as “purposeful actions taken voluntarily to achieve particular outcomes”. Similarly, Brown (1980) writes that strategies are “any deliberate, planful control of activities that gives birth to comprehension” (pp. 64-65). Others, however, leave the question of consciousness open. Barnett (1989) writes “the term strategy refers to the mental operations involved when readers purposefully approach a text to make sense of what they read. These may be either conscious techniques controlled by the reader or unconscious processes applied automatically” (p. 66).

## **4.2 Operational constraints**

There may be a straightforward explanation for the definitional conflict over consciousness. Those who define strategies as either conscious or unconscious conceive of them taking place in both normal reading and in reading interrupted in some way by a researcher conducting a study. In contrast, those who define strategies as necessarily conscious may be working from the standpoint that strategies can be tapped by the researcher through mentalistic data-gathering techniques. In other words, researchability may be the issue at the base of this conflict.

Indeed, there is plenty of evidence that definitions are often operationally constrained. At one extreme, there are very broad definitions, such as that of Lysynchuk et al. (1989): “any steps or actions that students could take to enhance comprehension” (p. 458). This was probably deliberately designed to be very wide as it served as the basis for a literature review of studies investigating strategy instruction.

At the other extreme, there are narrow definitions; for example, Kohonen (1992:24) says that strategies are problem-oriented in that learners use them to respond to a learning need. This notion of problem-orientation is further narrowed by other researchers. In a study looking at the syntactic and semantic parameters of cloze test responses, Clarke and Burdell (1977) define reading strategies as “a subject's use of context in responding to a mutilated text” (p. 133). Kletzien (1991), also using cloze, defines reading strategies as “deliberate means of constructing meaning from a text when comprehension is interrupted” (p. 69).

Clearly, when reading is interrupted, whether due to researcher intervention or a break in the interaction between reader and text, it is likely to make more salient any strategies deployed to deal with the break, and hence more researchable, than those used during uninterrupted reading. Nonetheless, not all definitions preclude on-going reading. For example, Palincsar et al. (1984) include both on-going and interrupted reading when they say that reading strategies are “processes for enhancing comprehension and overcoming comprehension failures” (p. 118). (The issue of problem-orientation is further discussed later in this thesis.)

More recently, Davies's (1995) definition, while not pursuing the idea that interruption needs to take place for there to be a strategy, adds two important elements to the debate. She writes that “a strategy is a physical or mental action used consciously or unconsciously with the intention of facilitating text comprehension and/or learning” (p. 50). Here, there is a recognition that a reading strategy may involve an observable physical action, such as highlighting or note-taking. (The issue of observability is pursued later in this thesis.) In addition, there is the important point that a reader's purpose may extend beyond comprehension to learning, thus underlining the close link between learning strategies (discussed above) and reading strategies. Indeed, perhaps to differentiate between comprehension and learning, some researchers investigating reading comprehension have opted for the term ‘comprehension strategy’ rather than ‘reading strategy’ (e.g., Sarig 1987a:110, Kern 1989:145).

Sarig (1987b) went even further and opted to use the term ‘comprehension move’ in her study; that is, she opted not to use the term ‘strategy’ in her study comparing L1 and L2 data. Working with ten female high school students aged 17 and 18, she



defined a move as “each separate action the reader took while processing the text” (p. 110). When analyzing her data (yielded by tasks requiring “main ideas analysis” and “synthesis of overall message” by subjects introspecting orally), she evaluated the effect of each reported move on the overall success in performing the reading tasks in her study. She then divided the moves into two categories: ‘comprehension promoting’ and ‘comprehension deterring’. Thus, Sarig’s concept of a reading move incorporates the notion of ‘good’ and ‘bad’ strategies but with a crucial difference. She did not regard strategies as good or bad in themselves but concerned herself with whether they yielded a successful or unsuccessful outcome in a given context. This contrasts with the earlier studies, cited above, of both learning strategies and reading strategies (e.g., respectively, Rubin 1975 and Hosenfeld 1977).

### **4.3 Section summary**

In answer to Research Question 2.2, which can be found at the beginning of the chapter, this section described the absence of consensus about what constitutes a reading strategy. It was suggested that one reason for the multiplicity of definitions is that many of them have been tailored for the purposes of individual research studies. In order to avoid the minefield of differing and often conflicting definitions, some researchers have opted to avoid using the term ‘strategy’ and have used terms of their own. Reflecting the previously mentioned paradigm shift in work on learning strategies, there was a similar shift in the reading strategies literature. This involved a move away from a focus on the strategies used by ‘successful learners’ towards a focus on whether strategies have ‘successful outcomes’.

## 5.0 STRATEGIES: THE PROBLEM OF CLASSIFICATION

The two previous sections have shown that the establishment of a consensus on the precise identity of a learning strategy and a reading strategy has proved elusive. It follows that it is no less hard to classify them, the issue raised in Research Question 2.3 at the beginning of this chapter. It is perhaps testament to this difficulty that many researchers categorize the strategies they find in their data *post hoc*. As will be shown, there is also the potential for confusion among researchers between the higher level of strategy *categories* developed from a data set and the lower level of strategy *descriptors*, the latter being the terminology used to describe individual strategies within a category. In addition, it is widely acknowledged that, since both the categories and the descriptors are often data-driven and based on small samples, they cannot possibly amount to a complete representation of the strategies readers use; the issue of completeness was raised in Research Question 2.4.

Another reason for the lack of consensus is that the study of reading encompasses more than one research tradition, the most pertinent being those of cognitive psychology and applied linguistics. The complex theoretical background to the study of reading is reflected in the way in which reading strategies have been classified and described in the literature. Cognitive psychologists tend to categorize both learning and reading strategies in terms of awareness of information processing while applied linguists, focussing more on reading, do so in terms of text features. These two ways of looking at the classification of strategies are now discussed, starting with cognitive psychology and moving on to applied linguistics. The discussion finishes with a review of some of the work that blends these two traditions.

### 5.1 The cognitive psychology perspective

Working from a cognitive psychology perspective, Dansereau (1985:210, 219) divided learning strategies into 'primary' and 'support' strategies. The former have "a direct impact on the target information" and include strategies for acquiring and storing information (or 'comprehension/retention strategies') as well as those for subsequently retrieving and using this stored information (or 'retrieval/utilization



strategies'). In contrast, the latter have "an indirect impact [on the target information] by generally improving the level of the learner's cognitive function"; these are divided into three subcategories, 'planning and scheduling', 'concentration management' and 'monitoring and diagnosing'.

Weinstein and Underwood (1985:241) group strategies into four types of behaviour, of which the second, third and fourth categories are similar to Dansereau's 'support' category with its three subcategories:

- 'cognitive information-processing' (e.g., "techniques for organizing and elaborating on incoming information to make it more meaningful")
- 'active study' (e.g., note-taking)
- 'support' (e.g., techniques for organizing study time)
- 'metacognitive' (i.e., monitoring)

Rubin (1987) identifies two "major" kinds of learning strategies: 'metacognitive' and 'cognitive' (p. 23). Metacognitive strategies are those that oversee and regulate or "self-direct" learning. There are both L1 researchers (e.g., Brown and Palincsar 1982) and L2 researchers (e.g., O'Malley and Chamot 1990) who believe that the use of metacognitive strategies can bring about more successful learning. Moreover, as was observed above, they have advocated 'strategy training' designed to raise learners' awareness and promote the use of strategies. Rubin's cognitive strategies, on the other hand, act directly on learning and refer to the steps or operations a learner uses in learning or problem-solving.

O'Malley and Chamot (1990:137-139) identify three categories of learning strategies in their data. Their L2 learners undertook four tasks (listening, reading, writing and cloze) and were told to think aloud as they did each one. O'Malley and Chamot's categories were similar to Rubin's: 'metacognitive', 'cognitive' and 'social and affective' strategies. The latter involved learners interacting with their peers "to assist learning or using affective control to assist a learning task" (p. 139).

Against the prevailing trend, Oxford's (1990) work is not directly data-driven; rather, she draws together categories that have evolved in the literature and then develops her

own framework. She has two main classes of strategies, ‘direct’ and ‘indirect’, each with three subclasses of strategies (pp. 16-21):

Direct strategies	Indirect strategies
Memory	Metacognitive
Cognitive	Affective
Compensation	Social

Each of these subclasses is then divided further into two more subsets. The strategies sometimes become specifically related to particular language skills in the first of these two subsets; for example, ‘compensation strategies’ are split into ‘guessing intelligently’ and ‘overcoming limitations in speaking and writing’.

Furthermore, Oxford uses this overall classification scheme as a framework for ‘strategy applications’ in the four language skills, listening, reading, speaking and writing. While the idea of pooling strategies in this way may be initially appealing, it has to be said that a substantial amount of work would have to be done to see if it is tenable. Many of the strategies Oxford lists are very general, even vague; for instance, the only two reading strategies listed under the rubric ‘compensation’ are ‘using linguistic clues’ and ‘using other clues’ (p. 322). Others, despite the vast number of definitions of strategies in the literature, might not qualify as strategies at all; for instance, two of Oxford's reading strategies listed under the rubric ‘affective’ are ‘rewarding yourself’ and ‘listening to your body’ (p. 323).

Wade et al. (1990) studied both learning and reading *tactics*, which they defined as ‘individual study techniques’ (p. 149); however, since *strategy* is the more usual term used in the literature, it will be used here to report on this study. Wade et al.'s subjects were L1 undergraduate adults in the age range 17-38. Subjects read a text that had eight marked points where they were asked to stop reading. At each marked point, they were interviewed about the ‘study methods’ they used to understand and remember important information in the text.

Wade et al.'s data yielded three main categories of strategies (each with four or five strategies listed under it). The main categories were ‘text-noting’, ‘mental-learning’, and ‘reading strategies’. (The nomenclature is confusing since the first two of these



categories would in themselves be termed ‘reading strategies’ by many researchers.) Wade et al.’s first main category is distinct in that the strategies are observable as they “produce artefacts” (p. 153) (cf. Davies’s 1995 definition of strategies, which includes observable behaviour and is cited above). The second category involves cognitive processes used to learn and store information while the third category covers adjustments to on-going reading, such as skipping, rereading and changes in reading speed.

There are a number of difficulties with taking an exclusively cognitive psychology perspective when investigating L2 reading strategies. An obvious one is that it does not really address the pertinent issues of language development and deficit among L2 readers (e.g., Bernhardt 1991). Further, it does not concern itself with text-related issues, such as what caused an interruption to the flow of reading or which part of the text was reread or skipped. Account can be taken of issues such as these when including an applied linguistics perspective in the categorizing of reading strategies.

## **5.2 The applied linguistics perspective**

For about a decade, there appeared to be a consensus among those working from a broadly applied linguistics perspective in favour of two categories, essentially reflecting what could loosely be described as ‘global’ and ‘local’ reading. This may have been a reflection of the earlier reading theory, discussed in Chapter 1, that reading involved either top-down processing (with readers leaping through text) or bottom-up (with readers plodding through text). Of the two, bottom-up processing clearly has the greater potential to be amenable to description in terms of particular linguistic features in the text. More recently, however, as will be shown below, there has been a trend away from this narrow, two-category, approach towards one that spawns four or five categories. Again, this may be a reflection of the shift in prevailing theory, in this case, a shift away from either a top-down or bottom-up view towards the more complex view that reading is interactive or bottom-up interactive.

Olshavsky (1976-77:665-666) developed two categories, ‘word-related’ and ‘clause-related’ strategies. These categories, which were developed after data collection, have

extremely limited scope in that they only account for very local reading. Further, they are clearly an artefact of the task, which involved L1 teenage readers ‘thinking aloud’ after reading each clause of a text. This task almost certainly distorted the reading process since it must have entailed atypically frequent interruption to the reading flow. As a result, some of the strategies placed under each category may reflect what readers did for this particular task more than what they would do in a more normal reading situation. For example, by far the most frequently reported strategy in Olshavsky's data was ‘synonym substitution’; intuitively, this would not seem to occur with such relative frequency in normal reading. (For a comparison of Olshavsky’s think-aloud task with similar tasks used in comparable studies, see Chapter 3.)

At about the same time, Hosenfeld (1977) conducted a pioneering study into L2 reading strategies in which she developed two data-driven categories. The first category was what she termed the reader’s ‘main meaning line’, or meaning retention, as revealed in think-aloud tasks. The second was ‘word-solving’ strategies, or word decoding, that is, what the reader does “when his [or her] main meaning line is interrupted” (p. 111). Each category comprised about ten ‘subsets’ of symbols used to describe her subjects’ strategies, and these were assigned to one or other category. Although both of Hosenfeld’s categories allowed for cognitive processes, the second presupposes very local, word-level, reading.

Block (1986b) developed two categories that were rather similar to Hosenfeld’s: ‘general comprehension’ and ‘local linguistic’. While the first described comprehension and included such mental processes as predicting, inferencing and using prior knowledge, it also included such text-specific strategies as recognizing text structure. (Block’s combining of extra-textual and textual factors into a single category is interesting, and I will make further reference to this feature of category-building in discussion of Sarig’s 1987b work below as well as in Chapters 5 and 6.) The second category focussed on the processing of “specific linguistic units” (p. 473).

Barnett (1988a:150) refers to the “binary division” of categories in the literature, and she adheres to it herself with her ‘text-level’ and ‘word-level’ strategies. This duality in the formation of categorization schemes closely follows that of several L1 and L2 reading researchers (including, in L2, Hosenfeld 1977, Barnard et al. 1980, Block



1986b, and, in L1, Olshavsky 1976-77, Fisher and Smith 1977). The 'text-level' category incorporates strategies related to the passage as a whole "or to large parts of the passage", including the use of background knowledge, titles and illustrations, skimming and scanning. The 'word-level' category, on the other hand, includes guessing word meanings and identifying the grammatical category of words.

A problem with taking a perspective on reading strategies that is mainly linguistic or text-based is that it may create artificial constraints that tend to portray reading processes as being either global or extremely local, the latter being exemplified in both of Olshavsky's categories and in Hosenfeld's 'word-solving' category. It seems to me that categories such as these would be unable to capture the rich complexity of mental processes readers actually use. Such a polarized perspective may ignore mental processes that fall between global and local reading.

### **5.3 A blending of perspectives**

So far in this discussion, I have suggested that the development of categories of reading strategies that ignore, or manifest an imbalance between, the contributions of either cognitive psychology or applied linguistics can serve to dilute the potential richness to be derived from readers' reports of their behaviour. I will now turn to a review of studies, most of them relatively recent, which incorporate into their categorization schemes greater equilibrium in their accounting of mental and physical behaviour as well as text features.

Sarig's (1987a and 1987b) work, being considerably more complex and detailed than the work cited so far in this discussion of reading strategies, in my view represents a watershed in the literature. Her classification of strategies, or 'moves' as she prefers to call them, goes a long way beyond the binary view noted and espoused by Barnett.

From her data, Sarig (1987b) developed four categories of strategies and thirty-five strategies. The first, 'technical-aid' strategies, are those that a reader uses "to facilitate text processing" (p. 111), and they include highlighting, writing summaries in the margin, skipping, scanning, and using a glossary. Using classroom observation

of each of her subjects in turn, Sarig appears to have noted their physical actions, actions that are similar to those in Wade et al.'s. (1990) 'reading strategy' category (described above). Sarig's second category are 'clarification and simplification' strategies in which readers utilize their linguistic competence to raise the level of linguistic and textual redundancy of the text; this is achieved by various types of paraphrase, such as the use of synonyms or circumlocution. 'Coherence-detection' strategies involve the utilization of macro-textual and extra-textual knowledge for the construction of a coherent meaning of the text. Finally, readers use 'monitoring' strategies to plan task performance and subsequently to monitor their processing for error detection.

Sarig's first category represents an addition to the literature on L2 reading strategies. Though physical actions had previously been accounted for at the lower level of strategy descriptor (in other words, as individual strategies within a wider category, e.g., Block 1986b), they had not been raised to the level of category. However, this had already occurred in the literature on learning strategies, for instance, Weinstein and Underwood's 1985 'active study' category, cited above. (As discussed in Chapter 5, Sarig's separate categorization for observable actions was to have an influence on the development of my own categories.)

Sarig's second and third categories to some extent reflect, respectively, the local and global categories used in previous binary classifications (e.g., Hosenfeld 1977, Barnett 1988a). Nevertheless, with her four categories, Sarig has clearly ventured beyond Barnett's "binary division", but it is worth noting that she did not create a fifth category for extra-text processing; rather, this is assigned to the same category as text processing (i.e., 'coherence-detection'). Though she doesn't mention her reasons for this decision, I see it as a principled stance as it seems to be in line with current reading theory. As discussed in Chapter 1, the most recent influential model suggests that reading is a bottom-up or bottom-up interactive process (Rayner and Pollatsek 1989). Hence, since extra-textual processing, arguably, does not take place in isolation, it may be preferable to categorize such top-down processing in combination with that at lower levels. (This notion of binding readers' processing beyond the text with that within the text was to be reflected in my own categorization scheme in Chapter 5.) Finally, Sarig's fourth category shares qualities of strategies categorized



from a cognitive psychological perspective (e.g., Weinstein and Underwood 1985, O'Malley and Chamot 1990).

In a study looking at the effects of cultural schemata on reading strategies, Pritchard (1990) asked sixty L1 and L2 subjects (half in each language group) to 'think aloud' after reading each sentence in two passages. The subjects worked individually in a room with the researcher in the background. From the data yielded in this way, Pritchard created five data-driven categories and a taxonomy of twenty-two strategies. His first category, 'developing awareness', covers monitoring of progress and problems encountered. Related to this is 'accepting ambiguity', a category in which readers realize they have not understood a point in the text and utilize a strategy that might overcome the problem. Two related strategies are 'establishing intrasentential ties' and 'establishing intersentential ties'; again, these reflect the original binary categories. The final category is 'using background knowledge'. Thus, while Pritchard's coverage is similar to Sarig's, he has split 'using intersentential ties' and 'using background knowledge' into two categories whereas Sarig uses one.

Against the norm, Kletzien (1991) gathered her verbal report data retrospectively rather than concurrently. Individual subjects were asked to complete three cloze tasks, each at a different level of difficulty, and report after each one. Also against the norm, Kletzien did not develop a classification scheme from her data; instead, based on work originally developed for a pilot study with L1 readers in 1986, she worked solely at the lower level of strategy descriptor.

When beginning analysis of her 1991 data, she included all the comprehension strategies she could find in the literature on reading strategies. She then listed the strategies she found in the twenty-one studies she reviewed and added the strategies reported by her 1986 pilot subjects that did not seem to be accounted for in the literature. At this point, she had a list of fifteen strategies (including one called 'non-use of strategies'). During the course of her analysis, Kletzien dropped several of the strategies developed in the 1986 pilot study. In the case of those trawled from the literature, in order to "provide for an economical and accurate statistical evaluation", she combined them with other strategies on the bases that they were either not reported at all or were "not often" reported (p. 77). Further, she combined three

strategies dealing with text, paragraph, and sentence structure. She ended up with these seven strategies:

- rereading preceding text
- reading subsequent text
- recognizing structure
- using prior knowledge
- using main idea
- making inferences
- focussing on vocabulary

Although, as has been noted, she did not group her strategies into categories, it is worth noting that Kletzien's strategy 'using prior knowledge' accounts for the same reading action as Pritchard's overall category 'using background knowledge', which had a subset of six strategies. In other words, there appears to be an enormous difference in the level of detail recorded in strategies across the Kletzien and Pritchard studies. An alternative interpretation of this situation might be that Kletzien's seven strategies record processes occurring at the same level as Pritchard's five strategy categories; that is, Pritchard has provided an additional level of detail to the strategies that fall under his categories.

Drawn from an extensive literature review of both L1 and L2 reading studies, Davies (1995:51-54) proposes "a broadly based categorization" of five different types of strategies. Her first category, 'control reading process', is observable; and this echoes Sarig's (1987a and 1987b) 'technical-aid' category and Wade et al.'s (1990) 'text-noting' category. The other four categories are non-observable: 'monitor reading process', 'interact with text', 'utilize source of information: textual', and 'utilize source of information: external/background knowledge'. 'Monitoring' involves readers consciously attempting to find ways to comprehend the text and evaluating such attempts. 'Interacting' represents responses to the text by "questioning, expressing feelings etc on the basis of at least an interim interpretation of the text" (p. 51). The category targeting text sources of information covers readers' attention to linguistic features of the text and/or their own textual/linguistic knowledge as evoked



by reading the text. The category covering non-text sources of information accounts for readers' own knowledge as evoked indirectly by reading the text.

Having formed her five categories, Davies reviewed nine recent studies of L1 and L2 reading strategies for 'strategy labels', which she describes as "broad summary labels based upon my own interpretation of different researchers' varying labels" (p. 55). (In the current study, these strategy labels, which are at a lower level than strategy categories, are referred to as 'strategy descriptors'.) Davies then placed her strategy categories, along with their respective strategy labels, on a grid (pp. 52-54). For example, under her category 'control reading process', she listed such strategy labels as 'adjust rate of reading', 'pause', 'regress/reread', 'go forward in text', 'skip' and 'write notes or summary'.

Davies's grid is very useful in that it reveals contrasting findings about the higher level of strategy category and the lower level of 'strategy label', or strategy descriptor. At the higher level, all nine studies reviewed yielded data that could be placed under each of Davies's five categories while, at the lower level, there was considerable variation among the studies. For example, while six of the nine studies recorded subjects 'regress/reread', only one recorded 'pause' and 'write notes or summary'. This sort of variation may be due to more factors, or combinations of factors, than can be listed here. However, they might include the level of detail recorded, task type, task difficulty, actual differences across various populations reading various genres, and, in the case of L2 readers, language proficiency level. Thus, Davies's grid shows what appears to be a consensus about what reading broadly entails. However, in answer to Research Question 2.4 about the completeness of our knowledge, it is obvious that we still have a long way to go before we attain a clearer idea of the detail involved or verify the considerable differences shown on the grid.

#### **5.4 Section summary**

This section has attempted to answer several of the research questions posed at the beginning of this chapter. In partial answer to Research Question 2.1 about the context of research done on strategies, it has shown how early work by cognitive

psychologists (mostly on learning strategies) and early work by applied linguists (on reading strategies) has converged in more recent research. Addressing Research Question 2.3 about how strategies have been classified, this convergence has rendered relatively complex sets of categories that combine description of mental actions reported and physical actions reported or observed, previously the domain of cognitive psychologists, with that of text features, previously the domain of applied linguists. In general, even when researchers develop categories that reflect current reading theory, they make no direct reference to it.

While this section has focussed on the level of strategy category, during the course of the discussion, reference was made to the lower and more detailed level of strategy descriptor. Depending on such factors as operational constraints and what behaviours particular readers happen to report, different studies have reported different strategy descriptors. It seems clear, therefore, that at this lower level, we still only have a limited idea of the range, variability and richness of reader behaviour. Thus, in answer to Research Question 2.4, our collective description of reading behaviour remains partial; nonetheless, investigating and describing it as ‘strategies’ seems like a promising avenue to explore.



## 6.0 CHAPTER SUMMARY

This chapter posed and sought to answer the following research questions:

- 2.1 What is the context behind research into reading strategies?
- 2.2 What are reading strategies and how have they been defined?
- 2.3 How have they been classified?
- 2.4 Can they provide at least a partial description of reading behaviour?

This chapter started with a historical outline of learning strategies. Research began in the 1970s with attempts to discover what ‘successful’ learners do and then train their ‘unsuccessful’ counterparts in the use of these strategies. In the study of reading strategies, this approach was refined with attempts to discover which strategies led to successful and unsuccessful outcomes in particular contexts.

The chapter then outlined the difficulty of establishing a generally recognized definition of both learning and reading strategies. Further, since a definition has proved elusive, I sought to show that attempts to classify strategies have proved at least as fraught with problems. Typically, definitions have been framed operationally, and classifications have been data-driven, though almost invariably based on data yielded from only one task.

Nevertheless, the research area is far from static and, in recent years, there has been some development in how strategies are categorized. While earlier classifications tended to be devised separately in broadly cognitive psychology and applied linguistics traditions, there has more recently been an increasing recognition, albeit implicit, that these two fields of study, taken together, offer a better foundation for describing the mental processes that take place in reading. However, L2 reading is still a very young field of study, and we only have a partial view of L2 reading behaviour. In order to deepen that view, it seems, at the present time, that there is plenty of scope for further work to be done in the area of reading strategies.

There are a number of ways in which this process of deepening our view of L2 reading through the study of strategies might be conducted. First, based on a clear but

broadly inclusive definition of what constitute reading strategies, there is a need to devise a classification scheme that is not simply based on the data from one study. Such a scheme might be developed from data yielded in two or more studies, each using a different data-gathering technique. The scheme should also either be consistent with current reading theory or provide grounds for theoretical modification. In addition, in order to add detail to our existing knowledge of the reading process in such areas as variation across genres and individuals, extensive work needs to be done at the lower level of strategy descriptor.

In view of their potential for this kind of research, the following chapter discusses mentalistic approaches to the collection of data.



## CHAPTER 3: MENTALISTIC APPROACHES

### 1.0 INTRODUCTION

In Chapter 1, several strands in the area of L2 reading research were identified, and one of these, that of reading strategies, was selected for review in Chapter 2. From this review, it was clear that this was potentially a fruitful way to study reading behaviour. In this chapter, I will discuss introspective research methods that have been used to obtain data from which reading strategies can be identified. Nunan (1992) has defined introspection as “the process of observing and reflecting on one’s thoughts, feelings, motives, reasoning processes, and mental states with a view to determining the ways in which these processes and states determine our behaviour” (p. 115).

As I will show, this is not an easy route for a researcher to take. However, as Eskey (1987) has made clear, L2 reading researchers, depending on the focus of individual studies, need to be very flexible in their approach to methods used:

“Given the obvious complexity of the conception of reading as an interactive process involving many kinds of reader variables (e.g., background knowledge, processing strategies) and text variables (e.g., cultural assumptions, linguistic density), it should be obvious that researchers will have to continue to employ a wide range of approaches, ranging from statistical analyses of texts to such person-oriented techniques as interviews, introspections, and think-aloud protocols if we are ever to develop a fuller understanding of the process as a whole.” (p. 189)

While Eskey’s advocacy of such methods as introspection and think-aloud protocols may seem eminently sensible today, it is only with the resurgence in the use of mentalistic data in the past two decades that their use in the social sciences, along with statistical analyses, has become generally accepted.

Conceptually, it is possible to classify much research into reading so as to reflect two distinct research paradigms. These paradigms can be described as analytic/deductive and heuristic/inductive (e.g., Seliger and Shohamy 1989). Broadly, the first is theory-driven and is used to seek causal relationships through the collection of empirical

evidence while the second is data-driven and is used to seek to understand processes, often through informants' self-report data. In the real world, where a substantial amount of research now being conducted, including the present study, uses aspects of both research traditions, these paradigms should be seen as extremes on a continuum. Seliger and Shohamy (1989) have used the term 'descriptive' research for studies that combine elements of experimental and qualitative research paradigms. Historically, however, the level of pragmatism apparent today has not always been accepted.

Reading researchers following the analytic/deductive tradition have investigated such directly observable aspects of reading as eye movements (e.g., Buswell 1922, Tinker 1939, Rayner and Bertera 1979, McConkie and Zola 1981, and some of the studies reported in Rayner and Pollatsek 1989). In Rayner and Bertera (1979), for example, subjects read from a computer screen which was programmed to mask a letter in foveal vision.

However, no such directly observable evidence is available to researchers interested in tapping the mental processes involved in reading (e.g., Cavalcanti 1987, Kletzien 1991, Davis and Bistodeau 1993, Jimenez 1997). Mentalistic data can only be obtained by having subjects report what is going on in their minds concurrently with, or subsequent to, their reading. Any observable evidence that may be used to supplement introspective data (e.g., eye movements, note-taking, text highlighting, text annotating) can only provide indirect confirmation of the mental processes that a reader might report. Nonetheless, in cases where the observable data are themselves valid and, in the same study, yield similar findings to the mentalistic data, it could be claimed that the mentalistic data set has convergent validity. (This issue is addressed in greater detail later in this chapter.)

Clearly, then, there are a number of problems involved in the collection of introspective data from readers; however, in order to provide this chapter with a manageable number of foci, I have developed the following four research questions:

- 3.1 Is there any theoretical support for mentalistic approaches?
- 3.2 When designing studies yielding mentalistic data, what practical concerns does the researcher need to consider?



- 3.3 How has introspective methodology been used to investigate reading strategies?
- 3.4 What are the advantages and disadvantages of various data-gathering techniques?

In order to answer these research questions, Sections 2.0 and 3.0 of this chapter contain literature reviews of certain theoretical and practical issues that researchers contemplating the use of mentalistic data need to consider. Section 4.0 provides a review of the methodologies that have been used to conduct research into L1 and L2 reading strategies and a discussion of the pros and cons of mode of introspection and of a number of data-gathering techniques.

## **2.0 THEORETICAL CONSIDERATIONS**

Historically, there have been serious doubts about the use of introspective data in any field of study. These doubts have been prevalent in two fields relevant to the study of reading, psychology and linguistics (including applied linguistics), where theoretical developments have had an enormous impact on what data-collection techniques are regarded as acceptable. The discussion that follows traces developments in theory, first in psychology and then in linguistics.

### **2.1 Impact of prevailing theories in psychology**

In psychology, the debate about the validity and reliability of mentalistic data has provided one of the main areas of controversy between behaviourists and cognitivists.

#### **Behaviourism**

Behaviourist psychology, looking at stimulus-response relationships, is relatively restricted to the study of measurable human performance and abilities; in contrast, cognitive psychology more ambitiously seeks to understand the detailed steps in series of related mental processes and utilizes introspective data. During the period in which behaviourism dominated the field, doubts about mentalistic data reached such a peak that misgivings resulted in a virtual halt to the use of such data. This left psychology bereft of data-collection techniques other than those used in the hard sciences, thus effectively turning the field, at least temporarily, into a hard science. However, before, and especially after, the era of behaviourism, introspective methodology in data-collection was widely accepted and utilized.

Ericsson and Simon (1987) state that verbal reports were “a mainstay of classical introspection” (e.g., James 1890, Titchener 1912) (p. 24). Introspective methodology was used by psychologists analyzing mental processes in problem-solving (e.g., Selz 1913, 1922) and by those conducting clinical analyses of thought (e.g., Freud 1914) and free-association tasks (e.g., Jung 1910). In related work, but specific to the early



study of reading, Huey (1908/1968) asked his subjects to introspect on word-level tasks. Henderson (1903) had his subjects recall short narrative and expository text in written protocols and Bartlett (1932) conducted a similar study but with a longitudinal element.

A central problem with some of the early work, as Ericsson and Simon (1987) observe, is that the investigators often asked their subjects to report specific types of information “without any regard for how cognitive processes generating such reports were feasible” (p. 25). As a consequence of this lack of rigour, findings from the early studies were often conflicting, and it was very easy for opponents of introspective methodologies to argue that they were totally unscientific (e.g., Watson 1920).

However, as Cohen (1987) says, in more recent times, mentalistic approaches have had an increasing amount of support from “a respectable group of cognitive psychologists” (p. 88). He cites, among others, Bakan (1954), Radford (1974), Lieberman (1979) and Ericsson and Simon (1980).

In a defining article about verbal report data, Ericsson and Simon (1980) state that behaviourism, with its prerequisite of the direct observability and measurability of data, was “schizophrenic” about the collection, analysis and interpretation of introspective report data (p. 216). However, in Ericsson and Simon’s view, this was an untenable position for several reasons, not least since subjects’ key punches (or card punches), so often used as behaviourist data in standard experimental research paradigms, are “psychologically indistinguishable from verbal reports” (p. 216). However, Ericsson and Simon (1980) criticize researchers using introspection for a tendency not to report fully on their procedures, making it impossible either to replicate these studies or otherwise substantiate any claims made.

### **Ericsson and Simon’s (1980) information-processing model**

Ericsson and Simon (1980) propose a model based on earlier work on information-processing in human cognition (e.g., Newell and Simon 1972). The model shows how

subjects, when asked to think aloud, verbalize the information they are attending to in short-term memory. An essential assumption in the theoretical framework of this work is that information is stored in several memories, each having different capacities and accessing characteristics. In addition to several sensory stores of very short duration (e.g., iconic and echoic) are short-term memory (STM) and long-term memory (LTM). While STM is hypothesized to have a limited capacity and to be of fairly short duration, it is relatively easy to access. Simon (1979) found that about 8-10 seconds are all that is required to assemble each chunk of new information in STM and store it as a new chunk in LTM. In contrast, LTM is seen to have an apparently infinite capacity and to allow relatively long-term storage, but it is relatively hard to access.

The diverse characteristics of STM and LTM, as posited in Ericsson and Simon's model, appear to be highly important in the understanding of both how mentalistic data are tapped and how valid they may be. Since retrieval of information in LTM is seen to be relatively difficult, introspecting on information stored therein is thought to require a higher cognitive load. It follows that, for a reader, whose attention is already given, at least in part, to interacting with a text, any effort needed to dip into LTM might impact on the normal reading process. Such a resort to LTM might take place where an informant is attempting to satisfy a researcher's probing for introspective report data, and it could be argued that this effort might have an impact on the very reading process the researcher is trying to study. On the other hand, since information still in STM is regarded as being relatively easy to retrieve, any resort to STM for the purpose of reporting mental processes would be less likely to distort the normal reading process and, being concurrent with the process, would not interrupt it in the same way.

Nonetheless, it seems to me that one could equally well argue a different case about the use of LTM in introspection. Since readers routinely resort to information stored in LTM, often referred to as the use of 'non-text information' (e.g., in order to access prior knowledge of text content), some reference to LTM may not interrupt or interfere with the quality of normal reading as much as some may fear. In fact, it may simply represent the utilization of a processing route that is frequently used during reading.



### **Brief review of other work on human memory**

As will become apparent in this thesis, Ericsson and Simon's model of human memory has been very influential on mentalistic approaches to data-gathering, but it is by no means the only one. Due to space constraints, however, only a brief review is possible here. What many of these theories have in common is a dual-system explanation. For instance, Shiffrin and Schneider's (1977) work posits that complex cognitive tasks that have not been mastered require 'controlled' processing whereas, once they are mastered, they only require 'automatic' processing. The former are cognitively demanding while the latter require virtually no cognitive effort.

From this, it seems to follow that readers may find it harder to introspect directly about processes that have become routine (i.e., automatic) than about those that remain problematic or challenging (i.e., controlled). However, some researchers (e.g., Tarone 1982, Hulstijn and Hulstijn 1984) have argued that data gathered while subjects are doing tasks that require 'controlled' processes may suffer from interference from performance anxiety. As McLeod and McLaughlin (1986) have said, skills such as typing or driving, which may be highly 'automatic', can be hampered if, by slowing them down for research purposes, subjects give them abnormally high levels of attention.

Anderson's work (e.g., 1983) on human learning has been very influential. He developed an information-processing theory of cognition and memory that differentiates between 'declarative' knowledge and 'procedural' knowledge. In essence, whereas the former refers to what we know about and may be acquired quickly, the latter refers to what we know how to do and may be acquired relatively slowly. According to Anderson, as we repeatedly utilize a particular piece of procedural knowledge, it gradually becomes automatic.

However, there is no consensus about the existence of such an interface within this dual-system explanation of learning. Reber (1993), for instance, has argued that learning takes place by means of two functionally separate systems. Specifically, 'implicit' learning occurs without conscious effort while 'explicit' learning takes place when learners consciously engage in the process. In the L2 acquisition literature,

Krashen (e.g., 1985) has been particularly influential in his view that an unconsciously acquired knowledge base develops independently of a consciously learned one. More recently, Schmidt (1995a) has proposed that conscious awareness of input at the level of ‘noticing’ is necessary to subsequent L2 development. (For a thorough review of attention and awareness in foreign language learning, see Schmidt 1995b.)

Looking specifically at short-term memory (STM), Baddeley (1990) presents empirical evidence suggesting that it is “in some sense a speech-based system” but that this is “only one component” of STM (p. 95). He concludes that, although we have little understanding of this component of STM (or of how information is retrieved from it), it is potentially important for learning to speak, read and understand spoken discourse. In Baddeley’s view, new developments in connectionism, which assumes parallel distributed processing of information, may hold the key to improving our understanding of memory. One such development is that computers can now simulate human memory better than hitherto. For instance, while conventional computer memories can give perfect recall if correct locations are addressed, they give no recall if they are not; now, however, as in human memory, partial recall and generalization can be achieved.

To conclude this discussion of the influence of prevailing theories in cognitive psychology on mentalistic approaches to data-gathering, it seems that the Ericsson and Simon model, with its distinction between characteristics of short-term memory and long-term memory, has been particularly influential.

## **2.2 Impact of prevailing theories in linguistics**

This discussion now shifts from a focus on influential theories in psychology to a focus on their counterparts in linguistics. After a consideration of the impact on the use of mentalistic data of Chomsky’s work, there is a review of the shifts that have occurred in methodological paradigms in the past three decades.



### **Chomsky's contribution**

Although changes in perceived respectability and actual use of data-gathering techniques in the field of psychology influenced and accelerated similar changes in linguistics and applied linguistics, the initial catalyst for this major shift of perception among linguists came from a famous linguist, Chomsky. Among the radical ideas in his groundbreaking book 'Syntactic structures' (1957), Chomsky argued that the study of language and the mind were intimately connected; in the same year, Skinner (1957), the psychologist, wrote an account of language from a behaviourist perspective. Chomsky (1959) responded to Skinner's work with a devastating attack, arguing that behaviourism was unable to account for language processes or for language learning. This attack ultimately proved so devastating that it effectively ended the dominance of behaviourism in the study of both psychology and linguistics, as well as the social sciences generally.

Once behaviourism had been discredited, a major reappraisal of approaches to the study of mental processes could take place, and indeed this is what occurred in the study of language learning and language use as well as in that of reading. As Grotjahn (1987:69) has observed, Chomsky's lead in a mentalistic conception of language opened the way for researchers to investigate, through introspection, the intuitions of learners themselves. The area of second language acquisition, with its relatively narrow structural focus (e.g., on natural acquisition orders or on teaching methods) broadened into one that embraced second language use. In their book on introspection in L2 research, Faerch and Kasper (1987) specifically state:

“We have chosen to refer to the field of study as second language (SL) research, thus avoiding the bias towards developmental issues implicit in the more common term ‘second language acquisition research’.” (p. 5)

In a similar vein, Rayner and Pollatsek (1989) note that, after a near standstill lasting four decades, work on the study of reading could finally progress (p. 6). During the interim, research had been scant, and what did take place was largely confined to narrow studies of eye movements that made no attempt to relate those movements to mental processes and to applied aspects of reading, such as studies into methods of teaching children in the L1.

### **Shifts in methodological paradigms**

Reflecting the shift from a focus on the exclusive use of observable and measurable evidence of linguistic structure to one that uses mentalistic evidence direct from learners, methodological paradigms used in L2 research have developed radically in the past half century. In the fifties and sixties, ‘contrastive analysis’ was used with a view to determining L1 influence on L2 development. Predictions were made about positive and negative transfer (often referred to, respectively, as ‘facilitation’ and ‘interference’) across languages. In order to do this, linguistic analyses were conducted on various languages, but without reference to learner data (e.g., Di Pietro 1971, James 1971; and, for a thorough review of this literature, see Odlin 1989).

In contrast, the procedures that followed, ‘error analysis’ (e.g., Richards 1974) and ‘performance analysis’ (e.g., Selinker 1972, Corder 1978) used data produced by learners. While the former made use of L2 learners’ errors alone, the latter permitted the use of all learner data, whether correct and erroneous. However, as Faerch and Kasper (1987) observe, this did not mean in practice that ‘performance analysis’ actually exploited all the information present in the data. In this respect, they cite Selinker and Gass (1984), whose performance analysis was conducted on data that had been “cleansed” of “most of its performance features” (e.g., self-corrections, temporal variables such as pauses, speed, rate and manner of articulation) (p. 8). Faerch and Kasper (1987) also point out that advocates of this type of analysis did not explain how the data should be analyzed.

Faerch and Kasper (1987) go on to note that “reconstructing unobservable phenomena from performance data will always entail situations where the ambiguity between product and process cannot be solved” (p. 9). The problem, then, was to find data-collection techniques that would yield data containing learners’ statements about the ways they organize and process information. In order to achieve this, techniques used in linguistics (e.g., metalinguistic judgments) and in other social sciences, such as sociology (questionnaires and interviews) have been adapted for research into L2 use. In addition and more recently, psychology has provided sustenance for the development of techniques for collecting introspective data (e.g., think-aloud tasks).



## **2.3 Section summary**

In answer to Research Question 3.1, which can be found in Section 1.0 of this chapter, there is considerable theoretical support, both from cognitive psychology and linguistics, for mentalistic approaches to research in the social sciences. Following the demise of behaviourism, with its insistence on measurement and causality and its attendant research methods, the way became clear to investigate cognitive processes in areas such as language use and reading through the collection of mentalistic data. In other words, developments in the theory of human cognition have led to a widespread reassessment of approaches to data-generation. In linguistics, Chomsky's work on the connection between the study of the mind and of language seemed to activate applied linguists, including those specializing in L2 work, and several methodological paradigms have emerged in the past three decades. An earlier focus on analysis of learners' errors has been followed by a preference for a more inclusive approach to learner data. In order to make full and justifiable use of mentalistic data, researchers contemplating their application need to design both their studies and their tasks with considerable care. In the following section, these practical issues are addressed.

### 3.0 PRACTICAL CONSIDERATIONS

In dealing with some of the practical considerations involved in the use of mentalistic data, many of which emanate from the discussion of theory above, this section seeks to answer Research Question 3.2, posed at the beginning of this chapter. The considerations are presented under three subheadings: task design, research design and naturalness.

#### 3.1 Task design

As suggested above, human information-processing theory has considerable potential to inform researchers about how best to design tasks collecting mentalistic data as well as about the likely content of the data gathered through different tasks. It is now argued that, in order to ensure that only information in STM is reported, tasks purporting to tap on-line processes should adhere to three key criteria. The first involves the temporal relationship between the mental process and its reporting, the second deals with the elicitation procedure and the third concerns the type of information elicited.

##### Timing

In order to tap heeded information, introspection should be *concurrent* (i.e., while the information is still in STM) rather than *retrospective* (i.e., when it has passed to LTM). Many researchers writing about mentalistic data have made this point; in fact, the temporal quality of the informants' reporting often serves as a way of defining categories of introspective data. For example, Table 3.1 shows how Cohen (1984, 1987), writing about L2 learners, divided mentalistic data into three categories.

Since these labels do not clearly indicate the temporal quality of the report, they are rather confusing. Seliger and Shohamy (1989) use the same categories as Cohen but attribute different labels to them. They call the first category 'introspection', the second 'retrospection' and the third 'thinking aloud'. Again, as the term



‘introspection’ could refer to the process an informant undergoes to produce any mentalistic data, irrespective of its temporal quality, there is potential confusion in the labelling of the first category.

**Table 3.1: Cohen’s (1984 and 1987) categories of mentalistic data**

Category	Description
Self-revelation	Informants think aloud while the information is in STM (i.e., it is being reported concurrently).
Self-observation	Subjects report on specific behaviours that may or may not be in STM (i.e., it may be reported concurrently or retrospectively).
Self-report	Learners report on their learning in general (i.e., the temporal relationship between the mental action and its being reported is immaterial).

Ericsson and Simon (1980), on whose work most subsequent work relating to the collection of mentalistic data has been based, refer to ‘concurrent’ and ‘retrospective’ reporting and to ‘thinking aloud’ (p. 218). Here, the third category is potentially confusing since it could be argued that the action of ‘thinking aloud’ takes place in all verbal (though not written) reporting of mental processes. In a follow-up article, Ericsson and Simon (1987) rename their three categories, respectively: ‘process observation’, ‘post-process observation’ and ‘performance observation’ (p. 30).

To conclude, the temporal relationship between the mental action and its reporting is one of the key ways in which the literature has sought to distinguish various types of introspection and, thus, the quality of the resulting mentalistic data. This is a reflection of current theory in information processing about crucial differences in the characteristics of STM and LTM. Thus, researchers interested in investigating mental processes in on-going reading should ensure that their research tasks tap *concurrent* mental activity (i.e., while it is in STM).

**Elicitation procedure**

The degree of researcher intervention during informants’ reporting of their mental processes is another important issue to emanate from information-processing theory. Since conscious mental processes are believed normally to remain in STM for only 8-

10 seconds, it follows that there is very little time for a researcher to probe directly before the required information has passed to LTM. It seems reasonable to assume from this, then, that self-initiated, as distinct from other-initiated, reporting is more likely to tap conscious processes. In other words, the collection of *concurrent* introspective data from informants, as recommended above, seems to need *minimal researcher intervention*.

Reinforcing this point, Ericsson and Simon (1987) claim that new information can be “maintained” in STM “until the corresponding verbalization is completed” and that such data “remains the same” as it would have been if the informant had been silent (i.e., without the verbalized introspection) (p. 32). As shown in Table 3.2, they then distinguish between two reporting processes. Following Ericsson and Simon’s paradigm, it is clear that any researcher interruption to the informants’ stream of consciousness is likely to impede or destroy the “talking aloud” process.

**Table 3.2: Ericsson and Simon’s (1980) reporting paradigm**

Category	Description
Talking aloud	From thinking directly to vocalizing ‘silent speech’
Thinking aloud	From thinking to verbal encoding and then to vocalizing ‘silent speech’

In sheer practical terms, if an informant is conscious of his or her mental processes, and is concurrently verbalizing them, there would be no need for researcher intervention. Where subjects fall into periods of silence, many researchers allow a given period of time (often between fifteen seconds and one minute) before gently reminding them to talk about what is currently going on in their minds.

There is an additional reason to minimize researcher intervention in the data-collection process. The less intervention there is, the less likely it is to suffer from problems with internal validity and, hence, the more objective the study is deemed to be. Furthermore, the less intervention there is, the easier it is to replicate the study, and, consequently, the less likely it is that the study will compromise itself in terms of external validity (e.g., Seliger and Shohamy 1989).

Finally, the language in which elicitation takes place should be considered. Some researchers have recommended that introspective studies into L2 reading should have



informants report in the L2 (i.e., the same language should be used for reading as for reporting). Rankin (1988), for instance, believes that this procedure would discourage “translation and other strategies where they might otherwise have been less evident” (p. 123). In contrast, Block (1992) cautions that reports in the L2 may be incomplete due to lack of language proficiency or additional processing demands. Bernhardt (1991) agrees and complains that, though this is a widely held view, it has gone largely unheeded (p. 200). Davis and Bistodeau (1993) report on work by Lee (1986) suggesting that written recall protocols should be conducted in the L1, but they were unable to find any research evidence about the language of verbal reports. (The question of the language of elicitation is discussed further with reference to selection of informants below.)

### **Information type**

One of the key aspects of Ericsson and Simon’s paradigm is that information needs to be ‘heeded’ in order for it to be reported. Crucially, some types of information may be more easily heeded than others. Flanagan (1954) proposed the ‘critical incident technique’, which entails asking subjects to report on specific events rather than on either general information or interpretations of events. In terms of L2 research, this view is echoed by Faerch and Kasper (1987), as follows:

“sudden breakdowns of automatic processing, such as when the learner is faced with a problem in reception or production due to a lack of relevant (declarative) linguistic or other knowledge, often initiate attended processing, e.g. the use of communication strategies. These attended processes are then available to introspective reports.” (p. 12)

Furthermore, in the introduction to their book ‘Introspection in second language research’, Faerch and Kaspar (1987) make two overall, and related, observations from their collection of a dozen or so L2 studies, all of which use introspection. Their first observation is that problem-solving (including that in reading) lends itself well to introspection; by their very nature, problems demand attention. Corroborating this point, Ericsson and Simon (1980) note that familiar tasks, unless they involve problem-solving, frequently “leave only the final product as trace in STM” (p. 236).

Faerch and Kasper's second observation is that higher-level issues relating to vocabulary (including lexical search and lexical comprehension) are more likely to be heeded than lower-level processes (including issues of linguistic competence), which are more likely to be automatized. In other words, as Seliger and Shohamy (1989) have pointed out, certain issues of L2 *use* lend themselves better to introspection than issues of L2 *acquisition*. In terms of reading, Block (1992) points out that it is precisely these problem-solving issues in L2 use (e.g., lexical comprehension) that will cause subjects to have to "repair" gaps in their understanding of a text (p. 320).

Seliger (1983), at the time a sceptic of the use of mentalistic data, raises the issue of whether verbal reports provide a description of actual processes or whether they are merely *post hoc* guessing or inferencing based on comprehension. Indeed, Ericsson and Simon (1987), well-known proponents of the use of introspective methodology, warn against researchers demanding reasons for, and causes of, their mental processes. They say that asking for explanations forces informants to report information not normally heeded, and that this leads to changes in the sequence of thoughts compared to those in the silent condition (p. 38). Ericsson and Simon (1987) also say that seeking such explanations is "likely to lead to additional inferential processing with no obvious relation to a particular observed cognitive process" (p. 46). Further, seeking explanations from informants may lead to a mixing of data emanating from STM (stream-of-consciousness) and LTM (explanations), thus potentially affecting the internal validity of the data in terms of its representativeness (Seliger and Shohamy 1989).

Related to the issue of consistency is that of completeness. Ericsson and Simon (1980) say that, within the framework of their model, they would predict three causes of incomplete reporting, as follows. The information is:

- not heeded, and hence is not in STM;
- not all available in STM at the time of reporting;
- not all in LTM, or retrievable from LTM.



One reason for the first case is evident from work by Schneider and Shiffrin (1977), where there were clear differences in reports of ‘automatic’ and ‘controlled’ processing. This suggests that it is hard to report on familiar or overlearned processes, probably because they leave no trace in STM. (One way to avoid this overfamiliarity is discussed below with respect to training.) Nonetheless, processes that have been ‘automated’ over long periods of time may appear as new information when they become ‘controlled’ in research studies. Anecdotally, immediately following both the pilot-testing and data-gathering sessions in the current research, some of the subjects spontaneously expressed the sense that, by participating in introspective research, they had learned more about what they do as readers.

Having considered some practical factors impacting on the design of introspective tasks, the discussion now turns to factors affecting design of mentalistic research studies.

### **3.2 Research design**

This section addresses three issues that should be considered in the design of research studies that use mentalistic data: triangulation, the selection of informants and training.

#### **Triangulation**

It is well known that exploratory/interpretive research methodology lends itself to describing or explaining, rather than proving, phenomena. However, the validity of this kind of research can be enhanced in various ways, one of which is through triangulation. This is defined by Cohen and Manion (1994) as “the use of two or more methods of data collection in the study of some aspect of human behaviour” (p. 233). They say the objective of triangulating is for the researcher to be confident that his or her data are not simply artefacts of one method of data collection.

Seliger and Shohamy (1989:105, 188) say that confirmability of data in heuristic research is an important part of establishing internal validity. They say that criterion,

or concurrent, validity can be secured by comparing report data with directly observable behaviour. However, they warn that it is not always possible to collect comparable data using different sources. With this in mind, they cite L2 studies using informants' reports of strategies or metacognition, saying that the collection of concurrent and retrospective reports on the same point (e.g., "a language error") does not amount to drawing on the same source.

With information-processing theory in mind, what Seliger and Shohamy say appears to be borne out. Nonetheless, Haastrup (1987) used concurrent and retrospective data in her study, though she found that "for this particular study retrospection did not add a great deal of information" (p. 211). The crucial point, though, is not whether information is *added* in this way but if it can be *complemented*. In other words, the question remains whether the researcher can justifiably feel more confident presenting findings based on the comparison of concurrent and retrospective reports than those based on either concurrent or retrospective reports alone.

Cavalcanti (1987) conducted a complex three-phase study using introspective methodology with both L1 and L2 readers. The first phase involved her L2 informants giving their own assessment of their proficiency in English. In the second phase, there were several procedures including both concurrent and retrospective verbal reports. The third phase involved a cloze test, given for the purposes of assessing the informants' language proficiency. In her conclusion, Cavalcanti made a fairly radical suggestion as to how she would modify her research design if she were to conduct a similar study again:

"The triangulation I have in mind should minimally add to my own research design the informant's view of the analyst's interpretation of the ... protocols, making the informant a research assistant." (p. 246)

In her L2 research using introspective methodology, Grotjahn (1987) interprets the term 'triangulation' more broadly than either Cohen and Manion (1994) or Seliger and Shohamy (1989), who were cited above. As Table 3.3 shows, Grotjahn describes three types of triangulation that she used in her study.



**Table 3.3: Three types of triangulation (Grotjahn (1987))**

Triangulation type	Description
Data triangulation	where a variety of data sources is used to study the same problem
Theory triangulation	where a single data set is interpreted from different theoretical perspectives
Methodological triangulation	where a variety of methods is used to study the same problem

Taking this broader view of triangulation, one might use data that were demonstrably valid (e.g., a standardized language proficiency test) with data which, alone, might not be said to be valid (e.g., informant report data). If the two data sets render similar findings, the researcher could claim convergent validity (Cohen and Manion 1994:281). One might also use distinctly different ways of collecting introspective data to investigate the same research question. (Examples of studies that have triangulated are given below and the pros and cons of verbal and written report data are discussed below in Section 4.0.)

**Selection of informants**

There are two principal factors to bear in mind when selecting L2 subjects for studies using mentalistic data: language proficiency level and age.

As described in Chapter 1, Bernhardt’s (1991) theory posits that L2 reading is developmental in relation to L2 proficiency level. Depending on the difficulty level of any given text, L2 readers with low language proficiency are likely to need to devote greater attentional capacity to the reading process than those with relatively high proficiency levels (e.g., Block 1992). There is therefore a danger of cognitive overload taking place, especially where the subject is at a low language proficiency level and is expected *concurrently* both to read and to introspect in the L2. As mentioned above, while Lee (1986) suggests that written protocols should be in the L1, Davis and Bistodeau (1993) found no research findings about the language of verbal protocols. (Both the written and verbal introspective tasks in the current

research were conducted in English, an L2 for the Hong Kong subjects who participated in the study; however, I will argue a case for this in Chapter 4.)

On the other hand, following information-processing theory, L2 readers may be more aware of what they do because their reading is less likely to be as ‘automatic’ as it is in the L1. The consequent increase in ‘controlled’ processing could mean that relatively large quantities of information are processed in, and are therefore retrieved from, STM. Apparently in support of this contention, Cavalcanti (1987) reports that “protocols based on L1 reading are less informative than protocols based on FL [foreign language] reading, something which was also found in the pilot studies, and which needs further investigation” (p. 245).

Irrespective of L2 proficiency level, there may be considerable individual differences in the ability to introspect. The only way for a researcher to determine individual capacity to introspect would be to monitor potential subjects in pilot tests or during training sessions. However, while this might be counterproductive in terms of a practice effect (discussed below with respect to training), for some L2 readers, the completion of a task designed to generate introspection during reading may actually, if only temporarily, raise awareness of reading processes.

Most L2 reading research has been conducted on adults or young adults already familiar with, and fully proficient in, L1 reading. Any researcher interested in reading processes in L2 children needs to consider very seriously whether the combination of youthful naïvety with reading *and* reporting in an L2 may be so onerous that it is incapable of yielding reliable data.

## **Training**

Most potential informants for studies using introspective techniques are likely to be unfamiliar with the process of reporting their mental processes, not least in combination with reading, which is usually a silent activity.



One way to foster stream-of-consciousness verbalizing is to provide a training phase for informants before the main study is conducted. Rankin (1988:122-123) believes that training is very important in order to explain and model introspective tasks. However, the quantity of any training needs to be planned very carefully. While informants who are inexperienced in introspecting may not verbalize sufficiently, there is a danger of over-training subjects, thus causing them to be so familiar with the processes the researcher is seeking to investigate that those very processes cease to be processed at the conscious level.

There are numerous calls in the literature for informants to be relatively inexperienced in any introspective task they are asked to do. Ericsson and Simon (1980) recommend that informants should be “novices” rather than “experts” (p. 235) while Seliger and Shohamy (1989) say they should be as “naïve” as possible in order to avoid triggering a “practice effect” or “test wiseness” (pp. 102-103). Likewise, Cohen (1987:91) warns of the dangers of too much, too little, or the wrong type of, training. As far as a practice effect is concerned, it seems feasible that this could occur during the course of a data-gathering session, though, in some cases, this might be obviated by having the informants working in pairs where the duo-dynamic might make them feel constrained to clarify their experiences. (The informants in Study 2 of the current research work in pairs.)

### **3.3 Naturalness**

In addition to the considerations of task design and research design, in studies that utilize mentalistic data, naturalness of context, text and task also need to be borne in mind. Taken together, these factors, discussed briefly below, are likely to create an atmosphere of positive affect, which, in turn, may improve the richness of the data yielded.

First, while the act of reading and reporting on the process cannot be said to represent normal reading, it is important to maximize the naturalness of the process within the bounds of an introspective study. Tarone (1982), among others, has warned of the distorting effects of artificial research settings. One way to avoid such distortions is to

provide as natural a research context as possible. For example, it may be possible to use a classroom, where student informants are used to reading, rather than, say, the impersonal atmosphere of a language laboratory with separate booths for each subject. Furthermore, unless necessary, the use of special equipment, such as tachistoscopes and even video or audio recorders, should be avoided (e.g., Cavalcanti 1987). In addition, in order to replicate the subjects' normal reading as far as possible, it is important to select a text that is likely to be of interest to them; also, the text should represent an authentic example of a genre that they are used to reading. Finally, as discussed above, tasks should be designed to intrude as little as possible on the normal reading process.

An individual's overall reading experience at any one 'sitting', formulated by such factors as context, text and task, can be plotted, albeit very subjectively, on a simple continuum representing readers' perceptions of naturalness and artificiality. One end of the continuum would represent a feeling of completely natural reading while the other end would represent an overwhelming sense of artificiality. At the natural end, the reader would be in his or her chosen context, would select the text and would not have any externally imposed task. Needless to say, under these conditions, where the reader is a completely free agent, it would be impossible for a researcher to capture mental processes *online*. At the artificial end, the researcher would intervene to the point where the reader felt as though he or she had no control over the process at all.

Between these two extremes, the continuum would represent reading situations where researcher and reader had varying amounts of control over contexts, texts and tasks. For instance, a situation that felt relatively, but not completely, natural might be one in which a reader chose the context and text, and was only required to retrospect (i.e., introspect *after* reading the entire text). Hastrup (1987) used retrospection in part of her study, but she chose both the context and text utilized in her research. However, as discussed above, there are disadvantages to retrospective data-collection, and these should be balanced against the need for naturalness. In contrast, a situation that felt relatively, but not overwhelmingly, artificial might be one in which a researcher made a sensitive choice of context (i.e., not a language laboratory) and text (i.e., a genre familiar to the readers). The text might be marked at intervals to indicate where the



reader’s task was to introspect on his or her mental processes (e.g., Olshavsky 1976/77, Block 1986b and 1992, Wade et al. 1990, Jimenez et al. 1996).

3.4 Section summary

In answer to Research Question 3.2, posed in Section 1.0 of this chapter, Table 3.4 summarizes practical considerations in the use of mentalistic data in L2 reading studies and consequent safeguards. Following the preceding discussion, these considerations are divided into those pertaining to task design and research design as well as some additional issues about naturalness of context, text and task.

Table 3.4: Considerations in task design and research design

Consideration	Possible safeguard
<i>Task design</i>	
Timing	<ul style="list-style-type: none"><li>- Tap information in short-term memory.</li><li>- Use concurrent introspection, rather than retrospective.</li></ul>
Elicitation procedure	<ul style="list-style-type: none"><li>- Minimize researcher intervention.</li><li>- Maximize informant-initiation.</li></ul>
Information type	<ul style="list-style-type: none"><li>- Select problem-solving.</li><li>- Avoid seeking explanations.</li></ul>
<i>Research design</i>	
Triangulation	<ul style="list-style-type: none"><li>- Conduct more than one study into the same question.</li></ul>
Selection of informants	<ul style="list-style-type: none"><li>- Select informants with fairly high L2 proficiency.</li><li>- Avoid children.</li></ul>
Training	<ul style="list-style-type: none"><li>- Train for task clarification only.</li><li>- Avoid over-training and consequent risk of automatic processing.</li></ul>
<i>Naturalness</i>	
Context	<ul style="list-style-type: none"><li>- Choose natural settings to promote positive affect.</li></ul>
Text	<ul style="list-style-type: none"><li>- Select an authentic text of manageable difficulty level, of interest and of familiar genre to informants.</li></ul>
Task	<ul style="list-style-type: none"><li>- Devise task to maximize normal processing.</li></ul>

From this, it is reasonable to conclude that, with careful research and task design, mentalistic data are suitable for research of a heuristic/inductive nature. This kind of research is exploratory, interpretive and/or descriptive, and the data may be used to discover patterns and so to generate hypotheses for use in further research. Follow-up research may use the findings of this kind of study to test hypotheses through the application of analytic/deductive techniques that will render data that can be measured and may permit claims of causality.

Nevertheless, however carefully researchers try to design their studies of mental behaviour, as Nunan (1992) acknowledges, “it may well be that the act of spelling out our thought processes alters those processes”. However, he goes on:

“This worry is not peculiar to introspection ... researchers, regardless of the methods they use or the traditions to which they adhere, will (if they are honest) often have to confront the possibility that their results are in some ways artifacts of the procedures they have used.” (p. 117)



**4.0 INTROSPECTIVE METHODOLOGY AND READING STRATEGIES**

Having sought to show the potential offered by introspective methodology for research into mental processes, it is time to move on to two rather more specific questions (i.e., Research Questions 3.3 and 3.4). These are: how has introspective methodology been used to investigate reading strategies and what are the pros and cons of particular introspective techniques. Where appropriate, the following discussion, which covers both L1 and L2 research, considers the extent to which individual studies meet the design considerations in the preceding section. The following discussion begins with a review of recent mentalistic studies of reading strategies and then focuses on some introspective data-gathering techniques that have been used in the spoken and written modes.

**4.1 Introspection in recent studies**

Table 3.5 below presents a summary of several features relating to recent studies of reading strategies; the studies are presented chronologically. The table includes information on ‘introspection type’, that is, whether the subjects’ introspection was concurrent with, or retrospective to, reading text. Given in brackets in the same column, the table includes the ‘introspection mode’, that is, whether informants reported verbally or in writing. Further, it summarizes the ‘task type’, or what the subjects were asked to undertake. Finally, ‘task mode’ refers to whether the task was conducted by subjects individually or in pairs and what role the researcher played during introspection.

**Introspection type and mode**

First, it is clear from Table 3.5 that all the studies utilized some kind of introspective methodology. Satisfying the temporal criterion of tapping reading processes while they are in STM, the majority of the researchers opted for concurrent introspection; that is, their subjects were asked to introspect while reading. Four of the studies in the table made use of both concurrent and retrospective introspection (Haastrup 1987,

Cavalcanti 1987, Feldmann and Stemmer 1987 and Li and Munby 1996). However, representing a clear breach of the temporal criterion to tap mental processes while they are in STM, one of the researchers (Kletzien 1991) opted exclusively for retrospection, that is, informants introspected immediately after reading a complete text. In all the studies, some or all of the introspection took place in the spoken mode and was audio-tape-recorded. With the exception of Li and Munby's (1996) second task, there is a notable absence of introspection through the written mode. (Mode of introspection is discussed further below.)



**Table 3.5: Studies of reading strategies using mentalistic approaches**

Study	Introspection type (mode)	Task type	Task mode
Olshavsky 1976/77 (L1)	Concurrent (Spoken)	Read marked text & think aloud after each clause	S alone with R in background
Hosenfeld 1977 (L2)	Concurrent (Spoken)	Read text & think aloud	S alone with R interviewing
Block 1986b (L1 & L2)	Concurrent (Spoken)	Read 2 marked texts & think aloud after each sentence and paragraph	S alone with R prompting
Sarig 1987b (L1 & L2)	Concurrent (Spoken)	Read 2 texts for main ideas and synthesis	S alone with R in background
Haastrup 1987 (L2)	Concurrent & Retrospective (Spoken)	Read text & guess word meanings	i) S pairs with researcher in background; ii) S alone with R interviewing
Cavalcanti 1987* (L1 & L2)	Concurrent & Retrospective (Spoken)	Read text in L2 and think aloud at reader-generated pauses; read same text again and think aloud at researcher-generated pauses; oral summary after each paragraph	S or Ss with R prompting (for all tasks listed)
Feldmann and Stemmer 1987 (L2)	Concurrent & Retrospective (Spoken)	Do <i>C-test</i> & think aloud	i) S alone without R ii) S alone with R interviewing
Wade et al. 1990 (L1)	Concurrent (Spoken)	Read marked text & think aloud	S alone with R interviewing
Pritchard 1990 (L1 & L2)	Concurrent (Spoken)	Read 2 marked texts & think aloud after each sentence	S alone with R in background
Kletzien 1991 (L1)	Retrospective (Spoken)	Do 3 rational cloze texts & think aloud after each one	S alone with R prompting
Block 1992 (L1 & L2)	Concurrent (Spoken)	Read marked text and think aloud after each sentence	Unclear, but probably Ss alone with R either in background or prompting
Davis & Bistodeau 1993* (L1 & L2)	Concurrent (Spoken)	Read text typed on a separate card for each sentence (single-sentence cueing)	Unclear but, for task listed, probably individual S alone with R
Li & Munby 1996 (L2)	Retrospective (Spoken), Concurrent (Spoken) & Retrospective (Written)	i) Unstructured interview with R ii) Read text and think aloud iii) Read text and write journals	i) S alone with R interviewing ii) S alone with R prompting iii) S alone
Jimenez et al. 1996* (L1 & L2)	Concurrent (Spoken)	i) Read unmarked texts ii) Read texts marked with 17 asterisks and think aloud at each one	Both tasks: S alone with R prompting
Jimenez 1997 (Low-literacy bilinguals)	Concurrent (Spoken)	Read texts in L1 and L2 and think aloud after each line of text	Ss alone with R prompting

**Key for Table 3.5**

S = subject

\* denotes additional features not listed here

R = researcher

**Note for Table 3.5**

The *C-test* is a test based on the principle of reduced redundancy in text (through the deletion of letters from words) and on the construct of an internalized pragmatic expectancy grammar (see Feldmann and Stemmer 1987:251).

## Task type and mode

Turning to task type, most of the studies used texts that were marked in some way (usually dots). The following studies all marked their texts to indicate where the readers should stop reading and introspect: Olshavsky (1976/77), Block (1986b and 1992), Wade et al. (1990), Pritchard (1990), Jimenez (1997) and part of the Cavalcanti (1987) and Jimenez et al. (1996) studies. Pritchard also instructed his readers to stop and introspect at any other time they wished. The frequency of the markers in the text varied considerably; clearly, the more densely the markers are inserted into the text, the more researcher intervention there is. This is something that varied considerably across the studies cited. While Olshavsky (1976/77) inserted a dot after each clause in her text, Wade et al. (1990) inserted a total of eight dots (to a text that covered fifteen double-spaced pages), each one after occurrences of what were deemed “major shifts in topic” (p. 151).

Four other studies used texts that were marked in diverse ways. Feldmann and Stemmer (1987) used a C-test, which involves deletion of letters from words in a text. Kletzien (1991) used cloze texts (discussed further below). Haastrup (1987) used a text with a given number of words that were unknown to her subjects and whose meaning they were instructed to guess. Davis and Bistodeau (1993) report using a technique developed earlier by Olson, Duffy and Mack (1984) to promote concurrent introspection. In order to ensure that mental processes relating only to the sentence currently being read are verbalized, they typed each sentence of their text onto separate cards. In this way, they sought to constrain their subjects to attend to one sentence at a time. Unfortunately, in my view, this ‘single-sentence cueing’ runs the serious risk of creating a very artificial reading task that could well affect the data yielded. Nonetheless, of all the studies discussed here, with respect to task type, two stand out most clearly as being in breach of the principle of minimal researcher intervention. These are a very early study (Olshavsky 1976/77, with a text marked after each clause) and a very recent one (Jimenez 1997, with a text marked at the end of each line).

Finally, Table 3.5 presents information on task mode. In every study where procedures are clearly described, individual subjects spent some or all the data-



gathering time alone with the researcher, though the level of researcher intervention varied widely. In only two studies were subjects reported to be reading without the researcher present. In the first part of their three-stage project, Feldmann and Stemmer (1987) reported that individual subjects were “left alone” in a room (p. 252). In the third part of their study, Li and Munby (1996) had their subjects do their journal writing in private. In some cases, the researcher merely remained in the background, available to help any subject in difficulty (Olshavsky 1976/77, Sarig 1987b, Pritchard 1990); this level of intervention also applied to the first stage in Haastrup’s (1987) two-stage study.

In other cases, the researcher prompted a subject who remained silent for more than a specified period or who appealed for help (Block 1986b and 1992, Cavalcanti 1987, Kletzien 1991, Li and Munby 1996, Jimenez et al. 1996, Jimenez 1997). The highest level of intervention took place in studies where the researcher interviewed the subjects *concurrently* with their reading (Hosenfeld 1977, Wade et al. 1990), a procedure that almost certainly affected any reading processes reported from short-term memory. In contrast, the researchers in two of the three multi-stage studies (Haastrup 1987, Feldmann and Stemmer 1987) confined their interviews to the second stage, that is, *retrospectively*, thus allowing the informants, at an earlier stage in the procedure, to report on their reading without researcher intervention.

Several more points need to be made about task mode. First, there seems to be no effect for language, there being no difference in the level of researcher intervention across the L1 and L2 studies cited in the table. Second, there appears to be a positive effect for training informants in introspection. All the L1 studies used subjects who were adults or young adults, an age group that does not normally read and introspect aloud on their mental processes. It must be assumed that, during the training process, researchers made pragmatic decisions as to the level of intervention needed during data collection for their particular group of subjects. Third, as already mentioned, three studies (Haastrup 1987, Feldmann and Stemmer 1987 and Li and Munby 1996) conducted some form of triangulation, thus seeking to confirm, or at least compare, their findings within studies. Finally, only one study reported subjects working in pairs (Haastrup 1987). (The issue of reading tasks conducted in pairs is addressed below and in Study 2.)

## **4.2 Data-gathering procedures**

Having shown how recent mentalistic studies have investigated L1 and L2 reading strategies, there follows a review of some of the data-gathering procedures available for this kind of research. As shown above in Table 3.5, only one of the studies cited (Li and Munby 1996) utilized written introspection; moreover, this three-task study used the spoken mode as well as the written mode. Reflecting this bias in the literature, the following discussion begins by surveying some techniques that have been used to elicit verbal reports of reading strategies and it ends with some techniques that have been used to obtain written reports. (A further discussion of introspection using the written mode, which reflects on its use in the current research, can be found in Chapter 7.) Since space does not permit a full review of each technique, I have used tables to summarize what appear to me to be the main advantages and disadvantages of each and then highlighted the most important of these in the accompanying discussion.

### **4.2.1 Some procedures for verbal report data**

There follows a discussion of three techniques that have been used to elicit verbal report data on reading strategies: oral reading, cloze and interviewing.

#### **Oral reading**

The main point to be drawn from Table 3.6 is that oral reading appears to be a much more useful technique for investigating children's reading than that of L2 adults and it has, indeed, been far more widely utilized to study the former than the latter. Working with adult L1 and L2 readers, Cavalcanti (1987) expressly rejected oral reading as it does not meet her stated requirement of preserving reading in her study as both "silent and private" (p. 231). Reviewing several L2 reading studies including one of her own, Bernhardt (1991) found that oral reading:

"consistently impedes comprehension for second language readers in general. This study [Bernhardt 1983] casts doubt on the appropriateness of using oral



reading for any investigation of the comprehension process as well as on the findings of the former studies [i.e., the L2 studies she cites that have used oral reading].” (p. 38)

Table 3.6: Pros and cons of oral reading

Some advantages	Some disadvantages
<ul style="list-style-type: none"><li>- for children, speeds in oral and silent reading may be similar (Rayner &amp; Pollatsek 1989:181)</li><li>- taps concurrent processes</li></ul>	<ul style="list-style-type: none"><li>- for adults, an “unnatural” activity (Rayner &amp; Pollatsek 1989:180); L1 and L2 adults usually competent at decoding graphics</li><li>- for adults, speeds in oral and silent reading may vary considerably (Rayner &amp; Pollatsek 1989:181)</li><li>- for adults, possible negative response due to lack face validity and ‘performance’ aspect of task</li><li>- may differ from normal silent reading (Cavalcanti 1987:231)</li><li>- impedes comprehension (Bernhardt 1991:38)</li><li>- ‘miscue analysis’ focuses on errors</li></ul>

In addition, since oral reading offers no insights directly from readers, the ‘performance data’ yielded are only available to ‘miscue analysis’, that is, investigating instances of subjects misreading the text. Focusing alone on readers’ errors seems to be an unlikely way of gaining a full understanding of the reading process. As mentioned above in discussion of shifting analytical paradigms, ‘error analysis’ has been superseded in applied linguistics by a broader view of analyzing learner data.

Cloze

Researching readability, Taylor (1953) first used the term ‘cloze’ to describe the principled deletion of words from a text. The name originates from Taylor’s assumption that, when using redundancy in the text to guess the deleted words, readers undergo a gestalt process of ‘closure’. As shown in Table 3.7, deletion may be either ‘random’ or ‘rational’. ‘Random’ cloze involves a fixed ratio deletion, conventionally one word in every five or seven; its proponents claim that a variety of textual constraints representative of normal reading is highlighted by this procedure.

‘Rational’ cloze, on the other hand, involves deletions being made with whatever criteria the researcher selects.

Table 3.7: Pros and cons of cloze

Some advantages	Some disadvantages
<ul style="list-style-type: none"><li>- easy to devise</li><li>- flexible format (‘random’ v ‘rational’) that is amenable to varying research aims and objectives</li><li>- concurrent reporting possible as reader deals directly with text (Jenkinson 1957)</li><li>- problem-orientation</li><li>- conceptually simple and familiar to subjects</li><li>- procedurally simple and flexible (silent individual v oral pair)</li><li>- reduced reading speed need not affect normal reading processes (Rankin &amp; Thomas 1980:54) and does not affect memory or comprehension (Stanovich 1980:58-59)</li><li>- may represent normal reading (Rankin &amp; Thomas (1980:54) and require both ‘local’ and ‘global’ reading Jonz (1990)</li></ul>	<ul style="list-style-type: none"><li>- artificial, researcher-generated gaps in text</li><li>- some content words are impossible to predict (Alderson 1980)</li><li>- may distort normal reading by encouraging ‘local’ reading and discouraging ‘global’ reading (e.g., Alderson 1978)</li><li>- familiarity with cloze tests to assess language proficiency may cause negative response</li><li>- ‘exact-word’ scoring may cause ‘right-answer mentality’ (Hedge 1991:180)</li></ul>

For instance, Clarke (1981) deleted every seventh to tenth word, saying that “an effort was made to delete words whose replacement would seem to require comprehension of the entire passage” (p. 71); usually, these were discourse markers, content words (nouns, verbs, adjectives and adverbs) and anaphoric and cataphoric referents. Levenston et al. (1984) developed another variant of ‘rational’ cloze, which they termed ‘discourse’ cloze. Based on Halliday and Hasan’s (1976) categorization of cohesion, they deleted items that marked relationships between propositions. More recently, Bensoussan (1990) developed what she called ‘cohesion’ cloze, in which the task is further constrained so that readers have to fill the cloze gaps with words already in the text.

Because of the ease with which a cloze passage can be devised in a range of different forms as well as the ease with which it can be completed (correctly or otherwise) and scored (for exact-word or acceptable-word replacements), the procedure became popular with researchers and teachers. In addition to its use for assessing readability, it was soon perceived as a reliable and valid measure of L1 reading comprehension



(e.g., Jenkinson 1957, Rankin and Thomas 1980, Clarke 1981) and L2 proficiency (e.g., Chihara et al. 1977, Clarke and Burdell 1977).

Despite these perceived advantages, there was a widespread belief that cloze required only 'local' rather than a combination of 'local' and 'global' reading. Space does not allow a full review of the vast literature on this, so I will take one recent and thorough review, that of Jonz (1990). (For extensive reviews of this aspect of cloze, see Bachman 1982 and 1985 and Jonz 1987 and 1990.) Jonz (1990) provided evidence that 'random' cloze may contain greater variety of constraints than had previously been assumed. He analyzed eight random cloze passages published in research studies between 1975 and 1990. Apparently but not explicitly using Halliday and Hasan's (1976) categories of grammatical and lexical cohesion, he found the following frequencies: 41.80% for lexical selections, 34.10% for textual cohesion and 24.10% for syntax. From this, Jonz concluded that 'random' cloze is sensitive to constraints on response ranging well beyond the level of local syntax. It is possible that 'rational' cloze could modify these figures and provide even more evidence of 'global' reading.

One of the few studies to use cloze for introspection is that of Kletzien (1991); for the purposes of the current study, it is therefore worth reviewing in relative detail. Investigating L1 reading strategies and reading difficulty, Kletzien used 'rational' cloze, making deletions on context-dependent content words. She divided her high school subjects (N=48) equally into 'good' and 'poor' comprehenders and asked them to read three cloze passages, each approximately 250 words long, on familiar topics at three levels of difficulty on the Fry readability scale: 'independent', 'instruction' and 'frustration'. Retrospectively, the subjects reported the strategies they used on the passages. She found that, on the easy passage, the two groups used the same types of strategies at similar frequencies but that, as passage difficulty increased, the good comprehenders used a relatively wide range of strategies and used them more frequently. In absolute terms, none of her subjects used a great variety of strategies, there being a total of only fourteen in her data-driven list.

When she analyzed differences in strategy use among the three texts, Kletzien found more evidence on the 'independent' level text of the following strategies: 'reading

previous text', 'reading subsequent text' and 'using prior knowledge'. From this, she concluded that, when readers are processing automatically, they are freed from lower-level decoding constraints and there is more capacity to make associations, integrate ideas and process higher-level information (e.g., LaBerge and Samuels 1974). Conversely, when processing is relatively hard, readers are constrained to use strategies for decoding the text (e.g., Paris and Myers 1981).

Kletzien's findings suggest that, depending on text difficulty, introspective 'rational' cloze tasks have the capacity to differentiate between local and global reading, an issue raised in Table 3.7. In terms of procedure, however, Kletzien could have used concurrent introspection (i.e., as subjects tackled each cloze item) rather than retrospective (i.e., after they had completed each item in an entire passage).

It will be recalled in the discussion of task design above that an element of problem-solving might help to bring 'automatic' processing to a more conscious level, and hence be more likely to be reported. As Table 3.7 shows, the cloze procedure, with its gaps, or simulated breakpoints, contains a problem-orientation. Moreover, verbal reports from readers doing a cloze activity are likely to be concurrent with any such conscious, or 'controlled', processing.

Table 3.7 also shows that there are some advantageous practical aspects to the cloze procedure. It is conceptually and procedurally simple, features that are particularly important if introspection is an unfamiliar activity for the subjects. In other words, it would seem desirable to use the subjects' very familiarity with cloze to introduce the novelty, and potential difficulty, of introspection. Counterbalancing this, as the table also shows, is the possibility that subjects who have been subjected to numerous cloze *tests* might have negative associations with the entire procedure. Such negative feelings are especially likely where scores have been calculated on an exact-word, rather than appropriate-word, basis; this is where only the replacement of the original word is accepted as 'correct'. However, such 'baggage' might be dispelled through training sessions that introduce subjects to the notion of introspective cloze tasks, that is, where the objective is the relatively unconstrained one of airing mental processes rather than producing correct products. Moreover, since training would be a necessary



part of any introspective research project, this phase of a research project would not entail extra work on the part of either researcher or subjects.

**Interviewing**

One of the main points to be drawn from Table 3.8 is that interviewing is a flexible activity. At one end of the spectrum, it can be highly structured, that is to say, interviewers have a list of prearranged questions from which they do not deviate; at the other end, an interview can be unstructured, giving interviewers the flexibility to respond to unanticipated information. Between these two extremes, there are ‘semi-structured’ interviews, which combine elements of their more extreme counterparts. In a sense, the semi-structured variant is the most flexible as it allows both structure and spontaneity. It has been suggested (e.g., Gordon 1980) that the more structured the interview, the more interviewees will feel like ‘subjects’ rather than ‘informants’. In addition, as Nunan (1992) has observed, L2 interviewees might feel more intimidated in interviews with rigid formats than they would where there is less structure. Related, as Dowsett (1986), among others, has noted, the data yielded in the relative freedom of an unstructured, or semi-structured, interview might be of a richer quality.

**Table 3.8: Pros and cons of interviewing**

Some advantages	Some disadvantages
<ul style="list-style-type: none"><li>- when ‘structured’, less subject to interviewer bias (Cohen and Manion 1994:281-282)</li><li>- when ‘unstructured’, flexible in dealing with unanticipated outcomes</li><li>- when ‘unstructured’, potential to balance ‘asymmetrical’ relationship, between interviewer and interviewee (especially L2 interviewee) (Nunan 1992:150)</li><li>- possible positive response, especially when ‘unstructured’ (Gordon 1980) and thus potentially rich data (Dowsett 1986:53)</li></ul>	<ul style="list-style-type: none"><li>- retrospective reporting</li><li>- analysis labour intensive, especially when ‘unstructured’</li><li>- possible negative response, especially when ‘structured’ (Gordon 1980)</li></ul>

Finally, unlike either oral reading or cloze, interviews cannot tap introspection concurrently with reading; however, they may be a useful device to use

retrospectively, thereby yielding data that can be used to triangulate with those collected in the same study through the use of other techniques.

4.2.2 Some procedures for written report data

The next part of the discussion moves from reviewing three procedures used to elicit verbal report data for reading strategies to three procedures for eliciting written report data. As noted above, to my knowledge, only one published study, that of Li and Munby (1996), has used the written mode (in this case, journals) to collect introspective data to investigate L2 reading strategies. The three procedures reviewed are: questionnaires, introspections on *breakpoints* (explained below) and journals.

Questionnaires

Since questionnaires are probably the most widely used format for gathering data, there is a vast literature on them; however, space constraints only allow a very brief review here, with the specific brief of their potential to study reading strategies. (For a comprehensive review of their potential in educational research, see, for example, Oppenheim 1992.) One of the most interesting points in Table 3.9 is the indirect potential they offer researchers investigating reading strategies. Kirby (1988) has noted that questionnaire data tend to be rather general and Crossley and Vulliamy (1984) have suggested that questionnaires can be used to obtain a good general idea of a research context, a picture that can be embellished through the use of other techniques.

Table 3.9: Pros and cons of questionnaires

Some advantages	Some disadvantages
<ul style="list-style-type: none"><li>- easy to administer</li><li>- flexibility in length</li><li>- potentially valid and reliable</li><li>- easy to analyze data from some question types</li><li>- can be used with other tasks (e.g., triangulate; also, Crossley and Vulliamy’s 1984 context v detail)</li></ul>	<ul style="list-style-type: none"><li>- hard to devise and need considerable pilot-testing</li><li>- collect rather general data (Kirby 1988:234-5)</li><li>- inflexible in face of unanticipated outcomes (Crossley and Vulliamy 1984:198)</li><li>- subjects’ indifference (Hedge 1991:38)</li></ul>



In the case of reading strategies, questionnaires might therefore provide a better method to probe readers for what they usually do than for what they did with a particular text. This is what Scott (1990) did in an unpublished PhD thesis. Scott conducted twelve studies of lexical inferencing among L2 readers (Portuguese-speaking Brazilians). For one of the studies, he focused on lexical inferencing strategies, using a 'delayed retrospection' technique in which he asked his subjects to complete a short questionnaire probing what they would generally do when faced with certain hypothetical reading situations. Scott was surprised to obtain richer data from this questionnaire than from a verbal think-aloud technique he used on different subjects in a similar study. Noting that "it is easier to speak than to write" about reading processes, he concluded that this difference in data quality was probably due to differences in the subjects participating in the two studies (p. 289).

Questionnaire data about reading behaviour is thus bound to be retrospective to reading a particular text and will probably be timeless and general. This would make it more likely that the data yielded would include reports of subjects' resorting to long-term memory, and possibly interpreting their reading behaviour in general rather than simply recalling their behaviour with a particular text.

There might be additional reasons why process-oriented data, needed to investigate reading strategies in ongoing reading, might suffer. As Hedge (1991:38) discovered, subjects can be indifferent to the rather impersonal nature of questionnaires; indeed, some populations, such as university students, are frequently asked to complete their peers' questionnaires, and a certain level of respondent fatigue can occur. While it is obvious that data elicited through the use of *any* introspective technique only reveals the behaviour subjects report, it seems there is a greater likelihood that subjects will have negative affective reactions to questionnaires (particularly to those that are dauntingly lengthy) than to other introspective data-collection techniques. If this is the case, it would have the potential to endanger the quality of data yielded by questionnaires.

**Introspections on breakpoints**

Davies (personal communication 1994) developed an innovative introspective task for use with postgraduate students at the University of Bristol. She asked her subjects to write reports of what they did whenever they experienced a breakpoint during their reading of a text. The term ‘breakpoint’, first used by Hedge (1991), refers to any interruption in the normal flow of reading. Though she did not analyze her data for reading strategies, Cavalcanti (1987) used a similar technique to elicit oral reports, or, as she called them, ‘pause protocols’.

Given the relative ease with which this kind of research task can be devised and administered, it is surprising that it has not been used (at least, to my knowledge) in any published work. It may be that the labour intensive nature of data analysis has discouraged its use. Table 3.10 shows that an additional advantage is that there is likely to be little or no researcher intervention during the course of the task. From the subjects’ perspective, the task is conceptually easy, and its free-reading element offers a high degree of reader autonomy. Further, researchers interested in maximizing reader autonomy are free to have their subjects select the texts to be used with the task.

**Table 3.10: Pros and cons of introspections on breakpoints**

Some advantages	Some disadvantages
<ul style="list-style-type: none"><li>- relatively easy to devise</li><li>- procedurally and conceptually easy</li><li>- problem-orientation</li><li>- concurrent</li><li>- subject autonomy</li><li>- minimal researcher intervention</li></ul>	<ul style="list-style-type: none"><li>- relatively complex to analyze data</li></ul>

This kind of introspective technique has theoretical support. Following information-processing theory, writing should begin within 8-10 seconds of the actual experience (Simon 1979, Ericsson and Simon 1980). Arguably, readers are likely to start writing introspectively about what precipitated their breakpoints concurrently with the experience. Information-processing theory allows for heeded information to remain accessible in short-term memory, thus allowing for lengthy written entries that may take several minutes to complete. In addition, this task is consonant with this theory



in that breakpoints are a kind of problem and, as such, represent information that is likely to be ‘controlled’ and therefore heeded. (Information-processing theory was discussed at length earlier in this chapter.)

**Journals**

To my knowledge, journal writing is the only procedure using introspection in the written mode that has been used in a published study of L2 reading strategies. As shown above in Table 3.5, this technique was used as one of three procedures in Li and Munby’s (1996) study, the other two being unstructured interviews and oral think-alouds. Although Li and Munby report that their two subjects agreed to write their journals “while reading” (p. 203), it is clear from the only sample they quote from one of their subjects that she was writing very generally about her reading strategies. Unfortunately, in presenting their results, the researchers do not specify which findings originate from which of the three techniques used.

**Table 3.11: Pros and cons of journals**

Some advantages	Some disadvantages
<ul style="list-style-type: none"><li>- inherently introspective in nature (Bailey and Ochsner 1981:189, Nunan 1992:123-4)</li><li>- reveal personal variables (Bailey and Ochsner 1981)</li><li>- subject-centred (Porter et al. 1990:237)</li><li>- informal setting of subjects’ own choice</li><li>- minimal pressure on subjects</li><li>- no researcher intervention</li><li>- potential for longitudinal element</li></ul>	<ul style="list-style-type: none"><li>- little or no researcher control</li><li>- learners feel so free that they may fantasize</li><li>- probably retrospective</li></ul>

Table 3.11 shows how proponents of journal writing promote its inherently introspective nature, though this is likely to be retrospective to a particular reading experience. While there is little or no researcher control, the trade-off is that subjects are likely to feel free of pressure and may therefore provide a detailed and revealing, if subjective, picture of their reading behaviour. It seems to me that this technique would be excellent for case studies with the specific objective of studying individual readers over time. It could also be used to triangulate with data drawn from other techniques.

### 4.3 Section summary

In answer to Research Question 3.3 (posed at the beginning of the chapter) about the ways in which introspective methodology has been used to study reading strategies, the first part of this section presented a review of recent studies in both the L1 and L2 literature. The most striking feature of this body of research is that there has been almost exclusive use of oral introspection at the expense of written introspection. I cannot immediately see an obvious reason for this. Certainly, the work on human information-processing conducted by such cognitive psychologists as Ericsson and Simon appears to assume use of the spoken mode. As far as the two cited studies of reading strategies that have used written introspection are concerned, Li and Munby (1996) make no comment about mode of reporting and Scott (1990), as reported, notes that it is “easier” to introspect about reading behaviour orally than in writing (p. 289). It is possible that this ease may apply more to Scott’s Brazilian subjects than to subjects from educational contexts, such as Hong Kong (described in Chapter 4), where there has traditionally been a particular emphasis on writing. Certainly, Li and Munby’s sample extract from the journal of a Chinese university student studying in Canada is very lucid (p. 214). It therefore seems worth investigating the use of written protocols for mentalistic studies of reading strategies.

Another feature of the introspective research studies reviewed is the extensive use of text marking, invariably justified in terms of encouraging subjects to introspect. With such researcher-intervention, there is the potentially negative trade-off that readers’ normal processes will be distorted. Training subjects to introspect prior to the conduct of studies may help familiarize them with the whole process of introspecting and so obviate the need for such text marking.

In answer to Research Question 3.4 about the pros and cons of various introspective techniques, it is clear that, since so little work has been done in the written mode, more research needs to be conducted before an informed judgment can be made about particular techniques. Nonetheless, at this stage, Davies’s procedure of tapping introspections on breakpoints, or subject-generated breaks in the flow of reading, looks extremely promising for concurrent processing. As far as the oral mode is concerned, it seems certain that oral reading and its associated miscue analysis are



unsuitable for adult subjects who are already literate. With its flexibility, the cloze procedure appears to have some potential, though, as with marked texts, cloze gaps are researcher-generated and so may have a distorting effect on reading processes as well as a negative affective response on the part of subjects. Interviewing seems to offer potential for follow-up, retrospective work whose data can be used to triangulate with data yielded through other techniques.

## 5.0 CHAPTER SUMMARY

This chapter posed and sought to answer the following research questions:

- 3.1 Is there any theoretical support for mentalistic approaches?
- 3.2 When designing studies yielding mentalistic data, what practical concerns does the researcher need to consider?
- 3.3 How has introspective methodology been used to investigate reading strategies?
- 3.4 What are the pros and cons of various data-gathering techniques?

There is theoretical support for mentalistic approaches from work conducted on human cognition; in essence, the distinctive natures of short-term memory and long-term memory determine how easy it is to retrieve information.

Since the theory posits that information in short-term memory is heeded, it follows that researchers interested in reading processes may prefer to attempt to tap their subjects' ongoing reading concurrently, rather than have them retrospect about it, though retrospection is a serious contender as a supplement to concurrent introspection. Also in relation to heeded information, problem-solving tasks may be relatively easy to tap. Another practical concern in task design is to allow reading to be as normal as possible by minimizing researcher intervention. Training subjects in introspection may help to avoid the need for such intervention. As far as research design is concerned, since no single set of mentalistic data can be deemed definitive, triangulation may help to validate any findings. Care in the selection of informants is also important, especially in L2 reading, where participants reporting in L2 need to have sufficient proficiency in the language to be able concurrently to read and report in it.

In terms of methodology, there has so far been almost universal use of introspection in the spoken mode; there being no obvious reason to avoid the use of the written mode, this seems to be a promising way forward. In addition, with a view to coaxing subjects to verbalize their reading behaviour, there has been extensive use of marked texts.



There is invariably a trade-off in introspective techniques. On the one hand, there are those that constrain the reader and so make data analysis relatively straightforward (e.g., oral reading, questionnaires, structured interviews). On the other hand, there are techniques that offer relative freedom to the reader but which entail greater effort on the part of the researcher at the analysis stage (e.g., introspecting on reader-generated breakpoints, journal writing, unstructured interviews).

The first three chapters have laid the foundation for the current research: Chapter 1 surveyed reading theory, Chapter 2 reviewed reading strategies and this chapter has investigated mentalistic approaches. In the following chapter, the rationale for the current research, its context and method are described.

## **CHAPTER 4: RATIONALE, CONTEXT AND METHOD**

### **1.0 INTRODUCTION**

Chapter 1 presented the current state of theory and knowledge in the field of reading, both in the first language (L1) and second language (L2). This involved a description of the evolution of reading models and a review of research evidence about the mental processes that take place during reading. It was concluded that a major difference between adult L1 and adult L2 reading is a matter of language proficiency and that the lower the level of proficiency, the more disruption there is likely to be to fluent and accurate decoding of text (Bernhardt 1991).

Chapter 2 discussed some of the difficulties in describing reading behaviour and showed that much work remains to be done in the area. Among the research gaps identified were the need for development of a classification scheme of reading strategies that could adequately describe reading behaviour from more than one data set and the need for consideration of how to classify integrated or fused processes.

Chapter 3 considered mentalistic approaches to data-gathering, both from theoretical and practical points of view. It was shown that there was theoretical support for the tapping of 'controlled' rather than 'automatic' processes, which, in practical terms, suggested recording subjects reading and concurrently responding to problems encountered. At the same time, there was a concern to make the research task as natural as possible. An additional concern was to avoid overloading non-native readers with tasks that combine reading texts and concurrently introspecting in the L2. Areas identified for further investigation included the use of the written mode to record readers' introspections.

Having established theoretical foundations, both in terms of reading and of data collection, and outlined the nature of the problems involved in categorizing and describing reading behaviour, this chapter seeks to piece together essential elements in planning the current research study. It begins with a discussion of the gaps that the research seeks to address and a presentation of the research aims and objectives



(Section 2.0). After that, the chapter turns to a description of the research context in Hong Kong, with particular reference to language issues in society and education (Section 3.0). Finally, there is a section describing the method adopted in the current research (Section 4.0). This includes a description of the research design and the essential similarities and differences in the tasks used in the two studies conducted, henceforth referred to as 'Study 1' and 'Study 2'. It also includes an outline of the pilot-testing conducted in the current research, including the subjects who took part in it, the training materials they field-tested and the procedures adopted in so doing. In addition, the section includes a description of the subjects who participated in Studies 1 and 2, the training processes they underwent, and the procedures conducted at the sessions gathering the data for both Studies 1 and 2.

## 2.0 RATIONALE

Human information-processing theory, spearheaded by the work of Ericsson and Simon (1980), has been influential in informing methodology in the study of reading behaviour (e.g., Haastrup 1987, Feldmann and Stemmer 1987, Pritchard 1990). Since information is assumed to remain in short-term memory (STM) for about 8-10 seconds (Simon 1979), the researcher using introspective methodology to gather mentalistic data on reading processes should normally seek to tap introspection on *concurrent* processing. Following this theory, reading tasks should be designed so that both the elicitation procedure and the type of information generated by the task promote on-line introspection. Consequently, care needs to be taken in the design of the tasks to minimize interference with, or distortion of, on-task processing in STM. Although, as shown in Table 3.5 in Chapter 3, many studies have used concurrent processing, most have used tasks that included some kind of researcher intervention.

Also relevant to the quality and completeness of the data yielded is the type of information elicited. When considering the options, one should remember that, even in the L2, many reading strategies are utilized automatically and so there may be no trace of them in STM. Since automaticity does not lend itself to revelation through introspection, but controlled, conscious processing does, it is important to focus on that part of the reading process most likely to contain the latter type of processing. As I now seek to show, behaviour following an interruption to the normal interaction between reader and text appears to provide just such a suitably researchable point.

### 2.1 Breakpoints and breakpoint cycles

Cavalcanti (1987) conducted a study whose task yielded 'pause protocols', that is, protocols of her subjects introspecting at points in the text where there was a pause in their reading. Using the same highly useful concept, Hedge (1991) coined the term 'breakpoint' to describe the impact of an interruption to the flow of reading. Taking Hedge's term, 'breakpoint' is defined in this study as 'the point when the interaction between reader and text is interrupted'. It is evident that breakpoints could provide saliency to readers by having the effect of switching them, at least temporarily, from





predicted that such processes might include reading strategies utilized in attempts to deal with whatever triggered the breakpoint rather than those shown in the figure as occurring at the point marked 'start or continue reading'. There is no claim, therefore, that the reading behaviour recorded in the data will be a complete representation of the reading process. They are more likely to represent an incomplete account of behaviour following breakpoints with the additional possibility that there will be a few reports of behaviour occurring in ongoing reading.

## 2.2 Defining reading strategies

In the present study, a reading strategy is defined as 'a mentally driven action or series of related actions, serial or in parallel, taken consciously or unconsciously, to facilitate text comprehension and/or learning and which may contain an affective or evaluative element'. This definition allows for both *single* actions and *multi-stage* sets of related actions; following models of the reading process that allow for parallel processing, these actions may take place sequentially or simultaneously. While Dansereau (1985), for example, defines a strategy as "a set of processes or steps" (p. 210), there appear to be no studies with data-driven categorization schemes that reflect this definition. Instead, researchers have restricted themselves to single-step strategies, and this seems an unnecessary limitation. Furthermore, and intuitively, it seems worth investigating the possibility that there are combinations or sequences of strategies used by readers to achieve particular outcomes with a view to providing evidence of actions, previously analyzed discretely, as connected.

This notion of strategies with multiple elements that are integrated or fused will henceforth be referred to as 'compound' strategies; where there are sequences of strategies that may not be integrated or fused, they will be regarded as 'single' strategies. Allowing for both single and compound strategies leads to consideration of whether to use a term other than 'strategy'. As discussed in Chapter 2, the term has macro-level connotations in its non-academic uses and terms such as 'tactic' (e.g., Wade et al. 1990) and 'move' (Sarig 1987a and Sarig 1987b) have occasionally been used in the literature for more micro-level actions. In particular, 'move' seems a good term to use because it is neutral sounding and, unlike terms such as 'strategy' and



‘tactic’, does not have any connotations of a position in a hierarchy. However, since the term ‘strategy’ is used by the vast majority of researchers in both the learning and the reading literature, I decided, in the end, to use it here.

Apart from allowing for compound strategies, the definition of reading strategies adopted for this study accommodates certain other features. First, while mental actions are given primacy, physical actions are not precluded (though it should be noted that the stimulus for such physical actions is likely to be mental). Second, the definition allows for both conscious and unconscious actions. While it is axiomatic that research subjects may only verbalize those strategies that are conscious, this does not preclude the existence of other strategies that are neither conscious nor verbalized (e.g., Barnett 1989). In fact, as discussed in Chapter 3, it seems likely that those strategies that are verbalized probably only represent the tip of the iceberg.

In addition, the definition of strategy specifically allows for utilization of both text and non-text sources of information used in text comprehension. This seems important since there is clear evidence of the use of non-text sources of information (e.g., Pritchard 1990) and, as shown in Chapter 1, this is supported by most reading models. Further, it is recognized that, in addition to comprehension, learning has a role as a motivating factor in the utilization of strategies. Yet, with few exceptions (e.g., Davies 1995), learning and language learning are excluded from definitions of strategies. This may be because learning does not seem obviously implicated among L1 readers (especially adult L1 readers). However, it is clearly not the case with students studying in the L2, where strategies for language learning may well be utilized during reading. Indeed, in some educational contexts (e.g., Hong Kong, described in Section 3.0 below), the examination system seems likely to encourage a link in students’ minds between reading and language acquisition. Moreover, Hedge (1991) found evidence that some of her pre-sessional L2 subjects read for language acquisition as well as for meaning.

Non-text sources of information that readers bring to the process include responses to the text that are affective and evaluative, responses that are often integrated or fused with readers’ use of text sources of information. This dimension is explicitly included in the current definition, something that is consistent with Ruddell and Speaker’s

(1985) and Mathewson's (1985) models of reading, outlined in Chapter 1; nonetheless, to my knowledge, affective responses do not appear in any published definition of reading strategies.

Since strategies used at breakpoints have been isolated as the focus for this research study, it is to be expected that many of the strategies reported will represent the result of an interruption to the reading process. Unlike Kletzien's (1991) definition, however, the definition used here does not specifically require that strategies result from an interruption. One reason for this is that it would be hard to ascertain whether all strategies reported were the result of disrupted comprehension; it is quite conceivable that, once disruption has taken place, subjects might reflect on strategies used prior to a breakpoint. Another reason is that this study seeks, broadly and without unnecessary restriction, to explore and describe the subjects' mental processes during reading.

Having defined the term 'reading strategy' very broadly, and shown how, in certain respects, it can be viewed in a new light, the way has been cleared for a discussion of how the current research design covers new ground.

### **2.3 Towards a research design**

As noted in Chapter 2, a challenge that remains is to investigate whether it is feasible to devise a classification scheme for L2 reading strategies that is viable across more than one set of data derived through *concurrent* introspection. There is the additional and related challenge of attempting, under each strategy category, to formulate identical (or very similar) strategy descriptors of mental processes that are identical (or very similar) but were driven by different tasks. In other words, there is work to be done both at the higher level of strategy category and the lower level of strategy descriptor.

These challenges also have design implications for both tasks and methodology. First, in order to attempt to develop a categorization scheme that is transferable across tasks, there is a need to conduct research that incorporates task variation. One way of



achieving this is to isolate a particular aspect of the reading process (as described with respect to breakpoints above) and investigate it using more than one task. There is also the potential to vary methodology; specifically, there seems to be no good reason why there is (at least to my knowledge) only one published study that uses introspection in the written mode. As described below in Section 4.0, the current research design comprises two introspective studies, which, while closely related in certain respects, are diverse in others.

**2.4 Research aims and objectives**

The preceding discussion outlined some of the considerations in conceptualizing the current research and pinpointed particular gaps that it seeks to fill. Emanating from this discussion are the overall research aims and objectives. These are presented in Table 4.1 and discussed below.

**Table 4.1: Research aims and objectives**

Research Aim 1	To explore online mental processes among first-year university students reading short academic articles in English as a Second Language (ESL)
Research Aim 2	To establish a categorization scheme of L2 reading strategies that is consistent with current theories of reading and with data evidence
Research Objective 1	To target the readers’ mental processes in short-term memory
Research Objective 2	To develop a categorization scheme of L2 reading strategies that is viable across variation in methodology and tasks
Research Objective 3	To develop descriptors of strategies that are viable across variation in methodology and tasks
Research Objective 4	To incorporate into the categorization scheme a mechanism for mapping combinations of related strategies
Research Objective 5	To compare the findings of two related studies (i.e., Study 1 and Study 2) through use of the categorization scheme

**Research aims**

Research Aim 1: *To explore online mental processes among first-year university students reading short academic articles in English as a Second Language (ESL)*

Although some of the reading done by non-native speakers may be relatively low level and routine in terms of language and content, students at tertiary level in many parts of the world have to read vast quantities of academic articles in order to attain

academic qualifications necessary for professional development. As Bernhardt (1991) says, “These readers require extremely sophisticated literacy skills – more sophisticated than those achieved by the majority of native speakers of a language” (p. 4). Not surprisingly, therefore, many of these students have trouble with the texts they are required to read; indeed, I have personally observed this difficulty in several tertiary-level settings (Hawaii, Micronesia, Hong Kong, Japan and Britain). It therefore seemed fruitful to study reading processes among non-native-speaker university students in an English-medium university.

As observed in Chapter 3, care in the selection of informants for research involving introspection is crucial, especially when both the reading and the introspection are to take place in the L2. As Scott (1990) found, choice of subjects can have a marked effect on data quality. When comparing the data on lexical inferencing strategies yielded by subjects on an English for Specific Purposes (ESP) course with those of advanced English as a Foreign Language (EFL) subjects, he noted that the latter were more explicit in their reports. From my own experience both as a lecturer and researcher (Mahoney, Hull and Shillaw 1997) at City University of Hong Kong, it was clear to me that the students, while universally non-native speakers of English, were used to conducting most of their studies in the language. Although proficiency levels in English are widely reported to be falling in Hong Kong (e.g., *Guardian Education* 1997), I felt that university students there would be able to cope with the demands required by the proposed research.

This confidence would be especially justified with students studying for a Bachelor of Arts degree in Teaching English as a Second Language (BA TESL), entry to which carried higher language requirements than to other degree-level courses. (Details of these requirements are provided in the following section.) Most first-year university students are young adults and it seemed to me that this would be a good age for subjects, with prior training, to introspect on their mental processes while reading. Yet, since these first-year students are likely to be new to academic articles as a genre, they would also be more likely to have reading problems than second-year and final-year students. Further, being new readers of the genre, their mental processes are less likely to have become automatic than their more senior counterparts’. This combination of the greater likelihood of reading problems and the novelty of the genre



is reasonably expected to have the consequence that a greater proportion of their reading processes may be controlled (i.e., heeded in short-term memory). Thus, as discussed in Chapter 3, these readers are more likely than their senior counterparts to be aware of these processes and so be in a better position to introspect and report on them.

Although L2 university students have often been used as subjects in reading research (e.g., Block 1986a and 1986b, Grotjahn 1987, Block 1992, Parry 1996 [her Chinese subjects only]), to my knowledge, no study exists with a comparable set of aims and objectives, and certainly none conducted in Hong Kong. (A description of the Hong Kong context for this research can be found in Section 3.0 below.)

*Research Aim 2: To establish a categorization scheme of L2 reading strategies that is consistent with current theories of reading and with data evidence*

Any classification of reading strategies needs minimally to have categories that *either* reflect current reading theory *or* can be justified from the data; alternatively, they may reflect *both* theory and data. In theoretical terms, as discussed in Chapter 1, the current prevailing view, and one based on a substantial body of data, is that reading is an interactive or bottom-up interactive process. Consequently, in developing the strategy categories and, under these categories, the descriptors of individual strategies, this study should draw on cognitive psychology and applied linguistics as well as the data collected. In sum, it should be reflective of both theory and data, and, failing that, it should provide an adequate explanation for such a failure.

## Summary

The research aims have specified the context (Hong Kong), population (first-year university students) and genre (academic articles) as well as the area to be investigated (L2 reading strategies). To complement these aims, the research objectives, described below, give a more specific view of what part of the reading process is to be targeted, how the research should be designed and how the data analysis has the potential to cover new ground.

## Research objectives

Research Objective 1: *To target the readers' mental processes in short-term memory*

As explained in Chapter 3 and referred to earlier in this chapter, human information-processing theory posits that mental processing in short-term memory is more likely to be retrievable than that in long-term memory. In recognition of the strong influence of this theory on the introspective study of reading, the current research makes it a key concern to tap the readers' concurrent introspections.

Research Objective 2: *To develop a categorization scheme of L2 reading strategies that is viable across variation in methodology and tasks*

As explained in Chapter 2 and earlier in this chapter, there seems to be no study that, observing variation in methodology, attempts to use the same categorization scheme from data yielded from more than one task; hence, there appear to be about as many schemes as there are studies. This objective seeks to investigate whether it is possible to move away from this proliferation of classifications. The current study hopes to establish a categorization scheme that is robust enough to describe mental processes reported through variation in methodology and task. While this objective addresses the higher level of strategy categories, the following objective addresses the lower level of strategy descriptors.

Research Objective 3: *To develop descriptors of strategies that are viable across variation in methodology and tasks*

Not only do strategy categories need to be viable across variation in research methodology and reading tasks, but, in order to reduce the current diffusion in the description of identical mental processes, the same applies, at the lower level, to strategy descriptors. This objective seeks to establish descriptors of the same mental processes that are generalizable across tasks.

Research Objective 4: *To incorporate into the categorization scheme a mechanism for mapping combinations of related strategies*

I sought to show in both Chapter 2 and earlier in this chapter that, while there are numerous definitions of reading strategies in the literature, the concept has become somewhat fossilized in terms of viewing all mental actions as discrete. The current



research seeks to move the research forward by identifying links between or among related mental actions and describing them in ‘compound’ strategies.

*Research Objective 5: To compare the findings of two related studies (i.e., Study 1 and Study 2) through use of the categorization scheme*

As can be seen from the foregoing aims and objectives, the main focus of the current research is on the development of a categorization scheme of reading strategies. However, the function of this final objective is to establish whether the data yielded by the two studies, which are designed to be closely related but with certain key differences, can provide evidence that supports Research Objectives 2, 3 and 4 above. It will also be interesting to see if, using the categorization scheme, the two studies yield similar findings. (The findings of the two studies are presented in Chapter 6.)

Having described where the current study seeks to cover new ground and presented its aims and objectives, the next section of this chapter describes the research context with particular reference to tertiary-level education. (The subjects’ biodata, and the pilot-testing and training undertaken prior to data collection are provided in the final section of the chapter.)

### 3.0 CONTEXT: HONG KONG

Hong Kong defies simple description when it comes to language issues. In part, this may be a consequence of its history as a British colony from 1841 till 1997 and of its geographical proximity to mainland China. For well over a century, English was the sole official language in the civil service as well as the legal domain in Hong Kong. It was not until 1974 that Chinese was given the status of an official language; between then and the end of British sovereignty in 1997, efforts were made to make Hong Kong officialdom bilingual.

Luke and Richards (1982) argue that English is neither a second nor a foreign language; rather, it is an auxiliary language. They say that Hong Kong is characterized by ‘societal bilingualism’; that is, it is a community in which two monolingual communities co-exist but are socially disjunctive. Chan and Kwok (1982) refer to “an ‘additive’ form of bilingualism, where the first language is maintained and the cultural identity associated with it strengthened while the second language is being learned, in this particular case, largely for practical reasons” (p. 11). Crewe and Tong (1988) agree, saying that the motivation for learning English “for the vast majority” is “purely instrumental”. They conclude, “in short, it would be true to say that a knowledge of English is essential to a full participation in the working life of Hong Kong” (p. 370). Even as recently as 1990, Lin (1990:5) characterized English as the language of power and prestige and Cantonese as the language of intimacy and solidarity.

With changes in language policy in some areas of law and public administration, there are early indications that, under the new government of Hong Kong as a Special Administrative Region of China, the language of power is shifting from English to written Chinese and to two varieties of spoken Chinese. These are Putonghua (i.e., Mandarin, the official spoken standard variety of Chinese in mainland China) and Cantonese (i.e., the spoken variety of Chinese prevalent in Hong Kong and much of that part of South China that is adjacent to the territory). In fact, 98% of Hong Kong people speak Cantonese as their L1, a surprisingly high proportion for such an international city. Nonetheless, as Lo (1988) observes, in the Hong Kong context, it is no simple matter to define the term ‘mother tongue’. While Chinese is often



designated as the mother tongue in the territory, there is no adequate definition of the Chinese language. Indeed, in a book on the Chinese language, Norman (1988) noted that “few language names are as all-encompassing as Chinese” (p. 1).

### **3.1 Language and tertiary education**

The first tertiary-level institution, Hong Kong University, was founded in 1911 and it has always been English-medium. Indeed, until the founding of Chinese University of Hong Kong in 1963, English was the sole medium of instruction at tertiary level in the territory. Even today, Chinese University is one of only two Chinese-medium universities, the other four being English-medium. The enduring use of English is even more remarkable since it has occurred in the face of an enormous expansion in tertiary-level education in the past three decades. In the early 1970s, only 2% of young adults studied at degree level. While, by 1994, this figure had risen to 5%, by 1999, it had jumped dramatically to 18%. This rapid expansion has had consequences affecting both English and Chinese.

First, with university education no longer being a preserve of a well-educated elite, students, having been exposed at secondary level to great variation in both quantity and quality of English, often find it very hard to cope with English-medium instruction. In recognition of this problem, some universities have recently established compulsory pre-sessional English for Academic Purposes (EAP) courses. For example, in 1994, City University established a ninety-hour pre-sessional course for all successful applicants with an overall Grade E in ‘Use of English’. (The ‘Use of English’ exam is described below.) In addition, many universities have developed self-access language learning centres (e.g., City University, Hong Kong Polytechnic University, Hong Kong Baptist University).

As Flowerdew (1999) reports, a further, specific, consequence of this rapid expansion in tertiary-level education was the development, in 1991, of City University’s three-year BA TESL programme, designed to prepare students to teach English in Hong Kong secondary schools. (It is from this programme that the subjects in this research project were selected.)

However, English is not the only beneficiary of the rapid expansion of tertiary education; Chinese has been receiving increasing recognition, too. For example, the use of Chinese is now officially sanctioned by the English-medium universities for instruction in certain subjects. For example, students studying for a Bachelor's degree in Social Work (BSW) at City University are lectured in Chinese (except when they have non-Cantonese-speaking lecturers), the argument being that, upon completion of their degrees, their work will be exclusively with Cantonese-speaking clients. Nonetheless, these students are expected to take all their exams in English.

Thus, it would be wrong to suggest that there are two poles at tertiary level, exclusively English-medium and exclusively Chinese-medium instruction and study. On the one hand, Chinese University (officially Chinese-medium) resorts to extensive use of English both because required readings are often only available in English and because the institution relies on non-Chinese-speaking academic staff. On the other hand, the use of Chinese is widespread in officially English-medium institutions. Gibbons (1979) refers to the 'campus language' of Hong Kong University and, a decade later, Liu (1988) refers to a 'cocktail language' used by students at tertiary level (p. 222).

Interestingly, City University, the institution where the current research was conducted, has a policy statement about the language of instruction that seems to acknowledge the reality in the classroom. City University of Hong Kong Academic Regulation 1 (1995) states:

“The normal medium of instruction and assessment at the university is English. However, individual staff members may use Chinese at their discretion to explain or clarify particular points.” (p. 73)

Indeed, as a former lecturer at City University, I personally witnessed, and this was confirmed to me anecdotally and informally by many students there, that there is extensive use of Cantonese in lectures and tutorials. Moreover, there is no sense that this is done surreptitiously; classroom doors, often left open during lectures, are not hurriedly closed as lecturers switch to Cantonese.



### **The 'Use of English' examination**

All applicants to universities are required to take the Hong Kong Examinations Authority's 'Use of English' examination. Until 1993, the exam had four sections, each one aimed at testing one of the four skills: speaking, listening, writing and reading. In 1994, the year most of the subjects in the current study graduated from secondary school, a fifth section, called 'Practical Skills for Work and Study', was added. To arrive at a student's overall grade, scores in each of the five sections of the exam are averaged.

The 'Use of English' exam is now discussed in terms of grades required for university entry and its components; particular reference is made to components purportedly testing reading ('Reading' and 'Practical Skills for Work and Study').

#### ***University entry requirements***

To qualify for entry to university for most subjects at degree level, applicants need to obtain at least an overall Grade E in the 'Use of English'. Grade E might be viewed as falling in the range of 4.5-5.0 on 'IELTS', the test used to assess proficiency in English among non-native-speaker applicants to universities in Britain. In terms of the equivalent test in the United States, Grade E might fall in the range of 470-520 on 'TOEFL' (Test Of English as a Foreign Language), though Flowerdew et al. (1998) suggest a more precise score of 515. However, for subjects where English is particularly important (e.g., Teaching English as a Second Language), some universities in Hong Kong (e.g., City University) require an overall Grade D in 'Use of English'. This is roughly equivalent to an IELTS score of 5.0-5.5 and a TOEFL score of 520-540.

To provide a context to these requirements, the University of Glasgow requires an IELTS score of 7.0 for applicants wishing to study subjects such as law and medicine, which require a considerable amount of reading, and scores of 6.0 and 6.5 for other subjects. In the United States, many universities require a TOEFL score of 550 of non-native-speaker applicants, whatever their proposed subject of study. This represents a Grade C in the 'Use of English'. The standard of English proficiency required in Hong Kong is, therefore, considerably lower than in either Britain or the

United States, particularly the former. Indeed, this difference may actually be wider than indicated in the discussion so far. According to Flowerdew et al. (1998), in order to fill quotas, universities in Hong Kong may admit some applicants who fail to show they have reached the required proficiency in English. Unfortunately, assuming this practice occurs, Flowerdew et al. provide no details about which institutions follow this practice or whether it is more common in the hard sciences than in the social sciences.

### *Reading skills*

Since the current research focuses on reading processes, three issues that transpire from the Hong Kong Examination Authority's stipulations on reading are now highlighted for discussion. First, reading (along with listening) is referred to as a 'receptive' skill. (In contrast, writing is seen as a 'productive' skill.) This seems to reveal a disregard for current reading theory.

Second, all the questions in the reading section of the exam are in a multiple-choice format. Although this format, due to its ease of marking, is commonly utilized in exams worldwide, in reading, it has a tendency to focus on text comprehension at a very local level. In L2 reading, this may result in a situation where exams purporting to test reading proficiency are in fact testing language proficiency. Further, multiple-choice tests leave little or no room for personal reactions to texts. In an exam-oriented educational culture, where teachers' success is often judged on exam results, these considerations are likely to have a pedagogic impact.

Third, although listening, writing and speaking each have their own dedicated section in the exam format (Sections A, B and D, respectively), each of which carries a weighting of 18%, the section purportedly on reading (Section C) is divided into two parts. The first part is, indeed, devoted to reading but carries a weighting of only 6%; in contrast, the second part, with a weighting of 12%, is devoted to what is termed 'language systems'. This overt linking of reading to language acquisition in general, and grammar in particular, is, again, likely to affect pedagogic outcomes, reinforcing the tendency mentioned in relation to the use of the multiple-choice format for test items. These outcomes could hamper the development of reading proficiency as they may create a perceived, but false, reality of L2 reading as a 'receptive skill'. At the



same time, they are likely to discourage readers from reading for meaning, from evoking personal reactions to a text or from reading for pleasure in English.

As already mentioned, a new section (Section E) was added to the 'Use of English' exam in 1994. This was called 'Practical Skills for Work and Study', given a weighting of 28% and divided into three parts. One of these parts was called 'Reading Skills' and, in the current educational culture in Hong Kong, represented an ambitious attempt to encourage a more dynamic approach to reading. This involved testing for the following skills:

- scanning
- skimming
- reading for detailed information
- reading for the writer's purposes, feelings, attitudes and assumptions
- reading for inferences or predictions based on the text

This is a welcome development and it has had an impact on that part of the secondary school syllabus concerned with preparation for the 'Use of English' exam. Indeed, with the specific brief of adding exercises promoting this new set of requirements, I was recently asked by a publisher in Hong Kong to draft a new edition of a series of textbooks for 15- and 16-year-old secondary school pupils learning English (Hull 1999). However, since the change to the exam syllabus was only introduced in 1994, the year most of the subjects in the current study graduated, it is unlikely, at least until their final year, to have affected their instruction in L2 reading.

### **Attitudes to language at tertiary level**

Attitudes are an important factor in driving language policy and language use in education in Hong Kong. There is clear evidence that career prospects create strongly positive instrumental attitudes towards English. Recent research looking at attitudes among secondary school pupils (e.g., Pennington and Yue 1994), teachers (e.g., Richards et al. 1992), university students (e.g., Hyland 1996) and university lecturers (e.g., Flowerdew et al. 1998) all indicate this positive instrumental attitude.

Focusing on tertiary level, Hyland (1996) surveyed 926 students at City University and found that the three attitude statements (out of a total of twenty-five) that students most strongly agreed with were the following:

- I wish I could speak fluent and accurate English.
- I believe I will continue to need good English skills after I graduate.
- The ability to communicate in English is very important for success in my subjects at university.

The three statements (out of ten) that these students most strongly disagreed with were as follows:

- English will be less important in Hong Kong after 1997 [the year Hong Kong reverted to Chinese sovereignty].
- After 1997, all government and legal business should be carried out in Putonghua.
- English should not be a medium of instruction in Hong Kong schools.

Also at City University, Flowerdew et al. (1998) conducted a survey of Cantonese-speaking lecturers' attitudes towards the use of English in instruction. Their research was carried out between 1993 and 1996, and their twenty subjects came from a spread of disciplines in the hard and social sciences. Fifteen of the twenty lecturers expressed general support for the university's English-medium policy (described above) and the same number wanted flexibility in its application. While six lecturers thought language choice should depend on the discipline, only one believed the university should consider Chinese as a medium of instruction.

The most frequently expressed views for supporting English as a medium of instruction in the Flowerdew et al. study were instrumental in that the language helped students with their course work, examinations and future careers. The greatest problems cited were the students' low proficiency levels in English and their unwillingness to participate and ask questions in English. As far as the use of Cantonese is concerned, all the lecturers said that they sometimes used it, and several of them said they believed it helped to promote interaction about the content of the



modules. Flowerdew et al.'s findings suggest that there is a large amount of code-mixing at City University. They concluded by predicting that, for the foreseeable future, this situation will continue to reflect the sociolinguistic tensions in Hong Kong society as a whole.

### **3.2 Section summary**

The well-documented code-mixing in Hong Kong's universities (something that also occurs widely at secondary level) indicates that, whatever the official medium of the institution, first-year university students, especially those qualifying to study for a degree in TESL, are likely to be used to operating in English in educational settings. However, although these students' L2 proficiency is likely to be sufficient to meet the requirements of research tasks calling on them to report their mental actions while reading, they may well lack confidence, not least because introspection is not a technique normally utilized in Hong Kong education. Since the educational system tends to treat students as 'empty vessels' who need to be filled with information for exam purposes, they are likely to be unfamiliar with the role of 'informant'; nonetheless, it is expected that training in introspection can modify this perception. From this, it seems reasonable to conclude that Hong Kong provides a suitable context for the current research.

## **4.0 METHOD**

So far, this chapter has presented a rationale for the current research and described its context. It now turns to the method adopted. First, research and task design are addressed. Following this, there are descriptions of how the pilot tests and Studies 1 and 2 were conducted; each of these descriptions includes biodata on the thirty-eight subjects who eventually participated in the studies (N=38), the training process they underwent and the procedures adopted during data-gathering. Overall, the section attempts to answer Research Questions 4.4 and 4.5 on, respectively, effective ways of using introspective data-gathering techniques and the ability of L2 subjects concurrently to read and introspect. (The research questions are listed at the beginning of the chapter.)

### **4.1 Development of the research design**

This section is divided into two, the first part focusing on the design of the research project and the second part on the design of the tasks. In both parts, reference is made to the theoretical and practical issues considered in Chapter 3 in relation to the use of mentalistic data.

#### **Research design**

There was an initial design of the current study that was far too ambitious in scope for a PhD thesis and several of the original research tasks administered to the Hong Kong subjects subsequently had to be dropped from the thesis. This initial design included a questionnaire whose main purpose was to obtain a general idea of the reading habits (including strategies) of first-year social science students at City University of Hong Kong. The idea was that the quantitative data yielded would complement the qualitative data yielded by the other tasks. The questionnaire was pilot-tested several times and its final version was administered to 298 students at the university, including the thirty-eight used in Studies 1 and 2. In addition, the research task devised for Study 1 had three sections, only one of which is described in detail in



Chapter 5; furthermore, this three-section task was followed by a short task asking the thirty-eight subjects to write retrospectively about their reactions to it.

Following the cloze task in Study 2, which is presented in Chapter 6, the original design had three additional tasks for the thirty-eight subjects. The first, conducted during the data-gathering sessions described below, sought the subjects' reactions to various reading tasks commonly used in L2 reading syllabi. The second and third were both administered on different days from the data-gathering sessions described below. One of these sought the subjects' reactions to participating in the entire research project while the other was Bigg's 'Study Process Questionnaire' (Biggs 1992). In addition, twelve of the thirty-eight subjects were interviewed retrospectively.

When it became clear to me that the original plan would have to be considerably reduced, my principal criterion for retaining tasks was the extent to which they focused on my main research interest, that is, the study of mental processes in L2 reading. This led me to select part of the written introspection task in what I am calling 'Study 1' and the oral introspective cloze task in 'Study 2'.

As shown in Table 4.2, text genre (academic articles), data type (mentalistic) and introspection type, that is, expected timing of introspection (concurrent with reading) are held constant across Studies 1 and 2. Care was taken to ensure that the informants would find the texts manageable and on topics of direct interest to their field of study, thus avoiding frustration and obviating the need for researcher intervention. The texts are discussed in detail in the respective studies in Chapters 5 and 6.

**Table 4.2: Features of research design**

Features	Study 1	Study 2
Text genre	Short academic article	Short academic article
Data type	Mentalistic	Mentalistic
Introspection type	Concurrent with reading	Concurrent with reading
Introspection mode	Written	Spoken
Task mode	Individual, silent	Pair, oral
Breakpoint type	Reader-generated	Researcher-generated

In addition to the features that are held constant, the table shows three that vary across the tasks in the two studies. This follows Research Objective 2, presented above. The first variation involves introspective mode. While the task in Study 1 employs written introspection, that in Study 2 calls for oral introspection. A further distinction concerns task mode. The Study 1 task replicates, as much as possible, normal uninterrupted reading, requiring the informants to read a text individually and silently; in contrast, the task in Study 2 requires the informants to do a 'talk-aloud' activity in pairs. Although the latter task entails audio tape-recording the subjects, it was hoped that any perceived intrusiveness of the recording equipment on their part would be mitigated by the fact that they were working in pairs. Third, the design incorporates two types of 'breakpoint' (defined above). Reflecting the relative naturalness of the reading task in Study 1, the readers are asked to report spontaneously on the breakpoints they experience; the cloze task in Study 2, on the other hand, presents the readers with what amount to simulated breakpoints. In other words, the breakpoints in Study 1 are *reader-generated* whilst those in Study 2 are *researcher-generated*.

At a wider level, the research design included two further key elements (described below). First, the informants, both in the pilot studies and in Studies 1 and 2, underwent a training process; second, all the tasks were field-tested.

In sum, the research was designed to comprise two parallel studies, with genre, data type and timing of introspection common to them both but with variation in introspective mode, task mode and breakpoint type. Other features of the tasks are now addressed.

### **Task design**

As already discussed in Chapter 3 and under Research Objective 1 above, human information-processing theory suggests that mental processes are stored in short-term memory for 8-10 seconds and that, while these processes remain there, they are more easily retrievable than when they have transferred to long-term memory.



This had three implications for the current design. First, both tasks were designed to tap concurrent introspection (i.e., while the mental processes are in short-term memory) as opposed to retrospection (i.e., when they have passed to long-term memory). In the case of Study 1, this involved writing about mental processes as they occur; and, in the case of Study 2, the informants interacted orally in pairs as they were completing a cloze task. Clearly, in neither case would it be possible to guarantee that retrospective information was excluded entirely.

Second, tapping mental processes that are concurrent with breakpoints presupposes that informants can manage the tasks with little or no researcher intervention. Hence, clear and uncomplicated tasks were adopted in both studies. As already stated, in Study 1, informants were asked to read silently and introspect in writing on their breakpoints. In Study 2, they were asked to talk aloud in pairs about a cloze task; cloze is a very familiar procedure for Hong Kong students, though the introspective aspect of the task would be new to them.

Third, mental processes relating to interrupted reading are, according to information-processing theory, more likely to be heeded than those relating to ongoing reading. Consequently, focussing the research on breakpoints (i.e., where there is a problem or issue to be confronted, or simply a need for reflection) is likely to promote heeded processing.

## **4.2 Pilot-testing**

Having described the design of the research and tasks, I will now move on to various aspects of pilot-testing. First, there is a description of the subjects who participated in the pilot studies; this is followed by descriptions of the training materials that were pilot-tested by some of the pilot subjects and the procedures adopted.

## Subjects

In all, there were four pilot studies, each with different participants. These are now described.

### *Pilot Study 1*

Pilot Study 1 consisted of thirty subjects (N=30). They were all attending a pre-session intensive programme run by the Language Institute at City University of Hong Kong in the late summer of 1994. They had all just completed secondary school and had been accepted by the Applied Social Sciences Department for three years of full-time study leading to a Bachelor's degree in Social Work (BSW). They were obliged to attend this ninety-hour programme before they could start their first year of study because they had scored Grade E in the 'Use of English' exam (discussed above). Students with Grades A-D were exempted from the programme; applicants failing the exam (i.e., with a grade lower than Grade E) were, at least officially, ineligible for study at degree level. As mentioned above, at a rough estimate, a 'Use of English' Grade E would be in the region of 4.5-5.0 on the 'IELTS' test and a 'TOEFL' score of 470-520, though Flowerdew et al. (1998) are more specific about the latter, estimating it at about 515.

### *Pilot Study 2*

Partly in an effort to attract more participants for further pilot-testing, I decided to offer a seven-week Language Enhancement Course on 'Reflective Reading' at City University of Hong Kong. Classes took place between the seventh and thirteenth weeks of the sixteen-week autumn semester of 1994, were held once a week and lasted two hours. (The Language Enhancement Course was offered for additional reasons not relevant to the current research; these included my wish to try out some instructional materials for L2 reading.). Language Institute lecturers at the university are invited to design courses targeting specific study and/or language development work for which students can volunteer. A key feature of these courses is the small class size, usually about ten students in each group. Students are accepted on a first-come first-served basis, typically resulting in courses composed of students studying both social sciences and hard sciences and in any year of study.



Posters advertising the proposed fourteen-hour course on 'Reflective Reading' were placed around City University and forty-five volunteers signed up. Notification that they could attend the course was sent to the first fifteen students; for research purposes, this group was designated as Pilot Study 2. Experience had shown high attrition rates from Language Enhancement Courses, so fifteen, rather than the usual ten, students were chosen; there were fourteen participants (N=14) in the first session.

Pilot Study 2 was neither a pedagogic nor a research success. As a voluntary course, it started too late in the semester (approximately at mid-point) to maintain reasonable attendance levels. Further, due to the first-come first-served selection basis, there was little cohesion to the group, the students' real and perceived study needs being very disparate.

### ***Pilot Studies 3 & 4***

At the beginning of the spring semester of 1995, I decided to run two more pilot studies. For Pilot Study 3 and Pilot Study 4, I contacted ten students who had signed up for the Language Enhancement Course but, because more students had volunteered than could be accepted, had not been selected to attend the original course. Learning from the problems experienced in Pilot Study 2, two selection criteria were established for Pilot Studies 3 and 4. The first criterion was that participants should be first-year or second-year students of the humanities, thus narrowing the subject range to one nearer to that of the intended subjects in Study 1 and Study 2. The second criterion was that they could volunteer to do eight hours as research subjects. The intention here was to have the students commit themselves to a given amount of time as research subjects. After much negotiation over the logistics, an arrangement was made for a total of eight students fulfilling these criteria to meet in two small groups (Pilot Study 3, N=6; Pilot Study 4, N=2), for two hours a week for four weeks.

### **Training materials**

The training programme comprised three stages. Two of the stages involved tasks taken from Hosenfeld et al.'s (1981) curricular sequence for teaching reading strategies, and, for the third stage, a follow-up rational cloze passage was developed. (This can be found in Appendix 4.1.)

In the first of Hosenfeld et al.'s two tasks, subjects were asked to add two sets of numbers and to be in a position afterwards to report *how* they had added each set. The numbers in one set were random whereas, in the other, there was a regular pattern. Hosenfeld et al. argued that the use of numbers instead of words provides a relatively easy start to the introspection process.

For the second task, 'nonsense words' (i.e., invented words, not real English words) replaced the original words in a short text. In a sense, this acted as a precursor to the cloze task since the nonsense words, by forcing breakpoints in the reading process, acted in a similar way to cloze blanks. I asked subjects to identify the six nonsense words, to try and work out an appropriate real word, and to report how they arrived at their answers.

The third and final stage in the introspective training programme was a twenty-item rational cloze task. Content words were deleted in the same way as for the cloze task in Study 2, and the text comprised 333 words, providing an average deletion frequency rate of one word in every seventeen (1:17). (The cloze task in Study 2 is described in Chapter 6 along with discussion of deletion rates in recent studies using the cloze procedure.)

Taken from Scollon and Wong Scollon (1995:122-123), the text was on the topic of intercultural communication and stereotyping. Hong Kong being an international city, this topic seemed likely to be one with which the subjects could readily identify. The text describes and discusses a fictitious, but ostensibly typical, misunderstanding between people from different cultures. Two businessmen, an American and a Hong Kong Chinese, are meeting for the first time. The text opens with a short conversation between the two men; this is followed by two paragraphs, the first describing the American's reaction to the meeting and the second contrasting this impression with that of the Hong Kong Chinese.



## Training procedures

The sequence of training materials described here was pilot-tested on the Pilot Study 2 subjects in the autumn semester of 1994 and again, following minor changes, on the subjects in Pilot Studies 3 and 4 at the beginning of the spring semester of 1995. No further changes were made to the materials following Pilot Study 4. For all pilot studies, subjects worked in pairs for the entire training sequence. This lasted two hours, including time for feedback on each of the three stages and a ten-minute break at midpoint, the first two stages being conducted in the first hour and the third stage in the second hour.

When reporting how they had approached the first task, it is clear that most of the pilot subjects employed a variety of strategies to add each set of numbers; in addition, a considerable range of strategies was used among the subjects to add the numbers with a regular pattern. Participants also reported using a number of different strategies in the second task.

The cloze task appeared to be at a suitable level of difficulty for these subjects, and they took about twenty minutes to complete it. In most cases, they became quite animated as they progressed through the passage. They were observed to consider alternative words for the blanks, to give reasons for their choices, and to discuss the pros and cons of one another's suggestions.

Overall, the training programme seemed to raise the participants' awareness of what introspection involved; however, they often reported simply *feeling* that a cloze item should be filled with a certain word and were unable to be more specific about their reasoning. Although this appeared to presage a similar loss of data in the main studies, it was clear that a considerable amount of data were nonetheless being generated.

Also field-tested by participants in the pilot studies were the research tasks specific to Study 1 and Study 2 (described, respectively, in Chapters 5 and 6). This part of the pilot-testing process clearly indicated that it was feasible to ask these subjects concurrently to read and introspect in the L2.

Having described the pilot-testing conducted in the current research, the way has been cleared for descriptions of the participants in Study 1 and Study 2, the training they underwent and the procedures undertaken during the data-gathering sessions in which the subjects completed the research tasks.

### 4.3 Study 1 and Study 2

This section moves from pilot-testing to aspects of Study 1 and Study 2 that were common to both studies, namely, the subjects and the procedures used for the training materials and for data-gathering.

#### Subjects

Thirty-nine first-year students (N=39) were selected as informants for the main part of the current research (i.e., Study 1 and Study 2). They were doing a BA TESL (Teaching English as a Second Language) degree at City University of Hong Kong. Their biodata, presented in this section, were gathered as part of the questionnaire survey whose entire findings were originally intended to be included in this thesis but were, as mentioned above, eventually excluded due to lack of space. The biodata were collected during the same week as the subjects were participating in the training programme already described.

The participants' age range was 19-38; most of them were between 19 and 23 years old, though there was an outlier aged 38 (S36). Since students entering the BA TESL degree course need a minimum overall Grade D in the 'Use of English' exam, their level of English was considerably higher than that of the pilot subjects. Frequencies of the BA TESL subjects' 'Use of English' grades are listed in Table 4.3 below. (Details of the 'Use of English' exam, including specific mention of the 'Reading' and 'Practical Skills for Work and Study' papers, are given above.)



As shown in Table 4.3, some scores were not available (N/A). Three of the four missing ‘Reading’ scores and one of the two missing ‘Practical Skills for Work and Study’ scores were a result of subjects having forgotten them. However, one missing score in each of the three categories in the table were a result of the oldest participant (S36) not having taken any part of the exam. Like the other subjects, his L1 was Chinese (Cantonese) and he began his secondary education in Hong Kong, but, unlike his peers, he completed his schooling in Canada and then went on to university there, taking a Bachelor’s degree in Mathematics and History. He then remained in Canada until shortly before he applied to take the BA TESL degree at City University of Hong Kong. His proficiency level in English was noticeably higher than that of the other subjects, yet he was clearly a non-native speaker of English.

**Table 4.3: Subjects’ ‘Use of English’ scores**

Grade	Reading	Practical	Overall
A	3	1	0
B	3	6	1
C	9	12	11
D	16	15	26
E	4	2	0
F	0	1	0
N/A	4	2	1
Total	39	39	39

During the spring semester of their first year at City University, BA TESL students attend a module called ‘Reading Skills’ (module number EN0305); this is one of four modules designed to improve language skills early in the degree course. Mindful that Hong Kong university students are already required to work very hard and so tend to resist extra demands made on them, the module lecturer, Mr Lindsay Miller, made a suggestion. He wanted the current research to be conducted in a manner designed to link it as closely as possible in the students' minds with the coursework; in other words, to avoid attrition, he felt that the research needed face validity for his students. Fortunately, since the module syllabus includes reading strategies, this was relatively easy to achieve.

Students were therefore told that, instead of a scheduled assignment reading and critiquing a novel, they would write a short paper giving their impressions of being subjects in the current research and evaluating the experience. (Due to space

constraints, these evaluations, as previously stated, are not reported in this thesis.) Moreover, though the research would require some work during out-of-class time, the preparatory training programme would be carried out during a regularly scheduled class period. It is this training that I now describe.

### **Training materials and procedures**

The training was conducted in the sixth week of the spring semester in 1995, and each subject attended one training session in their usual classroom during class time and in their normal class groups. There were two groups, one with twenty participants and one with nineteen (N=39). Each training session consisted of the same tasks as those described for the pilot tests and they were administered in an identical manner as described above. All the subjects worked in pairs except for three, who worked together as a small group. There was a ten-minute break at midpoint in the two-hour session.

In the first hour, I began by briefly eliciting from the subjects reading strategies they could remember having discussed with their regular lecturer during the first week of the semester; next, I gave the subjects, and asked them to complete, the two tasks from Hosenfeld et al. (1981). (These can be found in Appendix 4.1.)

In the second hour, I gave the subjects the training cloze task developed from the Scollon and Wong Scollon (1995) text, which is in Appendix 4.1. During the 20-25 minutes they took to complete the task, I observed that, after a brief initial period of silence, during which the subjects skim read the text, there was plenty of pair interaction. During the class-report phase, I elicited strategies which subjects had used for some of the items and wrote some of them on the board. It was evident that the subjects had really engaged with the task and had used both global and local strategies. For example, Item 6, at the beginning of one paragraph, could only be guessed after reading the beginning of the following paragraph. Many subjects reported changing their original answer for this item (based only on local reading) once they had read ahead (global reading).



### **Data-gathering procedures**

Starting the same week as the training outlined above, the subjects attended data-gathering sessions for Study 1 and Study 2. (Since the data for both studies were gathered in the same sessions, the procedures are described here rather than separately in Chapters 5 and 6, which describe the distinct aspects of the two studies.) The subjects came in four groups on consecutive Thursday afternoons during Weeks 6-9 of the spring semester of 1995. Three groups each consisted of ten participants (i.e., five pairs) and one group consisted of eight participants (i.e., four pairs) (N=38). One of the original thirty-nine subjects was lost due to illness. (His overall 'Use of English' score was Grade D and his scores on both the 'Reading' and 'Practical Skills for Work and Study' papers were Grade C.) In fact, since one of the tasks involved pair work (the oral introspection task for Study 2), it was logistically more convenient to have an even number of informants than an odd number.

The venue for the data-gathering sessions and the schedule of activities are now described.

#### ***Venue***

As the subjects were to be tape-recorded for the first time (for the oral introspection task) and as there was a logistical need to record several pairs simultaneously, the location used for data-gathering for Studies 1 and 2 was different from that used with these subjects in the training sequence. The sessions were conducted in a large room which, prior to the subjects' arrival, had been divided into four sections by soundproof curtains. These were drawn to enable each pair of subjects to work without distraction (i.e., out of earshot and out of sight of other pairs). City University students regularly attended classes in these rooms, which were usually used for video-taping various activities.

However, since the first three data-gathering sessions involved recording five pairs simultaneously for the oral introspective task, two of the pairs were accommodated in two curtained-off sections in a similar adjacent room, leaving three pairs in the original room. (In each case, these two pairs returned to the original room in the second hour of the session.) For the final data-gathering session, there were only four

pairs and so each pair was accommodated in one of the four sections of the original room.

The curtained-off sections prepared for conducting the oral introspective pair task for Study 2 resembled small recording studios. They each contained two chairs, set side-by-side, and a table, on which there was a microphone connected to a cassette recorder. The recorders were placed out of sight on the floor nearby. Before the subjects arrived, two copies of the text for the oral introspection task were placed on each table, one copy for each participant. As the subjects arrived, they sat down in pairs of their own choosing. Sitting side-by-side in this manner for pair activities is a familiar configuration for students at City University.

### *Schedule*

Table 4.4 shows the overall sequencing of each of the two-hour sessions. I began each of the sessions with a brief introduction to the schedule and the types of activity that the subjects would do. I reminded them of the training sequence of introspective tasks they had previously undertaken and told them that the tasks in hand were similar in nature.

**Table 4.4: Schedule of data-gathering sessions**

Duration (minutes)	Activity
5	Introduction to session and to first hour
35	Oral introspection task (Study 2) (Appendix 6.1)
5	Retrospective task (not reported in this thesis)
10	Break
5	Introduction to second hour
35	Written introspection task (Study 1) (Appendix 5.2)
25	Reading activities task (not reported in this thesis)

Next, there was an introduction to the oral introspection cloze task for Study 2. When the subjects were ready to start, the researcher went round from pair to pair to switch on the cassette recorders. Some pairs started talking immediately whereas others began by skim reading the text, and some pairs were far more animated than others. Except in the first few minutes, none of the subjects appeared concerned about being tape-recorded, and they all seemed comfortable with pauses in the oral interaction. At the end of the allotted time, all except one of the nineteen pairs had completed the



twenty-five cloze items; indeed, many pairs had had sufficient time to reread the text one or more times in an attempt to complete or rework some of the items.

At this point, I gave the subjects a four-item survey (not reported in this thesis) calling for their retrospective comments on the oral introspection task. This took them no more than five minutes to complete.

During the ten-minute break, I retracted the soundproof curtains and removed from sight the audio tape-recording equipment. I did this in order to facilitate administration of the written introspection task for Study 1. It also provided a more normal classroom environment, and subjects previously working in the adjacent room were able to join their peers for the second hour.

Following the ten-minute break, I gave a brief introduction to the second hour, which was devoted to two tasks. The first of these was the written introspection task for Study 1, and the second was a study on reading activities (not reported in this thesis). I administered the written introspection task as follows:

- I gave the subjects a copy of the task (Appendix 5.2) and the accompanying text (Appendix 5.1) and a minute to leaf through the task.
- I ran through the instructions to each section of the task and asked if there were any questions.
- I gave the subjects four minutes at the beginning for the 'skim-reading' section and four minutes at the end for the 'post-reading' section of the task, the remainder of the allotted time (27 minutes) being devoted to the 'study-reading' section (Chapter 5).
- After thirty-five minutes, I collected the subjects' copies of the task, and the subjects retained their copies of the text for a follow-up task on reading activities (not reported in this study).

### *Sequencing*

Apart from the duration of each stage in the data-gathering sessions, Table 4.4 above also shows their sequence, one aspect of which should be explained. There were several reasons why I administered the oral cloze task for Study 2 before the written

task for Study 1. First, following Seliger and Shohamy (1989:157), I conducted the task with relatively high explicitness (i.e., the cloze task with its simulated breakpoints) before the task with relatively low explicitness (i.e., the written task with spontaneous breakpoints). This created the possibility that the subjects would be relatively schematized for the individual and silent written task. Second, the cloze task resembled the third and final stage in the training programme, which was also a cloze task (described above), and therefore provided a better sequel to the training materials.

Further, getting the subjects to start with the interactive pair task, in which the pilot-testing and the training materials had shown they would quickly become focussed and animated, seemed a good way to prepare them to settle into the more intense and private written task. Fourth, and related, given the need to set up audio tape-recorders and microphones for the cloze task in advance of the subjects' arrival, it was logistically more convenient; also, as already stated, the recording equipment could be swiftly removed from sight during the break.

#### 4.4 Section summary

This section described the research design, which incorporates two studies (Studies 1 and 2) using introspective data-gathering methodology and outlined the pilot-testing and training in introspection that were undertaken.

The current design holds constant, across the two mentalistic studies, text genre (short academic articles) and timing of introspection (concurrent with reading). However, while Study 1 utilizes the *written* mode of introspection with an individual and silent reading task, Study 2 utilizes the *spoken* mode with a pair task. Also distinct across the two studies is the type of 'breakpoint' featured in the task, with Study 1 using a *reader-generated* variety and Study 2 using a *researcher-generated* variety. Many of the design decisions with respect both to the research as a whole (triangulation and training) and to its tasks (a focus on concurrent introspection and problem-solving along with the principle of minimal researcher intervention) were informed by



information-processing theory. It remains to be seen precisely how effective this design using mentalistic approaches turned out to be.

During the course of the four pilot studies and administration of the proposed training materials, it became clear that all the Hong Kong subjects were able to read and concurrently introspect in the L2. Details of the tasks used in Studies 1 and 2, and how they were modified during pilot-testing to improve their data-yielding capacity, are provided, respectively, in Chapter 5 and Chapter 6.

## 5.0 CHAPTER SUMMARY

This chapter has built on the first three, which, respectively, provided a theoretical foundation for the current research (Chapter 1), surveyed the literature on reading strategies (Chapter 2) and discussed mentalistic approaches to the study of reading behaviour (Chapter 3). The chapter began with a rationale for the research, outlining gaps in the L2 reading knowledge base. It described how interruptions to the normal flow of reading, or ‘breakpoints’, may provide an optimal point in the process to tap reading behaviour and provided a broad definition of the term ‘reading strategy’ that reflected both current reading theory and data evidence.

Following presentation of the research aims and objectives, there was a description of the research context, providing an outline of issues relating to language, including policy, use and attitudes in Hong Kong society; language issues at tertiary level, the setting of the current research, were also considered. I sought to show that Hong Kong university students, particularly those in their first year of study, would provide excellent subjects for the nature of the research projected.

The chapter continued with a description of the research and task design and how introspection could be effectively used to gather data. It ended with a description of the materials used to train subjects in introspection and of the subjects and procedures both of the pilot studies and the two main studies. This process gave clear confirmation of my hunch, mentioned above, that these Hong Kong university students would provide rich L2 mentalistic data, yielded concurrently with reading academic articles in an L2.

The following chapters describe the two studies, Study 1 in Chapter 5 and Study 2 in Chapter 6.



## CHAPTER 5: STUDY 1

### 1.0 INTRODUCTION

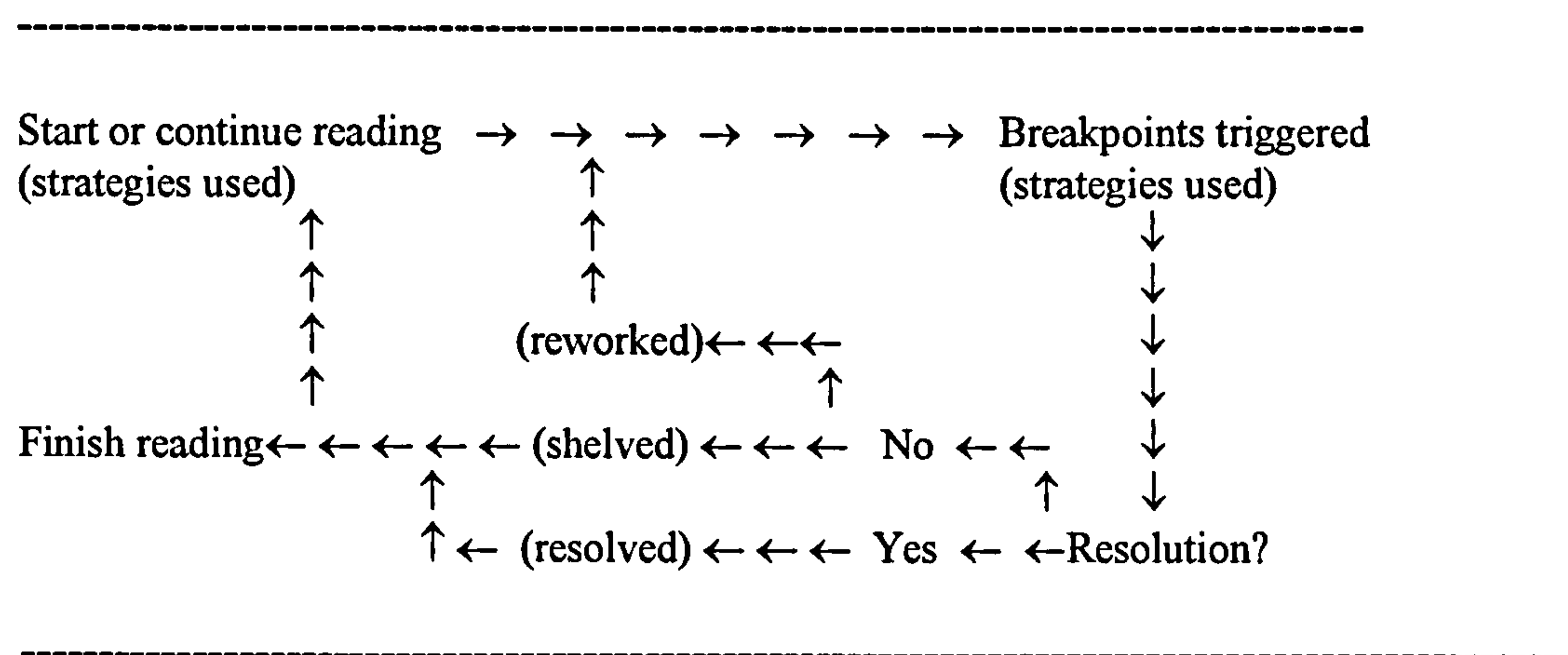
This chapter presents Study 1, the first of two related studies conducted in Hong Kong on a group of first-year students on a BA TESL degree course at City University of Hong Kong. (Study 2 is reported in Chapter 6.)

The research aims and objectives were presented and explained in Chapter 4. Also explained in Chapter 4 was the research design. This holds constant, across the two studies, the text genre utilized (short academic articles), the nature of the data collected (mentalistic) and the timing of introspection (concurrent with reading). However, there were three key differences between the two studies. In Study 1, the informants introspected in the written mode whereas, in Study 2, they did so in the spoken mode. In addition, Study 1 had an individual and silent reading task while Study 2 had an oral task conducted in pairs. Finally, the breakpoints in Study 1 were spontaneous (i.e., reader-generated); in contrast, those in Study 2 were simulated (i.e., researcher-generated).

In addition, Chapter 4 established a definition of the term reading strategies. They were defined as ‘mentally driven actions or series of related actions, serial or in parallel, taken consciously or unconsciously, to facilitate text comprehension and/or learning and which may contain an affective or evaluative element’. It will also be recalled from Chapter 4 that the concept of the ‘breakpoint cycle’ was posited to account for the reader’s behaviour between breakpoint trigger and breakpoint resolution, much of which is expected to involve the use of reading strategies. The original figure graphically representing the breakpoint cycle is reproduced here as Figure 5.1.

Essentially, the figure proposes that, once a breakpoint has been triggered, the reader, in an attempt to resolve it, uses a strategy or strategies. If there is a resolution to the breakpoint as a consequence of the use of a particular strategy or strategies, the reader either continues or finishes reading; alternatively, if there is no resolution, the reader can opt to rework the breakpoint or shelve it and either continue or finish reading. In

both cases, ongoing reading will involve the reader using strategies, but these strategies can only be tapped *retrospectively*; in contrast, strategies used following breakpoints can be tapped *concurrently*. As with all mentalistic data, it is highly unlikely that subjects' reports, irrespective of temporal considerations, will amount to a complete representation of all the processes that actually occurred. Further, since breakpoints represent the processing routes readers may take in responding to *needs*, such as time to reflect on the text, as well as *problems*, such as failure to comprehend it, resolution may be a matter of having reflected long enough or of having solved a problem.



**Figure 5.1: Proposed processing sequences in breakpoint cycles**

### Note for Figure 5.1

As will be shown, reading purpose, where known, will be attached to strategies used in breakpoint cycles.

In addition to a focus on reading strategies, my original plan for Study 1 was to include an analysis of factors reported by the subjects as triggering their breakpoints; however, due to space constraints, this part of the study was deleted from the final draft of the thesis. This chapter begins with a description of the development of the task used to collect the written data; this includes a description of the text selected for use with the task and the pilot-testing undertaken. Next, the chapter describes the related processes of data analysis and development of a classification scheme of reading strategies. (A comparison of findings for Studies 1 and 2 can be found at the end of Chapter 6.)



## **2.0 DEVELOPMENT OF THE TASK**

This section describes the text selected for this study and the development of the task used in conjunction with it. First, the criteria for selection of a text are presented; the text chosen is then assessed against these criteria. This is followed by a description of the pilot-testing of the task and a description of its final version.

### **2.1 Selection of the text**

Reference was made in Chapter 4 to a questionnaire survey which, though originally intended as part of this study, was dropped due to space constraints. This survey was conducted on 298 first-year students of the social sciences at City University, including the thirty-eight participants in the two studies reported in this thesis. Two of the survey's findings relating to genres are worth reporting here. First, the informants reported greater exposure to academic articles in their L2 (English) than in their L1; second, notwithstanding this exposure, the informants generally reported experiencing difficulty with academic articles in English. In sum, these findings provide clear evidence of Hong Kong university students' need, at least in the social sciences, to read a considerable number of academic articles and the related need to improve their reading skills with respect to this genre.

For the current research, I therefore decided to use an academic article. However, in order to ensure that the choice of text would not be too difficult for the intended subjects, and mindful of the difficulties they experience with the genre, the following selection criteria were developed. The article should:

- be on a topic of professional interest to the subjects;
- contain sections clearly marked by headings;
- contain elements typical of academic articles, such as figures or tables, bibliographic references and an appendix;
- be relatively free of long sentences, complex syntax and lexical density;
- be relatively brief, preferably under ten pages of A4 size paper.

These criteria are now discussed and justified in the sequence listed above.

An article by Farmer (1994) entitled ‘The limits of learner independence in Hong Kong’, which can be found in Appendix 5.1, was selected since it met most, or arguably all, of the criteria above. It appeared in a collection of articles on the topic of self-access language learning. On the back cover, the volume is described as “a practical guide and a source of reference for preparing students for self-access learning [and] of interest to those in the field of TEFL [Teaching English as a Foreign Language] and applied linguistics”.

Since the topic of the article is directly related to the subjects’ field of study, Teaching English as a Second Language (TESL), it clearly met the first criterion in the list above. In addition, I thought the highly local context of the research described in the article would enhance its interest and relevance to students so often asked to read academic texts written by westerners and set in western educational contexts. Using subjects quite similar to those in the current study, the article describes research on the development of a self-access centre at Hong Kong Polytechnic University, a former polytechnic with a similar history to that of City University. Indeed, City University has its own self-access centre, which some of the current subjects had used.

The article contains a range of salient textual features including headings and subheadings, brief quotations from university students describing their reactions to using a self-access centre, a table showing questionnaire findings, emboldened text, highlighting the content and chronology of a training programme, as well as references and an appendix. Intuitively, I felt that these features would tend to reduce the subjects’ sense of text density and, relative to some academic articles, would thus make the text appear to them to be a manageable challenge. In addition, there appears to be a preponderance of fairly high frequency vocabulary, and, since the content concerns the subjects’ field of study, I felt that many of the lower frequency vocabulary items in the text (e.g., *self-access centre*, *learner-centred approach*) might be familiar to them. It should be said that this impression of the vocabulary load the subjects would face in the text was based on my own professional judgment rather than on a more rigorous calculation of the lexical density of the text.



In terms of length, the article covers four pages of A4 size paper and consists of about 2,500 words; in addition, there are over three pages of references and appendices. L2 reading research has often been conducted on subjects at tertiary level using texts that are considerably shorter than the academic articles that university students are actually required to read, especially those reading the humanities. For example, working with pre-sessional and/or university-level subjects, Bachman (1985), Block (1986b, 1992), Sarig (1987a, 1987b), Jonz (1987), Hedge (1991), Davis and Bistodeau (1993), all utilized texts comprising well under 1,000 words.

Researchers often make use of very short texts in reading research for good reasons. For example, looking at certain linguistic features of text and the cognitive processes of language comprehension, Jonz (1987) used texts of about 750 words in length for cloze tests. Indeed, the cloze text utilized in Study 2 of the current research is of similar length. Clearly, using cloze with an academic article of average length would be unwieldy for both readers and researchers. It is worth adding that many researchers do not state the word counts of their texts. Also, two recent studies (Li and Munby 1996 and Jimenez 1997) asked their subjects to select their own reading materials, and in neither case were word counts for these texts recorded.

In sum, the article combines a familiar topic likely to be of both personal and professional interest to the subjects, with relative accessibility and relatively simple use of language. In this study, it was used with a research task designed to tap readers' introspections. The pilot-testing of this task and its final version are now described.

## **2.2 Pilot-testing**

As noted in Chapter 3, the written introspection task used with the Farmer text is a very close adaptation of one developed by Davies at the University of Bristol. (The task can be found in Appendix 5.2.) As described in Chapter 4, a slightly redrafted version of this task was pilot-tested on the subjects in Pilot Studies 3 and 4 early in the spring semester of 1995. Participants in Pilot Study 3 met on Thursdays and those in Pilot Study 4 on Fridays, thus enabling fine-tuning of the task in the twenty-four hours between the two sessions.

Using the Farmer text described above, the task was first run on Pilot Study 3; however, while participants said they enjoyed the text, the whole procedure took them well over an hour. An attempt was then made to shorten the task and combine it with a shortened version of another task concurrently being pilot-tested. (For space reasons, this other task, which focussed on reading tasks commonly used by teachers of L1 and L2 reading, was excluded from this thesis.)

However, this combination of the two tasks, when administered to the Pilot Study 4 subjects, proved too cumbersome; it required approximately the same length of time as the two tasks had taken separately. Bearing in mind the very limited contact time available for conducting research with the subjects, I decided to revert to the original single-focus task for the main study. The task was fine-tuned in an attempt to reduce and simplify what was required of the subjects.

Specifically, instead of asking the subjects in the 'study-reading' section of the task (described below) to introspect at ten breakpoints, as in the original, I reduced the number to only four breakpoints. In addition, I numbered every fifth line of the text. This feature meant that the subjects' writing was consolidated and confined to the task sheet; previously, the subjects had been required to mark the texts with the breakpoint number (i.e., 1-4) and write their introspections on the task sheet.

### **2.3 The written introspection task**

Following pilot-testing, the final version of the written introspection task comprised three sections: 'Read 1: Skim read'; 'Read 2: Study read'; and 'After reading' (Appendix 5.2). However, for space reasons, only the findings of 'Read 2: Study read' are given in this thesis. Nevertheless, as described under data-gathering procedures in Chapter 4, the subjects did all three sections of the task at one sitting. Therefore, I will now briefly describe the entire task.



### **The ‘skim-reading’ section**

Skim reading is defined by Lunzer and Gardner (1979, quoted in Davies 1995:137) as “a rapid style used mainly to establish what a text is about before deciding where to read”. The ‘skim-reading’ section of the task contained three questions. The first question was open-ended and asked the subjects, having skim read the text, for their initial reactions to it. The second question contained a five-point readability scale, followed by four lines for the subjects to write the reasons for their ratings. The third question asked the subjects if they expected to enjoy a more careful read of the text by ticking ‘Yes’, ‘Not sure’ or ‘No’ and giving reasons for their answers.

### **The ‘study-reading’ section**

Study reading is closely akin to Lunzer and Gardner's ‘reflective reading’. As reported by Davies (1995:137), this is “frequently interrupted by moments of reflection”. This may include periods when the reader is searching backwards or forwards within the text, or it may include periods of reflection that break the reader's interaction with the text (cf. Harri-Augstein and Thomas's 1984 ‘search read’ and ‘thinking read’).

The ‘study-reading’ section of the task began by defining ‘breakpoints’ as “places where you stop reading for any reason” and citing the following examples of possible reasons why readers might experience breakpoints: “you don’t understand something, you want to think or even rest”. While the term ‘breakpoint’ may be metalanguage, I considered it was necessary to have a single easily recognizable term to describe this crucial aspect of the research, and one whose meaning was unambiguous to the subjects. The subjects were asked to read the text at their own pace and report on four breakpoints they experienced while reading; specifically, they were asked to write what they were thinking and, if possible, to specify what strategy or strategies they used following each breakpoint.

My intention was that this section of the written introspection task would provide as natural a reading experience as possible. This is notwithstanding the fact that it was

actually part of a research project, that the text was researcher-selected and that the preceding skim-reading section of the task (described above) involved researcher-generated reading. The study-reading section of the task required individual and silent reading, the subjects were left completely undisturbed as they did it and both the topic and genre were familiar to them. While seeking to preserve a natural reading experience, this section of the written introspection task sought to conform to practical considerations in the design of tasks gathering mentalistic data.

Specifically, it sought to elicit from the informants introspection that was *concurrent* (i.e., while in short-term memory) rather than *retrospective* (i.e., when it has passed to long-term memory). It did this, in part, by adhering, as explained in Chapter 3, to the principle of minimal researcher intervention and by focussing on problem-solving in reading. Needless to say, since the task gave subjects free rein in conducting the task, there can be no guarantee that all the data recorded from it would qualify as concurrent.

### **The ‘after-reading’ section**

In order to avoid the possibility, though a remote one, of having to explain to the subjects the meaning of the term ‘post-reading’, to describe this section of the task, I used the term ‘after-reading’. It refers to the readers’ written responses immediately after finishing reading the text. Subjects were asked to consider their reactions to the first two sections of the introspective task (i.e., the ‘skim-reading’ and ‘study-reading’ sections) and to write about any changes in their reactions to the text since writing about their initial reactions while skim reading. They were also given four blank lines in which to record any other comments they might have had.



### 3.0 DATA ANALYSIS

In order to facilitate description of the lengthy process of data analysis, the following are presented at the outset: an initial grouping of reading strategies, the final version of the categorization scheme for the strategies and the coding principles adopted. There then follows a detailed description of the process of analyzing the data for reading strategies. Essentially, while bearing current reading theory in mind, this involved numerous passes through the data and the repeated ebb and flow between emerging higher-level strategy categories and lower-level strategy descriptors.

In the discussion, there are numerous illustrative extracts from subjects' protocols. Immediately preceding all such extracts, capital S followed by a number refers to subject number (e.g., S5 refers to Subject 5) and capital L followed by a number refers to a line number in the text (e.g., L286 refers to Line 286). Thus, 'S5: L286' would precede an extract from Subject 5's protocol that refers to Line 286 of the text. Where an extract is reproduced with omissions from the original entry, this is shown by three consecutive full stops (i.e., ...).

#### **Initial strategy groups**

Table 5.1 presents the five initial strategy groups, all subsequently to be modified in some way, which were developed at an early stage in data analysis. The 'reflecting' group contains strategies with no stated outcome and, as the sample descriptors indicate, involves text meaning, text language and extra-textual factors. In contrast, the 'making inferences' group contains strategies with a stated outcome (e.g., inferencing the meaning of a word); however, it encompasses the same textual and non-textual factors as the first group. The strategies in the 'reflecting' and 'making inferences' groups also have in common the fact that they are unobservable mental actions.

On the other hand, as the label suggests, the 'sensory' group contains strategies that require eye contact with the text (or, in the case of blind readers using Braille, finger contact) and are thus potentially observable; these include such physical actions as

rereading and skipping text. The ‘motivating’ group is distinct from the previous three groups in that it describes reading purpose and involves a minimum of two strategies, that is to say, the use of one strategy *in order to* use another strategy. Following Oppenheim’s (1992) recommendation, there is also a ‘miscellaneous’ group, containing strategies awaiting more specific categorization.

Finally, the table shows how variation in parts of text and in aspect of language reported are incorporated into strategy descriptors.

**Table 5.1: Initial strategy groups**

Initial strategy group	Sample descriptors
‘reflecting’	<ul style="list-style-type: none"><li>- Reflecting on meaning in [part of] text</li><li>- Reflecting on [aspect of] language in [part of] text</li><li>- Reflecting on own knowledge of content of [part of] text</li></ul>
‘making inferences’	<ul style="list-style-type: none"><li>- Making an inference about meaning in [part of] text</li><li>- Making an inference about [aspect of] language in [part of] text</li><li>- Making an inference from own knowledge of content of [part of] text</li></ul>
‘sensory’	<ul style="list-style-type: none"><li>- Rereading [part of] text</li><li>- Reading [part of] text</li><li>- Skipping [part of] text</li><li>- Skimming [part of] text</li><li>- Scanning [part of] text</li><li>- Underlining [part of] text</li><li>- Resuming reading [part of] text</li></ul>
‘motivating’	Using another strategy <ul style="list-style-type: none"><li>- to reflect on meaning in [part of] text</li><li>- to clarify meaning in [part of] text</li></ul>
‘miscellaneous’	<ul style="list-style-type: none"><li>- Memorizing information in [part of] text</li><li>- Predicting information in [part of] text</li></ul>

**Note for Table 5.1: Signalling working and final categories**  
For the sake of clarity, initial groups and interim categories will be indicated by the use of single quotation marks (‘’) and lower case initial letters; capitalization of initial letters will only be used for the final categories and subcategories.

Examples will be given of each of the three levels (category, subcategory and descriptor groups) during the discussion below.



**Final categorization of reading strategies**

Table 5.2 presents the final categorization of L2 reading strategies. It shows that, following an interim stage, the five initial groups presented above were reduced to two categories. *Reflective* Strategies, which accounts for reports of unobservable actions, has four subcategories. The subcategories ‘Text Meaning-Centred’ and ‘Text Language-Centred’ represent what are assumed to be distinct areas of text processing while the subcategory ‘Extra-Text-Centred’ captures evidence of top-down processing. The subcategory ‘Non-Specific’ was created for reports where there was a lack of clarity as to whether there was a focus on text meaning, text language or beyond the text.

**Table 5.2: Final categorization of L2 reading strategies**

Category	Subcategories	Descriptor groups
Reflective	Text Meaning-Centred	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Remembering (Study 1 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Text Language-Centred	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Extra-Text-Centred*	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Expressing personal reactions (Study 1 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Non-Specific (Study 2 only)	<ul style="list-style-type: none"><li>- Making inferences (Study 2 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
Sensory	Ongoing Reading	<ul style="list-style-type: none"><li>- Reading</li></ul>
	Recursive Reading	<ul style="list-style-type: none"><li>- Rereading</li></ul>
	Skipping Reading	<ul style="list-style-type: none"><li>- Skipping</li></ul>
	Annotated Reading*	<ul style="list-style-type: none"><li>- Underlining</li></ul>

**Note for Table 5.2**  
The subcategories marked with an asterisk (\*) occurred in ‘compound’ strategies only. These are strategies with two or more elements that are integrated or fused.

*Sensory* Strategies, which accounts for observable actions, also has four subcategories. These were based on reports and observations of how the subjects read the text. The subcategory ‘Ongoing Reading’ refers to evidence of normal reading whereas ‘Recursive Reading’ refers to rereading and ‘Skipping Reading’ refers to reading ahead while omitting some of the text, as in skimming and scanning. Finally,

‘Annotated Reading’ captures such written evidence as underlining or highlighting the text.

Among the key changes between the initial and final stages are the following:

- the final categories are based on a binary divide between reading behaviour which is unobservable (i.e., *Reflective* Strategies) and that which is observable (i.e., *Sensory* Strategies);
- processing of text meaning, text language and extra-textual factors become subcategories (instead of descriptors);
- ‘reflecting’ and ‘making inferences’ become descriptor groups (instead of initial strategy groups);
- the status of reading purpose is changed from strategy to *motivator* adjunct; that is, where known, purpose is attached to a strategy.

Furthermore, Table 5.2 shows variation across Studies 1 and 2. While the overall categories remain the same, the Non-Specific subcategory of Reflective Strategies is only reported in Study 2. As far as descriptor groups are concerned, there are two groups that were only reported in Study 1, ‘expressing personal reactions’ and ‘remembering’. In addition, ‘requesting clarification’ is driven by pair interaction; and, in the current research, this means it is confined to Study 2, which has a pair task. Finally, it should be noted that no formal analysis of Sensory Strategies was made in Study 2, though informal analysis revealed the same subcategories and descriptor groups as shown in the table.

The evolution of these strategy categories and the change in status accorded to reading purpose are described below. First, however, the coding principles that informed the development of the categorization scheme are presented.



**Coding principles**

In the following description of the process of developing a categorization of reading strategies, reference is made to five coding principles that were developed and refined during data analysis. These principles are presented in Table 5.3.

**Table 5.3: Coding principles for analysis of reading strategies (Study 1)**

5.1	The strategy categories should function at the same level as one another, as should the subcategories.
5.2	The strategy categories and subcategories should reflect current prevailing theory in reading; failing that, they should provide grounding for new theory.
5.3	All strategies should be coded as discrete processes unless elements are demonstrably integrated or fused into ‘compound’ processes or linked to a ‘motivator’ explicitly stated in the data [see Note below].
5.4	The principle of minimalism should be observed in the development of strategy descriptors. To help achieve this, ‘core’ strategy descriptors with variants indicating any detailed particulars reported, such as part of text or aspect of language or language use, should be established [see Note below]. In addition, efforts should be made to harmonize the wording of core strategies with the objective of representing the same reading behaviours in the same strategy descriptors.
5.5	The strategy descriptors should be a faithful representation of the data evidence. Any potential ambiguities in their wording should be avoided through definitions specified for the current study; conversely, any ambiguity in the data should be reflected in the coding.

**Note for Table 5.3: Terminology**

- ‘Compound’ strategies have two or more elements that are integrated or fused.
- ‘Motivators’ are reading purposes attached to strategies.
- ‘Core’ strategies are descriptors that allow for variation in terms of part of text and aspect of language reported.

The remainder of this section describes the process of developing a categorization scheme of L2 reading strategies. The discussion begins with the process of establishing strategy descriptors, that is, consistent terminology for the strategies emerging during repeated passes through the data (Section 3.1). Next, it turns to a description of how these descriptors were placed into initial groups of strategies (Sections 3.2 and 3.3). It continues with a description of how these groups were placed into interim categories (Sections 3.4 and 3.5) and how these, in turn, were then reorganized into the final categorization scheme (Sections 3.6 and 3.7).

3.1 Establishing strategy descriptors

First, each protocol was checked for breakpoint triggers, that is, factors reported by subjects as leading to breakpoints. (Due to space constraints, the study focusing on breakpoint triggers, as mentioned above, was deleted from the thesis.) As shown in the box below, each breakpoint trigger was summarized with its code in square brackets and reproduction of the protocol extract that began with a note of the subject number and line number. Corresponding strategy descriptors were then juxtaposed between the breakpoint triggers and protocol extracts. The example shows S7’s and S33’s protocols for the same line in the text (Line 1):

Breakpoint trigger	Strategy descriptor	Protocol extract
“embodies” [vocabulary item]	Rereading sentence to reflect on word meaning	S7: L1 “Because come across the word 'embodies' which I don't understand: I'll read the sentence once again.”
“embodies” [vocabulary item]	Rereading sentence to reflect on word meaning	S33: L1 “I am not sure about the meaning 'embodies', so I stop a while & reread from the beginning.”

As can be seen from the protocols, these two informants used different wording to describe the same breakpoint trigger and the same strategy. Nonetheless, I thought that, in order to facilitate the later stages of analysis, the strategy descriptors should have identical terminology for identical actions (Coding Principle 5.5). Thus, S7’s “read ... once again” and S33’s “reread” were both recorded as ‘reread’. Also, while S7 reread “the sentence” and S33 reread “from the beginning”, both were recorded as rereading the sentence. A quick look at the relevant point in the text, which can be found in Appendix 5.1, reveals that S33 can only have meant that she reread the sentence. In a similar vein, the term ‘reflect on word meaning’ was used for both entries above since the evidence suggests that the purpose behind both subjects’ reading was to think again about the meaning of the word ‘embodies’. (The notion of reading purpose will be discussed again at several points in this section.)

While an effort was made to harmonize the wording of the strategy descriptors in this way, at the same time, I sought to retain as many individual characteristics as possible. Nonetheless, although a balance between harmonizing terminology and reflecting



individual nuances needs to be observed, this is not always a straightforward process. One of the difficulties was that informants used certain terms that rather loosely describe strategies, and there was the distinct possibility that the meaning they attributed to these terms differed among subjects. In other words, two subjects may use the same terminology for different meanings or different terminology for the same meanings.

For instance, Hong Kong students tend to use the terms ‘skimming’ (i.e., rapid reading with no specific end in mind) and ‘scanning’ (i.e., rapid reading with a specific end in mind) interchangeably, sometimes collocating them as though they were indivisible strategies. In this case, I felt the only option was to assume that the informants were reporting their actions accurately and thus I coded their entries according to what they actually wrote (Coding Principle 5.5). Naturally, though, what informants report having done may not represent what they actually did; even so, it would clearly be entering a minefield to assume that a subject meant to report one of these strategies while actually reporting the other. Evidence to the contrary might be gained through retrospective interviews.

Another commonly used word in the protocols was ‘guessing’ (e.g., guessing word meaning). In some protocols, this was variously used *with a stated outcome* (which may or may not have been correct) or *without a stated outcome*. In other protocols, a verb other than ‘guess’, but covering some of its possible meanings, was used (e.g., consider; reflect on; think about; think of; catch, get or figure out the meaning). In still other protocols, the notion of guessing was conveyed through a question. For instance, when introspecting about the phrase “the body of learners” in the text, S9 wrote:

S9: L12

“Does it has [sic] another meaning other than ‘body’ (refers to biological meaning)?”

I eventually decided to differentiate in strategy descriptors between guessing with a stated outcome and guessing without a stated outcome, a differentiation illustrated, respectively, in the following extracts:

S8: L73

“‘momentum’ is a physical term. In physics, it means something like ‘impact between 2 things’. As it related [sic] to ‘speed’, I guess is as ‘pace’.”

S3: L22

“I am thinking why the word ‘service’ is marked by ‘ ’. So I try to guess if it has any special meaning behind the word. I reread it again.”

In the first extract, the subject is recorded as using the strategy ‘making an inference’ about meaning whereas in the second, the subject is recorded as ‘reflecting’ on language and meaning in text. Thus, the principle observed was that the maximum amount of explicit evidence of strategies to be found in the data was represented in the descriptors (Coding Principle 5.5).

Moreover, to ensure as far as possible that the terminology used to describe the strategies was reliable (i.e., that it was used consistently), potentially ambiguous or confusing terms, such as ‘guessing’, had to be defined for this study. Examples of definitions that were finally adopted are presented in Table 5.4 along with illustrative strategy descriptors. (Examples of strategy descriptors juxtaposed with the protocol extracts that inform them are given later in this discussion.)

**Table 5.4: Meanings covered in strategy descriptors**

Meanings covered	Terminology adopted	Strategy descriptor
Guessing, predicting, thinking about or considering with no stated outcome	Reflecting on	Reflecting on language in text
Guessing, predicting, thinking about or considering with a stated outcome	Making an inference	Making an inference about meaning in text
Clarifying or checking	Clarifying	*Using strategy to clarify meaning in text
Reading again, reading recursively	Rereading	Rereading text
Resuming normal reading (i.e., in the same manner as reading prior to the breakpoint)	Resuming reading	**Resuming reading
Main idea	Key point	Underlining key points in text
Paragraph, sentence, etc., currently being read	Paragraph, sentence, etc.	Reflecting on meaning in text [sentence]

**Note for Table 5.4**

- \* Reading purpose, shown in this descriptor, was later to be disqualified as a ‘strategy’ and was to be termed a ‘motivator’.
- \*\* This strategy descriptor was not included in the final categorization scheme.



The process of ensuring that the wording in the strategy descriptors was consistent required that the data were checked against the descriptors several times. To ensure continued accuracy of the wording, changes were only made with item-by-item reference to the data. During this process, patterns in wording evolved for coding the strategies. For example, words such as 'reflect' were used to the exclusion of other words with the same meaning; likewise, the term 'key points' was used to cover 'main ideas', the two terms having hitherto been used in free variation (Table 5.4).

### Reading purpose

Apart from individual terms such as those discussed above, I monitored emerging descriptors for whether they represented, on the one hand, mental or physical actions reported during reading or, on the other hand, the purpose behind those actions (Coding Principle 5.5). By far the majority of cases in the data were reports of actions, and these were described with gerunds (verb + *ing*); the small number that reported reading purpose were described with infinitives (*to* + verb). In the instance below, the descriptor begins with the action (rereading) and ends with the purpose (to clarify meaning):

- *rereading sentence to clarify meaning*

A case in point was S14, who wrote two entries with a very subtle difference in wording that was reflected in the corresponding strategy descriptors, as shown in the box below. In the first of these descriptors, the first stage of the strategy is followed by the conjunction 'and', which indicates some sort of connection between the discrete parts of the strategy, a connection that may, in reality, be extremely tenuous. In contrast, the second of the descriptors below has an infinitive, indicating that the second stage of the strategy describes the purpose of the first stage. Reference to the protocol extracts shows that, in the first breakpoint, the subject reported, albeit in the present tense and with no stated outcome, that she reread and guessed the meaning of the unfamiliar word. In contrast, in the second breakpoint, she reported that she

reread with the purpose (“in order to”) of guessing the word meaning; again, she did not actually report guessing the meaning, though it is possible that she did.

Strategy descriptor	Protocol extract
Rereading word and reflecting on meaning of word	S14: L17 “The word implemented. I skip it first and then I reread the whole sentence. I guess its meaning.”
Rereading text to reflect on meaning of word	S14: L23-34 “constitute – I don’t know its meaning. I also skip it and reread it in order to guess its meaning.”

Arguably, the behaviour S14 reported following these two breakpoints was, in fact, identical, but it would clearly be unsafe to assume so. Since I decided to avoid making assumptions from the data, this fine terminological distinction in strategy descriptors was retained (Coding Principle 5.5). (This distinction between reported action and purpose is pursued below.)

**3.2 From descriptors to initial strategy groups**

In order to establish a tentative categorization of reading strategies, the strategies were now grouped on the basis of the initial verb in each descriptor. The resulting lists of strategies yielded working categories that were essentially data-driven. The verbs that appeared in the initial list were: reflecting, making inferences, predicting, linking ideas, memorizing, rereading, reading, skipping, skimming, underlining, stopping reading, resuming reading, turning (i.e., the page).

Once grouped by initial verb in this way, the strategy descriptors in each group were then sequenced into logical patterns. For example, strategies grouped under ‘rereading’ were sequenced from relatively global text to relatively local text reported, as follows:

- rereading section
- rereading previous paragraph
- rereading paragraph
- rereading previous sentence
- rereading sentence



- rereading quotation
- rereading clause
- rereading unfamiliar word/phrase
- rereading unspecified part of text

From these lists of strategies grouped by initial verbs, it was immediately clear that there were several ‘core’ strategies (i.e., strategies with several specific variants). In the list of descriptors above, the core strategy was ‘rereading’ and the variants were parts of the text reported (Coding Principle 5.4). The nine strategies in the list above could therefore be consolidated into one core strategy, with variations in part of text utilized, thus:

- rereading text [section, previous paragraph, paragraph, etc.]

Subsequently, it became clear that the notion of the core strategy could be extended to language-centred strategies, whose variants in aspect of language reported could, in a similar way, be consolidated, thus avoiding unnecessary proliferation of strategies. Therefore, the following four strategies became two core strategies, the first pair in the ‘reflecting’ group and the second pair in the ‘making inferences’ group:

- reflecting on language [use of quotation marks] in text [sentence]
- reflecting on language [pronoun reference] in text [clause]
- making an inference about language [use of article] in text [phrase]
- making an inference about language [collocation] in text [phrase]

3.3 Initial strategy groups

The strategies listed by initial verbs were then placed into these five broad groups:

- reflecting
- making inferences
- sensory
- motivating
- miscellaneous

Below, strategy descriptors in each of the initial groups are presented in boxes, with variation in part of text and aspect of language reported shown in square brackets; these are juxtaposed with illustrative protocol extracts. Precise wording of some of the strategy descriptors was to be modified later.

*Initial strategy group: ‘reflecting’*

The first group contained strategies that involve ‘reflecting’ (Table 5.4). This group does not appear to have any exact precedent in the literature. It corresponds to parts of two of Sarig’s (1987b) categories (‘clarification and simplification’ and ‘coherence detection’) and to two of Jimenez et al.’s (1996) categories (‘text-initiated’ and ‘reader-initiated’). When contrasted with Davies’s (1995) categories, which were developed from an extensive review of the literature, it corresponds to parts of three of them (‘monitor reading process’, ‘utilize source of information: textual’ and ‘utilize source of information: external/background knowledge’). As shown in the box below, the strategies record evidence, respectively, of mental actions focusing on text meaning, text language and extra-textual factors (e.g., prior knowledge or personal experience of text content):

Strategy descriptor	Protocol extract
Reflecting on meaning in text [word]	S3: L14 “‘redress’ is a strange word to me. I am trying to think of its meaning ...”
Reflecting on language [pronoun reference] in text [unspecified]	S16: L34 “I am thinking of the meaning of ‘It’ ...”
Reflecting on relationship between personal experience and text topic	S25: L11 “I want to think that what [sic] type of learners I belong to.”



While the first extract is relatively straightforward, the second requires some explanation. Although S16 uses the word ‘meaning’, she is actually reflecting on the language issue of pronoun reference. In the third extract, S25 is referring to the issue raised in the text of variation in teacher dependence among learners in Hong Kong.

**Initial strategy group: ‘making inferences’**

The second group of strategies also records reports of mental actions, begins with the term ‘making an inference’ and involves guessing with a stated outcome (discussed above; also, Table 5.4). In all the studies reviewed, this group corresponds either to a strategy category (e.g., Jimenez et al.’s 1996 ‘interactive’ category) or at least one strategy descriptor (e.g., Kletzien’s 1991 ‘making an inference or drawing conclusions’ descriptor). As the examples in the box below show, there was clear evidence in the data of inferencing involving the same three-way distinction noted in the ‘reflecting’ group, that is, respectively, meaning, language and prior knowledge or experience of text content:

Strategy descriptor	Protocol extract
Making an inference about meaning in text [word]	S15: L66 “The word ‘laxity’ stops me from reading it. I guess that it means something very important or necessary.”
Making an inference about language [grammatical category] in text [word]	S8: L1 “‘embodies’ is a verb undoubtedly ...”
Making an inference from prior knowledge [personal experience] of content of text [unspecified]	S32: L2 “Reflecting: I try to think about how a small group learning [sic] is useful for studying English. To my experience, classmates always take the small group learning as a releasing chance as most of them think he/she can depend on the other one in team work.”

**Initial strategy group: ‘sensory’**

In contrast to the unobservable strategies in the first two groups (i.e., ‘reflecting’ and ‘making inferences’), the third group contains potentially observable actions. Several studies have the same, or a very similar, category (e.g., Sarig 1987a, Wade et al. 1990, Davies 1995); indeed, this seems to be the most stable category in the literature. However, because of the multiplicity of initial verbs describing such actions (rereading, reading, skipping, skimming, underlining, breaking, turning the page), I was constrained to depart from the idea of labelling initial strategy groups by their

initial verb. Thus, in order to differentiate this group from the others, I named it ‘sensory’; some of the strategies in the group are exemplified in the box below:

Strategy descriptor	Protocol extract
Rereading text [sentence]	S36: L1 “... Went back to the beginning of the sentence and started to read the second time.”
Reading text [unspecified]	S27: L95 “...reading ...”
Skim reading text [two sections (90 lines)]	S27: L95 “... go to the evaluation and conclusion [sections of the text]. I also skim some headings before evaluation part.”
Underlining text [key points in section]	S6: L18 “After reading a section (i.e. Introduction), I will break to underline some relatively important ideas.”

***Initial strategy group: ‘motivating’***

Fourth, the data yielded strategies expressing reading purpose; at this point, these were classified as the ‘motivating’ group. There are no direct parallels in the literature for this group, though Davies (1995:54-55) illustrates an example of her ‘utilize source of information: textual’ category which, arguably, could fall into this category. As shown in the box below, motivation came from two reported sources, respectively, text meaning and text language. However, there is an essential difference between strategies in this fourth group and those in the previous three groups (i.e., ‘reflecting’, ‘making inferences’ and ‘sensory’). While ‘motivating’ strategies need to be attached to a strategy in one of the other groups, the converse is not the case. In other words, ‘motivating’ strategies cannot stand alone whereas strategies in the other three groups can.

In order to distinguish strategies containing one discrete element (i.e., a physical or mental action) from those with two or more integrated or fused elements, such as those with an action and a purpose, the latter were considered ‘compound’ strategies. (The status of the ‘motivating’ group and the concept of the compound strategy are discussed later in this description of data analysis.) In the following instances of ‘motivating’ strategies, the part expressing purpose is highlighted in italics:



Strategy descriptor	Protocol extract
Rereading text [sentence] <i>to reflect on meaning in text [word and sentence]</i>	S3: L14 “‘redress’ is a strange word to me ... So I reread it again. After reading the whole sentence, I try to catch the meaning of this word [redress]. Or even if I can’t, I will get the meaning of the whole sentence.”
Rereading text [sentence] <i>to reflect on language [sentence length/complexity] in text [sentence]</i>	S33: L15 “The whole sentence is too long & too complicated. I need to read it again.”

**Initial strategy group: ‘miscellaneous’**

In addition, following Oppenheim (1992), I established a ‘miscellaneous’ group containing a small number of strategies, each of which occurred at low frequencies compared with the strategies in the other four groups and which would not fit into any of the other groups as currently constituted. Examples of strategies in this group are:

- memorizing or remembering key ideas in text
- making predictions about text
- expressing affective and evaluative reactions to text
- checking length of part of text [paragraph]
- stopping reading

**Summary**

The situation at this point in the process of data analysis is summarized in Table 5.5 below.

**Table 5.5: Initial strategy groups**

Initial strategy group	Description
‘reflecting’	Strategies with no stated outcome, focussing on unobservable mental actions concerning text meaning, text language and the readers’ prior knowledge or experience of text content
‘making inferences’	Strategies with a stated outcome, focussing on unobservable mental actions concerning text meaning, text language and the readers’ prior knowledge or experience of text content
‘sensory’	Strategies that involve potentially observable physical actions (usually eye movements, such as those involved in rereading) which are directed by associated, but unobservable, mental actions
‘motivating’	Strategies that express the purpose of other strategies, these discrete elements being contained in ‘compound’ strategies
‘miscellaneous’	A collection of strategies awaiting further analysis

Key concepts developed so far in the description of reading behaviour were ‘core’ strategies and ‘compound’ strategies. The former avoid potentially lengthy lists of very similar descriptors by incorporating reported variation in part of text and aspect of language while the latter allow, at this stage, for reading behaviour and related reading purpose to be described in the same strategy.

As will now be shown, some of these five initial strategy groups were to be placed into an interim set of categories and one group, following reference to the current study’s definition of reading strategies, was disqualified and given a new status.

### **3.4 From initial groups to interim strategy categories**

Having established the initial strategy groups, which were based on strategy descriptors devised from data evidence, it was time to consider whether there were grounds to modify them and, if so, whether the modification needed to be radical or relatively superficial. To provide anchor points for this process, I conducted it by checking the initial strategy groups against the following:

- the definition of reading strategies adopted in the current study, which is ‘mentally driven actions or series of actions, serial or in parallel, taken consciously or unconsciously, to facilitate text comprehension and/or learning and which may contain an affective or evaluative element’;
- the breakpoint cycle, which posits readers’ processing routes;
- the interface between current reading theory and data evidence.

As a result of this procedure, and with reference to the literature, the initial strategy groups were modified in several important ways. I will now describe these modifications and argue a case for each of them. The description begins with how reading purpose was coded and categorized and continues with a discussion of the need to distinguish between process and outcome.



## Reading purpose

The strategies in the ‘motivating’ group, one of the five initial strategy groups, describe reading purpose and, in so doing, these strategies are distinct from those in all the other groups, all of which describe actions. Moreover, the ‘motivating’ strategies are invariably ‘compounds’, that is, they are attached to other strategies.

However, reference to the current study’s definition of reading strategies shows that they do not in fact qualify as strategies at all. Specifically, reading purpose is qualitatively distinct from reading behaviour (or, as the definition puts it, “mentally driven actions or series of related actions”); in contrast, purpose motivates or drives reading behaviour, wholly or in part. As reading purpose does not function at the same level as reading behaviour, according them both the same status infringes the coding principle that all strategy categories should function at the same level (Coding Principle 5.1).

In order to avoid losing the information about reading purpose that had already been retrieved from the data, I accorded this information a new status by ‘attaching’ purpose, where recorded, to a strategy and considering it as identifiably separate from the strategy, or as an ‘adjunct’ to it. In this study, the term ‘motivator’ will henceforth be used to describe recorded purpose in reading and will replace the former ‘motivating’ strategy group. Thus, in the following descriptor, the strategy (rereading text) has a motivator (to clarify meaning in text) attached to it:

- rereading text [sentence] to clarify meaning in text [word]

As a consequence of this change of status for reading purpose, discrete strategies with motivators attached, such as the one above, could not qualify as *compound* strategies. (Compounds, which comprise more than one inextricably linked strategy, will be referred to again later in this discussion.)

Very little work has been done on investigating reading purpose. In her very broad ‘text-level’ category, Barnett (1988a) includes ‘reading purpose’ as one of the strategies along with such diverse behaviour as ‘considering background knowledge’,

‘predicting’, ‘skimming’ and ‘scanning’. One of the few researchers actively to have considered reading purpose as a distinct entity is Hedge (personal communication 1999). In order to describe various types of reading purpose, Hedge, in unpublished work following her 1991 study, developed the notion of what she called ‘drivers’. From her data, she identified four categories of driver: meaning, gist, language acquisition, and a ‘dual driver’ combining meaning and language acquisition. As Davies (1995:78) observes, the view that purpose can ‘drive’ the reading process has theoretical support as it is consistent with the role of the ‘message centre’ in Rumelhart’s (1977) model and that of the ‘thematic processor’ in Rayner and Pollatsek’s (1989) model.

At this point, I checked the data again for evidence of categories of motivators. While the former initial strategy group had had two categories (i.e., text meaning and text language), this more focused check yielded evidence of four categories of motivator:

- Text Meaning-Driven
- Text Language-Driven
- Text Meaning-Driven/Text Language-Driven (an integrated or fused combination, henceforth referred to as a ‘compound motivator’)
- Text Non-Specific

The use of the word ‘driven’ following all the categories except ‘Text Non-Specific’ is one that acknowledges the limitations of mentalistic data. As has been reiterated through the thesis, in reporting their mental processes, subjects are unlikely, either intentionally or unintentionally, to give an all-inclusive picture of what actually took place. Since these are final (though tentative) categories, I have presented them with initial capital letters.

The box below presents examples of each of the categories of motivator. In the first example, the subject’s purpose in skipping an unspecified amount of the text is clearly to understand the meaning of a specific part of it while the subject in the second extract cites a language-driven purpose in rereading part of the text. The third example illustrates a compound motivator in that the subject cited both meaning and



language as motivating her strategy of rereading. The last extract in the box shows an example of where an unclear purpose is reported (Coding Principle 5.5).

Motivator category	‘Strategy: Motivator’ descriptor	Protocol extract
Text Meaning-Driven	Skipping text ahead [unspecified] to search for meaning in text [clause]	S10: L95 “I stop reading at the phrases [sic] ‘have to date been drawn’ for I don’t know what it means. I skipped it the same reason as breakpoint 1 [I think there may be some explanations afterwards or it’s not an important word [phrase] for understanding the main ideas of this text].”
Text Language-Driven	Rereading text [sentence] to clarify language [sentence length] in text [sentence]	S33: L15 “The whole sentence is too long & too complicated. I need to read it again.”
Text Meaning-Driven/Text Language-Driven	Rereading text [sentence] to clarify meaning and language [sentence length/complexity] in text [sentence]	S31: L2 “rereading because the sentence is too long for me to understand.”
Text Non-Specific	Rereading text [sentence] to clarify unspecified factor(s) in text [sentence]	S3: L7 “I am thinking the sentence structure of this sentence. It is so strange to have ‘talking to speak a lot’. So I rereading [sic] the sentence again, see what happen to this student.”

Parenthetically, as the examples in the box indicate, all cases of motivators occurred with strategies in the initial strategy group labelled ‘sensory’. The lack of motivators recorded with other strategy groups will be taken up again later in this discussion. In addition, as will be shown in Chapter 6, no cases of motivators were recorded in the cloze data in Study 2.

While the Text Meaning-Driven motivator appears to equate with Hedge’s ‘meaning’ driver, the other categories of motivator appear somewhat different from her other categories of driver. While Hedge’s ‘dual driver’ links meaning and *language acquisition*, my ‘compound motivator’ links text meaning and *text language*. It may well be that Hedge’s subjects, some of whom were pre-sessional L2 postgraduates studying in Britain, were keen to use reading partly to acquire language and partly to understand the text. In contrast, the Hong Kong subjects, who were undergraduates studying in an English-medium institution within their own L1 environment, may

have had the more straightforward motivation of decoding language (i.e., English as a Second Language) whenever it impaired their understanding of the text. Clearly, there is plenty of scope in this area for future research.

In sum, to differentiate between reading behaviour and reading purpose, the former ‘motivating’ category of strategies, which indicated purpose, was given a new status and a new name. The status became that of ‘adjunct’ to reading strategies and the name became ‘motivator’. The data yielded four categories of motivator: Text Meaning-Driven, Text Language-Driven, Text Meaning-Driven /Text Language-Driven (a compound category) and Text Non-Specific. Thus, when I referred to the current definition of the term ‘reading strategy’, I realized that I had to discontinue one of the five initial strategy groups. In order to assess the viability of the other initial strategy groups, I now turn to the remaining anchor points mentioned above: breakpoint cycles and the theory/data evidence interface.

### **Process over outcome**

Breakpoint cycles propose various processing routes that readers might take following breakpoints in reading, the routes taken in the latter stages of the cycle being partly dependent on the outcomes of processing earlier in the cycle (Figure 5.1, Section 1.0). Juxtaposing this concept against the remaining four initial strategy groups (i.e., ‘reflecting’, ‘making inferences’, ‘sensory’ and ‘miscellaneous’) revealed an incongruity between *process* and *outcome* in two of the groups. While reading strategies were represented, both in the current definition and in their role in breakpoint cycles, as describing processes, the initial strategy groups ‘reflecting’ and ‘making inferences’ were based on differences in outcomes. The ‘reflecting’ group contained strategies without stated outcomes whereas the ‘making inferences’ group contained those with stated outcomes.

On reflection, I felt that, though this differentiation was important in terms of maximizing the amount of information recovered from the data and, ultimately, following Sarig (1987a), the potential to assess successful outcomes of strategy use, it



was not the best way to classify mental processes. What was needed, it seemed, was to preserve the information about outcomes at a lower level than strategy category.

At this point, reference to reading theory and current data evidence coincidentally pointed in the direction of categorizing through qualitatively distinct processing routes. Specifically, semantic, syntactic and schematic processing are distinguished in several of the most influential models of L1 reading (e.g., Rumelhart 1977, Rayner and Pollatsek 1989). Moreover, Bernhardt's theory of L2 reading highlights developmental issues related to language proficiency level. Reviewing the data evidence as represented in the strategy descriptors in the initial 'reflecting' and 'making inferences' groups, I found that each group contained strategies that were text meaning-centred, text language-centred and extra-text-centred.

This evidence of congruence between theory and data provided a compelling reason to use this three-way distinction to categorize strategies. Thus, the initial strategy groups 'reflecting' and 'making inferences' were reconstituted as the following three interim strategy categories:

- text meaning-centred
- text language-centred
- extra-text-centred

As with the use of the word 'driven' in the motivator categories discussed above, the word 'centred' is used for these strategy categories in explicit recognition of the fact that other, unreported, processes may be taking place concurrently with those reported and captured in the categories. Moreover, for the sake of clarity in this thesis, I am following the convention that initial capitalization is to be reserved for the final categories, so the interim categories now discussed are presented in lower-case letters.

### **3.5 Interim strategy categories**

Table 5.6 presents the situation at this interim stage in the development of a categorization of reading strategies. The table shows how, in each of the first three interim strategy categories listed, the names of two of the former strategy groups (i.e.,

‘reflecting’ and ‘making inferences’) have been transferred to the lower level of strategy descriptor. The ‘sensory’ group remains intact, though, following the review of reading purpose above, its descriptors may now have ‘motivators’ (i.e., purpose) attached to them. Finally, the ‘miscellaneous’ group awaits attention. Sample strategy descriptors are given for all five interim strategy categories.

**Table 5.6: Interim strategy categories**

Interim strategy category	Sample strategy descriptors
‘text meaning-centred’	<ul style="list-style-type: none"><li>- Making an inference about meaning in [part of] text</li><li>- Reflecting on meaning in [part of] text</li></ul>
‘text language-centred’	<ul style="list-style-type: none"><li>- Making an inference about [aspect of] language in [part of] text</li><li>- Reflecting on [aspect of] language in [part of] text</li></ul>
‘extra-text-centred’	<ul style="list-style-type: none"><li>- Making an inference from prior knowledge of content of [part of] text</li><li>- Reflecting on prior knowledge of content of [part of] text</li></ul>
‘sensory’	<ul style="list-style-type: none"><li>- Rereading [part of] text [+ motivator]</li><li>- Skipping [part of] text [+ motivator]</li></ul>
‘miscellaneous’	<ul style="list-style-type: none"><li>- Memorizing or remembering key ideas in [part of] text</li><li>- Making predictions about [part of] text</li><li>- Expressing affective and evaluative reactions to [part of] text</li><li>- Checking length of [part of] text</li><li>- Stopping reading</li></ul>

Parenthetically, the first three categories in the table are very similar to those developed for ‘breakpoint triggers’ in part of the study excluded from the final version of the thesis (i.e., *Meaning in Text*, *Language in Text* and *Reader Contribution*) and for two of the four motivators above (i.e., *Text Meaning-Driven*, *Text Language-Driven*). Moreover, four of the five interim strategy categories shown in Table 5.6 (i.e., excluding ‘miscellaneous’) have precedents in the L1 and L2 reading strategies literature, though, as will now be shown, these are sometimes realized at the lower level of strategy descriptor.

***Interim strategy category: ‘text meaning-centred’***

The processing of text meaning has often been classified into more than one category, though the manner in which it is split across categories varies. Sarig’s (1987a) ‘coherence-detection’ category, developed from L1 and L2 data, is very similar to my ‘text meaning-centred’ category, though she includes use of the reader’s prior knowledge as well as that of text meaning, a key point that I will return to below. In



addition, in cases where detection of errors involves text meaning, Sarig's 'monitoring' category is close to my category. Thus, meaning-centred behaviour is spread across two of Sarig's categories. In a similar way, in "broadly based" categories developed from an extensive survey of the literature, Davies's (1995:51) 'interact with text' and 'monitor reading process' categories code for meaning-related behaviours.

There are other ways in which the processing of text meaning is split across categories. For instance, in the Jimenez et al. (1996) study, meaning-centred strategies occur in all three categories: 'text-initiated', 'interactive' and 'reader-initiated'. The adoption of these categories results in the possibility that meaning-centred, language-centred and physically observable behaviour all fall under one category. Indeed, Jimenez et al. (1996) are not alone in using categories that combine diverse mental and physical actions.

Both of Barnett's (1988a) two categories ('text-level' and 'word-level') focus on whether the processing is global or local rather than on whether it is meaning-focused, language-focused or observable. Following Barnett, two of Pritchard's (1990) categories ('establishing intrasentential ties' and 'establishing intersentential ties'), developed from both L1 and L2 data, also focus on part of text utilized. He uses individual strategies within the two categories to account for text meaning. He also has an interesting and rather specific meaning-focused category called 'accepting ambiguity', which accounts for cases where his subjects report uncertainty over comprehension of the text.

As already mentioned, the concept of the 'core' strategy in the current study allows for coding of part of text utilized at the level of strategy descriptor. Thus, in contrast to the Barnett and Pritchard studies, I account for the extent to which reading is global or local at the lower strategy descriptor level and for the processing of text meaning at the higher strategy category level.

Finally, Kletzien (1991), working with data from L1 readers doing a cloze task, did not put her strategies into categories; thus, in the terminology I am using in this thesis, her data analysis is confined to the lower level of strategy descriptor. However, three

of her strategy descriptors ('using main idea', 'making inferences' and 'looking for key vocabulary') are meaning-centred.

***Interim strategy category: 'text language-centred'***

Sarig (1987a) has one language-centred category, 'clarification and simplification', as does Davies (1995), who calls her category 'utilize source of information: textual'. In contrast, Jimenez et al. (1996) have several language-centred strategies in their 'text-initiated' category, which, as noted above, also has meaning-centred and observable strategies. Likewise, Barnett (1988a) has no exclusively language-centred category; instead, she has several language-centred strategies in her 'word-level' category. By definition, researchers who do not develop strategy categories code for language-centred behaviour, along with all other reading behaviour, at descriptor level rather than category level. For instance, Kletzien (1991) has several language-centred strategies in her list, including 'using syntax or punctuation', 'recognizing structure of passage or paragraph' and 'recognizing structure of sentence'. In addition, from her L1 and L2 data, Block (1992) developed two language-centred strategies, 'grammatical approach' and 'using the base form of a word'.

***Interim strategy category: 'extra-text-centred'***

There are several recent precedents for classifying text and non-text sources of information separately. This situation is reflected in Davies's (1995) 'utilize source of information: external/background knowledge' category. Pritchard's (1990) 'using background knowledge' category, and some of Jimenez et al.'s (1996) 'reader-initiated' strategies are further examples of discrete categories for the processing of non-text information. In addition, among Kletzien's (1991) list of strategies is 'using prior knowledge'.

In contrast, there are only two instances in the literature where processing of text and non-text sources of information are coded under the same category, and, in neither case, does the researcher cite reading theory as contributing to the decision to combine the two. In justifying her 'coherence detection' category, Sarig (1987a) invokes text evidence, not reading theory, while Barnett (1988a) gives no reason for this combination in her 'text-level' category. As will be shown, despite the somewhat biased legacy in the literature, I was to come to a point of realization where I felt that



Sarig's and Barnett's minority stance in this matter was principled both in terms of theory and data evidence.

***Interim strategy category: 'sensory'***

This category is similar to Weinstein and Underwood's (1985) 'active study' category and to some of the strategies in O'Malley and Chamot's (1990) 'cognitive' category. As mentioned at the initial strategy stage, the 'sensory' category is a relatively stable feature in classification schemes in the literature. It is very similar, possibly identical, to, Sarig's (1987a) 'technical-aid' moves and to Davies' (1995) 'control reading process' strategies. In addition, there are two observable strategies in Kletzien's (1991) list of strategies ('rereading preceding text' and 'reading subsequent text') and at least one ('rereading') in Jimenez et al.'s (1996) 'text-initiated' category.

### **3.6 From interim to final categories**

Having established strategy categories that appeared to be reflective of both reading theory and data evidence, there were some remaining issues to be resolved. It will be recalled that Research Objective 4 was 'to incorporate into the final categorization scheme a mechanism for mapping combinations of related strategies'. Thus, one of the issues that needed resolution was differentiation of strategies that are fused or integrated from those, either within one breakpoint cycle or across several, which are part of a series of discrete actions. As described below, this distinction between integrated and discrete processes, which has been ignored in the strategies literature, necessitated another full check of the data, resulting in a full reassessment of the concept of 'compound' strategies. Further, individual strategies in the initial 'miscellaneous' group awaited categorization or disqualification. I also felt that work remained to be done on finalizing the precise wording of some of the strategy descriptors.

Thus, the following part of the discussion looks first at integrated processes and includes a detailed review of compound strategies. It then distinguishes integrated processes from discrete processes reported within a single breakpoint entry. Finally, it describes the process of dealing with strategy descriptors hitherto unclassified.

Reference will be made during the discussion to the same three anchor points mentioned above: the current definition of reading strategies, breakpoint cycles and the reading theory/data evidence interface.

### **Integrated processes**

Until this relatively late stage in the analysis of the data, extra-text-centred strategies, which account for top-down processing, were coded as a separate behaviour. However, following the principle that categories should, where possible, reflect current reading theory (Coding Principle 5.2), it became evident that these strategies could only function in combination with strategies that make direct reference to the text. In other words, the dominant theoretical position is that top-down processing in reading does not occur in isolation.

Moreover, justification for explicitly combining text and extra-text information in strategies need not be based merely on theory. A glance at the strategy descriptors in the ‘extra-text-centred’ category indicates that every one of them makes reference to the text. To highlight this point, the word ‘text’ is italicized in the following instances of strategy descriptors in this category:

- making an inference from own knowledge of content of [part of] *text*
- reflecting on own knowledge of content of [part of] *text*

The use of the word ‘text’ had not been done by engineering the descriptors to reflect reading theory but occurred naturally from close working with the data and attempting to represent them faithfully (Coding Principle 5.5). However, as will be shown, the precise wording of the above two descriptors was to be slightly modified later.

My realization of a clear convergence of theory and data evidence in the matter of top-down processing seems to have occurred curiously late in the process of analysis. In explanation, it seems likely that I simply assumed the viability of prevailing and widely accepted norms, thus delaying my insight into alternatives. As already mentioned, the default position among researchers has clearly been to categorize top-



down processing separately from other types of processing (e.g., Pritchard's (1990) category 'using background knowledge' and Kletzien's (1991) strategy 'using prior knowledge'). However, it now seemed to me that this discrete approach to describing top-down processing has the propensity, even if unintentionally, to obscure the interactive nature of the processing of text and extra-text information.

The link between text and extra-text processing has occasionally been made explicit in the reading strategies literature. Writing about her category 'utilize source of information: external/background knowledge', Davies (1995) writes that readers draw on "non-textual knowledge which may be evoked by, but is not given in the text, for example knowledge of topic, content area, culture etc." (pp. 51-54). However, when it comes to incorporating text and non-text processing into one category, as previously noted, there appear to be only two exceptions to the prevailing trend. These are Sarig's (1987a) 'coherence-detection' category, which contains 'macro-textual', 'pragmatic' and 'extra-textual' moves and Barnett's (1988a) 'text-level' category, which includes 'background knowledge', 'predicting', 'reading with a purpose', 'skimming' and 'scanning'.

One other study worth mentioning is that of Jimenez et al. (1996), some of whose subjects were bilingual. Their 'reader-initiated' category differs from those in the mainstream L2 reading literature in that it contains strategies invoking non-text sources of information as well as what they call 'bilingual strategies', such as searching for cognates, translating and code-switching. Juxtaposing readers' prior content knowledge and language knowledge echoes Eskey's (1988) observation that, in common with other knowledge, language is a schema.

As was shown in Chapter 1, top-down theories of readers leaping through text with little or no recourse to its decoding have been overtaken by a substantial body of research evidence. Thus, in order to reflect current interactive or interactive bottom-up theories of reading in this classification of reading strategies, it seemed to me appropriate to make explicit the link between the text and any prior knowledge or personal experience that the reader brings to the reading process. As a result of this convergence of reading theory and data evidence, all 'extra-text-centred' strategies

were coded as ‘compounds’ (Coding Principle 5.3); compound strategies are now discussed further with reference to data extracts and their corresponding descriptors.

**Compound strategies**

Compounds are strategies that incorporate integrated or fused elements. It will be recalled that they were originally developed at an earlier stage in data analysis and included a reading behaviour and a reading purpose; later, however, as has been shown, strategies describing purpose were disqualified as such. The discussion now describes how compound strategies were reassessed and used first to describe the integration or fusion between text and extra-text processing and then to describe other similarly bound combinations of reading behaviour. The discussion includes illustrations of how, in order to reflect the integration of processing, strategy descriptors were reworded.

*Compound strategies with ‘extra-text-centred’ processing*

As will be shown in the results section in Chapter 6, by far the most frequent combination of compound strategy involved the integration or fusion of ‘extra-text-centred’ and ‘text meaning-centred’ strategies. In order to describe this combination clearly, the precise wording of the new *compound* strategy descriptors required some modification from that used in the former *single* or discrete strategy descriptors. Thus, the first of the following strategies, which had hitherto been classified, discretely, as ‘extra-text-centred’, was reclassified as ‘extra-text-centred/text meaning-centred’ and its descriptor was modified explicitly to include the word ‘meaning’:

Strategy type	Interim strategy category	Strategy descriptor
Single	‘extra-text-centred’	Making an inference from prior knowledge of content of [part of] text
Compound	‘extra-text-centred/ text meaning-centred’	Making an inference from prior knowledge and meaning in [part of] text

This combination represents an interaction between the text and the reader, an interaction that might involve his or her contribution in terms of prior knowledge, personal experience, interest in or evaluation of the text. In the following example



from a protocol, S8 clearly goes beyond the text when, in his attempt to understand the meaning of the word ‘monopoly’ in the text, he cites his knowledge and understanding of the board game *Monopoly*:

Strategy descriptor	Protocol extract
Making an inference from prior knowledge and meaning in text [word] that is supported by text evidence	S8: L57 “Monopoly is some kind of games [sic]. As the winner of game is the ‘richest’ one, the word ‘monopoly’ may mean ‘domination’.”

As the strategy descriptor above indicates, S8’s inference was successful in that it was supported by the text; however, in some cases, text evidence conflicts with the reader’s inference. In the following example, S18’s personal experience, not surprisingly, made him doubt the suggestion made in a quotation in the text that an individual student could have the undivided attention of the teacher during an entire class session:

Strategy descriptor	Protocol extract
Making an inference from prior knowledge and meaning in text [sentence] that conflicts with text evidence	S18: L9 “I stopped reading because I doubt whether it occurred in English lesson. I have never experienced that case.”

*Additional compound strategies*

Convinced that the idea of compound strategies involving ‘extra-text-centred’ and ‘text meaning-centred’ processing was not only viable but also the most principled way of describing this kind of reading behaviour, I checked the data specifically for other combinations of processing that were either fused or integrated. Examples were found where ‘text meaning-centred’ and ‘text language-centred’ strategies could readily be combined into compounds rather than, as previously, coded as two singles. Here, S3 senses that the use of quotation marks with the word ‘service’ might somehow enhance its meaning, thus fusing her focus on text language and text meaning:

Strategy descriptor	Protocol extract
Reflecting on language [use of quotation marks] and meaning in text [word]	S3: L22 “I am thinking why the word ‘service’ is marked by ‘ ’ [i.e., quotation marks]. So I try to guess if it has any special meaning behind the word ...”

Hitherto, all the strategies exemplified as forming compounds have involved combinations of strategies that represent mental processes (i.e., ‘text meaning-centred’, ‘text language-centred’ and ‘extra-text-centred’). However, when I reviewed the data further, it seemed to me that one of the existing strategy descriptors under the ‘sensory’ category of strategies should be recategorized as a compound. The strategy ‘underlining key points in [part of] text’ combines the characteristics of an observable physical action (underlining) and a text meaning-centred mental action (selecting key points in text). Therefore, this strategy was reclassified as a compound: ‘sensory/text meaning-centred’. Two examples follow:

Strategy descriptor	Protocol extract
Underlining key points in text [section]	S6: L18 “After reading a section (i.e. Introduction), I will break to underline some relatively important ideas.”
Rereading text [paragraph?], and underlining key points in text [paragraph?]	S20: L57 “As I don’t know the exact meaning of the text, I try to reread it more slowly and underline the key word or phrases.”

The first example is a straightforward enough compound strategy. In contrast, the descriptor in the second example begins with a single strategy (rereading) and, following a comma, ends with a compound that has two integrated elements (i.e., underlining and selecting key points). Arguably, this is a triple compound, that is, ‘rereading’ is an integral part of ‘underlining’ and ‘selecting key points’; however, this connection is not as easy to justify as that between underlining and selecting key points. When in doubt about whether actions are integrated or discrete, I have erred on the side of caution by considering them discrete (Coding Principle 5.3).

In addition, the strategy descriptor ‘checking length of [part of] text’, which had previously been classified ‘miscellaneous’, was, on reflection, considered to represent both an observable physical action and a mental one. It was therefore recategorized as a compound: ‘sensory/text language-centred’. For example:

Strategy descriptor	Protocol extract
Checking length of text [paragraph]	S25: L19 “I want to see how long this paragraph is, i.e. half page or above.”



Finally, the data revealed that it is possible for a compound to consist of a strategy in the same category more than once. In the example below, S31 reports what are two highly integrated text meaning-centred processes, summarizing main ideas in text already read and making inferences about text yet to be read. Thus, this extract was classified as a compound: ‘text meaning-centred/text meaning-centred’:

Strategy descriptor	Protocol extract
Reflecting on key points in text [section] and making inferences about meaning in text ahead [unclear]	S31: L19 “When one section ends, I try to summarize what I read. I predict what the following sections contain.”

To summarize, compound strategies contain more than one integrated or fused element. In some cases, they are formed where there is supporting and converging evidence from theory and data; such is the case with the top-down processing described in ‘extra-text-centred’ strategies. In other cases, they are formed because a particular combination of actions, which might sometimes be a blend of the cognitive and the physical, must be integrated for the behaviour described to be achieved, as is the case with the strategy ‘underlining key points in text’. As the findings will show, some compounds occurred far more frequently than others in the current data.

**Discrete processes**

Discrete processes in a breakpoint cycle may be recorded either as a single strategy or as a series of discrete or loosely linked strategies. The former are, by definition, discrete and, since they were extensively discussed above (Sections 3.1 to 3.5), they will not be discussed further here. However, the latter, which have not yet been addressed, will now be discussed with examples from protocols.

As shown in Figure 5.1 at the beginning of this chapter, breakpoint cycles account for loosely linked strategies by showing that unresolved breakpoints may be reworked within a breakpoint cycle; thus, there may be separate attempts to resolve the same issue. However, the links between discrete strategies reported in the same breakpoint entry may be very tenuous, so tenuous in fact that, occasionally, as will be shown, the

only connection between them may simply be that they were recorded within a single breakpoint entry.

The data yielded numerous instances of discrete processes occurring within a breakpoint entry. In the three examples in the box below, the plus sign (+) following a strategy descriptor in the left column indicates that there is another, separate, strategy to come:

Strategy descriptor	Protocol extract
<ul style="list-style-type: none"><li>- Reflecting on information explicitly omitted from text [quotation] +</li><li>- Rereading text [unspecified] to reflect on unspecified factor(s) in text [unclear]</li></ul>	S5: L33 “– thinking what is missing in the quotation. – rereading & reflecting.”
<ul style="list-style-type: none"><li>- Making an inference about language [word class] in text [word] +</li><li>- Reading text [remainder of sentence] +</li><li>- Making an inference about meaning in text [word]</li></ul>	S8: L1 “‘embodies’ is a verb undoubtedly. After studying the next [part of the] sentence, I think it means ‘includes’ or somethings [sic] similar.”
<ul style="list-style-type: none"><li>- Rereading text [subordinate clause] +</li><li>- Making an inference about meaning [links between two phrases] in text [subordinate clause] +</li><li>- Making an inference from prior knowledge and meaning in text [sentence]</li></ul>	S32: L1 “Re-reading: I try to think about what is the meaning of ‘learner-centred approach’ and how does it related [sic] to self-access. I also try to guess it is a programme like the Language Institute provided by City U [City University].”

In the first example, S5 experienced a breakpoint at which he attempted to work out the content of information that was explicitly omitted from the text (indicated in Line 33 of the text by four consecutive full stops). His protocol, which he presented in two stages, each marked with a dash at the beginning of a new line, appears to show that he then reread an unspecified part of the text to try once again to reflect on the meaning and/or language of the missing text. S5’s breakpoint entry was coded ‘extra-text-centred/text meaning-centred + recursive reading: Text Non-Specific motivator’.

In the second example in the box above, S8 reports three stages in the breakpoint, focusing on language, reading ahead and, “after” that, focusing on meaning. His breakpoint entry was therefore coded ‘text language-centred + ongoing reading + text meaning-centred’.



There also appear to be three stages in the third example in the box above. S32 begins by reporting a physical action and then she reports linking two concepts in the text. These two stages appear to be linked typographically by the subject's use of a connecting colon. The third stage indicates integration or fusion between text meaning and her own prior knowledge of the university she is attending, knowledge that is related to the text content. This stage is marked from the previous one by the additive 'also'. Arguably, there is some integration between and among the three stages; however, I have only coded for such integration where the data make it incontrovertible (Coding Principle 5.3), not the case here. In fact, one could argue that the connection between the first two stages and the third stage in this extract is extremely tentative. It is an open question whether the additive 'also' serves to connect the stages or to indicate an additional and separate mental process. Thus, the breakpoint entry shown in S32's protocol extract was coded 'recursive reading + text meaning-centred + extra-text-centred/text meaning-centred'.

In sum, where subjects reworked a breakpoint trigger, and thus went through the cycle another time, my default position was to code the strategies separately, but showing them as being loosely linked (Coding Principle 5.3). In contrast and as discussed earlier in this section, where integration or fusion of strategies reported in a breakpoint entry was unarguable, they were bound together in 'compound' strategies.

### **Final check of strategy descriptors**

Prior to presenting the final version of the strategy categories, there remain, at the lower level of descriptor, a few hitherto unresolved issues. One of the initial strategy groups mentioned earlier in this discussion was termed 'miscellaneous', a small group of disparate strategies, all with low frequencies relative to most of the strategies already categorized. The box below presents these strategies, some of whose descriptors remained to be modified, along with illustrative protocol extracts.

In the first extract, S1 is very explicit about remembering key words in the text; however, since it is uncertain if this involved 'memorizing', this descriptor was modified, thus: 'remembering key ideas in [part of] text'. By deleting the word

‘memorizing’ in this way, I felt I was keeping faith with the subject’s own terminology without making assumptions about the precise mental process or processes involved (Coding Principle 5.5). Since the subject’s ‘remembering’ involved “important key words” (she may be making reference to the subheading in Line 20 of the text), this descriptor was classified as ‘text meaning-centred’.

Strategy descriptor	Protocol extract
Memorizing or remembering key ideas in text [subheading]	S1: L20 “... & I try to remember the words since I think they are important key words.”
Making predictions about [part of] text	S13: L7 “I’m think [sic] of the phrase ‘encourage me talking to speak a lot’. When I’m reading, I’d [sic] try to predict the following words. This time, I tried to predict the words following ‘me’ will be ‘to speak in English’.” S13: L13 “Again, a wrong prediction made me stop. The phrase ‘depend <u>to</u> ’, I think it should follow ‘depend <u>on</u> ’ [S’s own underlining].”
Expressing affective and/or evaluative reaction to [part of] text	S21: L94 “The further limitation arouse my interest to further analyse.”
Checking length of [part of] text	S25: L19 “I want to see how long this paragraph is, i.e. half page or above.”
Stopping reading	S2: L17 “I stopped to take a rest since the sentence is too long.”

The second example in the box above has two illustrative protocol extracts, which were consecutive breakpoint entries from the same subject (S13). The reason for presenting them together is that the first word of the second entry, “again”, appears to connect the two, an impression that is substantiated by the fact that both entries concern predictions. Further, since the subject states an outcome in both cases, I felt that the mental processes reported were essentially the same as those I had coded elsewhere as ‘making an inference’. There appeared to be only one difference between the predictions with stated outcomes and the inferences already established. This was that the former involved inferencing meaning about text *yet to be read* whilst the latter involved inferencing meaning about text *already read*. To avoid unnecessary proliferation of terminology, I used the same descriptors in both cases and recorded parts of text utilized (Coding Principle 5.4).



Having made this decision about terminology, the remaining question was to determine the nature of S13's inferences. Across the two entries, she reports that she is predicting "words" and a "phrase", in each case, citing the collocation she found problematic. However, while the entry for Line 13 of the text is clearly 'text language-centred' (i.e., 'depend to' versus 'depend on'), that for Line 7 was less easy to classify. There certainly appears to be a text language-centred element, as evidenced by the use of the word 'again' at the beginning of the second of S13's two entries, thus indicating some kind of link between the two 'predictions'. This feeling was substantiated by my sense that the subject was showing sensitivity to an inaccuracy spoken by a non-native speaker being quoted in the text (i.e., using a gerund after the verb 'encourage'). Like the entire group of subjects, she was hoping to be a teacher of English when she completed her degree. The descriptor was therefore modified to 'making an inference about [aspect of] language in [part of] text' and classified as 'text language-centred'; in this case, aspect of language was coded as 'collocation' and part of text as 'sentence'.

Naturally, there remains the possibility that there was also a text meaning-centred element to S13's processing in her entry for Line 7 of the text, though the only evidence of this was tenuous. While she predicted the phrase "to speak *in English*", the non-native speaker was quoted in the text as saying "talking to speak *a lot*" (my own italics). On balance, I felt that, especially at the very local level of reading being reported here, there was insufficient evidence that she was focusing on text meaning. There was also no meaning-related need for the subject to specify "English" since, three lines above in the text, the author established that the non-native speakers he was quoting were talking about the "English Language Study-Centre".

As has been stated at frequent intervals in this thesis, mentalistic data are, in their nature, incomplete and their analysis sometimes involves making difficult decisions about grey areas. This is why I felt it was important that the decisions I made in this study were based on stated coding principles; nonetheless, follow-up interviews might have helped illuminate this kind of issue.

With respect to the third extract above, S21 explicitly says that her interest in the content of the text inspired her to reflect further on the text. Following Mathewson

(1985) and Ruddell and Speaker (1985), who, as shown in Chapter 1, incorporated affective factors into their models of reading, I decided to retain expressions of personal reactions to the text as a separate group of descriptors. Since these entailed the reader processing beyond, though still interacting with, the text, they were assigned to the 'extra-text-centred' category of strategies. I then checked to see if the current definition of reading strategy required amendment explicitly to include personal reactions. The definition is 'mentally driven actions or series of related actions, serial or in parallel, taken consciously or unconsciously, to facilitate text comprehension and/or learning and which may contain an affective or evaluative element'. It needed no modification.

In contrast to the examples cited, however, there were other instances in the data set where personal reactions to the text were negative; for instance, boredom with the content of the text was reported to lead to skipping parts of it. Thus, since there were several qualities of personal response, including positive and negative affect, interest and evaluation, the rather broad strategy descriptor (shown in the third example in the box above) was split into several more specific descriptors. For the extract shown, the descriptor was modified to 'expressing positive evaluation of ideas in text [section]'.

In the fourth extract shown in the box above and previously cited, S25 was classified as using a 'sensory' or observable strategy. This probably happened because the beginning of a new section (with only the first line of text in that section) occurred at the bottom of a page and the subject wanted to have an idea of the length of the paragraph she was about to read.

The fifth extract in the box above, reporting an observable action, was coded 'sensory', but this was a temporary move. On further reflection, it was disqualified as a strategy since, I did not think it fitted the current definition of reading strategy. Specifically, 'stopping reading', with no evidence of reflection on the text, cannot be said to be a 'mentally driven action ... taken ... to facilitate text comprehension ...' (the full definition is reproduced above in discussion of the third extract). Furthermore, breakpoint cycles, as shown in Figure 5.1 (Section 1.0) allow for reading to stop. With hindsight, I felt that, at an earlier stage in the analysis, I had been overzealous in attempting to derive at least one strategy from each breakpoint entry.



## Summary

This section has described the processes involved in moving from the interim strategy categories towards the final categories. In doing so, reference was made to the definition of a reading strategy, the concept of the breakpoint cycle and the theory/data evidence interface. One of the main issues to be addressed was that of distinguishing between discrete and integrated strategies occurring within the same breakpoint entry at both the higher level of strategy category and the lower level of strategy descriptor. At the higher level, 'extra-text-centred' strategies were all classified as 'compounds', that is, they were integrated or fused with other strategies, almost invariably with those in the 'text meaning-centred' category; supporting evidence for this was cited from both reading theory and data evidence. Other compound strategies were created, including some that represented a fusion between mental and physical actions. At the lower level, in order to reflect the integrated nature of compound strategies, minor adjustments were made to some of the descriptors. Finally, a full review of strategies previously classified as 'miscellaneous' resulted in some being reworded and categorized in the evolving scheme while others were disqualified as strategies.

### 3.7 Final categorization

This section presents and discusses the final categorization of L2 reading strategies developed from the data in both the current study (Study 1) and that in Chapter 6 (Study 2) as well as from reference to reading theory. In the discussion, I will refer to three levels: category, subcategory and descriptor. Since this is the final categorization, initial capitalization will be used for the strategy categories and subcategories.

As Table 5.7 shows, at the highest level, there are two categories, *Reflective* and *Sensory*, essentially representing a division between mental and physical actions, respectively. In a very real sense, these categories indicate which reading behaviours can be investigated empirically through random control trials such as eye movement studies and which behaviours can only be researched through mentalistic studies such

as the current one. Clearly, there are severe limitations to what empirical studies alone can achieve, but, equally clearly and as reiterated through this thesis, mentalistic data have considerable limitations, too, particularly in terms of lack of completeness and the need to be very tentative about any findings.

**Table 5.7: Final categorization of L2 reading strategies**

Category	Subcategories	Descriptor groups
Reflective	Text Meaning-Centred	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Remembering (Study 1 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Text Language-Centred	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Extra-Text-Centred*	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Expressing personal reactions (Study 1 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Non-Specific (Study 2 only)	<ul style="list-style-type: none"><li>- Making inferences (Study 2 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
Sensory	Ongoing Reading	<ul style="list-style-type: none"><li>- Reading</li></ul>
	Recursive Reading	<ul style="list-style-type: none"><li>- Rereading</li></ul>
	Skipping Reading	<ul style="list-style-type: none"><li>- Skipping</li></ul>
	Annotated Reading*	<ul style="list-style-type: none"><li>- Underlining</li></ul>

**Note for Table 5.7**  
As discussed below, the two subcategories marked with an asterisk (\*) only occurred in compound strategies.

The table also shows that each of the two categories has subcategories and descriptor groups, the latter being groups of strategy descriptors with common initial verbs or verb phrases. Based on data evidence, the ‘Reflective Non-Specific’ subcategory and the ‘requesting clarification’ descriptor groups are exclusive to Study 2 (and are discussed in Chapter 6). Likewise, the ‘remembering’ and ‘expressing personal reactions’ descriptor groups are specific to Study 1. This evidence of the exclusivity of subcategories and descriptor groups to one study or the other in the current research is a graphic indication of how different reading tasks can create different reading behaviours, or at least reports of different behaviours. It may, in part, explain the vast differences Davies (1995:52-54) found in strategy use recorded across a number of recent studies. Thus, pervading the literature on strategies are categories,



subcategories and/or descriptors that are not only *data-driven* but, more narrowly, *task-driven*.

Below, the two categories, Reflective and Sensory, along with their respective subcategories and descriptor groups, are discussed in greater depth.

### **Reflective Strategies**

I chose the term ‘Reflective’ to describe mental actions during reading partly because it is self-explanatory and partly because it contrasts with the only other category, ‘Sensory’, which is observable. Previously, at the ‘interim’ stage of analysis, mental actions were divided into three *categories*, ‘text meaning-centred’, ‘text language-centred’ and ‘extra-text-centred’. The motivating reason for their change of status to *subcategories* was that, as discussed above, Extra-Text-Centred Strategies are invariably fused or integrated with other strategies in ‘compounds’. This resulted in a contravention against one of the coding principles, which stipulates that all categories should function at the same level (Coding Principle 5.1); specifically, the problem involved mixing a category that was dependent on others for its existence with categories that could reasonably stand independently of one another.

However, having changed the status of all the former categories describing mental actions to that of subcategory, it was clear that this also seemed to work well for both the other former categories (i.e., ‘text meaning-centred’ and ‘text language-centred’). It will be recalled that the word ‘centred’ had earlier been used as an explicit acknowledgment that other, related and/or supporting but unreported, processes might be occurring serially or in parallel. Likewise, it seemed appropriate now to reserve what was *certain* (i.e., some kind of mental processing) to describe the category and to give the status of subcategory to what was *less certain* (i.e., the precise sources of information used in that mental processing).

Specifically, the Extra-Text-Centred subcategory represents primarily, but (as has been noted repeatedly) not exclusively, top-down processing. In a similar way, the Text Language-Centred subcategory represents primarily, though probably not

exclusively, bottom-up processing. After all, the reader's own knowledge of the L2 is central to decoding the text (e.g., Eskey 1988, Bernhardt 1991). From this, one might speculate that, at least in some L2 readers, any deficiency in knowledge of the language used in the text merely serves to increase the readers' reference to their existing knowledge of both L1 and L2. Although this may occur particularly where the two languages have a common lexical base (e.g., French and English), there was evidence of this occurring among the current subjects, for whom there was no such connection (i.e., between Chinese and English). For instance, when confronted with the unfamiliar word 'transcendental' in the text used in Study 2, several subjects wondered if the word-ending, '-dental', indicated that it had a semantic connection with teeth.

Such behaviour is in line with Stanovich's (1980) interactive-compensatory model, described in Chapter 1. It is also consistent with Bernhardt's (1991) L2 reading theory, which posits that "problems in understanding related to sound and word-shape features quickly diminish as proficiency increases" (p. 170). Thus, Text Language-Centred Strategies may include top-down processing in the sense that readers utilize their prior knowledge of the language used in the text, though this knowledge relates to the *medium* rather than the *content* of the text.

Likewise, the Text Meaning-Centred subcategory, being primarily bottom-up/interactive or interactive, would almost certainly involve the processing of text language and may also involve readers' own prior knowledge of text content.

In Table 5.8, I posit a vertical continuum extending from top-down to bottom-up processing in reading, most (possibly all) of which, however 'top heavy' or 'bottom heavy' qualifies as interactive. At the top end are the reader's own sources of information while at the bottom end are relatively local text sources of information, with relatively global text closer to the middle. Also plotted against this continuum are the three subcategories of Reflective Strategies developed in the current research as well as Hedge's reading 'modes' (1991; personal communication 1999). While, in certain circumstances, readers may have default 'modes' (just as they may have default strategies), they should be seen as capable of switching modes, sometimes both fast and frequently.



Table 5.8: Vertical continuum of reading behaviours

Continuum	Source of information	Reflective Strategy	Hedge's (1991) reading mode
Top-down	Reader's prior knowledge of text content and language used in text	Extra-Text-Centred	Top-down, relative data exclusion
↑			Top-down, deferred interactive
			Interactive
	Text, relatively global	Text Meaning-Centred	
			Bottom-up, recursive, deferred interactive
↓	Text, relatively local	Text Language-Centred	Bottom-up, recursive
Bottom-up			Bottom-up, non-recursive

**Note for Table 5.8**  
None of the positions shown on the continuum for sources of information, strategies or Hedge's (1991) 'modes' should be seen as precise or fixed.

In presenting this table, my intention is to show the broad range of reading behaviours possible across and within individuals and to downplay precise points along the continuum for sources of information and my own subcategories of Reflective Strategies, points which could not be verified. In this, I might include Hedge's 'modes', which, presumably intentionally, include such imprecise terms as 'relative' and 'deferred' to refer, respectively, to level and timing of processing. Each of these 'modes', which were described in detail in Table 1.1 in Chapter 1, appear to be the most ambitious attempt so far at describing nuances in reading behaviour, thus tending to avoid specific labels that cannot be fully substantiated through data evidence. While labels are necessary to attempt to describe reading processes, they can easily create a false impression of precision.

In short, we lack the means, and therefore the knowledge, to enable us to be completely 'categorical' about reading behaviour, and my own subcategories of Reflective Strategies should be seen with this proviso in mind. Over time, research blending empirical and mentalistic data may add to our knowledge about the varying

qualities of mental processes used by individual readers. Interestingly, in some of her modes, Hedge uses the terms ‘recursive’ and ‘non-recursive’, thus directly rendering them potentially open to study through the combined use of empirical and mentalistic data. Inevitably, however, the further one goes towards the top end of the continuum (Table 5.8), the less tangible the processing.

**Descriptor groups**

Turning to the descriptor groups for Reflective Strategies (Table 5.7 above for an overview, and Tables 5.9 to 5.11 below for details), it is clear that there is a substantial amount of parallel terminology across the subcategories. The terms ‘making inferences’ and ‘reflecting’ are used in all three subcategories pertinent to Study 1. Each of the descriptor groups that occurs exclusively under one subcategory, ‘remembering’ (Table 5.9) and ‘expressing personal reactions’ (Table 5.11), were discussed at length in Section 3.6 above. All descriptors have the potential to code for part of text utilized and those in the Text Language-Centred subcategory (Table 5.10) can code for aspect of language reported.

**Table 5.9: Descriptors of Reflective Text Meaning-Centred Strategies**

<b><i>Making inferences</i></b> <ul style="list-style-type: none"><li>• making an inference about meaning in [part of] text</li><li>• making an inference about meaning in [part of] text that is supported by text evidence</li><li>• making an inference about meaning in [part of] text that conflicts with text evidence</li></ul> <b><i>Reflecting</i></b> <ul style="list-style-type: none"><li>• reflecting on meaning in [part of] text</li><li>• reflecting on key points in [part of] text</li></ul> <b><i>Remembering</i></b> <ul style="list-style-type: none"><li>• remembering key ideas in [part of] text</li></ul>
--

**Table 5.10: Descriptors of Reflective Text Language-Centred Strategies**

<b><i>Making inferences</i></b> <ul style="list-style-type: none"><li>• making an inference about [aspect of] language in [part of] text</li><li>• making an inference about [aspect of] language in [part of] text that is supported by text evidence</li><li>• making an inference about [aspect of] language in [part of] text that conflicts with text evidence</li></ul> <b><i>Reflecting</i></b> <ul style="list-style-type: none"><li>• reflecting on [aspect of] language in [part of] text</li></ul>
---



Table 5.11: Descriptors of Reflective Strategies in compounds

<b><u>Extra-Text-Centred/Text Meaning-Centred</u></b>
<i><b>Making inferences</b></i> <ul style="list-style-type: none"><li>• making an inference from prior knowledge and meaning in [part of] text</li><li>• making an inference from prior knowledge and meaning in [part of] text that is supported by text evidence</li><li>• making an inference from prior knowledge and meaning in [part of] text that conflicts with text evidence</li></ul>
<i><b>Reflecting</b></i> <ul style="list-style-type: none"><li>• reflecting on relationship between self and meaning in [part of] text</li><li>• reflecting on information explicitly omitted from [part of] text</li></ul>
<i><b>Expressing personal reactions</b></i> <ul style="list-style-type: none"><li>• expressing interest in ideas in [part of] text</li><li>• expressing lack of interest in ideas in [part of] text</li><li>• expressing positive evaluation of ideas in [part of] text</li><li>• expressing negative evaluation of ideas in [part of] text</li><li>• expressing agreement with ideas in [part of] text</li></ul>
<b><u>Text Meaning-Centred/Text Meaning-Centred</u></b> <ul style="list-style-type: none"><li>• reflecting on key points in [part of] text and making inferences about meaning in [part of] text</li></ul>
<b><u>Text Meaning-Centred/Text Language-Centred</u></b> <ul style="list-style-type: none"><li>• making an inference about meaning and [aspect of] language in [part of] text</li><li>• reflecting on meaning and [aspect of] language in [part of] text</li></ul>

As already explained, Reflective Extra-Text-Centred Strategies invariably occur in compounds; these and all other combinations Reflective Strategies occurring in compounds are shown in Table 5.11. (In the table, subcategories are underlined and descriptor groups are italicized.)

Finally, I would like to point out that, in contrast to Sensory Strategies, no explicit occurrences of reading purpose, or ‘motivators’, were found with Reflective Strategies. It seems likely that the reason for this is that the descriptors of mental actions often have purpose embedded in them. For instance, in the descriptor ‘reflecting on meaning in [part of] text’ (Table 5.9), it could be argued that, though not always at the conscious level, the quest for meaning motivated the strategy. Naturally, though, many strategies may occur with a level of spontaneity that precludes purpose as a motivating factor (e.g., some strategies representing an affective or evaluative reaction to the text).

**Sensory Strategies**

In contrast to Reflective Strategies, which underwent substantial changes between the initial and final stages of data analysis, Sensory Strategies have been remarkably constant. This stability, already noted, probably reflects the fact that they are physical actions and can be observed. Although, as reported above, some researchers combine physical actions and mental actions under the same category (e.g., Jimenez et al. 1996), the former are still clearly identifiable, even at the lower level of strategy descriptor. The subcategories are also very straightforward and potentially visible: Recursive Reading, Ongoing Reading, Skipping Reading and Annotated Reading, all of which, as shown in Tables 5.12 and 5.13, have clear-cut descriptor groups.

**Table 5.12: Descriptors of Sensory Strategies**

<i>Recursive Reading</i>
• rereading [part of] text
<i>Ongoing Reading</i>
• reading [part of] text
<i>Skipping Reading</i>
• skipping [part of] text
• skim reading [part of] text
• scan reading [part of] text

During the discussion of compound strategies above, reference was made to two that combined physical with mental actions. These are shown in Table 5.13. In the first descriptor, there is an integration of a physical action and a meaning-centred mental one, the former manifesting as marking the text with a pen or pencil. Like Reflective Extra-Text-Centred Strategies, discussed above, this subcategory of Sensory Strategies, Annotated Reading, only appears in compounds; in this case, mental action is required meaningfully to annotate text. Although not recorded in the current data, Annotated Reading might include such physical actions as using a highlighter to focus on key ideas and making notes in the margins. In the second descriptor in Table 5.13, a physical action is integrated with a language-centred one; in this case, since the action involves jumping forward in the text, the subcategory of Sensory Strategies is Skipping Reading.



**Table 5.13: Descriptors of Sensory and Reflective Strategies in compounds**

<u>Sensory Annotated Reading/Reflective Text-Meaning-Centred</u> <ul style="list-style-type: none"><li>• underlining key points in [part of] text</li></ul>
<u>Sensory Skipping Reading/Reflective Text Language-Centred</u> <ul style="list-style-type: none"><li>• checking length of [part of] text</li></ul>

As already observed, any one study can only report what emanates from the data and thus no one study could feasibly yield a definitive list of reading strategies. Here, it is clear that additional strategies, though not found in the current data, could be added to this list of two compounds blending physical and mental actions. For instance, ‘noting key points in [part of] text’ and ‘checking key information in appendix’.

**3.8 Section summary**

This section described the lengthy process of developing a categorization scheme of reading strategies, involving three main stages: initial, interim and final. At the initial stage, the process was data-driven as strategy descriptors were established. I then tried several ways of grouping these descriptors with a view to ensuring that the evolving categories were established on a principled basis. At the final stage, there were two categories: *Reflective*, which incorporated all reported mental processing and *Sensory*, which incorporated all reported or observed evidence of physical actions. In recognition of current reading theory, the scheme included ‘compound’ strategies, a mechanism for integrating processes, usually mental processes, that cannot stand alone. Evidence of reading purpose was recorded in ‘motivators’.

#### **4.0 CHAPTER SUMMARY**

This chapter has described the first of two studies (Studies 1 and 2) of L2 reading behaviour. Study 1 investigated the possibility of classifying this behaviour as reading strategies from evidence deriving from the subjects' written introspections while reading a short academic article. The bulk of the chapter described the interactive process of analyzing the data and establishing a categorization scheme of reading strategies. During this process, reference was also made to the data in Study 2, which were collected from the oral introspections of the same subjects working in pairs on a cloze task. The chapter that follows describes Study 2 in detail and ends with comparative findings from the two studies.



## CHAPTER 6: STUDY 2

### 1.0 INTRODUCTION

This chapter describes the second of two studies into L2 reading strategies. Instead of introspection using written protocols, as in Study 1, Study 2 investigates L2 reading strategies using oral introspection. Another key difference between the studies is that, while the breakpoints in Study 1 were *reader-generated*, in Study 2, they were *researcher-generated*, the breakpoints occurring as cloze gaps which the subjects were required to fill. For reference, the research aims and objectives are presented and discussed in Chapter 4. The subjects' biodata as well as the procedures observed in the training programme and the data-gathering sessions can also be found in Chapter 4.

The following three sections of the chapter comprise a description of the development and pilot-testing of the research task, a description of how the data were analyzed and some findings from both Study 1 and Study 2. A review of the entire research project can be found in Chapter 7.

## **2.0 DEVELOPMENT OF THE TASK**

This section, which describes the development of the oral introspection task, has three parts. First, it focuses on issues surrounding selection of the text to be used as a cloze task. It then outlines the pilot-testing of two interim versions of the task, the first with fifty items, the second reduced to thirty items, which eventually led to the final twenty-five-item version. The section concludes with a description of the final version of the oral introspection task.

### **2.1 Selection of the text**

In order to conform to the research design outlined in Chapter 4, in which text genre was to be held constant across Studies 1 and 2, a short academic article, was a prerequisite. However, interpretation of the genre was taken fairly loosely and included texts representative of those the subjects would encounter in the course of their studies. The criteria for text selection were similar to those outlined in the previous chapter for Study 1, but, in order to find a text suitable for cloze, modifications were made to some of these criteria. For Study 2, it was decided that the text should:

- be on a topic of professional interest to the subjects;
- have relatively short paragraphs and/or features to break up the text;
- be relatively free of long sentences, complex syntax and lexical density;
- contain a relatively easy vocabulary load;
- be relatively brief, preferably under three pages of A4 size paper when reproduced, double-spaced, in the cloze format.

These criteria are now discussed and justified in the sequence listed above.

As described below, the task developed for Study 2 used the same pilot subjects as those who participated in the creation of the Study 1 task. The participants in the pilot studies had unanimously complained that, of all the subjects they were currently studying, psychology was by far the most difficult. The subjects who participated in



the two main studies also had a psychology module they found difficult. Therefore, for pedagogic reasons as well as for the sake of face validity for all the subjects, whether taking part in one of the pilot studies or the main studies, I selected a text on psychology.

The text was a 'free standing' passage (Smith 1978) taken from the subjects' set text for their psychology module, Atkinson et al.'s 'Introduction to psychology' (1993). The passage selected for this study was taken from the beginning of Chapter 1 (pp. 3-4) (Appendix 6.1). Since the passage was extracted from one of their set textbooks, it is likely that the subjects had previously been asked to read it, probably at the beginning of the autumn semester. In the cases of the subjects in Pilot Studies 1 and 2, this would mean that they had previously read the passage in the same semester as their participation in the research. In the case of the subjects in Pilot Studies 3 and 4 as well as those in Studies 1 and 2, any previous reading of the text would have occurred in the semester preceding their participation in the research (i.e., at least six months earlier). Anecdotally, when asked if they remembered previously reading the text, none of the subjects had any recollection of it. This might indicate that they had simply not read it or it could mean that, since they were required to read so much, they were unable to recall reading this particular text.

The selected text explains in fairly straightforward terms how pervasively psychology affects people's lives. The first paragraph contains a string of disparate questions whose combined illocutionary force is that psychology is worth studying. The second, third and fourth paragraphs expand on the theme of the pervasive quality of psychology; the text is broken up in the third paragraph with a list of newspaper headlines marked with bullet points. The fifth and final paragraph outlines the contents of the remainder of the chapter.

Probably because this text is an introduction to a course book, albeit one designed for native speakers, the paragraphs and sentences are relatively short for an academic text, and the lexical load is relatively light. Indeed, in these respects, it appears to be an easier text than that used in Study 1. However, the main difference between the criteria for text selection in Study 1 and those for Study 2 involved text length. While the text in Study 1 contained about 2,500 words and was photocopied directly from

the original source, that used in Study 2 contained 770 words, and, in order to accommodate the cloze task, it was retyped, double-spaced. Including written instructions, it covered three pages of A4 size paper (Appendix 6.1).

Since it was envisaged that a cloze activity of about twenty-five items would be suitable for the final version of the research task, an initial version with fifty cloze items was created for pilot-testing from the 770-word passage, yielding an average of about fifteen words between each blank (1:15). For a comparison with cloze passages used in recent studies, see Table 6.1 below.

Table 6.1 shows that, in comparison to cloze passages used in recent studies, that used in the current study was, with the exception of Jonz (1987), relatively long. Furthermore, it was more than twice as long as the cloze text used in the training sequence, described in Chapter 4. However, it was an intentional aspect of this study to utilize a text that was, within the constraints imposed by a cloze task, as representative as possible of texts these subjects would normally encounter.

**Table 6.1: Profiles of some cloze passages in recent studies**

Study	Total words	Total cloze deletions	Average deletion rate
Bachman (1985)	330	30	1:11
Jonz (1987)	750	50	1:15
Kletzien (1991)			
■ training text	50	5	1:10
■ final texts*	250	12	1:21
Current study			
■ training text	333	20	1:17
■ pilot text	770	50	1:15
■ final text	770	25	1:30

\* Three passages of approximately 250 words each

The table also shows that, while the total number of cloze gaps in the final version of the task was well within the normal range, the average deletion rate was markedly below average. This is almost certainly a reflection of the distinctive purposes for which cloze was used in the studies cited. Specifically, the deletion rates that diverged most from those in the current study were those of Bachman (1985) and Jonz (1987), both of whom conducted *cloze tests*. The closest figure is that of Kletzien (1991), the only study cited that, like the present study, utilized *cloze for*



*introspection*. Furthermore, the difference between Kletzien’s deletion rate (1:21) and that of the present study (1:30) might be explained by the fact that her subjects were native speakers while those in the current study were non-native speakers. In other words, to compensate for language deficit, L2 readers might be expected to require more text than L1 readers successfully to inform their word choices in an introspective cloze task. Clearly, however, the degree to which this were true would vary with L2 proficiency level and the type of modified cloze used. (See Jonz 1987 on the marked impact of the loss of cohesive ties on L2 readers.)

There is abundant evidence that, when deleted, content words are not only harder to predict than function words but that they are more likely to require global reading (e.g., Rankin and Thomas 1980, Bachman 1985, and Jonz 1990). Rankin and Thomas (1980), for instance, claim that modified *rational* cloze, “far from being merely a test of sensitivity to local contextual clues, may well be considered as a slow motion simulation model for the reading process in general” (p. 54).

To foster a range of local and global reading, words with the following characteristics were selected for deletion in the pilot version of the task:

- content words;
- words whose restoration required the use of intrasentential clues;
- words whose restoration required the use of intersentential clues.

Table 6.2 below shows deletions by word class for the fifty-item pilot version of the task.

**Table 6.2: Cloze deletions by word class in the pilot version of the task**

Grammatical category	Details
26 nouns	Including parts of three compound nouns (underlining denotes deletion): newspaper <u>headlines</u> , <u>problem</u> solving and <u>research</u> methods
16 verbs	Functioning variously as main verbs, past participles, gerunds and infinitives
6 adjectives	
1 determiner	
1 preposition	

Having developed the cloze passage in this way, it was ready for pilot-testing. This process is now described.

## 2.2 Pilot-testing

The cloze passage described above was pilot-tested on four separate groups of subjects in Hong Kong. In this section, Pilot Study 1 is described with reference to the procedures adopted and the data analysis, which focussed on word scores and reading strategies reported. This is followed by a description of the other pilot studies. Finally, there is a description of the outcome of the pilot-testing.

In all extracts quoting from the pilot transcripts, italicized words represent subjects' word choices for cloze items. Individual pilot subjects are referred to by the prefix 'ASS' (for Department of Applied Social Studies) and a number (e.g., ASS11).

### Pilot Study 1

#### *Procedures*

Towards the end of the autumn semester of 1994, six of the subjects in Pilot Study 1 were asked to come to their usual classroom in pairs at half-hourly intervals. Since they had all agreed to the sessions being tape-recorded, an audio cassette-recorder was placed on the table where these subjects, sitting side-by-side, were to work. On the assigned day, one subject (ASS11) failed to materialize. In order not to exclude one subject (i.e., the absent subject's partner), three subjects worked together in one group (ASS02, ASS03 and ASS04), and, as planned, a pair (ASS06 and ASS09) worked together (N=5).

The subjects were told that each cloze blank represented one word and that the missing words were "important words for the meaning of the passage, such as nouns, verbs, adjectives and adverbs" and not "little words like *on*, *at*, *in*, *the*, etc.". They were asked to suggest words for the blanks giving reasons for their choices and to challenge their partners' ideas if they did not agree with them. They could tackle the



blanks in any order they chose and the researcher would answer any questions they had. They were asked to write the words they chose in the blanks, putting any second or third choices in brackets. Finally, they were informed that their papers would be collected at the end of the session.

When the subjects were ready, the audio cassette-recorder was switched on. Despite the initial instructions, the subjects tended to tackle the cloze items in linear fashion; however, especially later in the session, some of them were observed rereading and reading ahead, sometimes turning the pages. The researcher only intervened when subjects had been experiencing difficulties with a particular item for a while or when they requested help, and any strategy emanating from researcher intervention was excluded from the results. Subjects' papers were collected at the end of the session. It was found that, in the half-hour allocated, both the pair and the group of three subjects only covered the first twenty items out of the total of fifty items in this initial pilot version of the task. To ensure that all fifty items were pilot-tested, arrangements were made to meet in the same combinations (i.e., the pair and the group of three) for another half-hour the following week.

In the intervening period, I transcribed the cassette tapes of the subjects doing Items 1-20. It was clear that, in order to remind the subjects to give reasons for their word choices as well as to get them to challenge one another's word choices, I needed to intervene more frequently (especially with ASS06 and ASS09). This was done in the second recording session, which I administered in the same manner as the first session and, I hoped, would cover Items 21-50. In the event, due to insufficient time, one group failed to attempt nine of these items and the other group missed fourteen of them. At the end of the session, I collected the subjects' papers.

### *Data analysis*

The pilot subjects' responses to the cloze items were analyzed for both exact-word and acceptable-word choices; acceptability was judged by the researcher alone. Table 6.3 shows the average exact-word and acceptable-word scores based on the total number of cloze items attempted (this figure varied across individuals) and the total number possible (i.e., all fifty items in this version of the task). These scores are low even when bearing in mind that the pilot subjects left several items blank and had a

lower level of proficiency in English than the subjects who were to participate in the main studies.

**Table 6.3: Average word scores (%) for Pilot Study 1**

Average word scores	Exact	Acceptable
Based on total number of cloze items:		
• attempted	22.11%	39.47%
• possible	16.80%	30.00%

Furthermore, the scores are low, relative to those in other studies utilizing cloze. Jonz's (1987) subjects had exact-word scores of between 42% and 68% while Bensoussan's (1990) subjects scored an average of 81% (though she does not report if scoring was conducted on an exact-word or acceptable-word basis). In Kletzien (1991), the only study, along with the present one, to use cloze to tap readers' introspection, acceptable-word scores reported were 61.83% for 'independent' level texts, 52.58% for 'instructional' level texts and 18.75% for 'frustration' level texts. Using this terminology, which Kletzien drew from the Fry readability scale, the Pilot Study 1 subjects' average scores were somewhere between the average scores reported in the Kletzien study for the texts at the 'frustration' and 'instructional' levels.

Later, the tapes of the second recording session were transcribed. The data from both recording sessions were then analyzed and initial descriptors, such as the following, were ascribed to the strategies the subjects reported using:

- using knowledge/interpretation of the world
- using stylistic constraints
- using the subgenre [textbook introduction]
- using textual constraint [lexical, semantic, syntactic]
- inferring meaning from the topic of the passage
- inferring meaning from intrasentential lexical cue
- paraphrasing part of a sentence
- rereading intersententially
- seeking clarification from partner about syntax



Sometimes, there was evidence of two strategies being utilized together. For instance, data from ASS03 appear to show that his strategy for Item 11 in the fifty-item version used in Pilot Study 1 (which is identical to Item 5 in the final twenty-five-item version of the task) was to combine using paragraph topic and reading ahead:

ASS03	I think it's, uh, <i>mind</i> .
ASS02	Yeah.
ASS03	Because the whole paragraph's about ... yes, I think it's the <i>mind</i> under which individuals may not be same for their actions.
ASS04	Oh ... you're fluent!
ASS03	Yeah, and the last words is <i>actions</i> , so I think it's <i>mind</i> of thinking. Maybe not.

Examples of descriptors with two strategies being reported together follow:

- using prior knowledge and using a lexical cue
- using prior knowledge and using a known collocation
- using knowledge of genre and using a lexical cue
- using paragraph topic and reading ahead

In conclusion, it was obvious that the data drawn from the second recording session were richer than those from the first session. Turns tended to be longer, especially following the researcher's questions; specifically, while turns in the first recording session were usually only one or two lines long in the transcript, turns in the second session often reached six or seven lines in length. It is doubtful if such a high level of researcher intervention would have been necessary with the Pilot Study 1 subjects if they had been more familiar with introspective reading tasks. (It will be recalled from Chapter 4 that the introspective training programme was conducted only on the subjects in Pilot Studies 2-4.)

### Further pilot-testing

The subjects in Pilot Study 2 were given the same version of the introspective cloze task as those in Pilot Study 1. Subsequently, using a thirty-item version of the cloze passage, which had been devised to reduce task completion time, two further pilot

studies (Pilot Studies 3 and 4) were conducted at the beginning of the spring semester of 1995. None of these groups were tape-recorded. In addition, in order to obtain their *retrospective* affective comments on the introspective cloze task and their perceptions of its difficulty and usefulness, a four-item survey was given to the subjects in Pilot Studies 3 and 4. (For space reasons, this survey is not described or appended in this thesis, but reference was made to it in data-gathering procedures in Chapter 4.)

From the pilot-testing and the survey of retrospective comments, it was clear that (at least for the subjects in Pilot Studies 3 and 4, who had undergone the programme,) the introspective cloze task was at a reasonable training difficulty level. In addition, they had enjoyed it and felt it was useful. Moreover, since the English proficiency level of the subjects in the main study was higher than that of the pilot subjects, I felt it was reasonable to expect that the former would be able to complete the task more quickly. Nonetheless, even the thirty-item version of the task, completed by the participants in Pilot Studies 3 and 4, had taken far too long to complete.

### **2.3 The oral introspection task**

Following pilot-testing, the final version of the cloze task was formulated (Appendix 6.1). To achieve this, the Pilot Study 1 transcripts were checked with a view to reducing both the total number and the density of items. The number of items was reduced in two stages. As previously reported, it had already been reduced from fifty to thirty items for Pilot Studies 3 and 4; for the final version, it was now further reduced to twenty-five items. Details of the remaining deletions are shown by word class in Table 6.4 below (cf. Table 6.2 above).

Two criteria were used to reduce item density in the final version of the introspective cloze task. First, since the pilot-testing had shown that sentences with two or more blanks tended, not surprisingly, to be particularly difficult, I decided to insert no more than two blanks in any one sentence. Second, individual cloze items that had proved either too hard for the subjects or appeared to be introspectively ungenerative were abandoned; the latter were often items that appeared to be so easy that subjects simply



stated their word choices but didn't bother to comment. In other words, the most generative items appeared to be those that were neither too difficult to attempt nor too easy to warrant explanation.

**Table 6.4: Cloze deletions by word class in the final version of the task**

Grammatical category	Item reference and details
12 nouns	Items 1, 4, 5, 6, 8, 12, 14, 17, 19, 20, 24 and 25. These include parts of three compound nouns (underlining denotes deletion): newspaper <u>headlines</u> , <u>problem</u> solving and <u>research</u> methods.
8 verbs	Items 2, 3, 7, 10, 15, 16, 18 and 23. These function variously as main verbs, past participles, gerunds and infinitives.
2 determiners	Items 9 and 21.
1 adjective	Item 11.
1 adverb	Item 13.
1 preposition	Item 22.

Overall, the average number of words between cloze items in the text was increased from fifteen (1:15) in the original version to thirty (1:30) in the final version, representing, on average, a doubling of context between blanks (Table 6.1 above). This may seem a very low density of blanks compared with the original cloze deletion rate for cloze *tests* (typically 1:5 or 1:7, e.g., Taylor 1953) as well as for tasks in more recent studies (e.g., 1:11 in Bachman 1985, and 1:15 in Jonz 1990). Nonetheless, it should be remembered that the cloze task in this study was designed specifically to function as an L2 oral introspective data-gathering research task. Apart from Kletzien (1991), there are no studies that report using the cloze procedure in this way; and, as can be seen in Table 6.1 above, the Kletzien L1 study has a comparable deletion rate (1:21) to that in the present L2 study (1:30).

Moreover, apart from the fact that the subjects in the current study were required concurrently to read and introspect in a second language, there were two further compelling reasons for a relatively low item density in the cloze task in the current study. First, I felt intuitively that the deletion density selected was likely to be a closer approximation of the subjects' normal experience of reading academic articles in the L2. Second, with the exception of the brief training programme, described in Chapter 4, introspection was an entirely new experience for all the subjects.

No changes to the verbal instructions were deemed necessary, though supplementary written instructions were added to the subjects' worksheets in the final version of the introspective cloze task (Appendix 6.1).

To conclude, the initial analysis of the introspective cloze transcripts and protocols of the pilot subjects suggested that, with sufficient training and practice, cloze can reveal discernible patterns in reading strategy use following *breakpoints*. It also indicated that strategies could be viewed as both single and multiple mentally driven actions. As a result of the pilot-testing, a tentative categorization of reading comprehension strategies was developed (for space reasons, details are not given in this thesis), and I hoped that this could be developed during analysis of the data in the main study. Finally, it appeared that, through the pilot-testing process, an appropriate introspective cloze task had been developed in terms of genre, topic, language difficulty, task type and likely task duration. It also seemed that this task would yield reasonably rich data. The stage was now set to analyze the data.



3.0 DATA ANALYSIS

This section describes the process of analyzing the Study 2 data for reading strategies. In order to save space, information common to both Study 2 and Study 1 has been consolidated. Chapter 4 contains information on the subjects’ biodata, the training they underwent and the procedures observed in the data-gathering sessions. Chapter 5 contains complete information on the development of the categorization scheme of reading strategies, which was used in the analysis of both the Study 1 and Study 2 data. Description of the Study 2 analysis is in Sections 3.1 to 3.4 below.

First, however, in order to establish a framework for this description of data analysis, this section begins by highlighting the main differences in the design of Studies 1 and 2. It then summarizes the final categorization scheme for reading strategies and lists the coding principles adopted. It also outlines three anchor points used in discussion of the Study 1 analysis, which are also used here.

In all extracts quoted from the transcripts, italicized words represent subjects’ word choices for cloze items and underlined words or sequences of words represent subjects reading aloud from the text. Capital S followed by a number (e.g., S38) refers to subject number and capital T followed by a number (e.g., T165) refers to turn number in the transcript.

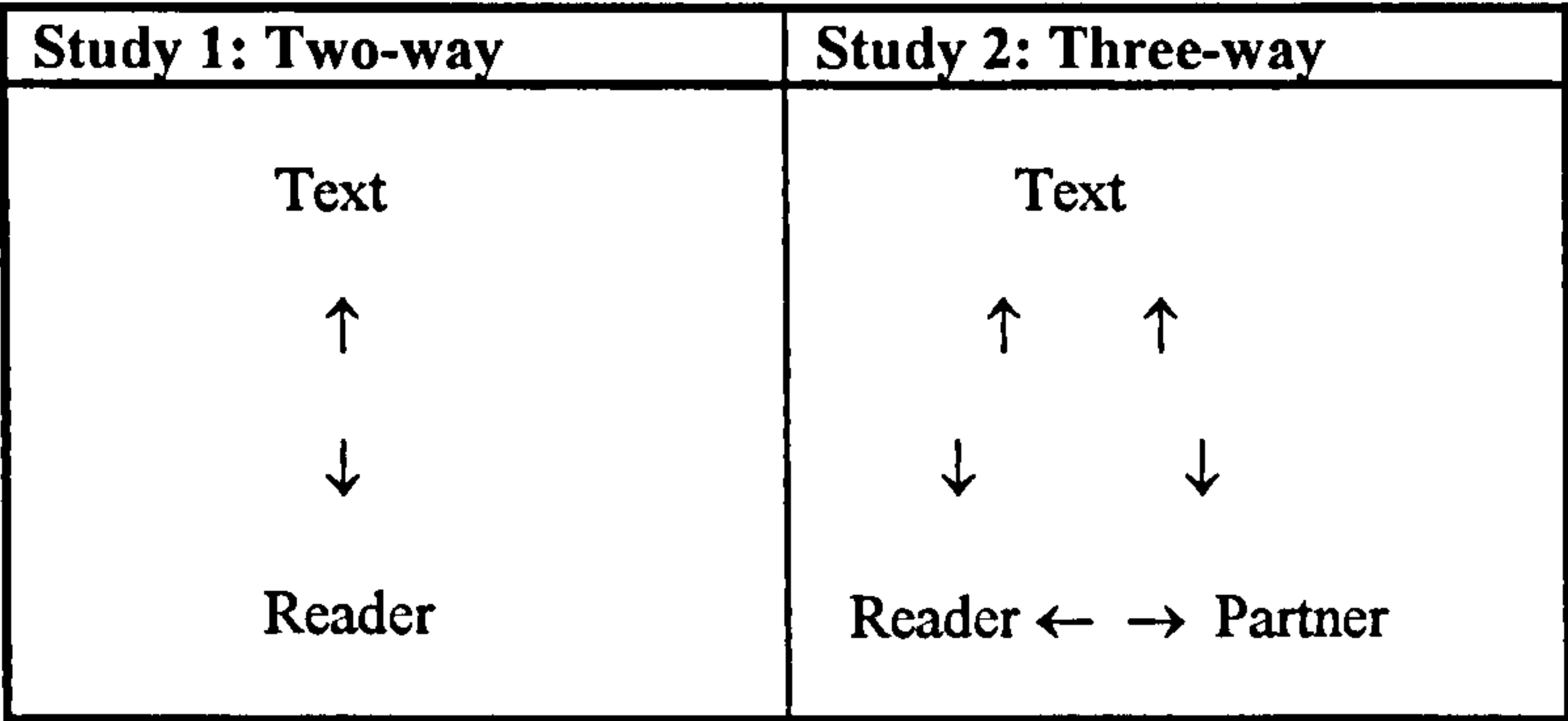
Design

It will be recalled from Chapter 4 that the design of the current research called for some key differences between Study 1 and Study 2. This was a reflection of Research Objectives 3 and 4, which sought, respectively, to develop strategy categories and strategy descriptors that were viable across variation in methodology and tasks. The differences between the two studies are summarized in Table 6.5 below.

Table 6.5: Design differences between Study 1 and Study 2

Task features	Study 1	Study 2
Introspective mode	Written	Spoken
Task mode	Individual, silent	Pair, oral
Breakpoint type	Spontaneous	Simulated

All three of the differences in task features given in the table made analysis of the Study 2 data relatively complicated. First, the fact that the data were drawn from spoken rather than written language meant that the subjects' reports of their mental processes were less likely to have been as considered, though they might have been more spontaneous. Furthermore, the combination of oral communication with a pair task meant that there were interruptions, resulting in partially expressed thoughts and strategies.



**Figure 6.1: Contrasting interactive dynamics across Studies 1 and 2**

In addition, and crucially, the pair element in Study 2 meant that, instead of a two-way interaction between text and reader, there was a three-way interaction among text, reader and partner (Figure 6.1). In Study 2, this three-way interactive dynamic resulted in data that were either *reader-generated* (i.e., the speaker reporting an original reaction to the text), *partner-generated* (i.e., the speaker reacting to a partner's utterance) or a combination of the two.

Lastly, Table 6.5 above shows that there was a difference in breakpoint type; in Study 1, they were spontaneous or *reader-generated* whereas in Study 2 they were simulated or *researcher-generated*.

Thus, the task in Study 1 was, in several ways, closer to normal reading than that in Study 2. Indeed, an initial pass through the Study 2 data confirmed the fact that the key differences between the tasks in the two studies had had a number of consequences on the data yielded. These consequences related both to the lower level of strategy descriptors and the higher level of strategy categories, key aspects of the current research that are discussed below.



First, however, in order to provide an idea of the end point in the coding frame used in Study 2, the final categorization scheme is presented below. In addition, to give an idea of how the process of data analysis was conducted, the coding principles adopted are listed below.

**Final categorization of reading strategies**

Table 6.6 presents the final categorization of reading strategies, a scheme that was developed from the data in both Study 1 and Study 2. (Identical tables are presented in Tables 5.2 and 5.7 in Chapter 5.) The table shows that one of the subcategories of Reflective Strategies (Non-Specific) was exclusive to Study 2. It also shows that, for all subcategories of Reflective Strategies, the descriptor group ‘requesting clarification’ was specific to Study 2, an indication of the interactive nature of the task (i.e., a pair task).

**Table 6.6: Final categorization of L2 reading strategies**

Categories	Subcategories	Descriptor groups
Reflective	Text Meaning-Centred	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Remembering (Study 1 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Text Language-Centred	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Extra-Text-Centred*	<ul style="list-style-type: none"><li>- Making inferences</li><li>- Reflecting</li><li>- Expressing personal reactions (Study 1 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
	Non-Specific (Study 2 only)	<ul style="list-style-type: none"><li>- Making inferences (Study 2 only)</li><li>- Requesting clarification (Study 2 only)</li></ul>
Sensory	Ongoing Reading	<ul style="list-style-type: none"><li>- Reading</li></ul>
	Recursive Reading	<ul style="list-style-type: none"><li>- Rereading</li></ul>
	Skipping Reading	<ul style="list-style-type: none"><li>- Skipping</li></ul>
	Annotated Reading*	<ul style="list-style-type: none"><li>- Underlining</li></ul>

**Note for Table 6.6**  
As discussed in Chapter 5 (Sections 3.6 and 3.7), the two subcategories marked with an asterisk (\*) only occurred in ‘compound’ strategies. These are strategies with two or more elements that are integrated or fused.

The description of the overall development of the categorization scheme was presented in Chapter 5, but factors specific to Study 2 that had an impact on the scheme are presented below in Sections 3.1 to 3.4.

Coding principles

Table 6.7: Coding principles for analysis of reading strategies (Study 2)

6.1	The strategy categories should function at the same level as one another, as should the subcategories.
6.2	The strategy categories and subcategories should reflect current prevailing theory in reading; failing that, they should provide grounding for new theory.
6.3	All strategies should be coded as discrete processes unless elements are demonstrably integrated or fused into ‘compound’ processes or linked to a purpose, or ‘motivator’, explicitly stated in the data.
6.4	The principle of minimalism should be observed in the development of strategy descriptors. To help achieve this, ‘core’ strategy descriptors with variants indicating any detailed particulars reported, such as part of text or aspect of language or language use, should be established. In addition, efforts should be made to harmonize the wording of core strategies with the objective of representing the same or very similar reading behaviours in the same strategy descriptors.
6.5	The strategy descriptors should be a faithful representation of the data evidence. Any potential ambiguities in their wording should be avoided through definitions specified for the current study; conversely, any ambiguity in the data should be reflected in the coding. In addition, faithful representation of data evidence should be reflected in the following ways: All utterances that contain strategies with stated outcomes should be differentiated from those that contain strategies without stated outcomes. For all utterances that contain one identifiable strategy but are spread over more than one turn in a sequence (due, for instance, to partner interruption), the strategy should be coded once. For all turns where both subjects simultaneously report using the same strategy, the strategy should be reported once for each subject. No strategies that demonstrably arise as a result of researcher intervention should be coded.
6.6	To reflect the interactive and oral nature of the task, specific codes that differentiate <i>reader-generated</i> strategies from various types of <i>partner-generated</i> strategies are required. Strategies are to be coded as follows: those that are wholly reader-generated, those that refer to partner’s word choice(s), those that refer to reader’s and partner’s word choice(s), those that involve responding to partner’s previous utterance (e.g., clarification request), and those that involve responding to partner’s clarification request about partner’s word choice.
6.7	For reference purposes, all turns that contain more than one discrete strategy should be divided into parts. For example, if three separate strategies are reported in Turn 50, the turns should be referred to as 50a, 50b and 50c.

Note for Table 6.7: Terminology

- ‘Compound’ strategies have two or more elements that are integrated or fused.
- ‘Motivators’ are reading purposes attached to strategies.
- ‘Core’ strategies are descriptors that allow for variation in terms of part of text and aspect of language reported.



In the following discussion, extensive reference is made to coding principles. These principles, presented in Table 6.7 above, are very similar to the coding principles adopted for analysis of reading strategies in Study 1; indeed, with only minor variations, Coding Principles 6.1 to 6.5 mirror Coding Principles 5.1 to 5.5 in Study 1. In contrast, Coding Principles 6.6 and 6.7, in dealing with the oral pairwork dimension specific to the cloze task, are exclusive to Study 2.

### **Anchor points**

During the following discussion of data analysis, the processes expected to occur as subjects do the Study 2 task should be related to the three anchor points referred to in the analysis of the Study 1 data.

First, the definition of reading strategies in the current research is ‘mentally driven actions or series of actions, serial or in parallel, taken consciously or unconsciously, to facilitate text comprehension and/or learning and which may contain an affective or evaluative element’. As will be shown, given the interactive pair task, a substantial number of utterances do not conform to this definition.

Second, as presented in Figure 4.1 in Chapter 4, the ‘breakpoint cycle’ posits processing routes for readers’ behaviour between whatever triggered the breakpoint and its resolution. In terms of breakpoint triggers and as was observed in the discussion of research design above, the difference between Study 1 and Study 2 is that, in the former, they were reader-generated whereas, in the latter, they were researcher-generated in the form of cloze gaps. Thus, in Study 1, the subjects might be expected to break for reflection at text breaks (i.e., paragraph and section breaks), but, in the current study, they might be more task-driven, that is, simply moving in linear fashion from one cloze item to the next.

The third anchor point pertains to the interface between theory and data evidence. While the data may be driving the analysis, reference should be made to current reading theory and any divergence between the two should be explained (Coding Principle 6.2).

The remainder of this section, which is divided into four parts, describes the modifications that were made to the categorization scheme for reading strategies developed and described in Study 1. First, the procedures adopted for transcribing the cloze data are explained. Then an initial phase exploring difficulties and potential in the data analysis is described. After that, there is a description of how the pair-work dimension of the cloze task was analyzed and coded. Finally, there is a review phase in which I describe how I ascertained the extent to which the strategy categories and descriptors developed in Study 1 were transferable to Study 2.

### 3.1 Procedures

To reduce the risk of loss of data, the audio-cassette tapes containing recordings of the nineteen pairs (N=38) doing the cloze task were immediately copied. Then, using a transcribing machine with foot controls and headphones, I transcribed the nineteen tapes. In order to assist in the transcribing process, reference was made to the subjects' protocols, which contained the final version of the cloze text and each subject's answers to the cloze items. Wherever it was unclear what a subject was saying, the transcript was marked as 'inaudible' or 'unclear'. Pauses in subject interaction and incomplete utterances were marked with three consecutive full stops (...):

#### Item 8

S8: T66            "Similar to you ... yes, I think *evidence* is a good answer."

When I had completed each transcription, I printed it out and checked it against the tape-recording. I resolved any anomalies by ensuring that the print-outs matched what was on tape. To facilitate subsequent reference and analysis, the hard copies were then marked for turn numbers and cloze item numbers. Wherever subjects reverted to a cloze item, having discussed other items in the intervening period, the item was coded with a capital R for each reversion (e.g., 6R, 19RR, 25RRR). Also to facilitate analysis, all subjects' word choices for cloze items were highlighted in yellow and all oral reading of the text was highlighted in green.



The bulk of the data analysis for Study 1 preceded that for Study 2. This meant that the emerging coding principles, strategy categories and descriptors in Study 1 could be utilized as starting points for data analysis in Study 2. A major aspect to the challenge was to see if the task-induced variation in the Study 2 data, such as the partner-generated utterances, would mesh, in a principled way, with the categories and descriptors developed in Study 1 (Research Objectives 2 and 3).

As with Study 1, the overall process of analyzing the data involved numerous passes through the entire data set, each pass resulting in a gradual emergence of the final strategy descriptors and strategy categories. These repeated passes were a reflection of the centrality of the data in the current research and the importance attached to its faithful representation (Coding Principle 6.5).

### 3.2 Initial phase

Before starting the data analysis process for Study 2, I identified three major concerns, all of which were later adopted as coding principles. These concerns, which were based partly on issues arising during analysis of the Study 1 data and partly on impressions formed from the initial pass through the data set mentioned above, were as follows:

- to retain the essence of every mental process reported while at the same time differentiating between *reader-generated* and *partner-generated* utterances (Coding Principle 6.6)
- to adhere as closely as possible to the wording used in the strategies reported in Study 1 and to aim for minimal proliferation in strategy descriptors through the development of ‘core’ strategy descriptors (Coding Principle 6.4)
- to differentiate between mental actions that involved a single element from those that integrated or fused more than one element in ‘compound’ strategies (Research Objective 4 and Coding Principle 6.3)

The transcripts for each of the nineteen pairs of subjects were analyzed with reference to their respective protocols (i.e., copies of the subjects' cloze passages with their handwritten answers). Analysis of the transcripts was conducted on every single numbered turn. Turns with more than one discrete strategy were subdivided, thus: Turn 50a, Turn 50b and so on (Coding Principle 6.7). Whenever possible, turns were coded with a reading strategy or combination of reading strategies, using descriptors and categories developed in Study 1. However, as the following description of the initial stage in the process makes clear, some of these descriptors and categories were easier to adapt to Study 2 than others.

### Some difficulties

In contrast to Study 1, almost all the strategies reported in Study 2 were, directly or indirectly, utilized because of a missing word (i.e., a cloze gap). In other words, reading purpose was reader-generated in Study 1 and might vary within and across individuals, while, in Study 2, it was constrained by the cloze task. Parenthetically, as if to reinforce this obvious constraint on reading purpose, there were no examples in the data of subjects specifying any kind of reading purpose or 'motivator', either task-oriented or otherwise. This contrasts with Study 1, where reading purpose is described in Chapter 5, Section 3.4.

Faced with this key, but deliberate, difference between the two studies, intuitively it seemed that the closest possible comparison of strategy use that could be made involved management of unknown vocabulary items. Specifically, an unknown word might be considered as acting like a cloze item. However, when this was done, it was immediately clear that some of the strategies used to manage vocabulary problems could only be utilized following a *spontaneous* breakpoint and not following a *simulated* breakpoint. Below are two examples of such strategy descriptors from Study 1:

- rereading text [word]
- rereading text [word] to check meaning



There were two problems in seeking to utilize these strategy descriptors for Study 2. First, since a cloze item contains a deletion rather than a word that is problematic to the reader, neither of these descriptors could be used to represent what readers do when attempting to complete a cloze item. At first, this did not seem to cause any difficulty since, in theory, thanks to the notion of the ‘core’ strategy descriptor (Coding Principle 6.4), only the part of text utilized (shown in square brackets in the examples above) needed to be altered. Thus, for a missing cloze item, the reader might be expected to report a strategy such as this:

- rereading text [adjacent to cloze item]

However, in practice, there was a second problem in attempting to mesh these strategy descriptors across the two studies. Taking the two examples from Study 1 above, the equivalent of *rereading* would, in Study 2, be coded as *reading aloud* even if the reader had previously (and unobserved) read the same part of the text. Thus, at least in terms of these *Sensory Strategies* (i.e., physically observable strategies), the attempt to devise strategy descriptors that were viable across data yielded in both the written mode (Study 1) and data yielded in the oral mode (Study 2) appeared confounded. Some strategies, while they no doubt took place in Study 2, were harder to trace than in Study 1, and, even where the breakpoints in the two studies were triggered by similar factors, there remained significant differences in the strategies utilized and/or reported.

**Some possibilities**

Despite these difficulties, there were many descriptors developed for Study 1 that could, with minor modifications, serve to describe reading behaviours reported in Study 2. As will now be shown, instances where compatibility was observed occurred with what were to become *Reflective Strategies* (i.e., mental actions). The italicized parts of the descriptors shown in Table 6.8 exemplify the way in which the precise wording in strategy descriptors might differ somewhat between the two studies. At the bottom of the table, the first of the two strategies is illustrated with extracts drawn from data in each of the studies.

**Table 6.8: Variation in strategy descriptors across the two studies**

Study 1 strategy descriptor	Study 2 strategy descriptor
Making an inference about meaning in text [word] <i>that is supported by text evidence</i>	Making an inference about meaning in text [word] <i>that supports own word choice</i>
Making an inference about language [collocation] in text [phrase] that conflicts with <i>text evidence</i>	Making an inference about language [collocation] in text [phrase] that conflicts with <i>own word choice</i>

Study 1 protocol extract	Study 2 transcript extract
S10: L66 “I stop at the word ‘laxity’. I guess it is a word used to describe/refer to something negative in sense by guessing from the context.”	<u>Item 3</u> S18: T43b “[reading aloud]... <u>psychotherapy</u> therapy, uh, have the meaning of ... has the meaning of, uh, <i>curing</i> [depression] ...”

In each extract from the data, the subject made an inference about meaning; however, variation in task across the two studies requires the minor variation in the strategy descriptors shown in Table 6.8. In Study 1, the subject inferred the meaning of the word that triggered the *reader-generated* breakpoint from text evidence; in Study 2, on the other hand, the subject, having read the context aloud, inferred the meaning of the *researcher-generated* cloze item from text evidence. Thus, these modifications reflect the intrinsic differences between the tasks in the two studies. Yet the principle of minimal variation in wording yields ‘core’ strategy descriptors, that is, descriptors that represent similar mental processes across the two studies (Coding Principle 6.7).

The issue of whether variation in task appears to lead to changes in the relative frequencies of particular mental actions reported is addressed in Section 4.0 below, where the findings of the two studies are compared.

**Summary**

In seeking to adapt strategy descriptors from Study 1 to Study 2, the initial experience was that there were insurmountable problems with *Sensory* Strategies but that the process appeared to be possible, though painstaking and labour intensive, with *Reflective* Strategies.



3.3 Describing the pair-work dimension

As already mentioned, one of the key differences between Study 1 and Study 2 was that, while the former had a silent task, the latter had a pair task. The implications of this on the process of analyzing the Study 2 data are now described. The main issues addressed are, in the following sequence, interruptions, partner-generated strategies, and utterances where no strategy was reported.

Interruptions

In order to preserve the integrity of subjects' utterances, care was taken to detect interruptions, a coding issue that did not arise in Study 1 with its individual silent reading task. Interruptions were defined as points in the interaction where subjects reverted, normally in their next turn, to what they had been talking about in a preceding turn, usually their immediately preceding turn. In some cases, however, an interruption acted as a tangent in the interaction between the two subjects; in such cases, the interruption lasted for more than one turn.

Coding sought to account for instances where a subject, when interrupted by his or her partner, was reporting on, or discussing the use of, a particular reading strategy or combination of strategies. In other words, coding was determined qualitatively rather than quantitatively by the number of turns a subject took to report a strategy (Coding Principle 6.5). In the example that follows, S21 is interrupted by her partner; however, since she reports only one strategy across Turns 46b and 48, she was coded for the strategy once:

Item 9

- S21: T46b      "... we need adjective ... something like ... *somewhat* to explain the level of violence ..."
- S22: T47      "Yeah, so it's adjective ... is it adjective?"
- S21: T48      "I think to emphasis [sic] the level of the violence, adjective is used to, uh, express the ..."

## Partner-generated strategies

Following the practice adopted in Study 1, the strategy descriptors were grouped by initial verb. This was done in order to harmonize further the wording in the descriptors and eradicate any strategies that appeared to be duplicated through more than one form of wording. However, since all the strategies in Study 1 were reader-generated, there were no available working examples of partner-generated strategies. It was therefore regarded as particularly important, at this stage in the analysis of the Study 2 data, to check the wording of all partner-generated strategies.

From this grouping process, it was found that there were three main groups of Reflective Strategies, involving ‘making inferences’, ‘reflecting’ and ‘requesting’. While the former two groups occurred in Study 1, the third group did not. Requesting information was a clear reflection of the interactive pair task in Study 2. Close inspection of the sample strategy descriptors below indicates that some of them are *reader-generated* and some are *partner-generated*, indicated in square brackets, respectively, as RG and PG:

- requesting clarification on meaning in [part of] text [RG]
- requesting partner's agreement on meaning in [part of] text [RG]
- requesting partner's opinion on semantic consonance between text evidence and own word choice [RG]
- requesting word with similar meaning to own word choice [RG]
- requesting clarification on syntactic category of cloze item [RG]
- requesting justification for syntactic category of partner's word choice [PG]
- requesting partner's opinion on syntactic category of cloze item [RG]
- requesting clarification on partner's inference or word choice [PG]

Most of the examples above clearly illustrate the reader taking the initiative, though the penultimate example is less obvious in this respect. Though the word ‘partner’ is used, this descriptor nonetheless represents a speaker’s original statement.

These examples from the ‘requesting’ group of strategies show the potential for ‘core’ strategy descriptors for requesting information. Specifically, the first four examples



are all meaning-centred and were later conflated into the one descriptor. While there were sixty-four occurrences of the first strategy, the next three meaning-centred descriptors each occurred only once. On checking the subjects' transcripts, I realized that, without losing any essential part of the mental processes pertinent to the current study, each of the three low-frequency strategy descriptors could be subsumed under the first one. For instance, while the following subject utterance was originally coded as the second descriptor listed above, it could equally well be coded as the first:

Item 15

S7: T156        “So you also agree that, uh, for the blank fifteen, should be a negative one ...?”

Thus, the first descriptor listed above acted as a core strategy (Coding Principle 6.4). As in Study 1, core strategies also incorporated part of text utilized and, in the case of Reflective Text Language-Centred Strategies, aspect of language reported. A similar process to that described here for minimizing the number of meaning-centred strategy descriptors took place for their language-centred counterparts.

Given that the Study 2 task involved pair interaction, deciding whether subjects' utterances were original (i.e., reader-generated) or reactive (i.e., partner-generated) was a key element in the quest to maintain a faithful representation of the data (Coding Principle 6.6). Following Research Objective 5 (discussed below in Section 4.0), this differentiation was also necessary in order to be able to compare the Study 1 findings, all of which were reader-generated, with the reader-generated findings in Study 2. I therefore decided to check the 'making inferences' and 'reflecting' groups of Reflective Strategies to see if they also represented both reader-generated and partner-generated utterances. Indeed, this check revealed plenty of evidence of pair interaction. For clarity in the examples that follow, the part of the wording describing the reactive aspect of the strategies is italicized:

- making an inference about meaning in text [sentence] *that supports partner's word choice*
- reflecting on meaning in text [word] *in response to partner's clarification request*

The box below provides an illustrated example of this type of strategy descriptor. S38 reactively makes an inference about language; in other words, without S37’s assertion in Turn 69, S38 almost certainly would never have made his utterance in Turn 70:

Strategy descriptor	Transcript extract
Making an inference about language [grammatical category] in text [word] <i>that conflicts with partner’s inference</i>	<div>Item 6</div> <div>S37: T69      “No, I don’t think so because <u>pertaining</u> is not a noun, so we shouldn’t put ...”</div> <div>S38: T70      “No, <u>pertaining</u> can be a noun!”</div> <div>S37: T71      “No, I don’t think so because it’s a verb ...”</div>

A further pass through the data set was conducted in order to verify all the partner-generated strategies and to harmonize the wording of their descriptors. The italicized parts of the strategy descriptors in the list and the box above show the wording most frequently used for this type of strategies. Essentially, these were responses to partners’ clarification requests and inferences that either supported or conflicted with a partner’s word choice or inference.

During this pass through the data, I also checked the coding I had given to the partner-generated utterances. In order to differentiate among various factors in the partner-generated utterances appearing in the data, I decided to establish separate codes for these factors. These factors were: reaction to partner’s word choice, reaction to partner’s and reader’s word choice(s), reaction to partner’s clarification request about partner’s word, and reaction to partner’s clarification request about text or reader’s word choice (Coding Principle 6.6).

**Interactional utterances**

Apart from yielding partner-generated strategies in the data, the pair task in Study 2 yielded utterances that did not contain information qualifying as reading strategies. There was considerable evidence of turns where subjects were responding to a clarification request without using a reading strategy or were merely communicating in a social or phatic manner. These turns were initially put into a ‘miscellaneous’ group of strategies, but, since they do not conform to the current definition of reading



strategies, they were later termed ‘interactional utterances’ (Coding Principle 6.5). Some examples follow:

- expressing inability to complete item
- expressing inability to answer partner’s clarification request
- confirming own word choice in response to partner’s clarification request
- complimenting partner
- thanking partner
- apologizing
- laughing
- joking
- interjecting

The box below illustrates an interactional utterance from the data:

Interactional utterance	Transcript extract
Expressing inability to complete item	<u>Item 4</u> S22: T14        “I don’t know.”

Summary

The interactive nature of the introspective cloze task meant that strategies recorded were both *reader-generated* and *partner-generated*. It also led to numerous ‘interactional utterances’, that is, instances of phatic communication between pairs of subjects. With the research aims and objectives in mind, in particular those seeking compatibility between Studies 1 and 2, as well as space constraints, interactional utterances were excluded from the findings. Moreover, since pair interaction was not a focus of the current research, partner-generated strategies were not considered in the findings.

### 3.4 Review phase

At this point, I conducted a review of the analysis process to find answers to the following questions:

- Are *Reflective* Strategies, with the subcategories and their respective descriptor groups developed in Study 1, transferable to Study 2? (As discussed above, *Sensory* Strategies were not included in the data analysis for Study 2.)
- Are key aspects of the final categorization scheme not yet covered, such as ‘compound’ strategies, transferable to Study 2?
- Is there any evidence of factors occurring in the Study 2 data that did not occur in the Study 1 data? If so, what are they?

These questions are now addressed in the sequence shown above.

#### **Reflective Strategies: Subcategories and descriptor groups**

The subcategories of Reflective Strategies are discussed and illustrated in the following sequence: Text Meaning-Centred (as single strategies), Text Language-Centred (also as singles), compounds with these subcategories as well as Extra-Text-Centred and, finally, Non-Specific Strategies. As shown in Table 6.6 above, every subcategory contained the descriptor group ‘requesting clarification’, which was not found in Study 1 because of its individual silent reading task; other variations are given separately for each subcategory. In all extracts from transcripts, I have used italics to show subjects’ word choices for cloze items and underlining to indicate sections read aloud from the text.

#### ***Single strategies***

In the Text Meaning-Centred subcategory, there were no examples in Study 2 of the ‘remembering’ descriptor group (which had occurred in low frequencies in Study 1). Table 6.9 illustrates the other groups, the first two of which were common to Studies 1 and 2:



Table 6.9: Reflective Text Meaning-Centred Strategies (singles)

Strategy descriptor	Transcript extract
Making an inference about meaning in text [clause] that conflicts with own word choice	<u>Item 12</u> S14: T100 “I think it is an action, yeah ... but <i>action</i> have a similar meaning to the word <u>reactions</u> .”
Reflecting on meaning in text [phrase: headline]	<u>Item 15</u> S6: T218 “Uh, not decided good effects or bad effects, not yet ...”
Requesting clarification on meaning in text [cloze item]	<u>Item 15</u> S7: T156 “So you also agree that, uh, for blank fifteen, should be a negative one ...?”

In the example of the ‘making inferences’ group, S14 infers that the meaning of the missing word contains the sense of ‘action’. However, her use of the word ‘but’ suggests that she feels that the actual word ‘action’ would be inappropriate for the cloze blank because the text contains the similar word ‘reactions’ in the same phrase. In the second example, S6 has not yet decided if Item 15 requires a word with a ‘good’ or ‘bad’ meaning. With reference to the same item, the third example illustrates S7 seeking clarification from her partner about whether the item requires a ‘positive’ or ‘negative’ meaning.

There were three descriptor groups of Text Language-Centred Strategies. In the first example in Table 6.10, S4 makes the inference that her previously stated word choice ‘related’ collocates with the word ‘closely’ in the text. In the second example, S10 is reflecting on the grammatical category of Item 21 but cannot make up her mind. In the third example, S2 makes a simple clarification request about the grammatical category of Item 20.

Table 6.10: Reflective Text Language-Centred Strategies (singles)

Strategy descriptor	Transcript extract
Making an inference about language [collocation] in text [phrase: headline] that supports own word choice	<u>Item 16</u> S4: T193 “... because of the word <u>closely</u> .”
Reflecting on language [grammatical category] in text [cloze item]	<u>Item 21</u> S10: T256 “I’ve changed my idea ... it must not be a noun, it must be an adjective or adverb ...”
Requesting clarification on language [grammatical category] in text [cloze item]	<u>Item 20</u> S2: T141 “It’s a noun, isn’t it?”

*Compound strategies*

Since one of the two combinations of compounds recorded in Study 2 was Text Meaning-Centred/Text Language-Centred, these are included here. Table 6.11 shows how S21, in the first instance, infers both the grammatical category of the cloze item and the meaning required and, in the second instance, seeks clarification on whether the determiner ‘such’ refers to the preceding list of newspaper headlines in the text.

**Table 6.11: Reflective Text Meaning-Centred/Text Language-Centred Strategies (compounds)**

Strategy descriptor	Transcript extract
Making an inference about meaning and language [grammatical category] in text [sentence?] that supports word choice	<u>Item 9</u> S21: T46 “Number nine I think is ... we need adjective ... something like ... <i>somewhat</i> to explain the level of <u>violence</u> ...”
Requesting clarification on meaning and language [reference] in text [word]	<u>Item 19</u> S21: T95b “I want to ask if the word <u>such</u> means that facts [sic] mentioned above.”

As discussed in Chapter 5, Extra-Text-Centred Strategies invariably occur in compounds. In contrast to Study 1, however, they co-occurred only with Text Meaning-Centred Strategies in Study 2. Another point of difference was that only the ‘making inferences’ descriptor group occurred in both studies. As shown in Table 6.11, the only other descriptor group in Study 2 was ‘requesting clarification’.

The descriptor groups missing from Study 2 were ‘reflecting’ and ‘expressing personal reactions’. In both cases, there was probably a task effect. It seems feasible that the combination of a pair task and a cloze activity led to fewer reports of simply ‘reflecting’ (i.e., without a stated outcome) and more of ‘making inferences’ (i.e., stating some sort of outcome in pursuit of the task requirement to complete gaps). In addition, the free-reading task in Study 1, albeit with a researcher-selected text, is more likely to give the reader the sense of freedom and control that might lead to personal reactions. In contrast, a cloze task, especially when there is a time limit, seems likely to foster an almost headlong and linear pursuit of the completion of numbered items. The variation in text topics across the two studies may also have been a factor. These issues are discussed further in Chapter 7.



**Table 6.12: Reflective Extra-Text-Centred/Text Meaning-Centred Strategies (compounds)**

Strategy descriptor	Transcript extract
Making an inference about meaning from prior knowledge and text [sentence] that supports own word choice	<u>Item 6</u> S1: T239-241 “I think about <i>thought</i> ... people thinking ... because people always <u>change</u> their mind.”
Requesting extra-textual information related to text topic	<u>Item 15</u> S7: T146b “... do you heard [sic] of any ... any news about the <u>drug</u> can help to improve the <u>memory</u> of somebody?”

The first example in Table 6.12 illustrates how prior knowledge can combine with a focus on text meaning to help a reader make an inference. In this case, S1 makes the word choice ‘thought’ on the basis that the text conveys the meaning of change and her real world experience tells her that people “always” change their minds. The second example illustrates a case where a reader seeks extra-textual information related to meaning in the text. Here, the text is the newspaper headline “New drug discovered to (15) \_\_\_\_\_ memory” and S7 is unsure as to whether the word ‘drug’ invariably has a negative meaning. (She later realizes that the word ‘discovered’ in the text probably means that it has a positive meaning and she and her partner choose the words ‘enhance’ and ‘facilitate’ for Item 15.)

*Non-Specific Strategies*

There are four subcategories of Reflective Strategies, one of which occurred only in Study 2 (Table 6.6 above). This subcategory, Non-Specific, is now described and illustrated. There were a substantial number of cases in the cloze data where there was insufficient information for me to conclude whether inferences reported were Text Meaning-Centred, Text Language-Centred or, occasionally, Extra-Text-Centred (Coding Principle 6.5). Therefore, this neutral subcategory of Reflective Strategies was established.

**Table 6.13: Reflective Non-Specific Strategies (singles)**

Strategy descriptor	Transcript extract
Making an inference from text [sentence] that supports own word choice	<u>Item 14</u> S6: T376            “ <u>Everyone has seen newspaper</u> ... <u>Newspaper article</u> is very natural ...”
Requesting partner’s opinion on own word choice	<u>Item 12</u> S17: T183           “Do you think that <i>thinking</i> is also OK?”

Table 6.13 presents an illustrated example of each of the descriptor groups. In the first case, S6 makes an inference that could be either meaning-centred, language-centred or both. As suggested in Chapter 7, follow-up interviews might disentangle these possibilities. The second instance illustrates S17 requesting his partner’s opinion of his word choice.

Interestingly, the lack of specific information about whether subjects were focussing on meaning, language or both sometimes seemed to occur in the items which subjects found most easy. For instance:

Item 17

S9	T133	“But we know <u>solving</u> ... <i>problem</i> <u>solving</u> ...”
S10	T134	“ <i>Problem</i> <u>solving</u> , yes ...”
S9	T135	“... just by guessing ...”

Here, Turns 133 and 135 were coded as ‘making an inference from text [phrase]’. Since it is unclear if S9 inferred her answer from meaning, language or a combination of both, the only safe option was to code her response as Non-Specific rather than Text Meaning-Centred or Text Language-Centred (Coding Principle 6.5).

**3.5 Section summary**

This section described the procedures undertaken to analyze and code the Study 2 data. Despite the differences in introspection mode, task mode and task design between Studies 1 and 2, only fairly minor modifications were needed for the categorization scheme for reading strategies (whose development was described in detail in Chapter 5) to be transferable to Study 2.

At the highest level, both categories of strategies in the scheme occurred in each study, though, due to variation in introspection and task mode, it would have been very hard to compare *Sensory* Strategies (i.e., those that are physically observable) across the studies. On the other hand, occurrences of *Reflective* Strategies (i.e., those that are exclusively mental) appeared to be comparable across the studies. At a lower level, all three subcategories of Reflective Strategies recorded in Study 1 were found



in Study 2 and only minor changes in the wording of descriptors were found to be necessary. Also occurring at this level were 'compound' strategies (i.e., those combining more than one element). An additional subcategory of Reflective Strategies, called Non-Specific, occurred in the Study 2 data. Since most of these were in the 'requesting clarification' descriptor group, this subcategory may have emerged in the data largely as a result of the pair dimension of the task. Also at the level of descriptor, the notion of 'core' strategies (i.e., those with the same overall wording but with variation in part of text and aspect of language recorded) was compatible with the Study 2 data.

#### **4.0 SOME TENTATIVE FINDINGS FROM STUDIES 1 AND 2**

In an ideal situation, presentation of detailed findings using the categorization scheme developed would help to demonstrate the capacity of the scheme to provide a comprehensive record of reports of reading behaviour. Indeed, this was my original intention with Research Objective 5, which was to compare the findings of Studies 1 and 2. However, due to the space constraints already noted, I had to make a decision about priorities for inclusion in this thesis. Given the small number of subjects (N=38) and the imbalance in the quantity of the data generated across the two studies, any findings would have to be interpreted very cautiously. Therefore, all findings, particularly those relating to individual variation, would be akin to those of a case study. Consequently, no firm conclusions could be drawn from them about the population under study; still less could generalizations to other populations be made.

In the end, I decided that it was more important to maintain a research focus on the development of the scheme than to dilute the focus by giving an incomplete description of the process and making room for findings that do not really make a contribution towards advancing our knowledge. Nonetheless, I felt that it would be a valuable use of space to present a few selected findings here. Such an exercise could serve to substantiate some of the claims made in this thesis for the categorization scheme; in particular, it could provide the evidence sought in Research Objective 2 that the scheme was viable across the two studies.

With this goal in mind, I had to decide which aspects of the findings across the two studies would best achieve it. First, to avoid a level of detail for which there was no space, I excluded results relating to the relatively low level of strategy descriptor group, focusing instead on the higher levels of strategy category and subcategory. However, as described above, given the difficulties I experienced in recording identical physical actions across the two studies, I excluded Sensory Strategies, restricting the findings to Reflective Strategies (i.e., the category describing mental actions). Consequently, since 'motivators' (i.e., evidence of reading purpose) were only recorded with Sensory Strategies, as noted in Chapter 5, these were excluded from the results. In addition, since the Non-Specific subcategory of Reflective Strategies was recorded in Study 2 only, they were excluded from the findings.



Finally, in order to maximize the potential to compare findings across the two studies, I included only ‘reader-generated’ strategies in Study 2, thus excluding all those marked as ‘partner-generated’ that were recorded with the cloze pair task.

Observing these exclusions, findings concerning the following are now presented: the use of Reflective Strategies, the use of various parts of the text with these strategies and evidence of individual variation. Discussion of the findings includes reference to part of Study 1, which, for space reasons, was dropped from this thesis. The focus of this part of the study was on what readers reported as triggering ‘breakpoints’ in their reading of the text. (Breakpoints and breakpoint cycles are explained in Chapter 4.)

**Use of Reflective Strategies**

Table 6.14 presents a comparison in the use of three subcategories of Reflective Strategies across the two studies: Extra-Text-Centred, Text Meaning-Centred and Text Language-Centred. The comparison should be interpreted cautiously, not least because the total number of strategies recorded was more than four times greater in Study 2 than in Study 1.

**Table 6.14: Use of Reflective Strategies compared across studies**

Strategy	Study 1	Study 2
ETC	22 (17.32%)	57 (10.77%)
TMC	91 (71.65%)	314 (59.36%)
TLC	14 (11.02%)	158 (29.87%)
Total	127 (100.00%)	529 (100.00%)

**Key for Table 6.14: Codes for Reflective Strategies**

ETC = Extra-Text-Centred

TMC = Text Meaning-Centred

TLC = Text Language-Centred

Nonetheless, what does seem clear from Table 6.14 is the primacy of meaning, at least as evidenced by the relative frequencies of Text Meaning-Centred Strategies. Across Studies 1 and 2, this subcategory comprises 71.65% and 59.36%, respectively, of the total. Since the task in Study 1 was more naturalistic than that in Study 2, the figure

for Study 1 is more likely to reflect the norm in silent reading of this genre among these readers.

In addition, Study 1 shows a higher figure for Extra-Text-Centred Strategies, which comprise 17.32% of the total, as against 10.77% in Study 2. This difference may be text-specific rather than task-specific. As has already been suggested, the topic of the Study 1 text, which was on learner independence among university students in Hong Kong, may have lent itself to reference to prior knowledge and the expression of personal reactions more than that of the Study 2 text. The latter was part of an introduction to a set book on psychology, a subject that many of the subjects had reported finding difficult.

Finally and not surprisingly, the cloze task in Study 2 appears to have triggered a considerably greater focus on language than the naturalistic task in Study 1, with Text Language-Centred Strategies comprising 29.87% as against 11.02%, respectively, of the total.

### **Evidence of global and local reading**

Partly in order to account for part of text utilized with reading strategies and thereby to reveal the extent to which the readers reported reading relatively globally or locally, I developed the concept of the 'core' strategy during data analysis in both studies. With the aim of clarifying the findings, five consolidated categories, shown in Table 6.15, were developed from an array of data-driven text categories; question marks set against data-driven categories indicate that I was almost, but not completely, certain of the category.

Across the two studies, the 'Global' category comprises references to the entire text and, in the case of Study 1, to sections and subsections. For the Study 1 text (Appendix 5.1), the 'Paragraph' category comprises references to the current paragraph and to an additional preceding or subsequent paragraph; since the writer took an entire paragraph to outline the macrostructure of the text (Lines 12-18), this was included in the category. For the five-paragraph text in Study 2 (Appendix 6.1),



this category included evidence of processing that was beyond the current paragraph but fell short of involving the entire text.

**Table 6.15: Text categories**

Consolidated category	From Study 1 data	From Study 2 data
Global	Entire text, Section/Subsection	Entire text
Paragraph	2 paragraphs, 2 paragraphs?, Paragraph, Paragraph?, Identification of macrostructure	Beyond paragraph, Beyond paragraph?, Paragraph, Paragraph?
Sentence	2 sentences, 2 sentences?, Sentence, Sentence?, Quotation(s)	Beyond sentence, Beyond sentence?, Sentence, Sentence?, Main clause, Main clause?
Local	Subordinate clause, Heading/Subheading, Phrase(s), Word(s), Word(s)?	Subordinate clause, Subordinate clause?, List of eight newspaper headlines, Headline, Phrase, Phrase?, Word(s)
Task-Specific	N/A	Own word choice, Cloze item

Since quotations referred to by the subjects in Study 1 were either a maximum of two sentences long (Lines 31-34 and 44-47) or listed as four separate quotations, each a sentence in length (Lines 5-11), they were included in the ‘Sentence’ category. In Study 2, this category included reports involving main clauses. The ‘Local’ category comprised anything from subordinate clauses to words and phrases. Headings and subheadings in the Study 1 text were included in this category as, notwithstanding the import of their content, they all consisted of words or phrases, thus requiring only local processing to decode. The Study 2 text contained a list of eight newspaper headlines in the third paragraph; in a grammatical sense, most of these headings were phrases, though one of them had a main verb. Finally, the ‘Task-Specific’ category pertained to Study 2 only and included references to readers’ own word choices and to cloze items.

Using the consolidated categories, Table 6.16 compares part of text recorded with all subcategories of Reflective Strategies across the two studies; the figures include instances that, for lack of complete certainty, were coded with question marks (Coding Principles 5.5 and 6.5). The main difference is that there were very few reports of processing beyond the sentence in Study 2. When combining the two highest categories (Global and Paragraph), these amounted to only 8.49%; the equivalent figure for Study 1 is 32.67%, about four times higher. Combining the two lowest

categories (Local and Task-Specific) reveals a total of 61.66% of all reports in Study 2; this contrasts with a somewhat lower Local figure for Study 1 of 46.53%. However, this latter figure is itself remarkably high and is perhaps indicative of the position held, among others, by Rayner and Pollatsek (1989), that fluent and accurate reading requires basic skills in bottom-up processing.

**Table 6.16: Part of text utilized with Reflective Strategies compared**

Consolidated text category	Study 1	Study 2
Global	8 (7.92%)	6 (1.08%)
Paragraph	25 (24.75%)	41 (7.41%)
Sentence	15 (14.85%)	165 (29.84%)
Local	47 (46.53%)	173 (31.28%)
Task-Specific	N/A	168 (30.38%)
Unclassified	6 (5.94%)	N/A
<b>Total</b>	<b>101 (100.00%)</b>	<b>553 (100.00%)</b>

**Note for Table 6.16**

- Since the data were drawn from a free-reading task in Study 1, there were no instances of 'Task-Specific' reference to part of text.
- The figures for Study 1 include 8 instances of compound strategies that combined Reflective and Sensory Strategies.

These findings may reveal evidence of a task effect, the cloze task requiring, or at least inducing, relatively large amounts of local processing. Nevertheless, they appear to conflict with Jonz (1990), who, as noted in Chapter 3, found compelling evidence that cloze has the potential to require both local and global processing. This evidence of conflicting conclusions between Jonz, who was investigating 'random' cloze, and the current study, which used a form of 'rational' cloze designed to encourage higher-level processing, might be explained as follows. Some readers, such as those in the Hong Kong study, may have a kind of reflex reaction to being asked to complete a cloze task (be it 'random' or 'rational') that engenders local reading despite evidence that relatively global reading can also be useful in filling cloze blanks. (This issue is pursued in discussion of individual variation below.)

Finally, Table 6.16 indicates that, whereas Study 2 shows progressively fewer instances as the text categories move towards the Global category, Study 1 shows no such trend, there being more reports in the Paragraph category (24.75%) than in the Sentence category (14.85%). Parenthetically, the latter finding is consistent with one



in the part of Study 1 that focused on 'breakpoints' which, for space reasons, was excluded from this thesis. This finding was that 59.12% of all breakpoints recorded occurred at the ends of paragraphs or sections in the text and gave substance to the notion that readers, at least in silent free-reading tasks, often use higher-level text breaks to reflect on what they have just read.

### **Evidence of individual variation**

The last point of comparison between the two studies is that of individual variation in the use of Reflective Strategies. In order to focus on a manageable number of individual subjects and to highlight any variation in the data, I decided to conduct further analyses on small groups that appeared to share particular characteristics. These characteristics were based on findings of that part of Study 1, originally intended to be part of this thesis, which focused on 'breakpoints'. Four categories of breakpoint were developed: 'Reader Contribution', 'Meaning in Text', 'Language in Text' and 'Extraneous Factors'. The first three of these, which have similar coverage, respectively, to the Extra-Text-Centred, Text Meaning-Centred and Text Language-Centred subcategories of Reflective Strategies, are relevant to the ensuing discussion.

From the analysis of four breakpoints reported by each subject, there emerged a group of six subjects (S11, S12, S17, S19, S24 and S38) who reported only 'Meaning in Text' breakpoints and a group of seven subjects (S2, S3, S4, S8, S10, S14 and S26) who reported only 'Language in Text' breakpoints. These groups of readers will now be referred to, respectively, as 'Text Meaning-Oriented' and 'Text Language-Oriented'. Further, in order to investigate strategy use among readers reporting extra-textual factors in breakpoints, I decided, for the purpose of investigating individual variation, to establish a third small group. To qualify for this group, subjects had to report use of two or more 'Reader Contribution' breakpoints. (Since none of the subjects reported exclusively extra-text-centred breakpoints, it would have been impossible to have formed a top-down-centred group on the same basis as the other two groups, that is, exclusive use of one breakpoint category.) Five subjects (S21, S25, S28, S32 and S37) qualified for this group, which was given the name 'Extra-

Text-Oriented’. Taken together, these three reader groups contained eighteen subjects out of the total of thirty-eight participants in the current research.

Table 6.17 presents a comparison of the Reflective Strategies used by these three groups across the two studies. Since the reader groups had varying N sizes and since the database was small, the figures need to be interpreted with caution. The most interesting finding here appears to be that, while the Extra-Text-Oriented and Text Meaning-Oriented readers showed considerable variation in strategy use across the two studies, the Text Language-Oriented readers showed almost identical use of all the strategies. This strongly suggests that the Text Language-Oriented readers approached the free-reading task in broadly the same manner as they approached the cloze task. They did so using mainly Text Meaning-Centred Strategies (about two thirds of the strategies they used) informed by a combination of Text Language-Centred strategies (about one in five) and Extra-Text-Centred Strategies (about one in ten).

**Table 6.17: Use of Reflective Strategies among selected reader groups compared**

Strategy	ETO readers		TMO readers		TLO readers	
	<i>Study 1</i>	<i>Study 2</i>	<i>Study 1</i>	<i>Study 2</i>	<i>Study 1</i>	<i>Study 2</i>
ETC	35.48%	11.36%	0.00%	6.33%	10.53%	12.50%
TMC	61.29%	52.27%	100.00%	49.36%	68.42%	65.44%
TLC	3.23%	36.36%	0.00%	44.30%	21.05%	22.06%

**Key for Table 6.17: Codes**

*Reader groups*

ETO = Extra-Text-Oriented (N=5)  
TMO = Text Meaning-Oriented (N=6)  
TLO = Text Language-Oriented (N=7)

*Reflective Strategies*

ETC = Extra-Text-Centred  
TMC = Text Meaning-Centred  
TLC = Text Language-Centred

In contrast, the Extra-Text-Oriented and Text Meaning-Oriented readers appeared to vary in the way in which they approached the two tasks. Both these groups used more Text Meaning-Centred Strategies in the free-reading task in Study 1 than in the cloze task in Study 2, the Text Meaning-Oriented readers actually reporting their exclusive use. Both groups reported using far more Text Language-Centred Strategies in the cloze task (up from 3.23% in Study 1 to 36.36% in Study 2 and from 0.00% to 44.30%, respectively). Moreover, the gulf between the Extra-Text-Oriented and Text Meaning-Oriented readers’ use of Extra-Text-Centred Strategies narrowed between Study 1 (35.48% and 0.00%, respectively) and Study 2 (11.36% and 6.33%,



respectively). There is no obvious reason for this, though one could speculate that the Extra-Text-Oriented readers were particularly interested in the Study 1 text topic (i.e., learner independence among university students in Hong Kong). They were clear outliers in their frequent use of Extra-Text-Centred Strategies in Study 1 (35.48% as against the corresponding Text Meaning-Oriented figure of 0.00% and the Text Language-Oriented figure of 10.53%).

The finding relating to the Text Language-Oriented reader group in Table 6.17 suggests the following questions:

- Are the Text Language-Oriented readers *actually* language-centred?
- Is their reading distinctive in any way from that of the other groups of readers?

In order to answer the first of these questions, these subjects' protocols were checked for evidence of patterns within 'breakpoint cycles' (i.e., between breakpoint triggers and strategies, as shown in Figure 4.1 in Chapter 4). I found that 12 of the 13 Text Meaning-Centred Strategies reported by the Text Language-Oriented readers followed 'Language in Text' breakpoint triggers that had been coded 'unknown vocabulary item or phrase'; the remaining case followed a 'Language in Text' breakpoint trigger coded 'typography'. A typical instance of the former, taken from Study 1, follows:

S14: L42        "I don't know the word subsequent. I just skip it and then reread the whole sentence and guess its meaning."

These findings suggest that the Text Language-Oriented group of readers, while language-centred in terms of factors reported to trigger breakpoints, are in fact meaning-centred (at least according to the current scheme), but at a relatively local level in the text. In other words, they appear to have a greater tendency than the readers in the other groups to process text information bottom-up. This seems to occur despite the fact that the subjects' language proficiency levels (as measured by their overall grades in the 'Use of English' exam, described in Chapter 4) are indistinguishable. L2 proficiency level, then, does not appear, on its own, to determine the extent to which these readers were reading bottom-up. Individual variation seemed to be the prevailing factor, something that tallies with one of Sarig's (1987b) findings. Working with ten L2 high school subjects, she found that 87% of

the ‘moves’ she recorded were ‘maximally unique’, by which she meant that they were “used by only four or fewer subjects with different intensity (i.e., frequency)” (p. 116).

The second question above asked if the Text Language-Oriented group read in a manner distinct from that of the other groups. In order to answer this, and in an attempt to substantiate the idea that the Text Language-Oriented group operates locally, I checked and compared part of text utilized with strategies among the three groups. Using the same consolidated categories of text units as those developed above in Table 6.15, Table 6.18 presents comparative figures for the two studies. Again, because the N sizes of the reader groups vary and the database is small, the figures should be interpreted cautiously.

**Table 6.18: Part of text utilized among selected readers compared**

Consolidated text category	ETO readers		TMO readers		TLO readers	
	<i>Study 1</i>	<i>Study 2</i>	<i>Study 1</i>	<i>Study 2</i>	<i>Study 1</i>	<i>Study 2</i>
Global	20.00%	2.17%	10.00%	2.35%	0.00%	5.51%
Paragraph	32.00%	0.00%	55.00%	1.18%	0.00%	4.72%
Sentence	28.00%	36.96%	25.00%	21.18%	21.86%	33.86%
Local	12.00%	21.74%	5.00%	42.35%	68.75%	36.22%
Task-Specific	N/A	39.13%	N/A	32.94%	N/A	19.69%
Unclassified	8.00%	N/A	5.00%	N/A	9.38%	N/A
Total (raw)	25	46	20	85	32	127

**Key for Table 6.18: Codes for reader groups**  
ETO = Extra-Text-Oriented (N=5)  
TMO = Text Meaning-Oriented (N=6)  
TLO = Text Language-Oriented (N=7)

In general, the evidence in Table 6.18 suggests that there was more processing at lower levels in Study 2 than in Study 1. In Study 1, the Text Meaning-Oriented readers were clustered in the Paragraph category and, to a lesser extent, in the Sentence category, together amounting to 80.00% of their reports of strategy use; in contrast, in Study 2, these same categories amounted to only 22.36%. The Text Language-Oriented readers, on the other hand, were clustered at a lower level in both studies; when combining the Local and Task-Specific categories, the figures amounted to 68.75% of their strategy use in Study 1 and 55.91% in Study 2. However, in Study 2, they did not report reading as locally as did either of the other two groups. The combined figures for the Local and Task-Specific categories show



that the Text Meaning-Oriented readers (at 75.29%) and the Extra-Text-Oriented readers (at 60.87%) both reported reading at a lower level in the text than the Text Language-Centred readers. It is the inclusion of Task-Specific instances that accounts for the relatively high percentages for the former groups of readers, with the Extra-Text-Oriented readers reporting twice as many cases as the Text Language-Centred readers. There doesn't seem to be an obvious explanation for this finding.

Interestingly, however, the Text Language-Oriented readers were recorded noticeably more often in the Sentence category in Study 2 than in Study 1 (33.86% against 21.86%, respectively). It could be that the cloze task actually had the effect of making these readers read at a higher level in the text than they would do in unconstrained reading. This echoes Jonz's (1990) finding, mentioned above, that higher-level processing can help to complete cloze items.

In further contrast, the only readers to report using a wide range of text units were those in the Extra-Text-Oriented group in Study 1. They reported more strategies in the Global category (20.00%) than the other two groups and yet they also reported more strategies than the Text Meaning-Oriented readers in the Local category (12.00%, as against 5.00%). Thus, the range of text they reported in Study 1 was spread relatively evenly and widely. Two observations about the Extra-Text-Oriented readers' behaviour in Study 1 can, therefore, be made. They appear to be processing top-down more than often than the readers in either of the other two groups. However, since they also appear to be considerably more flexible in part of text utilized during reading, this is not to say that they are top-down-dominant.

### **Difficulties in labelling groups of readers**

The findings in Tables 6.17 and 6.18 suggest that the names I have given these three groups of readers may not be optimal, though it needs to be stressed that the data informing *any* label for the groups are limited in at least three respects. First, the data are drawn from reading following breakpoints, and uninterrupted ongoing reading might yield different evidence (though its concurrent reporting is, naturally, inaccessible to researchers). Second, as has been reiterated, the N sizes of the three

groups varied and the database itself was very small. The third reason is the diversity in tasks observed across the two studies; while the data in Study 1 emanated from a relatively naturalistic task, those in Study 2 were constrained by the cloze task.

However, having acknowledged these limitations, it still seems worth considering other, more suitable, names for these three groups of readers, particularly with reference to the Study 1 data, which are more likely to represent 'normal' reading than the Study 2 data. Basing tentative labels for groups of readers on part of text reported, rather than on strategies recorded in a particular scheme, has the potential advantage that part of text is ultimately testable through the study of eye movements.

The Text Language-Oriented group might be 'Local' or 'Narrow Spectrum' readers. Though showing a marked preference for local reading, which manifested itself as a concentration on words and phrases and distinguished them from the other two groups, they nevertheless seem to attribute primacy to the extraction of text meaning (at least according to the current scheme). This is a characteristic that they share with the other groups. It appears, therefore, that what varies among the groups is the chosen or habitual route to meaning rather than the overall drive, or purpose, in reading. To complement the terms 'Local' or 'Narrow Spectrum', and consistent with the Study 1 findings, the Extra-Text-Oriented group might be renamed 'Global', 'Broad Spectrum' or 'Flexible' readers. Moreover, an intermediate term, such as 'Mid-Spectrum' readers, though problematic, might be adopted for the Text Meaning-Oriented group.

Naturally, especially with the above constraints in mind, all these possible labels should be seen as indicative of relative rather than absolute approaches to reading on the part of these subjects. Given texts other than that used in Study 1, they may report very different reading behaviour. A question that remains (but for which space constraints preclude further analysis) is whether it is possible to attribute even tentative labels to the other twenty readers who participated in the current research but whose data were not included in these findings on individual variation.



**Reader groups and the ‘Use of English’ examination**

Though based on very low figures, a comparison between the overall grades and the reading grades scored by subjects in the three reader groups in the ‘Use of English’ exam, discussed in Chapter 4, produces a finding that warrants further investigation with a much larger database. Table 6.19 shows that the Text Language-Oriented readers had a clear tendency (5/7) to obtain a higher reading grade than overall grade. In contrast, the Extra-Text-Oriented readers tended to obtain grades in the reverse direction (2 of the 3 known scores) while the Text Meaning-Oriented readers tended to score the same grade for reading as they did overall (3/6).

**Table 6.19: Comparison of reading and overall ‘Use of English’ grades**

Relationship between reading grades and overall grades	ETO readers	TMO readers	TLO readers
Higher reading grade	1	2	5
Same	0	3	1
Lower reading grade	2	1	0
Unknown	2	0	1

**Key for Table 6.19: Codes for reader groups**

ETO = Extra-Text-Oriented (N=5)

TMO = Text Meaning-Oriented (N=6)

TLO = Text Language-Oriented (N=7)

Although no firm conclusions can be drawn from findings that are akin to those of a case study, they leave a tantalising feeling that, given the nature of the ‘Use of English’ reading test, the Text Language-Oriented readers were most likely to score well in it. The reading test tended to focus on extraction of meaning through local reading. Consequently, it seems highly possible that these readers (i.e., those showing a tendency, irrespective of task, to be word-based and to read locally for meaning) should obtain higher scores than those who tend to read relatively globally for meaning.

## Summary

Using the current categorization scheme, there were several interesting findings. First, as might be expected, there was strong evidence from these readers across both studies for the primacy of meaning extraction from the text. There was also evidence that text content may have affected the quantity and quality of top-down, or extra-text, processing, though task may also have played a part in this. Task also appeared to affect the degree of bottom-up processing in that there were considerably more reports of language-centred strategies and local reading in the cloze task than in the free-reading task. Finally, there was evidence of variation among three reader groups. In particular, a group designated as Text Language-Oriented readers reported approaching the tasks in the two studies in an almost identical manner in terms of strategy use; the two other reader groups reported substantial differences in strategy use across the two tasks.



## 5.0 CHAPTER SUMMARY

This chapter has described the second of two studies designed to run in parallel and whose data with respect to reading strategies could be categorized in the same classification scheme. While the data from both studies were mentalistic, those in Study 1 were reported in the written mode and those in Study 2 were reported orally.

Study 2 used the cloze procedure with a short academic text and the readers' task was to complete the items in pairs. During extensive pilot-testing of the task, the number of cloze items was reduced from fifty to twenty-five. Administration of the final version of the task ran smoothly in data-gathering sessions in which the readers were audio-tape-recorded working in pairs.

Analysis of the data was extremely labour intensive, beginning with the transcribing of the tapes and moving on to repeated passes through the data. It was necessary to sift very carefully through the data in order to separate the dross, in the form of purely interactional utterances and partner-generated strategies, from the essence, in the form of reader-generated strategies. Each pass through the data, conducted with reference to a list of coding principles, also offered an opportunity to modify the emerging strategy classification scheme as well as the emerging strategy descriptors. The process of establishing the strategy categories and descriptors took place in turns through input from the Study 1 data as well as that from the study reported in this chapter. The purpose in so doing was to develop a classification scheme and descriptors that were applicable across the variation designed into the two studies (Research Objectives 2 and 3).

The following chapter contains an evaluation of both Study 1 and Study 2 and, with the current research in mind, a taxonomy for categorizing reading strategies and recommendations for future directions in L2 reading research. In addition, pedagogical implications are considered.

## **CHAPTER 7: CONCLUSIONS**

### **1.0 RESEARCH REVIEW**

This final chapter of the thesis begins with an evaluative review of the current research, considers ways in which the field might move forward from its present state and discusses the pedagogic implications of the project.

The review opens by drawing together our knowledge base in the areas covered in the first three chapters of the thesis. This knowledge comprises theoretical evidence, evidence from attempts at describing and categorizing reading behaviour as strategies and evidence about mentalistic approaches to the collection of data. On this foundation, I will then evaluate the research project within the framework of the aims and objectives and go on to discuss its limitations.

#### **1.1 Reading theory and the study of reading behaviour**

Chapter 1 described how the study of reading has developed, intermittently flourishing and languishing, over the past 130 years or so. An important factor in its protracted demise, from about 1910 till 1970, was the rise of behaviourist psychology, which led to an era when only observable and measurable evidence was deemed to have any value. With the precipitate fall of behaviourism, the way was open for the study of cognitive processes, and both L1 and L2 reading have been among the beneficiaries of this development in psychology. Concomitantly, this resurgence of interest in the study of reading through introspection seemed to precipitate a revitalization of the study of reading through the use of techniques designed to obtain empirical evidence, such as that of the tachistoscope in the study of eye movements (e.g., Rayner and Pollatsek 1989).

Over the past three decades, therefore, there have been two parallel research efforts in the study of reading, one using empirical evidence collected in relatively artificial laboratory contexts, the other using mentalistic evidence collected, at least potentially, in relatively naturalistic settings. For the study of an activity that, in the normal



course of events, is silent and private, and therefore hard to access, this diversity of evidence seems to me to provide a healthy foundation for the development of theory. Nonetheless, a substantial proportion of the published studies seem to have been conducted by a single researcher, and collaborative work, which is relatively common in the hard sciences, has the advantage that creative and insightful developments can take place through the exchange of ideas during the research process.

As discussed in Chapter 1, converging evidence from studies using diverse data-collection techniques, including the recording of eye movements and introspection, suggests that automatic and accurate word decoding is paramount in fluent reading. There is also evidence that, in order for readers to understand the text, their prior knowledge of text content and genre interacts, perhaps concurrently, with this decoding process. From our current state of knowledge, therefore, one might tentatively describe reading as a bottom-up interactive process.

Nevertheless, although the weight of evidence suggests this view of reading, it needs to be interpreted very cautiously. This caution is reflected in the contrasting levels of certainty that have accompanied the presentation of reading models over the past thirty years. Proponents of some of the earlier models (e.g., Goodman 1967) appeared convinced by the truths of their own claims, albeit on very little and very narrowly based evidence. In contrast, more recent models, based on far more and far more broadly based evidence, have been presented in a much more tentative manner (e.g., Rayner and Pollatsek 1989). It is as though the more we learn, the more we realize how little we know.

Being tentative in the face of incomplete evidence is clearly an essential *modus operandi*. However, the difficulty we face in obtaining a more complete picture is not simply a reflection of the difficulty in penetrating the ‘black box’; it pertains to such factors as the extensive variation that exists both in readers and in the reading process. For L2 readers, for instance, prior knowledge of the language (and in some cases, of the Roman alphabet) may vary considerably. If reading is a bottom-up interactive process, then language deficiency may seriously jeopardize the reader’s ability to decode text. With this in mind, Bernhardt (1991) initiated work in capturing variation in language proficiency in her developmental theory of L2 reading.

In addition, as seen in the findings presented in the previous chapter, and corroborated in pivotal work conducted by Sarig (1987a and 1987b), variation in reading behaviour occurs across individuals at similar language proficiency levels. Hedge (1991) noted how individual readers, in attempting to process a text, vary in their 'modes' of reading. Moreover, as Devine (1988) found in her case study of two readers, individuals may, at least in some circumstances, have different default modes or 'internalized models of reading' (p. 129).

However, even where, for instance, some subjects report reading more globally than others in a study, this does not necessarily indicate top-down dominance. There is evidence that good readers decode faster than poor readers, leaving capacity to draw on other sources in the text and on their own prior knowledge (e.g., Perfetti and Roth 1981). The 'Extra-Text-Oriented' (or 'Broad Spectrum') reader group, discussed in the findings section of the previous chapter, may represent examples of such readers. Though the notion of 'good' and 'poor' readers was not the focus of the current study, being, or having the capacity to be, a 'good' L2 reader may, among other factors, be a function of language proficiency level.

Indeed, L2 readers are known to cite unfamiliar words as being central to reading difficulty even when their language proficiency level might indicate that they are capable of coping easily with a particular text. For instance, in part of Study 1 of the current research, whose findings were excluded due to space constraints, 27.59% of all 'breakpoints' recorded (i.e., including those that were not problem-oriented) were attributed to 'unknown or unfamiliar vocabulary item or phrase'. Yet, in face of the substantial evidence that efficient bottom-up processing is a prerequisite to fluent reading, the prevailing climate in EFL/ESL pedagogy is a top-down one. It is as though both data evidence, in the form of L2 readers' collective voice, and that of current reading theory are unheeded by the profession. This may be a reflection of the difficulties faced by the research community in developing a complete and empirically based reading model.

Part of the reason for the current state of incompleteness is that mentalistic data evidence is hard to describe and classify. Chapter 2 showed how attempts to describe



reading behaviour as 'reading strategies' have been fraught with difficulties, including wide diversity in published definitions of strategies and similarly wide diversity in classifications of strategies. A continuing problem has been that individual researchers have derived their categories solely from their own data, almost invariably generated from only one task, most studies making no reference at all to reading theory. It is this unsatisfactory state of affairs that motivated me to investigate the fundamentals of classifying reading behaviour.

Ironically, though, while the central problem with definitions and descriptions of reading behaviour has been the unnecessary and undesirable level of diversity, the reverse pertains when it comes to methodological approaches that have been adopted; here, the problem has been one of excessive uniformity. As shown in Chapter 3, which restricted its scope of inquiry to that of introspective tasks, little creative thought appears to have been given to the use of introspection. As far as I can ascertain, only one published study (Li and Munby 1996) has utilized the written mode in tapping readers' introspections. Yet there appears to be no obvious reason for the research community to limit itself to use of the spoken mode in the collection of introspective data.

Another methodological issue that arises with mentalistic approaches is that of the timing of data collection. With one exception, recent studies have either used one task designed to collect data concurrently with reading or have used one task for concurrent data and one or more tasks designed to collect data retrospectively. The exception is Kletzien (1991), who asked her L1 subjects to do three cloze texts and to report on their strategies after completing each passage. This widespread adherence to concurrent reporting may be due to the influence of information-processing theory, which suggests that information in short-term memory is more easily retrieved than that stored in long-term memory. While the current work followed this norm by seeking to restrict itself to concurrent introspection, as I observed in Chapter 3, it seems likely that, in order to access prior knowledge, a process known to occur widely in reading, recourse to long-term memory in reading must be routine. Moreover, there are alternative theories of memory.

There is also a certain amount of uniformity in the type of task typically used in introspective studies. In order to indicate where subjects are supposed to break and introspect, texts tend to be marked with dots or asterisks, sometimes as frequently as at the end of every phrase or every line in the text. If the research objective is to improve our understanding of uninterrupted reading, then it seems self-defeating to impose fixed points for introspection. Such researcher-generated points may be inserted for the purpose of task clarification and to dispel fears that subjects, reading without these 'roadblocks', simply will not introspect. As became clear in Study 2 of the current research, the use of imposed 'breakpoints' may reveal more about reading behaviour resulting from that particular task than that in normal reading.

To summarize, although there are certain fundamentals about reading that can now be asserted with relative confidence (such as the need to decode words whatever the reader's language proficiency level), we still have a very incomplete picture of the process. The research effort has been fragmented when it comes to describing and categorizing behaviour but curiously regimented when it comes to selecting techniques for data collection. The current research has attempted to bring a modicum of order to the categorization of reading behaviour and to throw some new light on mentalistic approaches to the gathering of data. In the following discussion, during the course of which some of the issues already raised will be considered further, the research will be assessed in terms of its aims and objectives.

## **1.2 Review of research aims and objectives**

The main thrust of the current research has been the development of a categorization scheme of reading strategies across the two studies. This scheme, whose evolution was described in the two previous chapters, has a number of features, among which are the following:

- three levels: category, subcategory and descriptor
- at all three levels, the capacity to mark or code strategies for the manner in which they were reported, either spontaneously or reactively (i.e., reacting to



interaction external to that between text and reader, such as that among text, reader and partner or researcher)

- at all three levels, the capacity to account for integrated or fused actions in ‘compound’ strategies
- at descriptor level, ‘core’ strategies with the flexibility to incorporate variation in part of text and aspect of language reported

To provide a framework for this evaluation of the current research, the discussion makes reference to the research aims and objectives, which were originally presented in Chapter 4.

### **Research Aim 1**

In a broad sense, this aim, which involved an online exploration of reading behaviour, was met. The data yielded by the Hong Kong subjects reading and introspecting in their L2 were generally clear and rich. Both the texts selected were at an appropriate level of difficulty, hard enough to precipitate ‘breakpoints’, but not so hard as to be frustrating. The choice of genre seemed to be successful, and it appeared to have face validity with the subjects. In Study 1, the utilization of a much longer text than is normally used in L2 reading research did not appear to create any problems; indeed, many of the subjects spontaneously reported enjoying reading it. It appears that, as a community, researchers have been unnecessarily cautious in their propensity to select short texts. Unfortunately, as will be shown below, I did not fully exploit the fact that the subjects were reading a relatively long text in Study 1.

### **Research Aim 2**

This aim involved the creation of a categorization scheme of L2 reading strategies consistent with theory and data evidence, and much of the current thesis comprises detailed description of the lengthy process of establishing such a scheme. A distinguishing feature of the scheme was that, though the driving force behind its development came from data, these data were drawn from two sources. In addition,

throughout the analysis of the two data sets, reference was made to theory. For instance, the combining of integrated actions into ‘compound’ strategies is a reflection of the parallel processing that might take place in the ‘message centre’ in Rumelhart’s interactive (1977) model and in Rayner and Pollatsek’s (1989) bottom-up interactive model. (Compound strategies are discussed further under Research Objective 4.)

### **Research Objective 1**

Following information-processing theory, the first objective was for the data in both studies to be collected while in the readers’ short-term memories. In order to enhance the possibility that the data retrieved were from short-term memory, the tasks devised involved ‘breakpoints’ in the reading flow. To encapsulate this recurring but transitory aspect of the reading process, the concept of the ‘breakpoint cycle’, shown in Figure 4.1 in Chapter 4, posited that, before normal reading was resumed, breakpoints were resolved in some way, usually through the use of one or more reading strategies. Tapping the reading process during breakpoint cycles certainly seemed to yield data that were generally clear and comprehensible; however, as already suggested and discussed below, hindsight indicates that, while modifying the original research design, I over-relied on information-processing theory.

### **Research Objective 2**

This objective was to establish a categorization scheme of L2 reading strategies that is viable across variation in methodology and tasks. This was partially achieved with a scheme that, as mentioned above, had three levels: categories, subcategories and descriptors. The first two of these levels are reviewed under this objective while the third is reviewed below under Research Objective 3.

The two categories, Reflective and Sensory, appear to be reasonably robust in that they represent the palpable divide between, respectively, behaviour that is unobservable and that which is observable. This reflects the two research traditions in



the study of reading, mentioned above, one of which uses mentalistic data while the other uses empirical data.

At the level of subcategory, there seems to be no problem with those describing observable actions, such as 'Recursive Reading' and 'Ongoing Reading' (i.e., subcategories of Sensory Strategies). However, I feel considerably less confident about the subcategories that seek to describe mental actions (i.e., subcategories of Reflective Strategies). These were, from top-down to bottom-up, 'Extra-Text-Centred', 'Text Meaning-Centred' and 'Text Language-Centred'. The first of these is probably the easiest to deal with and, in order to give a clearer idea of what it represents, it could be renamed 'Knowledge-Centred' or 'Reader-Centred'. The two latter subcategories, both of which deal with the processing of text information, present a greater challenge, particularly with respect to text meaning, which is a highly elusive concept. One possibility would be to conflate the two subcategories into one, possibly named 'Text-Centred' (thus complementing the 'Knowledge-Centred' or 'Reader-Centred' subcategory mentioned above). Where there were reports of tangible text language issues (e.g., the grammatical category of a word in the text), strategies could be marked or coded for language. However, the use of the term 'meaning' to label a subcategory, which was perhaps an over-ambitious attempt to represent the readers' reports in the data, could be dropped. There is clearly room for advancement on the labelling of the subcategories in the current scheme.

Turning to the ability of the categories to represent the data from the two studies, there was greater success with Reflective than with Sensory Strategies. Three of the four subcategories of Reflective Strategies were evident in both studies; however, Study 2 provided evidence of a Non-Specific subcategory that did not appear in Study 1. The 80% of Non-Specific Strategies that were vague requests for clarification from a partner can be explained as artefacts of the oral pair task in Study 2; the remainder were non-specific inferences. (For space reasons, this finding was excluded from the thesis.) This difference between the data yielded in the two studies is a very minor one and it does not affect the overall scheme.

As far as data representation was concerned, Sensory Strategies proved to be the more problematic of the two categories. While, in Study 1, a reader might provide direct

evidence of rereading part of the text by stating it in writing, no such clarity was possible in Study 2. Here, subjects were talking in pairs and would not announce that they were rereading part of the text. The only evidence that would guarantee that rereading had taken place would be if a subject were tape-recorded reading aloud the same part of the text twice. In reality, of course, there was a combination of reading aloud (recorded) and silent reading (not recorded). This problem was one of recording physical actions in general, not just rereading. One way to resolve it would be to avoid this kind of relatively unguided oral introspection conducted in pairs; in addition, empirical evidence about such actions might be provided through eye movement studies.

### **Research Objective 3**

The third objective was to develop strategy descriptors that are viable across variation in methodology and tasks. With only two exceptions, the descriptor groups of Reflective Strategies appeared in both studies. The first exception was that the 'requesting clarification' group of strategies only occurred in Study 2. The explanation for this is that the pair task allowed subjects to seek assistance from one another but the silent individual reading task in Study 1 did not. Thus, adherence to a naturalistic reading task in Study 2 (i.e., avoiding a pair task) would have avoided this difference between the two studies.

The second exception was qualitatively very different from the first. Reflective Extra-Text-Centred Strategies in Study 1 contained a group called 'expressing personal reactions'; these reactions to the text involved interest, or lack of interest, in the topic, positive or negative evaluation of the text, and agreement with text content. Here, the explanation for diverging findings between the two studies almost certainly pertains to text topic. The text in Study 1 dealt with learner independence, or lack of it, among Hong Kong university students, something to which these learners might well have wanted to express personal reactions. In contrast, the text in Study 2 was on an academic subject (psychology) which the research participants had, as students, only just started studying and with which, prior to the research, they had expressed experiencing considerable difficulty. Their negative experience with this subject and



their sense of being novices in it may have reduced the likelihood that they would feel free enough to 'stand back' from the text and express personal reactions.

Another possibility is that a relatively naturalistic reading task, in this case with reader-generated breakpoints, might well be more likely to stimulate such affective and evaluative reactions than a cloze task where a given number of researcher-generated items have to be considered in a given amount of time. In other words, readers may engage with an 'unmutilated' text (i.e., free of deleted words) whose content resonates with their own life experience differently from how they might engage with a 'mutilated' text whose content, though relevant to their studies, they nevertheless find rather dry, academic and, possibly, even intimidating.

Neither of these exceptions poses a threat to the viability of the strategy descriptor groups developed in the two studies; in fact, as discussed below, this, the lowest of the three levels in the categorization scheme, was designed to be as flexible as possible. These descriptor groups are highly informal, and it would be unrealistic to expect uniformity in strategy descriptors developed from two such diverse tasks. From the reverse perspective, given the history of research into reading strategies, with its proliferation of descriptors, it is noteworthy that reading behaviours reported from two such diverse reading tasks and different texts can be represented by such similar descriptors. (Examples of fine differences in the wording of descriptors across the two studies are given in Chapter 6.)

#### **Research Objective 4**

This objective was the incorporation into the categorization scheme of a mechanism for mapping combinations of related strategies. In order to fulfil this objective, the concept of the 'compound' strategy was conceived. Some of these compounds occurred at category level, that is, they described reports of behaviour that combined mental and physical actions (i.e., Reflective and Sensory Strategies). For example, the strategy 'underlining key points in part of text' combines an observable action (i.e., the physical act of underlining) and a non-observable one (i.e., the mental process of differentiating between key ideas in the text and those deemed less important).

Similar combinations also occurred at subcategory level, in particular the integration of strategies utilizing readers' prior knowledge and text meaning, thus providing evidence of interactive reading (i.e., a combination of Reflective Extra-Text-Centred and Reflective Text Meaning-Centred Strategies). (Compound strategies are discussed further in Section 2.0 below.)

### **Research Objective 5**

The final objective was to use the categorization scheme to compare findings across Studies 1 and 2; these were presented at the end of the previous chapter. Although the development of the categorization scheme formed the main focus of the current research, the presentation of comparative findings using the same scheme represents new territory in research into reading strategies. Among the findings that are worth noting, there seemed to be a clear effect for task, with the free-reading task precipitating reflection on meaning at paragraph level and the researcher-generated cloze task precipitating relatively local reading. There were also fewer reports of use of prior knowledge in the cloze task; however, it is impossible to be sure if this was a task effect or a text effect. Any evidence of a task effect threatens the validity of the findings of studies that employ researcher-generated breakpoints where those findings purport to represent normal reading. A glance at the 'task type' column in Table 3.5 in Chapter 3 reveals that this might apply to virtually every study cited.

Using a small database, figures for selected groups of readers suggest that one group used similarly local strategies with the texts in both studies; in other words, there appeared to be no task effect for this group. Individual variation showing readers who process text flexibly, that is, including a broad spectrum of the text, and other readers who process it rigidly at the local level is echoed elsewhere in the literature (e.g., Sarig 1987a and 1987b, Devine 1988).



### 1.3 Limitations

This part of the review consolidates discussion of the limitations of the current research, some of which have already been alluded to. It begins by considering limitations that were potentially avoidable, including an assessment of the research design and, especially with respect to Study 2, the task design. It ends with some limitations that are inherent in this kind of research.

#### The original research design

As originally conceived in Chapter 4, the project was on far too large a scale for a thesis. The main criterion for deciding which of the several original research tasks to retain and which to jettison was the task's potential to address my main research interest, the study of mental processes in reading and the categorization of strategies. I was also interested in looking at the possibility of using the written mode to collect mentalistic data. Nonetheless, such *post hoc* manipulation of the original research design, which took place after data collection, is hardly an ideal way to conduct a research project. It resulted in a modified design that is far from optimal.

Moreover, despite the loss of several data sets following data collection, I was further compelled to jettison parts of the analysis during the write-up. For instance, for the purpose of obtaining an overview of the research context through the analysis of quantitative data, I developed a questionnaire that was pilot-tested and then administered to 298 first-year university students of the social sciences in Hong Kong. As a research instrument, however, the questionnaire, though useful for yielding quantitative data from large groups, is not ideal for gathering mentalistic data.

Also excluded from this thesis was a task that sought the views of the 38 BA TESL subjects on various reading tasks commonly used in textbooks, classrooms and exams. Apart from the space constraints already noted, this task was not sufficiently focussed on the main area of study, that of the development of a categorization scheme of L2 reading strategies.

Furthermore, retrospective semi-structured interviews were conducted on twelve of these subjects, but the data yielded were dropped from the project for two main reasons. First, these interviews did not adhere to Research Objective 1, which called on the research to “target the readers’ mental processes [while] in short-term memory” (though, as noted above, this objective was somewhat misguided). In addition and crucially, they were poorly conducted. Since the loss of interview data was, in my opinion, the most regrettable among the losses outlined above from among the jettisoned tasks, I have selected the interviews for further discussion below.

Finally, Study 1 was progressively reduced in size and scope. First, I reduced the number of breakpoints in the main section of the task from ten, as originally planned and as conducted by Davies (personal communication 1994), to four. This is the principal reason why Study 1 yielded fewer data than Study 2, so impeding comparison of the findings across the studies. Without the oral task in Study 2, sufficient time could have been allocated in the two-hour data-gathering sessions for the subjects to write about considerably more than four breakpoints. This would have offered the potential to strengthen the Study 1 findings, particularly with respect to individual variation. Second, I dropped the data from two of the three sections of the task; these data were drawn from pre-reading and post-reading activities. Third, towards the end of the write-up, I dropped one of the two analyses of the main section of the task; this was the analysis focusing on ‘breakpoints’.

### **The flawed nature of the interviews**

There were two main problems with the interviews I conducted in the study. First, due to scheduling difficulties, they were sometimes conducted within hours of the main data-gathering sessions whilst, on other occasions, they took place several weeks later. Second, the interviews were structured too narrowly on my immediate impressions derived from reviews of individual subjects’ data, an approach that would have been more suitable in a case study. However, in order effectively to structure the interviews for the purposes of the current study, findings from the entire data set should have been used in combination with those of individual readers. Needless to say, planning retrospective interviews in this way would have necessitated that the



data from Studies 1 and 2 had already been analyzed, a process that, in reality, stretched over a very long period of time following data collection. This would have enabled me to structure, for all interviewees, some key questions based on the main findings and, from the data set, to pinpoint questions that needed to be clarified.

Clearly, when a piece of research is being conducted for the first time, there are logistical problems in attaining this level of planning. Yet, on the one hand, there is a limit to how long or varied the lag time between original and retrospective data collection can justifiably be and, on the other hand, data analysis takes time, especially in exploratory and descriptive studies being conducted by one researcher.

Nonetheless, in the current study, there were two potential sources of information that could have informed the process of planning interviews. One source would have been to garner evidence from similar studies in the literature. Though there is little similar research, some studies might have helped inform this process (e.g., Sarig 1987a and 1987b, Hedge 1991, Block 1992, Li and Munby 1996). The other source, the preferred one, would have been gained from a far more rigorous analysis of the pilot study data than was actually conducted. With hindsight, it is clear that, during pilot-testing, there was an imbalance of attention in favour of the technical details of task design in both Studies 1 and 2 at the expense of data analysis and the development of strategies. A specific outcome of this was that, while tentative strategy categories were developed, insufficient work was done on the strategy descriptors at that stage. As a consequence, the problem of partner-generated strategies in the pair task in Study 2 was not realized when it should have been, that is, at the pilot stage.

### **The final research design**

Once the scale of the research had been reduced to a manageable size, its design was based on two parallel studies, as shown in Table 7.1. (This table is identical to Table 4.2 in Chapter 4.) It shows that three task features (text genre, data type and introspection type) were held constant and three (introspection mode, task mode and breakpoint type) were varied.

**Table 7.1: Features of research design**

Features	Study 1	Study 2
Text genre	Short academic article	Short academic article
Data type	Mentalistic	Mentalistic
Introspection type	Concurrent with reading	Concurrent with reading
Introspection mode	Written	Spoken
Task mode	Individual, silent	Pair, oral
Breakpoint type	Reader-generated	Researcher-generated

Although the design was successful in the limited sense that it was instrumental to the development of strategy categories, subcategories and descriptors across two studies, as has been reiterated, this was done at the expense of the findings. The problem with the design is that it has an uncomfortable mixture of experimental and non-experimental features. As I will now show, it would probably have been more successful if it had been either more or less experimental.

While genre was held constant, different texts on different topics were used in the two studies. This resulted, for instance, in the finding, presented at the end of Chapter 6, that extra-textual processing varied across the two studies. However, in order to ascertain if it was text content that caused this kind of variation, a more experimental design would be needed. Two texts, each with a different task, could still be used, but, departing from the current design, half the readers would do one task with one text while the other half would do the same task with the second text, and vice versa.

Introspection mode was varied, but the findings from the two studies cannot be compared and contrasted for a mode effect because there was variation in task across the two studies. If the objective of using both written and spoken introspection is to investigate whether there are any qualitative differences in the data recorded using these two modes, then all other design features would need to be held constant.

On the other hand, the quasi-experimental nature of the design hampered the more heuristic/inductive side of the study. For instance, holding introspection type constant precluded the potential, already mentioned, of collecting concurrent data yielded by a task and following this up with retrospective semi-structured interviews. As already noted, this kind of design would have the potential to pursue unclear aspects of the data, add concurrent validity to it, and enable one to obtain a deeper view of variation



among individual reader's behaviour. For instance, in the current research, missing information at descriptor level, such as part of text being utilized, could be pursued; in Study 2, questions about Reflective Non-Specific Strategies might be asked.

In addition to design problems affecting the entire study, there were some that were specific to Study 2. An argument was made in Chapter 3 for a task that would foster normal processing in reading; however, although this was heeded in Study 1, as is now explained, no such attention was given in Study 2.

### **The flawed nature of Study 2**

Texts that have been subjected to the cloze procedure have often been referred to as 'mutilated'; a way of expressing this in a manner relevant to the current study is to refer to a text with 'researcher-generated' breakpoints. However it is expressed, it is clear that doing a cloze task cannot represent 'normal' reading. Further, though subjects, both during pilot-testing and data-gathering for Study 2, spontaneously reported enjoying working in pairs on the cloze text, this positive affect can do nothing to promote normal reading. A cloze task done orally in pairs, though, may well be one of a number of ways successfully to train readers in introspection; ironically, it was partly its apparent success during training that was instrumental in making the technique my choice for the Study 2 task.

Although all the studies listed in Table 3.5 in Chapter 3 utilize oral introspection, this was usually done using texts marked at points where readers were expected to introspect. Though less contrived than a cloze task, this technique might still be considered contrived, especially when the marked points occur with such extreme frequency as at the end of every clause in the text (Olshavsky 1976/77) or at the end of every line (Jimenez 1997). There appears to be no study where an individual reader, alone in a room, is tape-recorded introspecting spontaneously. This is something that could fruitfully be considered.

## **The nature of reading research**

To complete this part of the discussion on limitations, some of the difficulties inherent in reading research are now considered. As was stated earlier in this thesis, researching reading behaviour is not the easiest of areas to study. Empirical data can be gathered in artificial contexts with complicated equipment such as the tachistoscope. Alternatively, mentalistic data can be collected, but these merely represent what particular readers said they did with a particular text on a particular occasion. There is no guarantee that anything reported is what actually occurred, though it is unlikely that everything reported failed to occur. In addition, however accurate reports may be in themselves, they are likely to represent a far from complete picture of what actually occurred. Focusing on the current project, there is no guarantee, for instance, that the data were all recorded concurrently with reading (defined here as within ten seconds). Thus, following information-processing theory, some of the data recorded may have been retrieved from long-term memory rather than from short-term memory, and hence they may be less reliable.

In addition, however well designed a study, mentalistic data can often be interpreted in various ways, and different researchers may develop different categories or subcategories from the same data set. Thus, with a principled approach to analysis and the best will in the world, research outcomes such as classification schemes seem likely to be stamped to some extent with individual researcher idiosyncrasy.



## **2.0 FUTURE DIRECTIONS**

Although, as has been shown, there are some clear limitations to the current research, there are also some useful outcomes. Based on the work conducted, these outcomes include recommendations for criteria to be observed in the development of categorization schemes of reading strategies and for a broader view of mentalistic approaches to data collection. There are also recommendations for future research in the area. These issues are now addressed.

### **2.1 Taxonomy for categorizing reading strategies**

The current research suggests that, in the development of a minimally reliable classification of L2 reading strategies, there are certain requirements. These are now discussed under the following rubrics: diversity of data input, structural flexibility and data collection.

#### **Diversity of data input**

There are several possible sources of data input which, when aggregated, have the potential to increase the credibility and stability of a categorization scheme of reading strategies. Though beyond the resources of an individual researcher, some of these sources involve reader and text variables that can be easy to manipulate.

To begin with the reader variables, the more diverse the populations of readers who report their behaviour, the more comprehensive and generalizable the classification can be. The obvious factor to consider is language proficiency level, which is central to Bernhardt's (1991) developmental theory of L2 reading; in addition, there are such factors as age, gender and L1. Moreover, several questions arise from the potential impact of reading purpose on reading behaviour. Does variation in purpose affect strategies and outcomes, and, if so, how? An additional question is that of commonly reported purposes, such as understanding text meaning, remembering text content, language learning as well as combinations of these. In the current classification

scheme, it was found possible to attach purpose, where reported, to physically observable actions (i.e., Sensory Strategies) but not strategies describing mental actions (i.e., Reflective Strategies). It may be that this absence is an artefact of the scheme; for example, a Reflective Text Language-Centred Strategy may be said to comprise a language-centred purpose. Nonetheless, the role of purpose in L2 reading behaviour, both at the levels of whole texts and parts of texts, remains one that might ultimately be included in a classification scheme.

Turning to the text variables, in order to capture as complete a range of reading behaviour as possible, it is essential to obtain data from subjects reading a wide range of genres. In addition, as Kletzien (1991) has shown with L1 readers, text difficulty can affect reading behaviour; thus, reports of L2 subjects reading texts of varying difficulty should be included. Also, since variation in content is likely to precipitate variation in affective response, texts with diverse content should be included.

Diversity of input can also be controlled by variation in task design and research design. As was suggested in discussion of the findings in Chapter 6, task can affect reading behaviour. If the objective of a classification scheme is that it has the potential to describe 'normal' reading behaviour, then, bearing in mind that concurrently tapping uninterrupted reading is, by definition, impossible, it is important to use a variety of tasks that are as naturalistic as possible. In Chapter 3, I raised the notion of a continuum with one end representing naturalness of reading task and the other end representing extreme levels of artificiality. Tasks that could be plotted towards the natural end would allow reader subjects to select the research context and the text, and the researcher would impose a task designed to intrude minimally on the reading process being reported. The more naturalistic of the two current tasks, the one which was described in Study 1 of the present project, concurrently tapped *reader-interrupted* reading. However, as shown in Table 3.5 in Chapter 3, most studies using introspection, including Study 2, seem to use relatively researcher-controlled tasks; these tend to precipitate *researcher-interrupted* reading.

On the other hand, the objective of a classification scheme may be to represent all reported reading behaviour, whether 'normal' or task-induced. It is quite conceivable, for instance, that a researcher might seek task-driven differences in reading behaviour.



In such cases, where a researcher selects the tasks he or she wishes to investigate, it is important that strategies generated by a particular task are marked or coded in some way. For instance, in Study 2 of the current research, strategies were marked in two ways. First, at category level, they were marked as being either *reader-generated* or *partner-generated* (among the latter of which I identified four subcategories). In order to facilitate comparison of strategies reported in Study 2 with those reported in Study 1 (which were all reader-generated), partner-generated strategies were excluded from the findings reported in the previous chapter. Second, at descriptor level, task-induced factors, such as word choices for cloze items, were included in the wording given to the strategies and coding for task sometimes surfaced in the variable slot for part of text utilized; these two possibilities are now illustrated:

- making an inference about meaning in text [clause] that conflicts with own word choice
- reflecting on language [grammatical category] in text [cloze item]

In addition to task design, research design can facilitate diversity of data input. As reported in Chapter 3, information-processing theory has been very influential on the study of reading behaviour, and this may have precipitated the norm in the published studies shown in Table 3.5 whereby report data are collected concurrently with reading. Nevertheless, there remains the possibility of triangulating through temporal variation in reporting. While greater credence may be attributed to concurrent reporting, with a gradual accrual of data, retrospective reporting may serve increasingly to validate it.

As mentioned with respect to task design above, in order to facilitate comparison of findings across individual studies, key aspects of variation in research design could also be recorded. For instance, strategy descriptors could be marked or coded for such factors as retrospective reporting, researcher intervention in the generation of data and mode of elicitation. Such marking or coding might manifest, respectively, as follows:

- making an inference about meaning in part of text [retrospectively]
- making an inference about meaning in part of text [following researcher input]
- making an inference about meaning in part of text [in writing]

## Structural flexibility

Marking strategies for particular research-induced features is one way of introducing flexibility into a categorization scheme, and this can be done at various levels. In my own scheme, the highest level, that of category, reflects the apparently unalterable reality of the divide between observable and unobservable behaviour. As shown in Chapter 2, other researchers have selected different distinctions; for example, Barnett's (1988a) 'text-level' and 'word-level' categories and Block's (1986b) 'general comprehension' and 'local linguistic' categories seek to classify strategies according to how globally or locally readers are processing text.

However, whatever categories are established, it is very easy for researchers in the social sciences to force their data to fit their discrete categories. While categories are necessary to provide a framework for classifying behaviour, there is a danger that researchers, probably unconsciously, may view their own artefacts as sacrosanct. Reading researchers need to be cognizant of the emphasis reading theory puts on the possibility of parallel processing. Allowing for the possibility of more than one process to be occurring simultaneously necessitates the creation of mechanisms for operationalizing categorization schemes in a flexible manner. In my own data, I found evidence of actions that represented *combinations* of my own categories, combinations that were almost certainly highly integrated and might even have occurred simultaneously rather than in linear fashion. Following Research Objective 4, discussed above, such actions were coded as 'compound' strategies.

In my view, such flexibility is required whatever categories are established. For example, in studies where there are 'global' and 'local' categories, when it comes to coding part of the text being reported, it is necessary to entertain the possibility that there are instances where there is a combination of both global and local processing. To my knowledge, though, no studies with such categories integrate these levels of text processing; it seems that, despite what reading theory has suggested about parallel processing since Rumelhart's (1977) interactive model, reading researchers have often shown a preference for excessively discrete coding.



In the current study, degree of global/local reading was recorded at the lower level of strategy descriptor, which included the concept of the 'core' strategy. This is a strategy that describes a particular action but allows for variation in part, or extent, of text being processed. In reality, this led to such an array of data-driven text categories that, in order to present clear and concise findings in Chapter 6, I decided to consolidate them into four broader categories. Where reported, aspect of language being processed was also recorded in core strategies. A further feature of data that might be recorded is meaning, though this is a very elusive area to investigate, especially at the local level, and one that is likely to be even more dependent on text content than is language. Nonetheless, such general semantic fields as positive/negative, right/wrong, and for and against, are among several that appeared in the current research; though not reported in this thesis, they might be interesting to investigate.

### **Data collection**

One of the most intriguing outcomes of the current research is that, as reported in Table 3.5 in Chapter 3, only one published reading study (Li and Munby 1996) has utilized introspection in the written mode. Thus, the use of written introspection in Study 1 represented territory barely explored by L2 reading researchers and yet it seems to have considerable potential, as well as a number of advantages, over oral introspection.

Li and Munby (1996), in common with the current project, utilized a combination of spoken and written introspection. Their written task was journal writing, which seemed to induce data about reading in general rather than reading a particular text, whereas Study 1 of the current research had a task requiring readers to write introspectively about their behaviour following breakpoints experienced while reading a text. The Study 1 task was adapted from that originally devised (but never published) by Davies for both L1 and L2 readers (personal communication 1994). A quick glance at some of Davies' data reveals an array of reading behaviour, including reports of rereading, reading word by word, skipping, understanding individual words in a sentence but failing to understand the sentence as a whole, using prior knowledge

and expressing various evaluative reactions to the text. My own use of Davies' technique in Study 1 yielded similarly rich data, though it was marred by my decision, mentioned above, drastically to limit the number of breakpoints I required my subjects to write about.

Reading is normally a silent activity in which the reader interacts with the text, and it seems intuitively appealing to use a written rather than an oral task to tap such behaviour. Writing about reading, rather than speaking about it, seems to maintain both the individual and the silent nature of the activity under study and to avoid third-party intervention. Researchers gathering oral data tend to interrupt silences and, when they do not, subjects, especially if they are non-native speakers, are likely to feel embarrassed about them, especially if they are being audio- or video-tape-recorded. Nevertheless, as already noted, Scott's (1990) experience of studying lexical inferencing among Brazilians reading in English led him to conclude that "it is easier to speak than to write" about such processes (p. 289). It certainly seems likely that written data will tend to be less spontaneous than oral data, and the fact that they may be more considered could add a burden on the informant, one that might require retrospection and resort to long-term memory.

On the other hand, as observed in Chapter 3, asking subjects introspectively to write reports on their reading behaviour is a relatively easy task to devise and is procedurally straightforward for the researcher. There is no need for recording equipment, which some subjects, at least initially, find intimidating. In addition, for the reader, it is a conceptually easy task and renders autonomy. It also avoids researcher intervention, so common in reading research. It is likely to tap concurrent reading of a particular text, rather than retrospections or comments on reading in general, though, as noted in the previous paragraph, it may include the latter.

Oral introspective data require transcribing, which makes them labour intensive; in contrast, written data are immediately accessible to the researcher. This means their content is potentially available much more quickly for follow-up retrospective interviews than that of oral data. Related, from my own experience in analyzing the data from Studies 1 and 2, I observed that written data are more likely to be concise than transcriptions of oral data, which, depending on task, may contain long stretches



of phatic communication between pairs of subjects or between subject and researcher. Thus, even when transcriptions have finally been completed, the data are less rich and dense and there is considerably more dross to sift through, making them harder to analyze.

In addition to the issues raised in the foregoing taxonomy for categorizing reading strategies, the outcome of the current research suggests some additional avenues for future research. These are now considered.

## **2.2 Future research**

This part of the discussion makes various recommendations for future research in the area of L2 reading in general and L2 reading strategies in particular. It begins by considering specific questions awaiting investigation which arise from the current research; some of these have already been mentioned. It ends with a discussion of issues arising from the literature; particular reference is made to Bernhardt's (1991) theory of L2 reading.

### **Issues arising from the current research**

The following issues about the design and administration of studies derive directly from the current research: the creative application of introspective methodology, the training of subjects, and the pilot-testing and selection of tasks that are as naturalistic as possible and utilized with texts of authentic length. In addition, there is scope for developing the idea of fused or integrated categories. Finally, following the current findings, there is a recommendation to investigate individual variation in reading behaviour.

The first issue is that, since my wish to investigate whether there might be essential qualitative differences in data yielded through written and spoken introspection was not achieved, this methodological question remains for future research; this recommendation is consonant with my call above for more research using the written

mode. Second, training subjects in introspection is also likely to be necessary and, for this purpose, cloze tasks conducted orally in pairs appear to be eminently suitable, though there are obviously other possibilities, including practice with the same task as that to be used subsequently in the research.

The next three recommendations pertain to task. First, it is essential that tasks are pilot-tested both to ensure that they are usable and that the data yielded can readily be analyzed. Second, tasks should be as naturalistic as possible. One of the arguments for using mentalistic rather than empirical data is that, while the latter require artificial conditions or equipment, the former can be collected in relatively naturalistic contexts. Consequently, it is recommended that this attempt at tapping reading processes under conditions that are as normal as possible should in future be complemented by naturalistic tasks. For instance, as most mentalistic studies have used oral introspection in the presence of the researcher and with the use of marked texts, a more naturalistic alternative might be to have subjects reading unmarked texts alone in a room and introspecting while being tape-recorded. Clearly, such a procedure would require training. Although an oral cloze task may reveal information about how readers approach that particular task, its findings cannot be generalized to 'normal' reading. Third, and related to the recommendation that tasks should be naturalistic, I suggest that texts whose lengths are authentic should be selected where possible; at least for advanced L2 readers, the choice of short texts may obscure or simply fail to reveal actual behaviour that typically occurs with longer texts.

Although there is no suggestion that the categories developed in the current project are definitive, as mentioned earlier in this section, it seems to me imperative that, as a field of study, we pursue the notion of combining whatever categories are developed in a scheme. Any serious attempt to describe and categorize the reading process must have the capacity to represent its very complexity. The evidence in the current data for such integrated or fused actions suggested the combined use of prior knowledge and information in the text.

Finally, the current research found considerable variation in the reading strategies reported across subjects. Individual variation presents researchers with an apparently endless array of possibilities. To what extent is L2 reader variation specific to



particular populations, L1 groups, educational contexts and so on? To what extent do particular patterns of variation aid comprehension and/or affect examination results? To what extent does flexibility in reading behaviour in the face of variation in text difficulty, genre, reading purpose, task, and so on, aid comprehension? To what extent does L2 proficiency level affect readers generally? To what extent does it affect individual readers? To what extent does the teaching of reading affect outcomes in individual reading behaviour? To what extent does the explicit teaching of reading strategies in general, or that of particular reading strategies, favourably affect reading outcomes?

As reported elsewhere, using her own classification scheme, Sarig (1987b) found considerable evidence of variation among her readers. Hedge (personal communication 1999), using 'modes' of reading, also found variation among her readers. Hedge's modes, shown in full in Table 5.8 in Chapter 5, seem to point to a possible way forward in the study of individual variation as they provide a framework for recording sequences in reading behaviour (e.g., the modes 'top-down, deferred interactive' and 'bottom-up, recursive, deferred interactive'). It seems to me that the concept of modes, such as these, would be well worth pursuing in future research focusing on individual variation as well as sequences of behaviour in reading.

Apart from recommendations for future research arising from the current project, there are others that emerge from the field generally and specifically from Bernhardt's (1991) developmental theory of L2 reading.

### **General issues**

It is beginning to look as though the revival in L2 reading research in the 1980s and early 1990s, which followed a similar revival in L1 reading in the previous decade, is losing momentum. Among the possible reasons for this may be the perception that research using introspective techniques is complex, labour intensive and therefore expensive, involving, as it does, the extensive pilot-testing of tasks and the training of subjects in introspection before data collection and analysis can take place. Moreover, given the 'black box' nature of reading processes, no single study is likely to have the

capacity to do more than scratch the surface of our knowledge of L2 reading processes. Findings from even the best individual studies can be no more than mere pieces of a vast jigsaw that may never be entirely completed.

Yet, there are clear rewards for researchers who are fascinated by the somewhat elusive quality of introspective data and the consequent challenges involved in their collection and interpretation. Moreover, as discussed below, many questions remain to be investigated in reading, an area that indisputably affects increasing numbers of non-native speakers studying in English at tertiary level.

The issue of quality of investigation and reporting is important in any research, but, in order to ensure it in the area of L2 reading strategies, the researcher needs to be open to input from perhaps an unusually large number of sources. For the conceptual stage of studies and for the interpretation of results, relevant disciplines include L1 as well as L2 reading; for the development of tasks and the description of reported behaviour, they include cognitive psychology and applied linguistics. In addition, the researcher using mentalistic data needs to be aware of whether findings generated by other methodologies, such as those used in eye movement studies, complement or conflict with those generated by introspection.

Hitherto, much L2 reading research, notably that on strategies, has been conducted with little or no reference to theory or, in some cases, to evidence from other studies. In future studies, therefore, an effort needs to be made to change the L2 reading research culture in such a way that reference to current theory and other research into reading processes is routine. However, over-adherence to a theoretical position (as with that in the present study, described above, relating to information-processing theory and the collection of concurrent introspection) can stultify creativity in research design. Nevertheless, much work remains to be done in investigating current theoretical positions, not least that of the only theory of L2 reading yet posited, that of Bernhardt (1991).

Though nearly a decade old, Bernhardt's developmental theory, which was described in Chapter 1, remains to be superseded in theoretical terms and to be verified or challenged through a concerted research effort involving different L2 populations.



Her theory begs a number of questions; for instance, she posits that ‘error rates’ in word recognition decline steeply as language proficiency advances. However, in the part of Study 1 focusing on breakpoints (deleted from this thesis for space reasons), unknown words were reported by the relatively advanced Hong Kong subjects to trigger considerably more breakpoints than any other single factor. Thus, along with L2 proficiency level, researching such variables as text difficulty, genre and reading purpose would help to test Bernhardt’s prediction about word recognition and L2 proficiency level.

In a similar vein, Bernhardt’s prediction that ‘error rates’ in syntactic feature recognition reach a peak at higher intermediate levels of language proficiency needs to be investigated. Contrary to this prediction, there were very few reports in the current research of problems of this nature. It is possible that the methodology used, that of concurrent introspection, tended not to generate such reports and that retrospective interviews might have revealed qualitatively different findings. It would be possible, for instance, to explore the possibility that, when readers report strategies for dealing with ‘vocabulary problems’, they may sometimes actually mean complex syntax. Anecdotally, I have been told by a Hong Kong teacher that Hong Kong students, participating as research subjects, are likely to be more willing to acknowledge deficiency in vocabulary than in grammar. Again, future research, using variation in methodology and context, could investigate this question.

Another area for future investigation is the role of readers’ prior knowledge in reading. Use of content knowledge is predicted in Bernhardt’s theory to decrease slightly as language proficiency increases. Nevertheless, using the classification scheme in the current research, the Hong Kong readers reported quantitative and qualitative variation in extra-textual processing across two texts of the same genre; moreover, there were substantial quantitative differences in top-down processing across individuals. The current project was not designed specifically to investigate the role of prior knowledge and these findings may be the result of factors such as introspection mode and task mode. Intuitively, it seems appealing to speculate that readers’ use of content knowledge might vary less with L2 proficiency level than with the level of their existing knowledge. As reported in Chapter 1, use of prior

knowledge has been contentious in the study of reading since its revival began about thirty years ago; yet it might still serve as a useful variable to investigate.

Factors such as interest, or lack of it, in text content may also affect levels of top-down processing. Indeed, affective factors in reading, cited in Mathewson's (1985) and Ruddell and Spence's (1985) models, both reviewed in Chapter 1, represent additional under-researched variables, ones that might be relatively straightforward to investigate through both concurrent and retrospective introspection.



### **3.0 PEDAGOGIC IMPLICATIONS**

In an educational context, the pedagogic implications of a piece of research are usually drawn from its findings. As the focus of the current research has been on the development of a categorization scheme rather than on findings, this section includes both implications derived from the findings, such as they are, and those derived from a more general view of the project. Issues arising from the findings, presented at the end of Chapter 6, include the overall importance of word decoding notwithstanding a substantial amount of individual variation in reading behaviour and evidence of effects for task and, possibly, text. Wider issues considered include variation in reading behaviour in concert with L2 proficiency level, responding to students' needs, training and awareness-raising, and process-oriented and student-centred instruction.

#### **Implications arising from the findings**

The findings, using the current categorization scheme for reading strategies, suggest that word decoding plays an essential part in reading. This is backed by recent reading theory, which suggests that reading is essentially a bottom-up interactive process. Grabe (1991) cites several L1 studies which have found that some 80% of content words and 40% of function words are directly focused on in reading.

This evident and pervasive need for 'bottom-upping' in reading does not mean that our reaction should be one of denial and of returning to the top-down orthodoxy which demands that we encourage students to guess and predict at the expense of recognizing words. Neither does it suggest that we should encourage students to reach for their dictionaries or seek assistance from friends or teachers every time they come across an unknown word. Nor should we resort to language-based reading classes. As Block (1992) has observed, conventional wisdom is "to predigest printed material for students[,] believing that comprehension depends on understanding of all the language features of the text" (p. 337). Rather, it appears that we should focus the reader on the need to approach unknown words flexibly and adapt strategy use to reading purpose. Where words cannot be decoded quickly, teachers should encourage their students to make judgments as to whether they are essential for their purpose in

reading that particular part of the text. This may require practice accompanied by follow-up pair and/or class discussion. Indeed, Block noted that one of the strengths of her 'proficient L2 readers' was their ability to decide which language problems they could ignore and which they needed to solve.

The research findings also suggest a pervasive quest for meaning across the two studies. This may seem obvious, but there are alternatives to a focus on meaning. These include a focus on the grammatical category of an unknown word (though this may actually be an indirect route to decoding meaning), the memorizing of text content and, as Hedge's (1991) subjects revealed, the use of L2 reading for language acquisition and a combination of meaning and language acquisition. However, the question remains how a focus on meaning can most effectively be achieved; again, one way might be through awareness-raising, an issue I return to below.

Evidence from Study 1 (with its free-reading task, the more credible of the two studies when it comes to 'normal' reading) suggests that many readers varied the depth of attention they gave to various parts of the text. They appeared to do this according to their assessment of how important a particular part of the text was for their particular needs. If detailed comprehension of a particular part of a text is required (as with the paragraph in Lines 12-18 of the Study 1 text, which described the text macrostructure), unknown words in that part of the text are likely to require greater attention than those elsewhere in the text. While this seems obvious and many students may know this, either consciously or unconsciously, it is clear that the reading behaviour of the group that treated the free-reading task (Study 1) in the same manner as a cloze task (Study 2) did not reflect this. As an exercise in awareness-raising, therefore, it may need to be stated explicitly, discussed in class and practice given in varying approaches to text and to particular parts of texts.

Another finding, which was related to the finding about part of text utilized, indicated substantial individual variation in reading strategies reported. Again, it appeared that one group of readers used the same profile of strategies for two very different tasks (free-reading and cloze) whereas two other groups of readers used very different strategy profiles for each of the tasks. This suggests that there was a task effect (and, perhaps, a text effect) for some but not all of the readers.



This evidence of individual variation among a relatively homogeneous group of subjects strongly suggests that syllabi relating to advanced L2 reading classes need to reflect the fact that readers are likely to have various 'reading styles', that is, habitual reading behaviours. (Reading styles were discussed in Chapter 2.) However, such styles may not be stable either within a single text or across such variables as genre, task, purpose and so on. This leads to wider questions, such as whether some reading behaviours are better than others in all circumstances or in certain specific circumstances. It also leads to questions about how the richness in reading behaviour almost certainly represented in the average classroom can be tapped for pedagogic purposes. These and other questions that arise from beyond the current findings are now explored.

### **Implications arising from beyond the findings**

Teachers have to work within the constraints of their students' needs. For example, increasing numbers of L2 university students need to be able to read academic articles, some of them quite lengthy. In such cases, it makes sense for teachers to provide practice with representative samples of those genres. Depending on the range of content areas required by the students in any one class, this may require having the students themselves select texts they wish to read.

On the other hand, if students' needs include the use of specific tasks, perhaps for examination purposes, such as cloze or comprehension questions, then reading strategies for those tasks can be explicitly taught, though success rates appear to vary. Grabe (1991) has observed, for example, that strategy instruction is often unsuccessful and that important predictors of success are duration of training, clarity of training procedures and student responsibility. Strategy training tends not to transfer from one reading task to another and not to persist over time.

Nonetheless, there is evidence that awareness-raising in a more holistic sense may be effective and that enhancing metacognition may be a key factor in improving effectiveness in L2 reading. In many contexts, Hong Kong among them, there is a

pervasive sense that reading is a 'passive' or 'receptive' skill, yet, as indicated above, current theory and data evidence clearly indicate that reading is far from passive. In a study investigating metacognitive awareness in L2 reading, Carrell (1989) concluded from her findings: "Adding instruction in 'awareness' or knowledge about a strategy's evaluation, rationale and utility should greatly increase the positive outcomes of instruction." (p. 129).

Such process-oriented instruction might be done through the use of introspection as a pedagogic tool (e.g., Grabe 1991). During both pilot-testing of the tasks and data-gathering for the current research, the subjects frequently made unsolicited comments about how the tasks requiring introspection had made them more aware of what they were doing during reading. When, during pilot-testing, these tasks were conducted in pairs with time allocated for follow-up discussion, the subjects added that they felt they had had insights into reading from listening to each other talk about it. In addition, the fact that the students, participating as research subjects, manifestly enjoyed exchanging information about reading in this way may well have encouraged many of them to go on to experiment in the way they approached reading and given confidence to those whose strategies were reportedly used by others. Though this evidence is merely anecdotal, it seems to suggest that such methodology may have breathed new life into an activity that previously seemed to these readers to be rather dry. Sadly, this exchange of information about approaches to reading was apparently completely new to these students.

Indeed, some of the tasks considered for research purposes in this thesis might be useful for practice in introspecting about reading. For instance, class practice might be given to students reading and writing about the breakpoints they experience and then, in pairs or groups, comparing their experiences. Out-of-class practice might involve journal-writing, the content of which could also serve as the basis for subsequent class discussion.

Another benefit of such student-centred instruction is that it can increase teachers' awareness of reading processes and so inform future instruction as well as the development or revision of reading syllabi. In contrast, a traditional language-oriented approach effectively denies a clear identity to L2 reading instruction,



debasing it and subsuming it under language instruction. Classroom discussion based on readers' varying experiences in dealing with, or ignoring, language problems might well have the effect of bolstering the confidence of excessively language-centred readers sufficiently for them to try new approaches. At the other end of the spectrum of reading behaviour, excessively top-down readers might be encouraged to engage more with the text (Hedge 1991).

Having touched on how readers at the more extreme margins of reading behaviour might be encouraged through information exchange following classroom activities to experiment with new approaches, or at least with shifts in approach, I believe that tolerance of individual variation should also be observed. What needs to be fostered is not an over-critical attitude to whatever readers are currently doing but an insight into the possible range of reading strategies and circumstances in which particular strategies may be more effective than others. Linked to this is the need to encourage flexibility in approach both among and within texts.

In sum, the teaching of L2 reading should be characterized by responding to students' needs within a particular educational context and, more broadly, by raising their awareness of the availability of a broad range of reading strategies. Using appropriate genres, practice in their use should be provided with a view to their becoming more flexible in their approaches to reading. In this way, reading in English should have the potential to be a skill that facilitates non-native speakers' immediate educational and professional requirements and enables them to read texts that may not be available in their L1. It may also be an enduring source of pleasure.

## 4.0 CHAPTER SUMMARY

This chapter has reviewed the research presented in this thesis. The overall goal of the project was to describe and categorize reading behaviour in a principled manner, that is to say, one that drew on different data sources and maintained a watchful eye on reading theory. The outcome, which in no sense is intended to be definitive, involves a scheme with three interconnected levels: category, subcategory and descriptor.

Instrumental to the evolution of the scheme were introspective data drawn from written and oral introspection data. Though both data sources were flawed in some way, the fact that there were two sources, instead of one, necessitated a degree of structural flexibility, especially at the lower levels of the categorization scheme. Accommodating data from diverse sources into one scheme and yet maintaining their potential for comparison with further data sources requires that the data are described or otherwise marked or coded in such a way that the task which generated them can be traced. While the reality has hitherto been that most researchers use tasks that manipulate the reading process in some way, the more naturalistic the task, the less need there is for such marking or coding of data.

Structural flexibility is needed in a classification scheme of reading strategies, not simply to allow for data generated in various ways but also to reflect or drive reading theory. There is now a considerable body of evidence that success in reading requires fast and accurate word recognition. Moreover, this bottom-up processing interacts with various levels of top-down processing which draw on the reader's prior knowledge both from within and beyond the text as well as on predictions about information in parts of the text yet to be read. Where readers report such interaction in their reading, and there were numerous reports of it in the current study, a categorization scheme should be able to represent it. However, this interactive quality in reading requires that discrete categories, which are necessary building blocks in such a scheme, have the potential to combine. Through the concept of the 'compound' strategy, the current scheme made a modest start with this need to represent integrated or fused actions.



Since development and consolidation of reading theory goes hand in hand with evidence, and the evidence is that reading involves highly complex behaviour involving interacting processes, the manner in which that evidence is described and represented in a categorization scheme requires a fair degree of sophistication. It is to be hoped that, ultimately, a scheme will evolve that has the capacity to account for reading behaviour, not simply from one or even two data sources, as hitherto, but for reading behaviour in general. Although this prospect seems to be an elusive one, there remains plenty of scope for creative work among reading researchers along the way.

APPENDIX 4.1: TRAINING SEQUENCE

Below is the three-stage training sequence administered to the subjects in the pilot studies and Studies 1 and 2. The tasks numbered 1 and 2 are adapted from Hosenfeld et al. (1981). The third task is adapted from Scollon and Wong Scollon (1995:122-123). The cloze items are numbered and the deleted words are shown underlined. The version given to the subjects contained blank lines of equal length for each item and was typed double-spaced.

- 1
- Add the following numbers.
- a) 1 + 7 + 4 + 26 + 17 + 1 + 8 + 5 + 93 = \_\_\_\_\_
- b) 1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 + 10 = \_\_\_\_\_

*Pair work* Did you add the two groups of numbers in the same way? What strategy or strategies did you use?

- 2
- There are several ‘nonsense words’ in the following passage. As you read it, try and do two things:
- a) identify the meaning of each nonsense word;
- b) make notes below on how you guessed the meaning of each word.

There was me, that is, Alex, and my three droogs, that is, Pete, Georgie and Dim, Dim being really dim; and we sat in the Korova Milkbar making up our rassodocks what to do with the evening, a flip dark chill winter... The Korova Milkbar was a milk-plus mesto, and you may have forgotten what these mestos were like, things changing so skorry these days and everybody very quick to forget, newspapers not being read much either.

(Adapted from Burgess, A., 1978, Clockwork Orange, New York: Ballantine)

Word 1

Word 2

Word 3

Word 4

Word 5

Word 6

*Pair work* Compare your strategies for the nonsense words with your partner. Which strategies worked best?

*Class activity* Report your findings to the class.



APPENDIX 4.1 (continued)

- 3

a)Imagine you’re wondering whether to read the passage below, so you give it a quick read. Imagine the blanks are unknown words.  
  
What strategies did you use?  
Do you want to read it more carefully?

c)*Pair work* Now fill in the blanks. As you do so, try to do three things:
  - tell your partner what words you suggest;
  - give reasons for your suggestions
  - react to your partner’s suggestions giving reasons for agreeing, disagreeing or hedging.

What is Culture? Intercultural Communication and Stereotyping

Two men meet on a plane from Tokyo to Hong Kong. Chu Hon-fai is a Hong Kong exporter who is returning from a business trip to Japan. Andrew Richardson is an American buyer on his first business trip to Hong Kong. It is a convenient (1) meeting for them because Mr Chu’s company (2) sells some of the products Mr Richardson has come to Hong Kong to (3) buy. After a bit of conversation, they introduce (4) themselves to each other.

Mr Richardson	By the way, I’m Andrew Richardson. My friends call me Andy. This is my business card.
Mr Chu	I’m David Chu. Pleased to meet you, Mr Richardson. This is my card.
Mr Richardson	No, no. Call me Andy. I think we’ll be doing a lot of business together.
Mr Chu	Yes, I hope so.
Mr Richardson	(reading Mr Chu’s card) “Chu, Hon-fai.” Hon-fai, I’ll give you a call tomorrow as soon as I get settled in my hotel.
Mr Chu	(smiling) Yes, I’ll expect your call.

When these two (5) men separate, they leave each other with very (6) different impressions of the situation. Mr Richardson is very (7) pleased to have made the acquaintance of Mr Chu and (8) feels they have gotten off to a very good start. They have established their (9) relationship on a first-name basis and Mr Chu’s smile seemed to (10) indicate that he will be friendly and easy to do (11) business with. Mr Richardson is (12) particularly pleased that he had (13) treated Mr Chu with respect for his Chinese background by (14) calling him Hon-fai rather than using the (15) western name, David, which seemed to him an (16) unnecessary imposition of western culture.

In contrast, Mr Chu feels quite (17) uncomfortable with Mr Richardson. He feels it will be (18) difficult to work with him, and that Mr Richardson might be rather insensitive to (19) cultural differences. He is particularly (20) bothered that Mr Richardson used his given name, Hon-fai, (21) instead of either David or Mr Chu. It was this (22) embarrassment that caused him to smile.

This short dialogue is, unfortunately, not so unusual in meetings between members of different cultures.

(Taken from Scollon, R. and Wong Scollon, S. 1995. *Intercultural communication*. Oxford. UK: Blackwell.)

APPENDIX 5.1: STUDY 1 TEXT

This text is taken from Farmer (1994). For the research project, it was photocopied from the original source, omitting Appendices 2 and 3; here, it has been retyped to show the line numbers (as far as Line 95, beyond which no references are made in the thesis).

Chapter 2

**The Limits of Learner Independence in Hong Kong**  
*Richard Farmer, English Language Study-Centre, Hong Kong Polytechnic*

INTRODUCTION

Since self-access embodies a learner-centred approach, it seems appropriate to start with some of the initial reactions which learners had to the introduction of learner independence in the programmes of study at the English Language Study-Centre, Hong Kong Polytechnic.

‘I prefer a teachers to teach me in the whole programme because  
he/she can know what I need and the progress in learning.’ 5

‘I prefer the teacher encourage me talking to speak a lot.’

‘I would like to have a teacher with me in the whole lesson but  
not always goes away and let us alone.’

‘Small group learning with needs analysis are useful for studying  
English.’ 10

This paper describes the body of learners attending programmes at the Study-Centre and discusses why they depend to such an extent on teacher direction. It explains how we have attempted to redress this through introducing, very gradually, certain features of learner independence and how far we have had to limit this objective in the light of student response. The extent to which learner independence has been successfully implemented is briefly considered and the paper concludes with an outline of plans for further development. 15

THE LEARNERS

Intake

The Study-Centre offers two categories of programme.

**Referral programmes.** Students requiring supplementary tuition are identified by their regular English-class teachers during the first few weeks of the ‘service’ course and are referred to the Study-Centre in pairs or 20



### APPENDIX 5.1 (continued)

small groups. Referred students constitute roughly the lowest 20% of a class.

Each referral drive is followed by a compulsory twenty-hour programme of instruction. Groups are expected to attend twice weekly. A booking system is in operation in order that they may attend at their own convenience.

**Summer programmes.** These are open to all Polytechnic students who would like to practise their English. Attendance is optional. 30

## A learner-independent approach?

Sheerin (1989:3) states that the primary aim of self-access facilities is 'to enable learning to take place independently of teaching. Students are able to choose and use self-access material on their own ... students are able to direct their own learning'. It needs stressing perhaps that this should be viewed as the ultimate aim of a self-access centre, rather than as an objective which may be immediately realized. However, independent learning has successfully been achieved amongst a number of users in many centres by providing initial learner training programmes (Harding-Esch 1982; Riley and Zoppis 1985; Dickinson 1987; Sheerin 1992). Learners are thereby given assistance with the definition of learning aims, the identification of learning needs and with the selection of relevant materials. Subsequent, occasional, access to a counsellor is sufficient to enable a substantial number of learners to pursue their studies independently. At the same time, Riley (1987:77) stresses that 'the person responsible for setting up a self-access system should have a good, hard look at his prospective clients – their needs, objectives, interests, practical constraints and motivations ...'

The majority of centres referred to above are in Western European countries in which one of the policies in education is to develop autonomy in the learning process. Initial investigations at Hong Kong Polytechnic indicate that this is not the case in Hong Kong, where, certainly at secondary level, and in many cases at tertiary level, tuition is highly structured and where learners are expected to, and themselves expect to, adopt a highly passive role. Indeed, it might be said that formal education in this context teaches the need to be taught: learners are conditioned to believe that in order to learn one must be taught and that the teacher holds a monopoly over the transmission of knowledge.

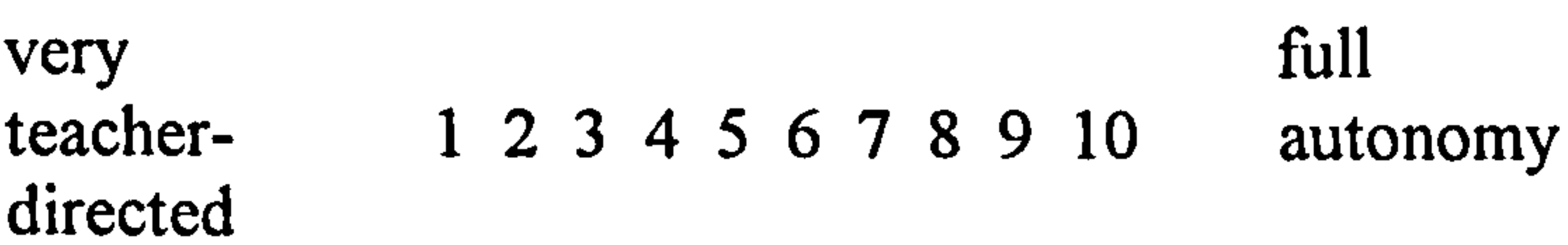
Such a situation results in two major limitations:

1. Learners have no foundation on which to develop autonomy and many have difficulty in even accepting the notion of independent learning. Initially, many students attending the Study-Centre showed open resistance to the idea of assessing their own needs, planning their own programmes and selecting relevant materials. Many were bewildered by the idea of accepting responsibility for their learning.



APPENDIX 5.1 (continued)

	They continue to feel very strongly that the absence of a teacher is a laxity on our part.	65
2.	Students lack confidence in using English. An analysis of the needs of students attending the Study-Centre indicates that the students' greatest perceived need is in the area of spoken English. However, students lack the confidence to participate in speaking activities without a very large measure of help and guidance from teachers; they need considerable direction and encouragement in working through activities and in maintaining momentum.	70
	The expectations of learners in Hong Kong and a lack of self-reliance impose two sets of limitations on implementing an independent approach to language learning in Hong Kong. There are further limitations specific to the Study-Centre:	75
-	The Centre caters to students who are particularly weak, some with a level of English little more than that of a post-beginner. These are learners who have already had little success in learning English and therefore require substantially more guidance and counselling.	80
-	The programmes are compulsory. Initially, a number of students appear little motivated to learn English by any means, let alone independently.	
-	Students have to date been drawn from among the first-year intake to the Polytechnic. They have, therefore, had little opportunity to develop study skills within their own specialisms which might be transferred to a language-learning context.	85
-	The programmes offered by the Centre are necessarily short. Given the large number of students for which the Centre caters, the programmes are restricted to twenty hours in length. Students expect to spend a very substantial amount of that time on language learning activities. If programmes are to have face validity, the amount of time that can be spent on learner training is extremely limited.	90
	The utmost caution is required in introducing self-directed learning to these students. It should not be thrust upon learners who are not ready for it. As enthusiastic advocates we should not 'confuse the idea, or our enthusiasm to introduce it, with the learner's ability or willingness to undertake it' (Dickinson 1987:2). However, this is not to say that we should abandon all attempts to instil and foster independent learning. In these situations the way forward is to introduce into the learning programme elements which train learners towards greater autonomy and aim towards a gradual development to full autonomy (Dickinson 1987). Dickinson represents this gradual move towards greater learner autonomy as a continuum:	95



(after Dickinson 1987)

Different learners have differing starting points; it is therefore important to identify an appropriate starting point and begin at where the learners are, even if this



## APPENDIX 5.1 (continued)

means beginning at almost 'degree zero'. We need similarly to accept that different learners will advance at differing rates and have differing degrees of success. There is therefore a need to recognize and accept the limitations of our learners before embarking on a project aimed at introducing self-directed learning.

The Centre at Hong Kong Polytechnic has, therefore, adopted a 'shallow-end' approach to self-access, one which we hope will satisfy the students' perceived need for teacher contact while introducing students to the rudiments of learner independence and learner responsibility.

## THE STUDY PROGRAMME: SEMESTER 1

### A group orientated approach

In preparing learners for learner autonomy, it is important to take account of cultural variations (Riley 1988). Hong Kong appears to be a group-oriented society; consequently, the programmes are as far as possible group driven and group negotiated. This provides peer support, offers greater opportunities for communication, and thereby helps to increase confidence. Learners feel less exposed than they would if working individually or when called upon to perform in front of a large class. It should be noted that the term self-access has led perhaps, to too much of a focus on the individual to the exclusion of the group. As Dickinson and Carver (1980:3) have pointed out, 'Autonomous learning does not imply learning in isolation, and many aspects of language practice are best done with others'.

### Programme structure

***Session 1: group profile.*** Groups have a short, informal talk with a teacher in the Centre. The purpose is to introduce the Centre, set students at ease, build up a rapport and provide an opportunity for them to voice their own concerns and interests.

***Sessions 2-5: foundation component.*** Groups work through a set menu of activities giving practice in a range of skills. This helps teachers identify more precisely the students' areas of weakness and enables students to make their own, informed choices as to the areas of language and skills which they need to practise later in the programme.

***Session 6: needs assessment.*** Groups have a short discussion with a teacher to identify the areas of the language in which they feel they need practice.

***Sessions 7-20: core component.*** This centres on those areas of difficulty which groups of students tend to have in common and takes account of weaknesses identified by class teachers as well as of students' own interests. Groups choose an area and activity type at the beginning of each session, based on those identified during the needs assessment. In session 15, groups are asked to complete a questionnaire in order to give their evaluation of the study programmes and the learning/teaching approaches.



## APPENDIX 5.1 (continued)

After the programme had been running for some time, forty-seven students (within one department of the Polytechnic) complained to their department that the Study Centre programmes were inadequate. A hastily conducted survey indicated that a significant number of other students were dissatisfied with the programme offered. Their grievances, among others, were that they were given inadequate teacher contact and supervision and that the programmes were unstructured.

In order to avoid a wholesale exodus of students, a much higher degree of guidance was immediately provided, with many groups receiving 100% teacher contact. At the same time, parts of the programme design were revised to provide a tighter structure.

### THE STUDY PROGRAMME: SEMESTER 2

The following components were amended:

***Session 6: needs assessment.*** Groups complete an extensive needs analysis questionnaire (Appendix 1) with guidance from a teacher. Part one of the questionnaire asks groups to assess the importance of broad areas of the language and rate these using a five-point scale; they then select three or four areas which they would like to work on. Part two is more focused: groups are provided with a series of picklists from which they choose specific language structures, functions, writing formats and topic areas which they consider relevant. The group, with the aid of the completed needs questionnaire, and again guided by a teacher, draw up a detailed and often extensive programme of study to include the specific materials they will use (Appendix 2). The reaction of students has been very positive: many have reported that close involvement at the planning stage has led to a greater commitment to the programme and thereby to a higher degree of motivation towards the learning of English.

***Sessions 7-20: core component.*** The programme invariably provides far more material than the group will be able to cover in the time available. The group selects from their programme at the start of each session. Student choice therefore remains an integral part of the study programme.

At the end of each visit, groups are expected to complete a record of work and provide a brief evaluation of the session (Appendix 3).

### EVALUATION

A survey, in the form of a questionnaire, was conducted in April 1993 asking students to evaluate the study programmes, the teaching/learning approaches, and their levels of confidence. The survey was made using a questionnaire in the form of a five-point Likert scale<sup>1</sup>, with the categories strongly agree, agree, neutral, disagree and strongly disagree; for the purposes of this discussion, these have been conflated to three: agree, neutral and disagree. 171 students answered the questionnaire after completing

---

<sup>1</sup> For a discussion of this form of evaluation see Oppenheim 1992:133-42.



APPENDIX 5.1 (continued)

roughly three-quarters of their study programme. The responses relating to learner independence can be seen in Table 1.

Table 1   Learner Independence in the Study-Centre

Item	Agree	Neutral	Disagree
<i>Identification of needs and selection of materials</i>			
The needs questionnaire in session 6 helped me decide which areas of the language I should practise.	62%	32%	6%
The study programme has given me practice in the areas which I consider important.	68%	31%	1%
I enjoy being able to choose which areas of the language to practise.	82%	16%	2%
The foundation programme (sessions 1-5) was a valuable introduction to the Centre.	35%	45%	20%
<i>Independence from teachers</i>			
It is important to have a teacher for the whole session when studying in the Centre.	80%	17%	3%
There is enough guidance from teachers in the Centre.	61%	34%	5%
I like having different teachers in different sessions.	61%	25%	14%
I feel better able to work on my own now since studying in the Centre.	31%	56%	13%

It should be stressed, firstly, that this is not a rigorous piece of research and any conclusions drawn must therefore be tentative. The discussion below suggests possible trends which emerge from an analysis of the data.

*Identification of needs and selection of materials.* The positive response to the needs analysis is encouraging: a large proportion of students (62%) appear not only to have recognized the value in analysing their own needs but also feel that they have succeeded in the exercise. There is also a fairly close match between those who felt that the questionnaire helped them initially to identify suitable areas and those who believed (68%) that the subsequent programme was relevant. Furthermore, few students appear to have reacted negatively.

There appears also to be a positive reaction (82%) to the self-selection of materials which formed an integral part of the ‘Core Component’. In contrast, there was only limited enthusiasm for the ‘Foundation Component’ which offers little choice. There are, of course, many factors which could have influenced this reaction, but it is interesting to speculate whether this resulted from the lack of choice.



## APPENDIX 5.1 (continued)

*Independence from teachers.* Predictably, the majority of students (80%) believe that 100% teacher-contact is important but a large proportion (61%) feel that the amount

of time spent with a teacher in the Centre (roughly 80% of a session) is acceptable and certainly few disagreed. A surprising number (61%) enjoyed the changeover of teachers from session to session. Students are therefore not dependent on a single teacher which suggests at least a small step in the direction of a more independent approach to learning. Overall, 31% of respondents felt that they had developed some independence, a figure which, though small, is encouraging.

## CONCLUSION

It seems that the format of the study programmes offered to students does provide some, albeit very limited, opportunity to develop the basic elements of independence: students are given responsibility for determining their own needs, drawing up their own programmes and selecting relevant materials. Students have reacted positively to this shift in control. However, groups remain heavily dependent on teacher direction while working through activities. Although there is no reason why a self-access approach should preclude some teacher direction, it is desirable to limit intervention as far as possible without alienating the learner. Now that we have a programme structure offering a modicum of independence that appears acceptable to students, we should try (cautiously) to extend that autonomy. There are a number of possible developments which might address this issue.

1. A learner training package which may be used to supplement the 'Foundation' component is currently being piloted.
2. The programmes based on the present needs analysis give little overall direction and lack cohesion. As one student pointed out, 'It is not a systematic learning programme.' Students need training in the formulation of a set of objectives, perhaps embodied in a form of learner contract. They also need training in self-assessment.
3. Through further analysis of student evaluation, it may be possible to determine whether some categories of student need more direction than others. Different levels of direction could then be provided for the different categories.
4. Perhaps the Study-Centre should be regarded as a half-way house between a classroom-based approach and a more orthodox self-access mode. We are now planning to establish a 'self-study' area for use by students who have completed a regular study programme at the Centre. Students would have access to a full-time counsellor for guidance in the analysis of needs and the design of a further programme, but thereafter would be expected to work through activities with greater independence. The regular programme would thereby act as a vehicle leading to fuller autonomy and subsequent programmes would provide a measure of sustainability which at present is lacking.

Finally, we should be realistic as to how much independence our students can cope with initially – where on the continuum they should start – and how much autonomy we can expect of them within a limited programme of study. We need to



## APPENDIX 5.1 (continued)

recognize and respect the expectations of learners in Hong Kong and the needs and expectations specific to the target learners of particular institutions. We ignore these expectations at our peril.

## REFERENCES

- Blue, G.M. 1988. 'Self-assessment: the limits of learner independence', in A. Brookes and P. Grundy (eds.) *Individualization and Autonomy in Language Learning*. ELT Documents 131: The British Council.
- Dickinson, L. 1987. *Self-instruction in Language Learning*. Cambridge: Cambridge University Press.
- Dickinson, L. and Carver, D. 1980. 'Learning how to learn: steps towards self-direction in foreign language learning in schools.' *ELT Journal* 35(1):1-7.
- Harding-Esch, E. 1982. 'The Open Access Sound and Video Library of the University of Cambridge: progress report and development.' *System* 10(1):13-28.
- Oppenheim, A.N. 1992. *Questionnaire Design, Interviewing and Attitude Measurement*. Pinter Publications.
- Riley, P. 1987. 'From self-access to self-direction', in J.A. Coleman and R. Towell (eds.) *The Advanced Language Learner*. London: CILT.
- . 1988. 'The ethnography of autonomy', in A. Brookes and P. Grundy (eds.) *Individualization and Autonomy in Language Learning*. ELT Documents 131: The British Council.
- Riley, P. and Zoppis, C. 1985. 'The Sound and Video Library', in P. Riley (ed.) *Discourse and Learning*. Harlow, Essex: Longman.
- Sheerin, S. 1989. 'Self-access'. *Language Teaching* 24(3):143-57

APPENDIX 5.1 (continued)

APPENDIX 1

Your Learning Needs

Group: Course:

The purpose of this questionnaire is to help you decide which language areas you would like to practise.

Part A\*

1. For each area listed, ask yourself how necessary or important it is for you and circle one of the numbers (1=least important); 5= most important).

<i>Language Area</i>		<i>How necessary is it?</i>				
		<i>Least</i>			<i>Most</i>	
<b>1</b>	<b>Listening</b>					
1.1	Conversational (social English)	1	2	3	4	5
1.2	Academic (e.g. lectures, taking notes)	1	2	3	4	5
1.3	Listening for pleasure (e.g. radio, film)	1	2	3	4	5
<b>2</b>	<b>Speaking</b>					
2.1	Conversational (social English)	1	2	3	4	5
2.2	Academic (e.g. oral presentations)	1	2	3	4	5
<b>3</b>	<b>Reading</b>					
3.1	Increasing speed	1	2	3	4	5
3.2	Improving comprehension	1	2	3	4	5
3.3	Note-taking	1	2	3	4	5
3.4	Related to specialism	1	2	3	4	5
3.5	Reading for pleasure	1	2	3	4	5
<b>4</b>	<b>Writing</b>					
4.1	Personal (e.g. letters to friends)	1	2	3	4	5
4.2	Formal (e.g. reports, business letters)	1	2	3	4	5
4.3	Organisation of ideas (e.g. linkwords, punctuation)	1	2	3	4	5
<b>5</b>	<b>Vocabulary</b>	1	2	3	4	5
<b>6</b>	<b>Pronunciation</b>	1	2	3	4	5
6.1	Individual sounds	1	2	3	4	5
6.2	Stress and intonation	1	2	3	4	5
<b>7</b>	<b>Grammar</b>	1	2	3	4	5

\* Part A is adapted from G.M. Blue (1988)



## APPENDIX 5.1 (continued)

- 2 Look again at the questionnaire and list the three or four activities you would most like to practise.**

1. \_\_\_\_\_
2. \_\_\_\_\_
3. \_\_\_\_\_
4. \_\_\_\_\_

## Part B

**Refer to the lists which the teacher will give you.**

1. Are there any grammatical areas you would like to practise?

---

---

---

---

2. Are there any functions words that you would like to practice?

---

---

---

---

---

3. Are there any topic areas which interest you?

---

---

---

---

### Which would you prefer:

- a. *some* of the materials to be related to your own degree subject;
- or
- b. *all* materials related general topics (as in the foundation programme)?

\_\_\_\_\_

4. Are there any particular types of writing you would like to practise?

---

---

---

---

5. Any other requests?

---

---

---

---

**APPENDIX 5.1 (continued)**

**PICKLISTS FOR NEEDS ANALYSIS**

**1. Grammatical areas**

**Tense**

past simple, past continuous  
present simple, present continuous  
present perfect simple  
(with for and since)  
future  
passives  
reported speech

**Miscellaneous**

articles a, an and the  
relative clauses  
comparatives and superlatives (for example:  
taller, tallest)  
conditionals (If clauses)  
conjunctions (for example: and, but, so ...)  
countable/uncountable nouns, how much/  
how many  
gerunds (nouns with ing)  
infinitives (to + verb)  
phrasal verbs  
prepositions  
punctuation  
questions – direct/indirect questions  
‘wh...’ questions  
yes and no questions  
word order

**2. Functions**

**Information, opinion and ideas**

agreeing and disagreeing  
giving reasons  
comparing  
complaining  
describing experiences  
describing objects  
describing people  
describing places  
describing processes  
instructions/directions  
likes and dislikes/preferences  
obligation (for example: have to  
must, should, ought to)  
opinions  
possibility, probability (for example:  
could, may, might, will)

**Social Conventions**

(for example: apologizing, forgiving,  
complimenting, congratulating, introduction  
and greetings, saying goodbye, thanking)

**Telling/Asking To Do Things**

invitations  
offers  
orders, commands  
permission  
persuading  
requests  
suggestions and advice



## **APPENDIX 5.1 (continued)**

### **3. Topics**

advertising	holidays
animals	housing/accommodation
art	marriage
biography	media
business	war
clothing	money
law and crime	politics and government
disaster	religion
customs/festivals	science
education	shopping
employment	social issues (for example: poverty, old age)
entertainment	sports
environment	supernatural and superstitions
family and friends	technology
food and drink	transport
future	travel/different cultures
medicine/health/illness	weather/climate
history	
hobbies	

### **4. Writing**

curriculum vitae	memos
essays	notices
form-filling	speeches
instructions	stories
letters (informal)	summaries
letters (business)	

APPENDIX 5.2: STUDY 1 TASK

The original task, adapted from Davies (personal communication 1994), comprised the three stages given here: ‘Read 1 Skim read’, ‘Read 2 Study read’ and ‘After reading’. The subjects did all three stages; however, due to space constraints, the process of data analysis and the findings of ‘Read 2 Study read’ only are presented in the thesis.

Reflective Reading – Evaluation of a text

Student's name: \_\_\_\_\_

In this session, you will (1) *skim read* and (2) *study read* an academic article. Please give your frank personal answers to the questions; negative and neutral comments are just as welcome as positive ones.

Read 1: Skim read

1      What are your initial reactions to the text? Please write anything that comes into your mind.

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

2      How do you rate the ‘readability’ of the text? Tick (✓) one of the following. Please give reasons below for your rating.

- \_\_\_ very difficult
- \_\_\_ difficult
- \_\_\_ reasonably readable
- \_\_\_ very readable
- \_\_\_ very easy

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

3      Do you expect to enjoy a more careful read of the text? Please tick (✓) one of the following and then explain your prediction.

\_\_\_\_\_ Yes    \_\_\_\_\_ Not sure    \_\_\_\_\_ No

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_



**APPENDIX 5.2 (continued)**

**Read 2: Study read**

*Breakpoints* are places where you stop reading for any reason (e.g., you don't understand something, you want to think or even rest). Readers use various strategies at breakpoints (e.g., rereading, reflecting, flagging a word or idea, skipping).

**Your task**

*As you read the text more carefully, try and do three things:*

- *note below the line number of each breakpoint (up to 4);*
- *write what you are thinking at each breakpoint;*
- *specify, if possible, what strategy or strategies you use following each breakpoint.*

Breakpoint 1 (line \_\_\_\_\_):

---

---

---

Breakpoint 2 (line \_\_\_\_\_):

---

---

---

Breakpoint 3 (line \_\_\_\_\_):

---

---

---

Breakpoint 4 (line \_\_\_\_\_):

---

---

---

**APPENDIX 5.2 (continued)**

**After reading**

Please consider your reactions so far to (a) the reading tasks you have just done and (b) the text. Refer to your earlier answers if necessary.

- 1

What are your reactions to the *skim-read* tasks?
- 2

What are your reactions to the *study-read* tasks?
- 3

Has your reaction to the *text* changed from your initial reaction? Tick (✓) yes or no. If yes, please briefly explain why.

Yes  No
- 4

Do you have any other comments to make?

(Task adapted from an idea by Davies, University of Bristol, UK, 1994)

(Text used with this worksheet: Farmer, R. 1994. The limits of learner independence in Hong Kong. In Gardner, D. and Miller, L. (eds.), *Directions in Self-Access Language Learning*, pp. 13-27. Hong Kong University Press: Hong Kong.)



## APPENDIX 6.1: STUDY 2 TEXT AND TASK

The text is adapted from Atkinson et al.(1991:3-4). Here, the cloze items are numbered with the original words underlined. The version given to the subjects contained blank lines of equal length for each item and was typed double-spaced.

### Reflective Reading – Introspective cloze task

Student's name: \_\_\_\_\_

The passage below is taken from the beginning of Atkinson et al.'s book entitled *Introduction to Psychology*. The book is used in first-year psychology modules at City University.

*Work in pairs: Your task is to fill in each blank with a word. If you can't decide between two or more words, write down all your suggestions (e.g., way/method); leave blanks unfilled if you're not sure what to suggest. You can deal with the blanks in any sequence you like.*

*As you discuss the blanks, try and do three things:*

- *talk to your partner about every idea you have;*
- *give reasons for your ideas;*
- *react to your partner's suggestions, giving reasons for agreeing, disagreeing or hedging.*

### Chapter 1: Nature of Psychology

No one today can afford *not* to know psychology; it touches virtually every aspect of your life. For example: How does the way your parents raised you affect the way you raise your own children? What is the best way to break a drug dependency? Can a man care for an infant as ably as a (1) woman? Can you recall childhood experiences in more detail under hypnosis? How should instruments in a nuclear power plant be (2) designed to minimize human error? What effect does prolonged stress have on your immune system? How effective is psychotherapy in (3) treating depression? Can learning be improved by the use of drugs that facilitate neural transmission? (4) Psychologists are working on these and many other questions.

Psychology also affects our life through its influence on laws and public policy. Psychological theories and research have influenced laws concerning discrimination, capital punishment, pornography, sexual behavior, and the (5) conditions under which individuals may not be responsible for their actions. For example, (6) laws pertaining to sexual deviancy have changed markedly in the past 40 years as research has (7) shown that many sexual acts previously classed as perversions are "normal" in the sense that most people engage in them. Consider also the effect of television violence on children. Only since psychological studies provided (8) evidence of the harmful effects of such programs has it been possible to modify television programming policies. Programs designed for children now contain (9) less violence, and an effort is made to restrict particularly brutal television to late-evening viewing hours.

## APPENDIX 6.1 (continued)

Because psychology (10) affects so many aspects of our life, it is essential that even those who do not intend to specialize in the field know something about its basic facts and research methods. An (11) introductory course in psychology should give you a better understanding of why people think and act as they do, and provide insights into your own (12) attitudes and reactions. It should (13) also help you evaluate the many claims made in the name of psychology. Everyone has seen newspaper (14) headlines like these:

- New drug discovered to (15) improve memory
- Anxiety controlled by self-regulation of brain waves
- Proof of mental telepathy found
- Hypnosis effective in the control of pain
- Emotional stability closely (16) related to family size
- Homosexuality linked to parental attitudes
- Transcendental meditation facilitates (17) problem solving
- Multiple personality (18) linked to childhood abuse

You can judge the validity of such (19) claims in part by knowing what psychological facts have been firmly established; you can then assess whether the new claim is compatible with these established facts. You can also judge the (20) validity in part by knowing the kind of evidence necessary to give credence to a new "discovery", so that you can determine whether the arguments in support of the new claim meet the usual standards of evidence. This book reviews the current state of knowledge in psychology – that is, it tries to present the (21) most important, established facts in the field. It also examines the nature of research – that is, how a psychologist designs a research program that is capable of providing strong evidence for or (22) against a hypothesis.

In this chapter, we first consider the kinds of *topics* that are studied in psychology. Next we (23) discuss the *perspectives* that psychologists adopt in investigating these topics. Then we describe the (24) research methods used in psychological investigations, and after that discussion, we turn to the various specializations or (25) fields of psychology. Finally, we briefly overview the contents of the rest of the book.

(Taken from Atkinson et al. 1991, *Introduction to psychology* (11th ed.). Harcourt Brace.)



## REFERENCES

- Abrams, S. and Zuber, B. 1972. Some temporal characteristics of information processing during reading. *Reading Research Quarterly* 8:40-51.
- Afflerbach, P. 1990. The influence of prior knowledge on expert readers' main idea construction strategies. *Reading Research Quarterly* 25:1:31-46.
- Alderson, J.C. 1978. Cloze procedure as applied to reading. In Buros, O. (ed.), *The eighth mental measurements yearbook*, vol. 2, pp. 1171-1174. Highbury Park, NJ: Gryphon.
- Alderson, J.C. 1980. Native and nonnative speaker performance on cloze tests. *Language Learning* 30:1:59-76.
- Allington, R. and Strange, M. 1977. Effects of grapheme substitutions in connected text upon reading behaviours. *Visible Language* 11:285-297.
- Anderson, J. 1983. *The architecture of cognition*. Cambridge, MA: Harvard University Press.
- Armbruster, B., Echols, C. and Brown, A. 1982. The role of metacognition in reading to learn: A developmental perspective. *Volta Review* 84:5:45-56.
- Atkinson, R., Atkinson, R., Smith, E. and Bem, D. 1993. *Introduction to psychology*, 11<sup>th</sup> ed. Orlando, FL: Harcourt Brace.
- Bachman, L. 1982. The trait structure of cloze test scores. *TESOL Quarterly* 16:1:61-70.
- Bachman, L. 1985. Performance of cloze tests with fixed-ratio and rational deletions. *TESOL Quarterly* 19:3:535-556.
- Baddeley, A. 1990. *Human memory: Theory and practice*. Hove: Lawrence Erlbaum.
- Bailey, K. and Ochsner, R. 1981. A methodological review of the diary studies: Windmill tilting or social science? In Bailey, K., Long, M. and Peck, S. (eds.), *Second Language Acquisition Studies*, pp. 188-198. Rowley, MA: Newbury House.
- Bakan, D. 1954. A reconsideration of the problem of introspection. *Psychological Bulletin* 51:2:105-118.
- Baker, L. and Brown, A. 1984. Cognitive monitoring in reading. In Flood, J. (ed.), *Understanding reading comprehension: cognition, language and the structure of prose*, pp. 21-44. Newark, DE: International Reading Association.

- Barnard, R., Brookes, H., Pozzo, G. and Rowe, T. 1980. Reading strategies. In Cortese, G. (ed.), *Reading in a foreign language*, pp. 400-415. Milan: Franco Angeli.
- Barnett, M. 1988a. Reading through context: How real and perceived strategy affects second language comprehension. *Modern Language Journal* 72:2:150-162.
- Barnett, M. 1988b. Teaching reading strategies: How methodology affects language course articulation. *Foreign Language Annals*, 21:2:109-119.
- Barnett, M. 1989. *More than meets the eye: Foreign language reading, theory and practice*. Englewood Cliffs, NJ: Prentice Hall.
- Bartlett, F. 1932. *Remembering*. Cambridge: CUP.
- Bensoussan, M. 1990. Redundancy and the cohesion cloze. *Journal of Research in Reading* 13:1:18-37.
- Berent, I. and Perfetti, C. 1995. A rose is a REEZ: The two cycles model of phonology assembly in reading English. *Psychological Review* 102:146-184.
- Bernhardt, E. 1983. Three approaches to reading comprehension in intermediate German. *Modern Language Journal* 67:111-115.
- Bernhardt, E. 1991. *Reading development in a second language*. Norwood, NJ: Ablex.
- Bernstein, S. and Carr, T. 1996. Dual-route theories of pronouncing printed words: What can be learned from concurrent task performance? *Journal of Experimental Psychology: Learning, Memory and Cognition* 22:1:86-116.
- Biemiller, A. 1970. The development of the use of graphic and contextual information as children learn to read. *Reading Research Quarterly* 6:1:75-98.
- Biggs, J. 1992. *Why and how do Hong Kong students learn? Using the Learning and Study Process Questionnaires*. Education Paper 14, University of Hong Kong.
- Block, E. 1986a. Comprehension strategies of non-proficient college readers. In J. Niles and R. Lalik (eds.), *Thirty-fifth yearbook of the National Reading Conference*, pp. 344-352. Rochester, NY: National Reading Conference.
- Block, E. 1986b. The comprehension strategies of second language readers. *TESOL Quarterly* 20:3:463-494.
- Block, E. 1992. See how they read: Comprehension monitoring of L1 and L2 readers. *TESOL Quarterly* 26:2:319-343.
- Bransford, J. and Franks, J. 1971. The abstraction of linguistic ideas. *Cognitive Psychology* 3:331-350.



- Brown, A. 1980. Metacognitive development and reading. In Spiro, R., Bruce, B. and Brewer, W. (eds.), *Theoretical issues in reading comprehension*, pp. 453-481. Hillsdale, NJ: Erlbaum.
- Brown, A. and Palincsar, A. 1982. Inducing strategic learning from texts by means of informed, self-control training. *Topics in Learning and Learning Disabilities* 2:1:1-17.
- Buswell, G. 1922. *Fundamental reading habits: A study of their development*. Chicago: Chicago University Press.
- Carr, T. 1981. Building theories of reading abilities: On the relation between individual differences in cognitive skills and reading comprehension. *Cognition* 9:73-114.
- Carr, T. 1982. What's in a model: Reading theory and reading instruction. In Singer, M. (ed.), *Competent reader, disabled reader: Research and application*.
- Carrell, P. 1984. Evidence of a formal schema in second language comprehension. *Language Learning* 34:2:87-112.
- Carrell, P. 1988. Introduction: Interactive approaches to second language reading. In Carrell, P., Devine, J. and Eskey, D. (eds.), *Interactive approaches to second language reading*, pp. 1-7. Cambridge: Cambridge University Press.
- Carrell, P. 1989. Metacognitive awareness and second language reading. *Modern Language Journal* 73:121-133.
- Carrell, P. and Wallace, B. 1983. Background knowledge: Context and familiarity in reading comprehension. In Clarke, M. and Handscombe, J. (eds.), *On TESOL '82*, pp. 295-308. Washington, D.C.: TESOL.
- Carrell, P., Pharis, B. and Liberto, J. 1989. Metacognitive strategy training for ESL reading. *TESOL Quarterly* 23:4:647-678.
- Casanave, C. 1988. Comprehension monitoring strategies in ESL reading: A neglected essential. *TESOL Quarterly* 22:2:283-302.
- Cattell, J. 1886. The time it takes to see and name objects. *Mind* 11:63-65.
- Cavalcanti, M. 1987. Investigating FL reading performance through pause protocols. In Faerch, C. and Kasper, G. (eds.), *Introspection in second language research*, pp. 230-250. Clevedon: Multilingual Matters.
- Chan, M. and Kwok, H. 1982. A study of lexical borrowing from English in Hong Kong Chinese. Hong Kong: Centre of Asian Studies, University of Hong Kong.
- Chihara, T., Oller, J., Weaver, K. and Chavez-Oller, M. 1977. Are cloze items sensitive to constraints across sentences? *Language Learning* 27:1:63-73.

- Chomsky, N. 1957. *Syntactic structures*. The Hague: Mouton.
- Chomsky, N. 1959. A review of *Verbal Behavior* by Skinner, B. *Language* 35:26-58.
- City University of Hong Kong. 1995. *Calendar 1995-96*. Hong Kong: Author.
- Clarke, M. 1979/1981. Reading in Spanish and English: Evidence from adult ESL students. *Language Learning* 29:121-150. (Reprinted in Shuy, R. (ed.), *Papers in Applied Linguistics: Linguistics and literacy, series 1: Learning to read in different languages*, pp. 69-92, Hudelson, S. (ed.). Washington, D. C.: Center for Applied Linguistics, 1981.)
- Clarke, M. 1980/1988. The short-circuit hypothesis of ESL reading – or when language competence interferes with reading performance. *Modern Language Journal* 64:203-209. (Reprinted in Carrell et al. 1988, pp. 114-124.)
- Clarke, M. and Burdell, L. 1977. Shades of meaning: Syntactic and semantic parameters of cloze test responses. In Brown, H., Yorio, C. and Crymes, R. (eds.), *On TESOL '77: Teaching and learning trends in research and practice*, pp. 131-143. TESOL: Washington, D.C.
- Clarke, M. and Silberstein, S. 1977. Toward a realization of psycholinguistic principles for the ESL reading class. *Language Learning* 27:135-154.
- Coady, J. 1979. A psycholinguistic model of the ESL reader. In Mackay, R., Barkman, B. and Jordan, R. (eds.), *Reading in a second language*, pp. 5-12. Rowley, MA: Newbury House.
- Cohen, A. 1984. Studying second-language learning strategies: How do we get the information? *Applied Linguistics* 5:2:101-112.
- Cohen, A. 1986. Mentalistic measures in reading strategy research: Some recent findings. *English for Specific Purposes* 5:2:131-145.
- Cohen, A. 1987. Using verbal reports in research on language learning. In Faerch, C. and Kasper, G. (eds.), *Introspection in second language research*, pp. 82-95. Clevedon: Multilingual Matters.
- Cohen, L. and Manion, L. 1994. *Research methods in education* (4<sup>th</sup> edition.). London: Routledge.
- Coltheart, M. 1978. Lexical access in simple reading tasks. In G. Underwood (ed.), *Strategies for information processing*, pp. 151-216. London: Academic Press.
- Coomber, J. and Hogie, D. 1974. *Perceptual and decoding factors in reading disability*. Paper presented at the annual meeting of the American Educational Research Association, Chicago.



- Corder, S. 1978. Language-learner language. In Richards, J. (ed.), *Understanding second and foreign language learning*, pp. 71-93. Rowley, MA: Newbury House.
- Crewe, W. and Tong, K. 1988. Handling the cultural component of ESL reading texts. In Bickley, V. (ed.), *Languages in education in a bi-lingual setting*, pp. 369-385. Institute of Language in Education: Hong Kong Education Department.
- Crossley, M. and Vulliamy, G. 1984. Case-study research methods and comparative education. *Comparative Education* 20:2:193-207.
- Cziko, G. 1978. Differences in first- and second-language reading: The use of syntactic, semantic and discourse constraints. *Canadian Modern Language Review* 34:473-489.
- Cziko, G. 1980. Language competence and reading strategies: A comparison of first- and second-language oral reading errors. *Language Learning* 30:101-114.
- Dansereau, D. 1985. Learning strategy research. In Segal, J., Chipman, S. and Glaser, R. (eds.), *Thinking and learning skills*, vol. 1, pp. 209-239. Hillside, NJ: Erlbaum.
- Davies, F. 1995. *Introducing reading*. In Carter, R and Nunan, D. (eds.), Penguin English Applied Linguistics series. London: Penguin.
- Davis, J. and Bistodeau, L. 1993. How do L1 and L2 reading differ? Evidence from think-aloud protocols. *Modern Language Journal* 77:4:459-472.
- DeFrancis, J. 1989. *Visible speech: The diverse oneness of writing systems*. Honolulu: University of Hawaii.
- Derry, S. and Murphy, D. 1986. Designing systems that train learning ability: From theory to practice. *Review of Educational Research* 56:1-39.
- Devine, J. 1987. General language competence and adult second language reading. In Devine, J., Carrell, P. and Eskey, D. (eds.), *Research in reading in English as a second language*. Washington, D.C.: TESOL.
- Devine, J. 1988. A case study of two readers: Models of reading and reading performance. In Carrell, P., Devine, J. and Eskey, D. (eds.), *Interactive approaches to second language reading*, pp. 127-139.
- Di Pietro, R. 1971. *Language structure in contrast*. Rowley, MA: Newbury House.
- Doff, A. 1988. *Teach English: A training course for teachers*. Cambridge: Cambridge University Press.
- Dowsett, G. 1986. Interaction in the semi-structured interview. In Emery, M. (ed.), *Qualitative research*. Canberra: Australian Association of Adult Education.

- Dubin, F. and Bycina, D. 1991. Academic reading and the ESL/EFL teacher. In Celce-Murcia, M. (ed.), *Teaching English as a second or foreign language*, 2<sup>nd</sup> edition, pp. 195-215. New York: Newbury House.
- Ericsson, K. and Simon, H. 1980. Verbal reports as data. *Psychological Review* 87:3:215-251.
- Ericsson, K. and Simon, H. 1984/1993. *Protocol analysis*. Cambridge, MA: MIT Press. (Reprinted under the same title in 1993.)
- Eskey, D. 1973. A model program for teaching advanced reading to students of English as a foreign language. *Language Learning* 23:2:169-184.
- Eskey, D. 1987. Conclusion. In Devine, J., Carrell, P. and Eskey, D. (eds.), *Research in reading in English as a second language*, pp. 189-192.
- Eskey, D. 1988. Holding in the bottom: an interactive approach to the language problems of second language readers. In Carrell, P., Devine, J. and Eskey, D. *Interactive approaches to second language reading*, pp. 93-100. Cambridge: Cambridge University Press.
- Faerch, C. and Kasper, G. 1987. From product to process – introspective methods in second language research. In Faerch, C. and Kasper, G. (eds.), *Introspection in second language research*, pp. 5-23. Clevedon: Multilingual Matters.
- Farmer, R. 1994. The limits of learner independence in Hong Kong. In Gardner, D. and Miller, L. (eds.), *Directions in self-access language learning*, pp. 13-27. Hong Kong: Hong Kong University Press.
- Feldman, U. and Stemmer, B. 1987. Thin\_\_\_ aloud a\_\_\_ retrospective da\_\_\_ in C-te\_\_\_ taking: diffe\_\_\_ languages – diff\_\_\_ learners – sa\_\_\_ approaches? In Faerch, C. and Kasper, G. (eds.), *Introspection in Second Language Research*. Clevedon: Multilingual Matters.
- Fisher, D. and Smith, M. 1977. An information processing approach. In Stanley, F. (ed.), *Linguistics and Reading Series*, 2, pp. 23-30. Arlington, VA: Center for Applied Linguistics.
- Flanagan, J. 1954. The critical incident technique. *Psychological Bulletin* 51:327-358.
- Flowerdew, J. 1999. The practicum in L2 teacher education: A Hong Kong case study. *TESOL Quarterly* 33:1:141-145.
- Flowerdew, J., Li, D. and Miller, L. 1998. Attitudes towards English and Cantonese among Hong Kong Chinese University Lecturers. *TESOL Quarterly* 32:2:201-233.
- Fox, B. and Routh, D. 1976. Phonemic analysis and synthesis as word-attack skills. *Journal of Educational Psychology* 68:70-74.



- Fries, C. 1972. Learning to read English as part of the oral approach. In Croft, K. (ed.), *Readings on English as a second language: For teachers and teacher-trainers*, pp. 168-173. Cambridge: Winthrop Publishers.
- Garner, R. 1987. *Metacognition and reading comprehension*. Norwood, NJ: Ablex.
- Gelb, I. 1952. *A study of writing*. Chicago: University of Chicago Press.
- Gibbons, J. 1979. U-gay-wa: A linguistic study of the campus language of students at the University of Hong Kong. In Lord, R. (ed.), *Hong Kong Language Papers*, pp. 3-43. Hong Kong: Hong Kong University Press.
- Gleitman, L. and Rozin, P. 1977. The structure and acquisition of reading I: Relations between orthographies and the structure of language. In Reber, A. and Scarborough, D. (eds.), *Toward a psychology of reading: The proceedings of the CUNY conferences*, pp. 1-54. Hillsdale, NJ: Lawrence Erlbaum.
- Golinkoff, R. 1975-76. A comparison of reading comprehension processes in good and poor comprehenders. *Reading Research Quarterly* 11:4: 623-659.
- Goodman, K. 1967. Reading: A psycholinguistic guessing game. *Journal of the Reading Specialist*, 6:4:126-135.
- Goodman, K. 1969. Analysis of oral reading miscues: Applied psycholinguistics. *Reading Research Quarterly* 5:1:19-30.
- Goodman, K. 1975/1988. The reading process. In Gollasche, F. (ed.), *Language and literacy: The selected writings of Kenneth Goodman*, pp. 5-16. London: Routledge and Kegan Paul. (Reprinted in Carrell, P., Devine, J. and Eskey, D. (eds.), *Interactive approaches to second language reading*, pp. 11-21. Cambridge: CUP. 1988.)
- Goodman, K. 1976. Behind the eye: What happens in reading. In Singer, H. and Ruddell, R. (eds.), *Theoretical models and processes of reading*, 2<sup>nd</sup> edition, pp. 470-496). Newark, DE: International Reading Association.
- Goodman, K. 1986. *What's whole in whole language?* Portsmouth, NH: Heinemann.
- Goodman, K. 1992. I didn't found whole language. *The Reading Teacher* 46:3:188-199.
- Gordon, R. 1980. *Interviewing: Strategy, techniques and tactics*. New York: Dorsey Press.
- Gough, P. 1972. One second of reading. In Kavanagh, J. and Mattingly, I. (eds.), *Language by eye and by ear*, pp. 331-358. Cambridge, MA: MIT Press.
- Gough, P. 1985. One second of reading: Postscript. In Singer, H. and Ruddell, R. (eds.), *Theoretical models and processes of reading* (3<sup>rd</sup> edition), pp. 687-688. Newark, DE: International Reading Association.

- Gough, P. and Hillinger, M. 1980. Learning to read: An unnatural act. *Bulletin of the Orton Society* 30:171-176.
- Grabe, W. 1991. Current developments in second language reading research. *TESOL Quarterly* 25:3:375-406.
- Grabe, W. 1996. *Textual coherence and discourse analysis*. AILA conference notes, July 1996.
- Grellet, F. 1981. *Developing reading skills: A practical guide to reading comprehension exercises*. Cambridge: Cambridge University Press.
- Grotjahn, R. 1987. On the methodological basis of introspective methods. In Faerch, C. and Kasper, G. (eds.), *Introspection in second language research*. Clevedon: Multilingual Matters.
- Grundin, H. 1994. Who's romancing reality? A response to Keith Stanovich. *The Reading Teacher* 48:8-10.
- Guardian Education. 1997. *The writings on the wall*, 4<sup>th</sup> February 1997, p. 6. London: Guardian Newspapers.
- Haastrup, K. 1987. Using thinking aloud and retrospection to uncover learners' lexical inferencing procedures. In Faerch, C. and Kasper, G. (eds.), *Introspection in second language research*. Clevedon: Multilingual Matters.
- Halliday, M. and Hasan, R. 1976. *Cohesion in English*. London: Longman.
- Harri-Augstein, S. and Thomas, L. 1984. Conversational investigations of reading: The self-organized learner and the text. In Alderson, J. and Urquhart, A. (eds.), *Reading in a foreign language*. New York: Longman.
- Hedge, N. 1991. *Reading and learning: An exploration of perceptions, approaches and strategies of overseas postgraduates using English for Academic Purposes*. Unpublished doctoral thesis: University of Liverpool.
- Henderson, E. 1903. A study of memory for connected trains of thought. *Psychological Monographs* 5:6:1-94.
- Hilgard, E. 1980. The trilogy of mind: Cognition, affection and conation. *Journal of the History of the Behavioral Sciences* 16:107-117.
- Hockberg, J. 1970. Components of literacy: Speculations and exploratory research. In Levin, H. and Williams, J. (eds.), *Basic studies on reading*. New York: Basic Books.
- Hong Kong Advanced Level Examinations Syllabus. 1994. Wan Chai, Hong Kong: Hong Kong Examinations Authority.



- Hosenfeld, C. 1977. A preliminary investigation of the reading strategies of successful and unsuccessful second language learners. *System* 5:110-123.
- Hosenfeld, C., Arnold, V., Kirchofer, J., Laciura, J. and Wilson, L. 1981. Second language reading: A curricular sequence for teaching reading strategies. *Foreign Language Annals*, 14:5:415-422.
- Hudson, T. 1982/1988. The effects of induced schemata on the "short circuit" in L2 reading: Non-decoding factors in L2 reading performance. *Language Learning* 32:1:1-31. (Reprinted in Carrell et al. 1988, pp. 183-205.)
- Huey, E. 1908/1968. *The psychology and pedagogy of reading*. Cambridge, MA: MIT Press.
- Hull, J. 1999. *Oxford English* (3<sup>rd</sup> ed.). Hong Kong: Oxford University Press.
- Hulstijn, J. and Hulstijn, W. 1984. Grammatical errors as a function of processing constraints and explicit knowledge. *Language Learning* 34:23-43.
- Hyland, K. 1996. *Language attitudes at the handover: Communication and identity in 1997 Hong Kong*. Paper presented at World Englishes Conference, Honolulu, HI, December 1996.
- James, C. 1971. The exculpation of contrastive linguistics. In Nickel, G. (ed.), *Papers in Contrastive Linguistics*. Cambridge: Cambridge University Press.
- James, M. 1987. ESL reading pedagogy: Implications of schema-theoretical research. In Devine, J., Carrell, P. and Eskey, D. (eds.), *Research in reading in English as a second language*, pp. 175-189. Washington, D.C.: TESOL.
- James, W. 1890. *The principles of psychology*. New York: Holt.
- Javal, E. 1879. Essai sur la physiologie de la lecture. *Annales d'Oculistique* 82:242-253.
- Jenkinson, M. 1957. *Selected processes and difficulties in reading comprehension*. Unpublished doctoral dissertation. Chicago, IL: University of Chicago.
- Jimenez, R. 1997. The strategic reading abilities and potential of five low-literacy Latina/o readers in middle school. *Reading Research Quarterly* 32:3:224-243.
- Jimenez, R., Garcia, G. and Pearson, P. 1996. The reading strategies of bilingual Latina/o students who are successful English readers: Opportunities and obstacles. *Reading Research Quarterly* 31:1:90-112.
- Jonz, J. 1987. Textual cohesion and second language comprehension. *Language Learning* 37:3:409-438.
- Jonz, J. 1990. Another turn in the conversation: What does cloze measure? *TESOL Quarterly* 24:1:61-83.

- Juel, C. 1980. Comparison of word identification strategies with varying context, word type, and reader skill. *Reading Research Quarterly* 15:112-126.
- Jung, C. 1910. The association method. *American Journal of Psychology* 21:219-269.
- Just, M. and Carpenter, P. 1980. A theory of reading: From eye fixations to comprehension. *Psychological Review* 87:4:329-354.
- Kerlinger, F. 1969. *Foundations of behavioral research*. New York: Holt, Rinehart and Winston.
- Kern, R. 1989. Second language reading strategy instruction: Its effects on comprehension and word inference ability. *Modern Language Journal* 73:135-149.
- Kintsch, W. and Vipond. 1979. Reading comprehension and readability in educational practice and psychological theory. In Nilsson, L. (ed.), *Perspectives on memory research*. Hillsdale, NJ: Erlbaum.
- Kirby, J. 1988. Style, strategy, and skills in reading. In Schmeck, R. (ed.) *Learning strategies and learning styles*, pp. 229-274. London: Plenum Press.
- Kletzien, S. 1991. Strategy use by good and poor comprehenders reading expository text of differing levels. *Reading Research Quarterly* 26:1:67-86.
- Kohonen, V. 1992. Experiential language learning: Towards second language learning as learner education. In Nunan, D. (ed.), *Collaborative language learning and teaching*, pp. 14-39. Cambridge: Cambridge University Press.
- Krashen, S. 1985. *The input hypothesis: Issues and implications*. Oxford: Pergamon.
- LaBerge, D. and Samuels, S. 1974/1976. Toward a theory of automatic information processing in reading. *Cognitive Psychology* 6:293-323. (Reprinted in Singer, H. and Ruddell, R. (eds.), *Theoretical models and processes in reading* 2<sup>nd</sup> ed., pp. 548-579. Newark, DE: International Reading Association.)
- Larsen-Freeman, D. 1991. Second language acquisition research: Staking out the territory. *TESOL Quarterly* 25:2:315-350.
- Lee, J. 1986. On the use of the recall task to measure L2 reading comprehension. *Studies in Second Language Acquisition* 8:201-211.
- Lee, J. 1991. On the dual nature of the second-language reading proficiency of beginning language learners. In Teschner, R. (ed.), *Assessing Foreign Language Proficiency of Undergraduates*. Boston: Heinle.
- Levenston, E., Nir, R. and Blum-Kulka, S. 1984. Discourse analysis and the testing of reading comprehension by cloze techniques. In Pugh, A. & Ulijn, J. (eds.), *Reading for Professional Purposes*. London: Heinemann.



- Levin, J. 1986. Four cognitive principles of learning-strategy instruction. *Educational Psychologist* 21:3-17.
- Lieberman, D. 1979. Behaviorism and the mind: A (limited) call for a return to introspection. *American Psychologist* 34:4:319-333.
- Li, S. and Munby, H. 1996. Metacognitive strategies in second language academic reading: A qualitative investigation. *English for Specific Purposes* 15:3:199-216.
- Lin, A. 1990. *Teaching in two tongues: Language alternation in foreign language classrooms*. Research Report 3. Hong Kong: Department of English, City University of Hong Kong.
- Lo, T. 1988. Using mother tongue as the teaching medium in Hong Kong: But whose? In Bickley, V. (ed), *Languages in education in a bi-lingual setting*, pp. 210-219. Institute of Language in Education: Hong Kong Education Department.
- Lukatela, G. and Turvey, M. 1990. Automatic and pre-lexical computation of phonology in visual word identification. *European Journal of Cognitive Psychology* 2:325-344.
- Luke, K. and Richards, J. 1982. English in Hong Kong: Functions and status. *English World Wide* 3:1:47-64.
- Lundberg, I., Olofsson, A. and Wall, S. 1980. Reading and spelling skills in the first school years predicted from phonemic awareness skills in kindergarten. *Scandinavian Journal of Psychology* 21:159-173.
- Lunzer, E. and Gardner, K. (eds.) 1979. *The effective use of reading*. London: Heinemann.
- Lysynchuk, L., Pressley, M., d'Ailly, H., Smith, M. and Cake, H. 1989. A methodological analysis of experimental studies of comprehension strategy instruction. *Reading Research Quarterly* 24:4:458-470.
- Mackworth, J. 1972. Some models of the reading process: Learners and skilled readers. *Reading Research Quarterly* 7:701-733.
- Mahoney, D., Hull, J. and Shillaw, J. 1997. Storing simple stories: Narrative recall and the Chinese speaker. *Language, Culture and Curriculum* 10:1:66-87.
- Mandler, J. 1978. A code in the node: The use of a story schema in retrieval. *Discourse Processes* 1:14-35.
- Mandler, J. and Johnson, N. 1977. Remembrance of things parsed: Story structure and recall. *Cognitive Psychology* 9:111-151.

- Massaro, D. 1975. *Understanding language: An information-processing analysis of speech perception, reading and psycholinguistics*. New York: Academic Press.
- Mathewson, G. 1985. Towards a comprehensive model of affect in the reading process. In Singer, H. and Ruddell, R., *Theoretical models and processes of reading*. Newark: DE: International Reading Association.
- McClelland, J. 1986. The programmable blackboard model of reading. In McClelland, J., Rumelhart, D., and the PDP research group (eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*, vol. 2. Cambridge, MA: Bradford Books.
- McConkie, G. and Zola, D. 1981. Language constraints and the functional stimulus in reading. In Lesgold, A. and Perfetti, C. (eds.), *Interactive processes in reading*. Hillsdale, NJ: Erlbaum.
- McLaughlin, B. 1978. *Second language acquisition in childhood*. New Jersey: Lawrence Erlbaum.
- McLaughlin, B. 1990. Restructuring. *Applied Linguistics* 11:113-128.
- McLeod, B and McLaughlin, B. 1986. Restructuring or automaticity? Reading in a second language. *Language Learning* 36:2:109-123.
- Miller, G., Bruner, J. and Postman, L. 1954. Familiarity of letter sequences and tachistoscopic identification. *Journal of General Psychology* 50:129-139.
- Mitchell, D. 1984. An evaluation of subject-paced reading tasks and other methods for investigating immediate processes in reading. In Kieras, D. and Just, M. (eds.), *New methods in reading comprehension research*. Hillsdale, NJ: Erlbaum.
- Mitchell, D. and Green, D. 1978. The effects of context and content on immediate processing in reading. *Quarterly Journal of Experimental Psychology* 30:609-636.
- Naiman, N., Frohlich, M., Stern, H. and Todesco, A. 1978. *The good language learner*. Research in Education Series 7. Toronto: Modern Language Center, Ontario Institute for Studies in Education.
- Neely, J. 1977. Semantic priming and retrieval from lexical memory roles of inhibitionless spreading activation and limited capacity attention. *Journal of Experimental Psychology: General* 106:226-254.
- Neville, M. and Pugh, A. 1975. An exploratory study of the application of time-compressed and time-expanded speech in the development of the English reading proficiency of foreign students. *English Language Teaching Journal* 29:320-329.



- Newell, A. and Simon, H. 1972. *Human problem solving*. Englewood Cliffs, NJ: Prentice-Hall.
- Nicholson, T., Lillas, C. and Rzoska, M. 1988. Have we been misled by miscues? *The Reading Teacher* 42:6-10.
- Norman, J. 1988. *Chinese*. Cambridge: CUP.
- Nunan, D. 1992. *Research methods in language learning*. Cambridge: Cambridge University Press.
- Nuttall, C. 1982. *Teaching reading skills in a foreign language*. London: Heinemann.
- Oakhill, J. and Garnham, A. 1988. *Becoming a skilled reader*. Oxford: Blackwell.
- Odlin, T. 1989. *Language transfer: Cross-linguistic influence in language learning*. Cambridge: Cambridge University Press.
- Olshavsky, J. 1976-77. Reading as problem solving: An investigation of strategies. *Reading Research Quarterly* 12:4:654-674.
- Olson, G., Duffy, S. and Mack, R. 1984. Thinking-out-loud as a method of studying real-time comprehension processes. In Kieras, D and Just, M. (eds.), *New Methods in Reading Comprehension Processes*, pp. 253-286. Hillsdale, NJ: Erlbaum.
- O'Malley, J. and Chamot, A. 1990. *Learning strategies in second language acquisition*. Cambridge: Cambridge University Press.
- Oppenheim, A. 1992. *Questionnaire design, interviewing and attitude measurement*. London: Pinter Publishers.
- Oxford, R. 1990. *Language learning strategies: What every teacher should know*. New York: Newbury House.
- Oxford, R. and Nyikos, M. 1989. Variables affecting choice of language learning strategies by university students. *Modern Language Journal* 73(2):291-300.
- Paap, K. and Noel, R. 1991. Dual-route models of print and sound: Still a good horse race. *Psychological Research* 53:13-24.
- Padron, Y. 1985. *Utilizing cognitive reading strategies to improve English reading comprehension of Spanish-speaking bilingual students*. Unpublished doctoral dissertation, University of Houston.
- Palincsar, A. and Brown, A. 1984. Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction* 1:2:117-175.
- Paran, A. 1996. Reading in EFL: Facts and fictions. *ELT Journal* 50:1:25-34.

- Paris, S., Lipson, M. and Wixson, K. 1983. Becoming a strategic reader. *Contemporary Educational Psychology* 8:293-316.
- Paris, S. and Myers, M. 1981. Comprehension monitoring, memory, and study strategies of good and poor readers. *Journal of Reading Behavior* 13:5-22.
- Paris, S., Wasik, B. and Turner, J. 1991. The development of strategic readers. In Barr, R., Kamil, M., Mosenthal, P. and Pearson, P. (eds.), *Handbook of reading research*, pp. 609-640. New York: Longman.
- Parry, K. 1996. Culture, literacy and L2 reading. *TESOL Quarterly* 30:4:665-692.
- Pennington, M. and Yue, F. 1994. English and Chinese in Hong Kong: Pre-1997 language attitudes. *World Englishes* 13:1-20.
- Perfetti, C. 1985. *Reading ability*. New York: Oxford University Press.
- Perfetti, C. 1995. Cognitive research can inform reading education. *Journal of Research in Reading* 18:2:106-115.
- Perfetti, C., Bell, L. and Delaney, S. 1988. Automatic phonetic activation in silent word reading: Evidence from backward masking. *Journal of Memory and Language* 27:59-70.
- Perfetti, C., Goldman, S. and Hogaboam, T. 1979. Reading skill and the identification of words in discourse context. *Memory and Cognition* 7:4: 723-282.
- Perfetti, C. and Roth, S. 1981. Some of the interactive processes in reading and their role in reading skill. In Lesgold, A. and Perfetti, C. (eds.), *Interactive processes in reading*, pp. 269-297. Hillsdale, NJ: Lawrence Erlbaum.
- Perfetti, C. and Zhang, S. 1995a. The universal word identification reflex. In Medlin, D. (ed.), *The psychology of learning and motivation*, vol. 33, pp. 159-189. San Diego, CA: Academic Press.
- Perfetti, C. and Zhang, S. 1995b. Very early phonological activation in Chinese reading. *Journal of Experimental Psychology: Learning, Memory and Cognition* 21:24-33.
- Perfetti, C., Zhang, S. and Berent, I. 1992. Reading in English and Chinese: Evidence for a 'universal' phonological principle. In Frost, R. and Katz, L. (eds.), *Orthography, phonology, morphology, and meaning*, pp. 227-248. Amsterdam: North-Holland.
- Plaister, T. 1968. Reading instruction for college level foreign students. *TESOL Quarterly* 2:3:164-168.



- Porter, P., Goldstein, L., Leatherman, J. and Conrad, S. 1990. An on-going dialogue: learning logs for teacher preparation. In Richards, J. and Nunan, D. (eds.), *Second language teacher education*, pp. 227-240. New York: Cambridge University Press.
- Pritchard, R. 1990. The effects of cultural schemata on reading processing strategies. *Reading Research Quarterly* 25:4:273-295.
- Radford, J. 1974. Reflections on introspection. *American Psychologist* 29:4:245-250.
- Rankin, J. 1988. Designing think-aloud studies in ESL reading. *Reading in a Foreign Language* 4:2:119-132.
- Rankin, E. and Thomas, S. 1980. Contextual constraints and the construct validity of the cloze procedure. In Kamil, M., and Moe, A. (eds.), *Perspectives on reading research and instruction: Twenty-ninth yearbook of the National Reading Conference*, pp. 47-55. Washington, D.C.: NRC.
- Rayner, K. and Bertera, J. 1979. Reading without a fovea. *Science* 206:26: 468-469.
- Rayner, K. and Pollatsek, A. 1989. *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.
- Reber, A. 1993. *Implicit learning and tacit knowledge*. Oxford: Clarendon Press.
- Richards, J. (ed.) 1974. *Error analysis*. London: Longman.
- Richards, J., Tung, P. and Ng, P. 1992. The culture of the English language teacher: A Hong Kong example. *RELJ Journal* 23:1:81-103.
- Rivers, W. 1968. *Teaching foreign language skills*. Chicago: University of Chicago Press.
- Rowe, D. and Rayford, L. 1984. *Instantiating background knowledge in reading comprehension assessment*. Paper presented at the annual meeting of the National Reading Conference, St. Petersburg Beach, FL.
- Rubin, J. 1975. What the good language learner can teach us. *TESOL Quarterly* 9:41-51.
- Rubin, J. 1981. Study of cognitive processes in second language learning. *Applied Linguistics* 11:117-131.
- Rubin, J. 1987. Learner strategies: Theoretical assumptions, research history and typology. In Wenden, A. and Rubin, J. (eds.), *Learner strategies in language learning*, pp. 15-30. New York: Prentice Hall.
- Ruddell, R. and Speaker, R. 1985. The interactive reading process: A model. In Singer, H. and Ruddell, R. (eds.), *Theoretical models and processes of reading*, pp. 751-793. Newark, DE: International Reading Association.

- Rumelhart, D. 1975. Notes on a schema for stories. In Bobrow, D. and Collins, A. (eds.), *Representation and understanding: Studies in cognitive science*, pp. 211-236. New York: Academic Press.
- Rumelhart, D. 1977. Toward an interactive model of reading. In Dornic, S. (ed.), *Attention and language performance VI*, pp. 573-603. Hillsdale, NJ: Erlbaum.
- Rumelhart, D. 1984. Understanding understanding. In Flood, J. (ed.), *Understanding reading comprehension*, pp. 1-20. Newark, DE: International Reading Association.
- Rumelhart, D., McClelland, J. and the PDP Research Group. 1986. *Parallel distributed processing: Explorations of the microstructure of cognition*, vol. 1. Cambridge, MA: Cambridge University Press.
- Samuels, S. 1977. Introduction to theoretical models of reading. In Otto, W. (ed.), *Reading problems*. Boston: Addison-Wesley.
- Samuels, S. and Kamil, M. 1984/1988. Models of the reading process. In Pearson, P. (ed.), *Handbook of reading research*. New York: Longman. (Reprinted in Carrell et al.)
- Sarig, G. 1987a. *Comprehension-promoting strategies: The sum of the parts and the whole*. Paper presented at the Second International Congress of Applied Psycholinguistics, University of Kassel, July 1987.
- Sarig, G. 1987b. High-level reading in the first and in the foreign language: Some comparative process data. In Devine, J., Carrell, P. and Eskey, D. (eds.), *Research in Reading English as a Second Language*. Rowley, MA: Newbury House.
- Schmidt, R. 1995a. Consciousness and foreign language learning: A tutorial on the role of attention and awareness in learning. In Schmidt, R. (ed.), *Attention and awareness in foreign language learning*, pp. 1-63. Second Language Teaching and Curriculum Center: University of Hawaii.
- Schmidt, R. (ed.) 1995b. *Attention and awareness in foreign language learning*. Second Language Teaching and Curriculum Center: University of Hawaii.
- Schneider, W. and Shiffrin, R. 1977. Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review* 84:1-66.
- Scollon, R. and Wong Scollon, S. 1995. *Intercultural communication*. Oxford: Blackwell.
- Scott, M. 1990. *Demystifying the jabberwocky: A research narrative*. Unpublished PhD thesis. University of Lancaster.



- Seidenberg, M., Waters, G., Barnes, M. and Tanenhaus, M. 1984. When does irregular spelling or pronunciation influence word recognition? *Journal of Verbal Learning and Verbal Behavior* 23:383-404.
- Seliger, H. 1983. The language learner as linguist: Of metaphors and realities. *Applied Linguistics* 4:3:179-191.
- Seliger, H. and Shohamy, E. 1989. *Second language research methods*. Oxford: Oxford University Press.
- Selinker, L. 1972. Interlanguage. *IRAL* 10:209-231.
- Selinker, L. and Gass, S. 1984. *Workbook in second language acquisition*. Rowley, MA: Newbury House.
- Selz, O. 1913. *Über die Gesetze des geordneten Denkverlaufs*. Stuttgart: Spemann.
- Selz, O. 1922. *Zur Psychologie des produktiven Denkens und des Irrtums*. Bonn: Friedrich Cohen.
- Serrano, N. 1984. Patterns of reading in L1 and L2. In Singleton, D. and Little, D. (eds.), *Language learning in formal and informal contexts*. Dublin: Irish Association for Applied Linguistics, 165-176.
- Shiffrin, R. and Schneider, W. 1977. Controlled and automatic human processing: II. Perceptual learning, automatic attending and a general theory. *Psychological Review* 84:127-189.
- Silberstein, S. 1987. Let's take another look at reading: Twenty-five years of reading instruction. *English Teaching Forum* 25:28-35.
- Simon, H. 1979. *Models of thought*. New Haven: Yale University Press.
- Singer, H. 1981. Instruction in reading acquisition. In Tzeng, O. and Singer, H. (eds.), *Perception of print*, pp. 291-311. Hillsdale, NJ: Erlbaum.
- Skehan, P. 1989. *Individual differences in second-language learning*. London: Edward Arnold.
- Skinner, B. 1957. *Verbal behavior*. New York: Appleton-Century-Crofts.
- Smith, F. 1971/1988. *Understanding reading: A psycholinguistic analysis of reading and learning to read* (1<sup>st</sup> edition). New York: Holt, Rinehart and Winston. (4<sup>th</sup> edition published in 1988)
- Smith, F. 1973. *Psycholinguistics and reading*. New York: Holt, Rinehart and Winston.
- Smith, M. 1981. Consciousness-raising and the second language learner. *Applied Linguistics* 11:2.

- Smith, W. 1978. Cloze procedure as applied to reading. In Buros, O. (ed.), *The eighth mental measurements yearbook*, vol. 2, pp. 1176-1178. Highbury Park, NJ: Gryphon.
- Stanovich, K. 1980. Toward an interactive compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16:1: 32-71.
- Stanovich, K. 1982. Individual differences in the cognitive processes of reading. *Journal of Learning Disabilities* 15:485-493, 549-554.
- Stanovich, K. 1986. Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21:4:360-406.
- Stanovich, K. and Cunningham, A. 1991. Reading as constrained reasoning. In Sternberg, S. and Frensch, P. (eds.), *Complex problem solving: Principles and mechanisms*, pp. 3-60. Hillsdale, NJ: Erlbaum.
- Stanovich, K., Cunningham, A. and Cramer, B. 1984. Assessing phonological awareness in kindergarten children: Issues of task comparability. *Journal of Experimental Child Psychology* 38:175-190.
- Stanovich, K. and Stanovich, P. 1995. How research might inform the debate about early reading acquisition. *Journal of Research in Reading* 18:2:87-105.
- Stanovich, K. and West, N. 1981. The effect of sentence context on ongoing word recognition: Tests of a two-process theory. *Journal of Experimental Psychology: Human Perception and Performance* 7:658-672.
- Stern, H. 1975. What can we learn from the good language learner? *Canadian Modern Language Review* 31:304-318.
- Stern, H. 1983. *Fundamental concepts of language teaching*. London: Oxford University Press.
- Strother, J. and Ulijn, J. 1987. Does syntactic rewriting affect English for Science and Technology (EST) text comprehension? In Devine, J., Carrell, P. and Eskey, D. (eds.), *Research in reading in English as a second language*. Washington, D.C.: TESOL.
- Tarone, E. 1982. Systematicity and attention in interlanguage. *Language Learning* 32:1:69-84.
- Taylor, W. 1953. 'Cloze procedure': A new tool for measuring readability. *Journalism Quarterly*, pp. 415-433.
- Tinker, M. 1939. Reliability and validity of eye-movement measures of reading. *Journal of Experimental Psychology* 9:135-147.



- Titchener, E. 1912. The schema of introspection. *American Journal of Psychology* 23:485-508.
- Tung, P. 1990. Why changing the medium of instruction in Hong Kong could be difficult. *Journal of Multilingual and Multicultural Development* 11:6:523-534.
- van Dijk, T. and Kintsch, W. 1983. *Strategies of discourse comprehension*. New York: Academic Press.
- van Orden, G., Pennington, B. and Stone, G. 1990. The effects of instructional bias on word identification. In Rieben, L. and Perfetti, C. (eds.), *Learning to read: Basic research and its implications*, pp. 189-203. Hillsdale, NJ: Lawrence Erlbaum.
- Venezky, R. 1984. The history of reading research. In Pearson, P. (ed.), *Handbook of reading research*. New York: Longman.
- Wade, S., Trathen, W. and Schraw, G. 1990. An analysis of spontaneous study strategies. *Reading Research Quarterly* 25:2:147-166.
- Wallace, C. 1992. *Reading*. Oxford: Oxford University Press.
- Watson, J. 1920. Is thinking merely the action of language mechanisms? *British Journal of Psychology* 11:87-104.
- Weber, R. 1970. A linguistic analysis of first grade reading errors. *Reading Research Quarterly* 5:427-451.
- Weinstein, C. 1978. Elaboration skills as a learning strategy. In H. O'Neill Jr. (ed.), *Learning strategies*. New York: Academic Press.
- Weinstein, C. and Underwood, V. 1985. Learning strategies: The *how* of learning. In Segal, J. Chipman, S. and Glaser, R. (eds.), *Relating instruction to research* pp. 241-258. Hillsdale, NJ: Erlbaum.
- Wenden, A. 1987. Conceptual background and utility. In Wenden, A. and Rubin, J. (eds.), *Learner strategies in language learning*. Hemel Hempstead: Prentice Hall.
- Wesche, M. 1979. Learning behaviors of successful adult students on intensive language training. *Canadian Modern Language Journal* 35:415-427.
- West, R. and Stanovich, K. 1978. Automatic contextual facilitation in readers of three ages. *Child Development* 49:717-727.
- Williams, E. and Moran, C. 1989. Reading in a foreign language at intermediate and advanced levels with particular reference to English. *Language Teaching*, 22:4:217-228.

Yorio, C. 1971. Some sources of reading problems for foreign language learners.  
*Language Learning* 21:1:107-115.

Yuill, N. and Oakhill, J. 1991. *Children's problems in text comprehension: An experimental investigation*. Cambridge: Cambridge University Press.