You're never really gonna be eliminating them:

A PHENOMENOLOGICAL STUDY OF ESL INSTRUCTORS' PERCEPTIONS
OF WRITTEN ERRORS AMONG ADULT ESL LEARNERS

by

Anthony Kin Kwok Tong

A thesis submitted in conformity with the requirements
for the degree of Doctor of Philosophy
Department of Curriculum, Teaching and Learning
Ontario Institute for Studies in Education of the
University of Toronto

© Copyright by Anthony Kin Kwok Tong (2000)
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-49868-9
You're never really gonna be eliminating them: A Phenomenological Study of ESL Instructors' Perceptions of Written Errors among Adult ESL Learners

Doctor of Philosophy, June 2000
Anthony Kin Kwok Tong
Department of Curriculum, Teaching and Learning
University of Toronto

ABSTRACT

This thesis examined, in the context of their practices for language assessment, the nature of language error and the assignment of gravity to different error types perceived by 4 English as a Second Language (ESL) instructors teaching an ESL Technical Writing course to undergraduate engineering students at a major university in Canada. Other participants were the 4 instructors' undergraduate students who had granted me the permission to use their written assignments as data and to record their classroom interactions with the 4 instructors.

The study adopted a phenomenological approach, using human perceptions as the starting point for investigating the nature and meanings of language errors in the context of language assessment. The data included 2 1-hour stimulated recalls (SRs) from each of the 4 instructors. In these recalls, the instructors reviewed their ESL students' previously assessed written reports and commented to me on the errors and their gravity. In another 1-hour interview, each of the 4 instructors explained their own beliefs, assumptions and knowledge systems (BAKs) related to language teaching. Other data included classroom observations and students' written reports.

I transcribed and performed 2 kinds of analyses on the SRs and BAKs. First, I coded the errors each of the instructors commented on in the SRs, based on the type and the gravity they had assigned. Second, I coded the themes (e.g., serious errors, minor errors, comprehensibility, irritating and disturbing errors, course objectives and
nature) which had emerged in the BAKs data to identify the essence of the instructors' perceptions.

Findings of the study refute the feasibility of a universal error hierarchy and support a constructivist perspective for language error evaluation, depicting error evaluation as interactive, context-sensitive activities greatly influenced by the judges' beliefs and various contextual factors. Findings also suggest that the meaning of error is embedded in the larger context of teaching, of which assessment is a part. I attempted to explain the nature of error in terms of Activity Theory (Leon'tev, 1981).
ACKNOWLEDGEMENTS

My interaction with my supervisor, Professor Alister Cumming, has proved to be most valuable and crucial to the successful completion of this thesis project; I wish to express my gratitude to him. I have very much appreciated his interpretative teaching style, his enthusiasm and skills in counseling and his pertinent and timely advice at key stages in the project.

Professors Merrill Swain and Patrick Allen, members of the thesis committee, and Professor James Heap, a former member of the thesis committee, have provided me with valuable comments and suggestions and I thank them. I thank also teachers at OISE/UT from whom I have taken courses but particularly Professors Alister Cumming, Birgit Harley, Merrill Swain, David Olson, Gordon Wells, and Joel Wesis. I have benefited a lot from the reading, writing and discussions that were a part of their classes.

A number of people have helped make this thesis project possible and I wish to express my gratitude to them. They are Usman Erdosy, Carmen Mok, Winnie Or, Alison Orszak, Carsten Quell, Mitzi Sakamoto, Alex Schorah, Katherine Teide, Marian Tyacke, and Howard Zuckernick. I am particularly grateful for the generous support of Usman and Katherine.

During the two years of my coursework at OISE/UT and in the process of completing this project, I have gained tremendous support from many of my family members and relatives. I would particularly like to mention my parents, Y. S. Tang and Chau-ying Chan, and my sisters Alice Yuen and Pat Thompson.

Finally, I wish to dedicate this project to my son Gary and daughter Katherine. Without them and their love, the melon of life would never have tasted as pleasant — a taste to which the completion of a humble thesis such as this one cannot compare.
<table>
<thead>
<tr>
<th>TABLE OF CONTENTS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT .................................................................</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS ........................................................</td>
</tr>
<tr>
<td>TABLE OF CONTENTS ........................................................</td>
</tr>
<tr>
<td>LIST OF TABLES ..................................................................</td>
</tr>
<tr>
<td>LIST OF FIGURES ..................................................................</td>
</tr>
<tr>
<td>LIST OF APPENDICES ................................................................</td>
</tr>
<tr>
<td>CHAPTER ONE: INTRODUCTION ..................................................</td>
</tr>
<tr>
<td>1.1. Where is this coming from? ..........................................</td>
</tr>
<tr>
<td>1.2. The organization of this thesis ....................................</td>
</tr>
<tr>
<td>CHAPTER TWO: THEORY AND RESEARCH RELATED TO THE PRESENT STUDY</td>
</tr>
<tr>
<td>2.1. The notion of error ....................................................</td>
</tr>
<tr>
<td>2.1.1. Error as a powerful notion in L2 classrooms and L2 assessment</td>
</tr>
<tr>
<td>2.1.2. Error as a useful construct for SLA research ..............</td>
</tr>
<tr>
<td>2.1.3. Error and learner error research ...............................</td>
</tr>
<tr>
<td>2.2. Error evaluation and error gravity ..................................</td>
</tr>
<tr>
<td>2.2.1. The tradition of error evaluation research ..................</td>
</tr>
<tr>
<td>2.2.2. Some key findings in error evaluation research ............</td>
</tr>
<tr>
<td>2.2.3. Various error perception indices ...............................</td>
</tr>
<tr>
<td>2.2.4. Error gravity and native speakerhood .........................</td>
</tr>
<tr>
<td>2.2.5. Error gravity and judges’ age ....................................</td>
</tr>
<tr>
<td>2.2.6. A survey of selected error evaluation studies ..............</td>
</tr>
<tr>
<td>2.3. Proposing a constructivist approach to error evaluation ........</td>
</tr>
<tr>
<td>2.3.1. The limitation of a conventional approach to error evaluation</td>
</tr>
<tr>
<td>2.3.2. Error and context ...................................................</td>
</tr>
<tr>
<td>2.3.3. Context of error production and context of error perception</td>
</tr>
<tr>
<td>2.3.4. Expert speakers’ context-sensitive norm .....................</td>
</tr>
<tr>
<td>2.3.5. Establishing a constructivist perspective for error evaluation</td>
</tr>
<tr>
<td>2.3.6. Error perception: the conventionalist and the constructivist</td>
</tr>
</tbody>
</table>
CHAPTER THREE: APPROACH AND DESIGN OF THE STUDY

3.1. The phenomenological approach ......................................................... p. 48
   3.1.1. What is a phenomenological approach? 48
   3.1.2. Justifying a phenomenological approach for this study 50
   3.1.3. Phenomenological research methods 51

3.2. Methodological issues ........................................................................ p. 60
   3.2.1. Data collection methods 60
   3.2.2. Reliability and phenomenological research methods 61
   3.2.3. Validity and phenomenological research methods 62
   3.2.4. The form of the study 65

3.3. The context of data collection ............................................................... p. 66
   3.3.1. Justifying the selected context of the study 67
   3.3.2. The participants 68
   3.3.3. Data types 71

3.4. Data collection procedures and data analyses ...................................... p. 75
   3.4.1. Classroom observations 75
   3.4.2. Marked written work 76
   3.4.3. Recalls of specific serious errors 78
   3.4.4. General descriptions of instructors’ beliefs, assumptions and knowledge 81
   3.4.5. Data treatment 83
   3.4.6. Analyses of descriptions of instructors’ beliefs, assumptions and knowledge 85
   3.4.7. Analyses of the instructors’ recalls of specific errors 88
   3.4.8. Reliability and validity of the theme analysis 94

3.5. Summing up ....................................................................................... p. 95

CHAPTER FOUR: FINDINGS

4.1. Research question No. 1 ...................................................................... p. 98
   4.1.1. Edward 98
   4.1.2. Kelly 100
   4.1.3. Marge 102
   4.1.4. Anna 104
   4.1.5. Summing-up 105
4.2. Research question No. 2 .................................................................................. p. 108
  4.2.1. The marking scheme 108
  4.2.2. The specific purpose of assessment and error perceptions 110
  4.2.3. Genre as context acting on error perception 113
  4.2.4. Change of percepts resulting from repeated engagement with scripts 116
  4.2.5. Summing-up 118

4.3. Research question No. 3 .................................................................................. p. 118
  4.3.1. What kind of errors did the instructors consider serious? 118
  4.3.2. Comprehensibility as a criterion for evaluating error gravity 121
  4.3.3. Irritation as a criterion for evaluating minor errors 124
  4.3.4. Frequency as a criterion for evaluating anomaly gravity 126
  4.3.5. Teaching advice as a criterion for anomaly gravity 128
  4.3.6. Learner as context and error gravity 130
  4.3.7. Genre as context acting on errors 134
  4.3.8. Summing-up 142

4.4. Research question No. 4 .................................................................................. p. 143
  4.4.1. Content anomalies 143
  4.4.2. Grammatical errors 155
  4.4.3. Summing-up 160

CHAPTER FIVE: AN INTERPRETATION OF THE INSTRUCTORS’ PERCEPTIONS

5.1. Summary of instructors’ perceptions of anomalies ......................................... p. 161
    5.1.1. Learner as context 161
    5.1.2. Assessment purpose as context 162
    5.1.3. Genre as context 162
    5.1.4. Teacher as context 163

5.2. Re-defining the scope of error ....................................................................... p. 166

5.3. Error and genre as context ........................................................................... p. 170

5.4. Error evaluation and the contexts of teaching ............................................. p. 174

5.5. The activities of assessment and error evaluation ........................................ p. 178

5.6. The affordances of the notion of error ......................................................... p. 183

5.7. The constraints of the assessment scheme ................................................. p. 185

5.8. Theoretical relevance of the study .............................................................. p. 188

5.9. Practical implications ............................................................................... p. 192

5.10. Limitations of the study ............................................................................ p. 197

REFERENCES ...................................................................................................... p. 200

APPENDICES ..................................................................................................... p.219
**LIST OF TABLES**

Table 2.1. Error classification system in James (1998) ........................................ p. 14
Table 2.2. Summaries of some major empirical research on error evaluation .......... p. 24
Table 3.1. Empirical phenomenological research ...................................................... p. 53
Table 3.2. Hermeneutics ............................................................................................ p. 54
Table 3.3. Heuristic research ..................................................................................... p. 55
Table 3.4. Transcendental phenomenology .............................................................. p. 56
Table 3.5. Summary of teacher participant’s consent to participate in the study ... p. 69
Table 3.6. Profile of the teacher participants in the study ......................................... p. 70
Table 3.7. Summary of data types, methods of collection, treatment, analysis and their uses ................................................................. p. 74
Table 3.8. Summary of classroom observations by instructor ................................ p. 76
Table 3.9. Samples of cue questions eliciting instructors’ instructors’ recalls of specific errors ................................................................. p. 79
Table 3.10. Frequency of assessed scripts and errors recalled in the instructor interviews ................................................................. p. 81
Table 3.11. Sample questions used to elicit descriptions of instructors’ beliefs, assumptions and knowledge about language learning and their teaching ................................................................. p. 82
Table 3.12. Error type labels generated based on the comments by the instructors on specific errors ................................................................. p. 92
Table 4.1. Summary of views of the instructor participants towards scope of errors, mistakes, slips and content and organization anomalies ................................................................. p. 107
Table 4.2. Frequencies and percentages of anomaly types considered serious by the instructors by anomaly category ................................................................. p. 119

- viii -
Table 4.3. Ranked order of serious anomaly categories recalled by the Instructors p. 120

Table 4.4. Examples of error quantification in evaluating anomaly gravity by instructor p. 126

Table 4.5. Frequency count of the use of restrictive relative clauses by Edward’s students p. 140

Table 4.6. Frequency count and percentages of the uses of the Know-your-reader notion to explain anomaly gravity by instructor p. 147
LIST OF FIGURES

Figure 5.1. The activity systems of assessment and error evaluation .................................. p. 182
Figure 5.2. The socioliteracies model proposed in Johns (1997) ........................................ p. 195
## LIST OF APPENDICES

<table>
<thead>
<tr>
<th>Appendix</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Letter to a potential teacher participant for the study</td>
<td>219</td>
</tr>
<tr>
<td>B</td>
<td>Letter to a potential administrator/co-ordinator of an ESL program</td>
<td>221</td>
</tr>
<tr>
<td>C</td>
<td>Letter of a potential ESL learner participating in the study</td>
<td>223</td>
</tr>
<tr>
<td>D</td>
<td>Summary of students' written texts collected for this study</td>
<td>225</td>
</tr>
<tr>
<td>E</td>
<td>The assessment scheme the four instructors were using</td>
<td>226</td>
</tr>
<tr>
<td>F</td>
<td>A sample of a prompt for an assignment task</td>
<td>227</td>
</tr>
<tr>
<td>G</td>
<td>A sample page of a student's written work</td>
<td>228</td>
</tr>
<tr>
<td>H</td>
<td>The transcription convention used in this study</td>
<td>229</td>
</tr>
<tr>
<td>I</td>
<td>A teacher's feedback to the transcription of her interviews</td>
<td>230</td>
</tr>
<tr>
<td>J</td>
<td>Sample pages of my interpretation of a teacher's percepts and my follow-up questions for the teacher to respond to</td>
<td>231</td>
</tr>
<tr>
<td>K</td>
<td>A sample page of a teacher’s responses to my interpretation of her percepts and my follow-up questions</td>
<td>233</td>
</tr>
<tr>
<td>L</td>
<td>Project summary and a list of themes used in the Nvivo 1.0 program</td>
<td>234</td>
</tr>
<tr>
<td>M</td>
<td>The error record sheet Anna’s students all used</td>
<td>237</td>
</tr>
<tr>
<td>N</td>
<td>A prompt for writing an incident report</td>
<td>239</td>
</tr>
<tr>
<td>O</td>
<td>Sample of a prompt for an investigation report used in Anna’s classes</td>
<td>240</td>
</tr>
</tbody>
</table>
CHAPTER ONE

INTRODUCTION

1.1. Where is this coming from?

Researchers come to their inquiries with their own interests and subjectivity (Peshkin 1988). Hence, there is a good reason to begin my present discussion with myself, my own personal experience to explain my biases (Hunt 1987). I have been a learner of English as a second language (ESL) since my primary school days. I made errors and I still make them. For the past twenty some years, I have been teaching English as a second language in an Asian city in which English is a second language. I therefore meet with errors and deal routinely with them.

One experience of encountering errors I can recall happened in my first secondary school year when I had to write my first composition in English. Before this in primary school, we were writing only sentences in English and even that I could not do well. The topic given this time was "My family". It was an ordeal to try to put things together to make it look like an essay. It was our first year in middle school. I had a new English teacher and it was my first English composition. I took it seriously and tried hard.

A week later, my English teacher came to class with the marked compositions. He picked up the first copy from the pile and said he was going to read to the class "the best piece". The class went quiet, for from his tone, one could tell that he was being sarcastic about it. Then he started reading the essay. Very soon, I realized that it was my essay he was reading aloud. However, the teacher was not just reading aloud; he was picking out the errors and commenting on them mockingly: "What a great writer! He could not even tell the difference between numbers and members. "There are six numbers in my family!" The class burst into laughs and I could feel my face burning. I was very upset but I was
even more worried that my classmates might soon find out who the writer was. The teacher stopped reading from time to time to make his mocking comments, but all the time he never looked at me as if he did not know who the writer was. When the composition book finally came back to me, I saw my writing all covered with red ink. At that point, my tears began to roll.

I guess even today I still do not understand why my teacher had to do that to me, or for that matter, to any student in that way. What I can be sure of is that it was one of the longest lessons in my life, and that was the most horrid, intense experience I have ever had with errors in English. When I became a teacher, I knew that I would never treat any of my students as I was treated, and although now I think that a balanced treatment of meaning and accuracy should be adjusted according to the teaching situation and the objectives of the teaching, the question of what language error is and how it should be dealt with in teaching and in a language curriculum, I maintain, is an important question.

While I report my most intense personal experience with error I also note that my experience with this thesis study signifies a new learning stage for me, as a language teacher and as a research student. I would like to document this experience. My initial interest in research on language errors actually has to do with a paper I submitted to Professor Birgit Harley for the Aspects of Second Language Learning course I took in the Spring of 1996. For that course I completed a small study on the English collocational knowledge of some young Cantonese-speaking immigrants. Linguistic collocation is defined as "words that often occur together" (Nation 1990, p. 33) as, for example, sunny collocates with day and disposition, and too collocates with much and late. The study involved comparing the collocational knowledge of this group of ESL young learners with that of some native-speaking Canadian and American adults. The collocations used in this study were phrases formed by a set of aesthetic adjectives of similar meaning (e.g.,
beautiiful, good-looking, pretty, handsome, attractive, lovely} pre-modifying a set of animate or inanimate nouns (e.g., flower, day, woman, man, voice, picture).

The collocation study required that in order to evaluate the ESL children's responses to my collocation questions, I first had to establish a set of norms for native speakers of English. While the test on ESL learners pointed to some common learner errors (e.g., a good-looking flower, a good-looking picture, an attractive day), the test on the adult native-speakers of English actually showed very mixed responses to the acceptability of many of the phrases I used in the study, e.g., an attractive flower, a pretty day, a beautiful man, a handsome voice. During the time of the study, I did not see that that the divergent responses of the native speaking adults themselves was itself an interesting phenomenon; in the study I treated that as a methodological hurdle and got round the problem by using a scoring method to determine the native speaking adults' acceptability level. When I showed this paper to my then faculty advisor Professor Alister Cumming for comments, he drew my attention to the mixed responses of the native speakers. Alister also suggested that I study the notion of error to see how people actually look at error as a phenomenon. I appreciated very much this suggestion and started working on a proposal for this thesis study.

The course Linguistic Analysis in Educational Settings I took from Professor Gordon Wells also influenced me in the formulation of an interactionist perspective on the issue of error evaluation. I report the use of this perspective in Chapter Two in detail. But the idea that I should consider an interactionist perspective for error evaluation actually goes back to an observation Professor Merrill Swain made in a class in 1995, when I was attending her course, Second Language Classroom Research. Merrill said in the class that celebrities like former U.S. Secretary of State Henry Kissinger's and Canadian Prime Minister Jean Chrétien's slips and errors in English in public are often given great
allowance by their audience, partly because they are speakers of ESL but also partly because they have certain social status and command certain respect. When I was working on a draft of the proposal for this study in 1998, what Merrill had said came back to me and actually prompted me to believe that error evaluation is basically a context-sensitive activity, acted upon by contextual elements such as the speaker or the writer, the milieu in which the errors are perceived, and so forth.

1.2 The organization of this thesis

Determining the organization of this thesis was an interesting experience, a decision I struggled with for some time. I considered adopting the genre of a narrative for this thesis. I have read some interesting phenomenological research and theses (e.g., Freire, 1997; Van Manen, 1982) written in narratives, and I suppose that a narrative genre is compatible with a phenomenological approach, which emphasizes the processes of coming to an understanding.

Eventually, however, I have made my decision based on a selfish reason. I feel that a narrative, reflective, process-focused thesis, although it may enjoy the advantage of providing a thick, context-rich description, may not be particularly easy for some people to write or to read — people who, like myself, have not a lot of literary exposure. Hence, I have finally decided to adopt a conventional format of a research report, as I feel more confident writing in it. In retrospect, the format that I have chosen seems to fit well with the kind of research questions I had asked and my findings. So a question for myself is: How much can form (genre) influence the content of my writing if most of the time I tend to believe that it is the content that determines the form? If I had chosen a different genre to write this thesis, would I have come to the same findings? Nevertheless, the structure
of my thesis is as follows: Introduction, Theory and research related to the present study, Approach and design of the study, Findings: Descriptions of teachers’ perceptions, and Discussion and implications: An interpretation of the teachers’ perceptions.

In Chapter Two I argue for the importance of the notion of error, report the tradition of error evaluation research, and present a review of the literature to arrive at a theoretical base for the approach I take in this study; there I also argue for the establishment of my research questions. In Chapter Three, I defend the research approach I have chosen for this study, discuss the various methodological issues I have considered as well as other considerations related to the data collection procedures, and report the actual procedures of data collection and analyses I performed. In Chapter Four I present my findings in a structure that answers one by one the four research questions I posed. Finally, in Chapter Five, I present my own interpretation for my findings. Before I move on to Chapter Two, I would like to note that all the names of teacher and student participants in this thesis study are pseudonyms. This means that any names appearing in the data of this study are false names.
CHAPTER TWO

THEORY AND RESEARCH RELATED TO THE PRESENT STUDY

In this chapter I discuss the published literature related to the present study. I begin with a discussion on the important role that error has played in second language (L2) teaching, L2 assessment and theory development in second language acquisition (SLA). I argue that in spite of its important role and considerable inquiry already devoted to it, there is the need for a further examination of the concept of error. I report the understandings of error evaluation from conventional studies, pointing out that although there is an increasing awareness of the linguistic contexts errors appear in, very few studies on error and error evaluation have adopted a naturalistic research approach to examine error and error gravity. I also present a survey of selected empirical error evaluation studies, deriving implications from them for the present study. Finally, I draw the conclusion that a social-interactionist approach to language error evaluation, taking into consideration the context of the error production and the context of error perception, reveals better an understanding of the meanings of error than do previous orientations to this issue.

2.1 The notion of error

2.1.1. Error as a powerful notion in L2 classrooms and L2 assessment

In this section, I wish to establish two points. First, the notion of error playing an important role in language teaching has been a subject for many studies on classroom and classroom interactions. Second, it is also an important notion in L2 assessment, but our knowledge of how error works in the process of language assessment is still very limited.
We know that children make errors in first language acquisition (Snow, 1972, 1977) and many adult native speakers' language production is often not totally free from the most obvious errors (Fromkin, 1980). Hence, it really seems strange that native speaking children's errors are considered "transitional" and that adult native speakers' errors are treated as "slips of the tongue", given "new meanings" in psychological jargon, and discussed in depth (Freud, 1924/1969), whereas L2 errors could at one time be even equated to sin. During the heydays of behaviourism in L2 teaching, it was said that "[L]ike sin, error is to be avoided and its influence overcome" (Brooks, 1964, p.58). However undesirable errors may seem in the learning of language in school, they are part of the everyday reality in an L2 classroom. 

Traditionally, language instructors have always given considerable attention to the errors students make. In many second language or foreign language classrooms, errors are identified, explained and corrected. Because of these intentional endeavours, errors in students' oral or written production often become either the target of teachers' evaluation, the topic of teacher talk, or the focus of teaching activities. Some classroom discourse analysts (Chaudron, 1977; Sinclair & Coulthard, 1975), for example, have found that quite a large portion of classroom talk has to do with corrective feedback on learner errors. Some researchers pursuing Sinclair and Coulthard's (1975) Initiation-Response-Follow-up model to analyse classroom discourse have shown that an error can be the cause of a teacher's Initiation move in an exchange (e.g., "Say it correctly"); it can also become the topic of a Follow-up/Evaluation Move (e.g., No, we don't eat sheep, we eat mutton or lamb.) (Sinclair & Brazil, 1982, p. 45). In sum, in classroom interaction errors serve the learners as elicitors of helpful feedback and the teachers as a source of teaching (Chaudron, 1988, p.134) and insight into the meanings that their pupils are making (Nicholls & Wells, 1985, p.18).

When assessing learners' work in subjects such as science, physics, history, or
social studies, what is usually focussed upon and evaluated is the positive learner production or performance; error remains a notion in the background and seldom appears in assessment or marking schemes.

The assessment of mathematics requires some special consideration for the obvious reason that mathematics can be considered a language — the "language" of modern science (Olson, 1994); it is a rule-based discipline that highly values precision and accuracy. In many mathematics classroom assessment exercises, quite a portion of the credit is given to valid operations and inferences, not just the correct answers. These seem justified practices. This is because in a formal mathematical system (e.g. Euclidean geometry, an algebraic Ring) the premises are usually assumed or given. In such a formal system, when the operations or the inferences are logically carried out, a correct answer (conclusion) is always deductively (and hence unavoidably) arrived at. Hence, in a mathematics test, the premises are always assumed to be true, and once a learner's reasoning is correct, there should not be any errors, so to speak. This justifies an emphasis on the correctness in the logic and the reasoning.

In many mathematics classroom assessments, errors certainly provide evidence and bases for teachers to understand individual students' uses of reasoning rules and conceptual understanding. By attending to the kind of errors students made, teachers are able to evaluate their students' learning and their own teaching. While this is the case, it should also be noted that the classroom assessment schemes most mathematics teachers use often do not explicitly or specifically mention error, i.e., the assessment schemes teachers use do not draw attention to error, error type, error frequency or error gravity.

Similar to assessments in many other school subjects, in many large-scale mathematics assessment practices, the notion of error is obscure and remains a hidden notion. The mathematics component in the Scholastic Achievement Test (SAT) provided
by the Educational Testing Service, USA, for example, contains only multiple-choice items. The focus of assessing these items is often the correct answer (option); the logic and reasoning involved is ignored in the scoring. The assessment of secondary mathematics in the Hong Kong Certificate of Education Examination (HKCEE) provided by the Hong Kong Examinations Authority (HKEA) contains components that use non-multiple-choice problems. (This is a subject examination all secondary school students have to take on exit from secondary education in Hong Kong.) According to C. K. Lai, Mathematics Subject Officer at the HKEA, whom I interviewed on January 18th, 2000 for this thesis study, the assignment of scores in a typical assessment scheme for HKCEE Mathematics is usually based on three criteria: (a) Accuracy, (b) methods and operations and (c) proof-argument-deduction. Although the evidence for making judgements on a test-taker’s performance is based on “correct” performance (and hence the errors) in these three domains, the assignment of scores is never based on error, error type or seriousness of specific errors. In other words, errors are only interpreted by a rater as evidence for the absence of “correct” or “positive” performance, and test-takers’ errors made in one way or the other do not affect the scoring in this particular examination.

One observation on the distinction of errors and “careless slips” in the classroom, however, provides a hint to the relativity of the notion of error in mathematics assessment. Consider this example. In a Primary Three mathematics class where addition and subtraction of numbers are being taught, a pupil’s failure to compute a simple sum such as “7 + 19 = ?” would often be perceived by the teacher to be a genuine error and perhaps subsequently given sufficient pedagogical attention. But consider the very same miscalculation coming from a university student in his or her advanced calculus test. In this situation, the teacher might just consider this miscalculation a careless slip. Marks may be deducted but the teacher would not treat the miscalculation as a genuine error since this is not taken to be a display of conceptual misunderstanding. This example suggests that error perception in school mathematics assessment is basically context-dependent and that
a learner mathematical operation violating a rule is not necessarily considered an error.

To sum up the role of assessment in many school subjects, one very possible reason for the obscure, undefined role error plays in many educational assessment practices is that teaching syllabuses or curriculum guidelines often specify only what learners are expected to learn (content specifications) and in what way they are expected to perform (performance criteria), but not how and what errors learners are expected (or not expected) to make or in what ways they are expected to make them.

Unlike the assessment of many other subjects, in L2 assessment, however, the notion of error is often foregrounded. For example, in the scoring guide for Hamp-Lyons’ (1991) Michigan English Language Assessment Battery (MELAB) scheme, under the trait “Mechanics” we see “spelling and punctuation errors” as descriptors, and under the trait “Sentence Structure” one sees quantification of error frequency such as “very frequent errors” or “occasional errors” (p.269). Other studies on language assessment have also substantiated the important role that error perception plays in rating language performance in spite of the frequently proclaimed importance of communicative value and meaning in many assessment schemes. Cumming’s (1990a) protocol study into the process of rating academic L2 writing, for example, reported that "classifying errors" accounted for large proportions (20%) of the total number of behaviours of expert raters. Other studies (Harris, 1977; Kline, 1976) also showed that errors were a much greater determining factor for scoring language production than teachers or curriculum planners would like to admit. In spite of the recognition of the important role error plays in assessment, very few published studies to date have looked systematically at the role of error in L2 assessment, nor tried to reveal its relationships with assessment. Cumming’s (1990a) protocol study of raters, for example, though it identified a high percentage of “classifying errors” behaviour by the raters, did not discuss the role of error evaluation in the assessment activities.
2.1.2. Error as a useful construct for SLA research

I have observed that error has served as a useful, powerful notion in the practice of L2 teaching and assessment. However, the usefulness of the notion of error is not limited to educational practice. In L2 education research, and especially SLA research, the construct of error has also shaped the development of L2 learning theories.

One line of SLA research focusing on the L2 classroom has afforded direct evidence for a relationship between language development and teachers' treatment of errors. Studies in and outside L2 classrooms have recently advanced the idea that learners who received explicit corrective feedback on the errors they made outperformed in general those who were given no or minimal corrective feedback (Aljaafreh & Lantolf, 1994; Carroll, Roberge & Swain, 1992; Carroll & Swain, 1993; Lightbown & Spada, 1990; Spada & Lightbown, 1993; Tomasello & Herron, 1989). Although the concept of error involved in these studies is usually restricted to grammatical inaccuracy, findings in this line of research no doubt point to the importance of error-related studies and the significance of the construct of error for SLA theory development.

Another group of SLA researchers who have exploited the construct of error for theory construction have maintained that there is a universal development sequence in L2 learning. These researchers have argued for a link between learner language (including the erroneous morpho-syntactic structures learners produce) and the stages of language development of the learner in relation to these specific syntactic structures. They maintained that the acquisition of certain specific L2 morpho-syntactic structures (e.g., English copula “be”, English question forms) proceeds in a stage-wise, systematic manner (Dulay, Burt & Krashen, 1982; Mackey, 1995; Pienemann & Johnston, 1987; Ravem, 1968; Spada & Lightbown, 1993; Wode, 1976). An implication for understanding learner errors from this research is that by judging learners’ erroneous forms of a specific morpho-
syntactic structure we are teaching, we would be able to tell, in Corder’s (1967/1974) words, “how far towards the goal the learner has progressed and what remains for the learner to learn” (p.25).

2.1.3. Error and learner error research

The line of SLA research which has directly and intensively studied error is learner error research (Ellis, 1994, p. 47). This research began with error analysis: collecting, identifying, describing and explaining errors. The proclaimed aim of this research is both pedagogical and theoretical. Corder (1967/1974) argued that by conducting systematic error analysis, the teacher can determine how far towards the goal the learner has progressed and what remains to be learned. The learner, on the other hand, can regard the making of errors as a hypothesis testing device to improve language development. To researchers like Corder (1967/1974) himself, however, errors are worthy of attention because “they give evidence of the system of language the learner is using” (p.25). Corder’s work has exerted great impact on three important branches of studies in SLA in addition to learner error research: error analysis (Jain, 1974; Odlin, 1989; Richards, 1971; Wardhaugh, 1970), interlanguage studies (Andersen, 1984; Selinker, 1972, 1995) and learner process and strategy research (Færch & Kasper, 1983; Oxford, 1989; Tarone, 1980; Wenden, 1986). The literature survey in this chapter does not cover the last two areas, interlanguage studies and learner process and strategy research, because these two areas focus on the language learner, rather than error as a phenomenon, which is the focus of this thesis study.

Learner error researchers in their attempts to describe and explain errors have expanded our understanding not so much on what error is but rather about the variety of errors and the complexity of inferred causes of errors. These studies have proposed a wide
variety of different descriptive systems. Corder (1967/74) suggested three main types of errors: *pre-systematic errors*, when a learner is unaware of the rules concerned, *systematic errors*, when a learner has discovered a wrong rule and *post-systematic errors*, when the learner has learned the correct rule but used it inconsistently. Corder (Ibid.) also proposed *covert errors* (linguistically correct but contextually inappropriate) and *overt errors* (appropriate intent but linguistically incorrect). Richards (1971) discussed *interference errors* and *transfer errors* and *intralingual, extralingual and developmental errors*. Burt & Kiparsky (1975) suggested *global* and *local errors* and argued that *global errors*, which are errors affecting more than one constituent in the sentence, are more serious than *local errors*, which affect only one constituent. Svartvik (1973) and Stenson (1975) offered examples for *induced errors*, which are errors caused by the kind of instruction learners have received. Dulay, Burt and Krashen (1982) reviewed a number of prior studies and described three different systems of errors: errors according to (1) linguistic categories: *morphology* and *syntax*, (2) surface strategy taxonomy: *omission, misordering, misinformation, addition, double markings, regularization*, etc. and (3) *developmental errors, interlingual errors, ambiguous errors* and *other errors*.

One more recently proposed error descriptive system which is worthy of attention is James (1998). In this book entitled *Errors in language learning and use: Exploring error analysis*, which is to this date the most thorough discussion on errors and error related research, James outlined an error descriptive system for both spoken and written errors, which I present in Table 2.1.
Table 2.1. Error classification system in James (1998)

<table>
<thead>
<tr>
<th>Substance errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Misspellings</td>
</tr>
<tr>
<td>Punctuation errors</td>
</tr>
<tr>
<td>Typographic errors</td>
</tr>
<tr>
<td>Dyslexic errors</td>
</tr>
<tr>
<td>Confusibles</td>
</tr>
<tr>
<td>Misspellings proper</td>
</tr>
<tr>
<td>Mispeluncliations</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Text errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lexical errors</td>
</tr>
<tr>
<td>Formal errors of lexis</td>
</tr>
<tr>
<td>Formal misselection</td>
</tr>
<tr>
<td>Misformations</td>
</tr>
<tr>
<td>Distortions</td>
</tr>
<tr>
<td>Semantic errors in lexis</td>
</tr>
<tr>
<td>Confusion of sense relations</td>
</tr>
<tr>
<td>Collocational errors</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grammar errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Morphology errors</td>
</tr>
<tr>
<td>Syntax errors</td>
</tr>
<tr>
<td>Phrase structure errors</td>
</tr>
<tr>
<td>Clause errors</td>
</tr>
<tr>
<td>Sentence errors</td>
</tr>
<tr>
<td>Intersentence errors (cohesion)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Discourse errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coherence</td>
</tr>
<tr>
<td>Topical coherence</td>
</tr>
<tr>
<td>Relational coherence</td>
</tr>
<tr>
<td>Sequential coherence</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pragmatic errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Taboos</td>
</tr>
<tr>
<td>Size of the imposition</td>
</tr>
<tr>
<td>Values</td>
</tr>
<tr>
<td>Power and social distance</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Receptive errors</th>
</tr>
</thead>
</table>
The classification system given here is comprehensive, but as James (Ibid.) himself admitted (p. 206), his discussion on error gravity was restricted to formal errors only, and the examples he gave for discourse errors or pragmatic errors in Chapter Five of his book were mostly spoken errors; James provided no thorough discussion on text or genre errors and the gravity of these errors.

Studies using descriptive systems of errors have met with some success in identifying errors and explaining the causes of learner errors. It is now believed that, for example, a large number of learner errors are in fact intralingual rather than transfer errors (Ellis, 1994, p. 62; Galloway, 1980); on the other hand, transfer errors are more common in learners' phonology and lexis than in their grammar and more common in adult learners than in child learners (Ellis, Ibid.).

One major criticism on earlier learner error research, however, is the element of arbitrariness in error classifications and error explanations. Many of the proposed systems contain overlapping categories. In Dulay, Burt and Krashen (1982), ambiguous errors are those errors that can be considered at the same time as developmental errors and interlingual errors; and the category of other errors is in fact “a grab bag” (p.172). Ellis (1994) pointed out that an inherent problem with many classification systems is that very often it is difficult to tell whether an error is a transfer error or an intralingual error if a particular distinction does not exist in the L1. For example, the error of using ‘make’ instead of ‘do’ by Italian learners can be considered either because the ‘make/do” distinction is non-existent in Italian (Ellis, 1994, p. 59).

The more serious weakness of earlier error research, however, is that some error analyses were based on errors totally stripped from their contexts such as learner background, learner’s L1, task type, etc. Take an error from Richards (1971) as an
example. Richards (Ibid.) labelled “I am interesting in that.” as an error of *distribution of verb groups*; this is not a classification with which many others would agree. As James (1998, p. 54) observed, an erroneous form in a target language can be made by two different learners from two different L1s (Spanish and Russian) for two very different reasons, i.e., resulting from very different cognitive processes. Richards (1971) reported that the errors he analysed were collected from a variety of L1 backgrounds: Japanese, Chinese, Burmese, French, Czech, Polish, Tagalog, and Maori, Maltese. It was not obvious either whether these errors were produced by learners of same or similar background in the same or comparable contexts, e.g., the same writing task. Finally, Schachter and Celce-Murcia (1977) and Schachter (1974) also criticized error analysis because it cannot show why a learner did not choose to use a certain language feature because the learner may have adopted an avoidance strategy.

### 2.2. Error evaluation and error gravity

2.2.1. The tradition of error evaluation research

The developing tradition of error evaluation research in SLA, with which the present thesis study is aligned, can be seen as a supplement to, as well as a continuation of, earlier learner error studies. The learner error studies in the sixties and the seventies focussed upon the language of the non-native speaking learner. However, "by their very nature ... these studies look at only one half of the communicative event" (Khalil, 1985, p. 336). In the last two decades, increased interest in communicative language teaching has prompted researchers to try to capture “the other half of the communication” by evaluating errors from the perspective of native speaking readers or hearers. This body of research has been labelled "error gravity", “error perception”, “error judgement” or “error

Eisenstein (1983) defined the scope of error gravity studies: "The literature considers perceptions of the relative seriousness of particular errors, often referred to as 'error gravity', along with more holistic characteristics including intelligibility, acceptability, and listener irritation." (p. 159) Rifkin and Roberts (1995) summarized the assumption, value and purpose of this line of research:

This research has assumed that some linguistic errors are more serious than others in terms of disrupting a NS's comprehension of a nonnative speaker's (NN's) message and that these error types can be identified. The apparent goal has been to establish hierarchies of L2 error types so that L2 teachers might focus on areas of language production judged by native speakers to be most disruptive to communication. (p. 512)

The study of error criteria and error evaluation is a theoretical and an empirical pursuit although the contribution these studies have attempted to make, as Rifkin and Roberts (Ibid.) suggested, is in teaching — the consideration of pedagogical priorities in the development of communicative competence. The past two decades saw this literature attempt to relate error gravity to a number of other constructs: types of errors (Burt, 1975; Johansson, 1978; Richards, 1971; Santos, 1987), acceptability (Chastain, 1980; Vann, Meyer & Lorenz, 1984), comprehensibility and intelligibility (Johansson, 1978; Khalil, 1985; Olsson, 1977; Varonis & Gass, 1982), grammaticality (Ervin, 1979;

¹ There are subtle differences among terms like “evaluation”, “judgement” and “perception”. The term “evaluation” tends to assume a purpose; the term “judgement” carries laden values and likes and dislikes, and the term “perception” is a much more inclusive term. In this thesis study, I will only use the two terms “perception” and “evaluation”, and I use them interchangeably.
Galloway, 1980), irritation level (Piazza, 1980; Santos, 1987), markedness of the language (Santos, 1987), and linguistic context (Burt, 1975; Lennon, 1991; Taylor, 1986). In the following section, I will discuss some key findings from this research.

2.2.2. Some key findings in error evaluation research

Our present knowledge about L2 errors is rather limited and uncertain in spite of much previous research. In particular, I will show that:

1. There is evidence to indicate that some errors are perceived (by native or non-native speaking judges) to be more serious than others.

2. There is no conclusive evidence for a universal hierarchy of error.

3. There is no conclusive evidence for a relationship between irritation level and error gravity, but there is some evidence to show that error gravity is related to judges' age or gender.

4. Although error gravity (seriousness), comprehensibility, acceptability, irritation, intelligibility and grammaticality all seem to be distinct constructs, the relationships among them are obscure and difficult to determine.

Nearly every error evaluation study reviewed here provided some evidence to support the assumption that some errors are considered more serious than others. In many studies, judges responded, to a statistically significant extent, differently to errors or different error types in terms of gravity, irritation or acceptability, regardless of whether the instrument used was matched guise (Politzer, 1978), Likert scale (Albrechtsen, Henriksen, & Færch, 1980), or paired sentence comparison (Delisle, 1982; Magnan, 1983). However, evidence for a "universal hierarchy of errors" is inconclusive. In the following I will give two groups of studies as examples, one in ESL and the other in
German as an L2. I do this to show that the inconclusiveness of evidence for a universal error hierarchy is manifested regardless of the target language under study.

McCretton and Rider (1993) identified in ESL the order of seriousness as follows: Lexis (most serious), spelling, negation, word order, prepositions, verb forms, concord (least serious). Although the observation that lexical errors seem the most serious found support in Santos's (1988) ESL study, a study by Vann, Meyer and Lorenz (1984) identified a different order of seriousness for ESL errors: Word order errors were the most serious, followed by It-deletion, tense, relative clause and then word choice. On the other hand, James (1977) identified a hierarchy as follows: transformations, tense, concord, case, negation, articles, word order.

In Delisle (1982), 180 high school students of 10 to 17 year-old in Germany considered gender error in L2 German the most serious, followed by verb morphology, word order, vocabulary, case ending and spelling, and in that order. In Politzer's (1978) German as L2 error study, however, 146 German native-speaking teenagers found vocabulary the most serious, followed in order by verb morphology, word order, phonology, then case ending.

2.2.3. Various error perception indices

Inconclusiveness is also associated with research findings regarding the concepts of irritation level and error gravity. Albrechtsen, Henriksen and Færch's (1980) ESL study found that irritation is related to the number of errors, "regardless of error type or of other linguistic aspects of the text" (p. 394). This is contrary to Santos' (1988) observation that her 178 professor judges felt most irritated by a Korean student's "double negation" error (i.e., This particular error is: *"They wouldn't get nowhere unless they used a
translator.”). This particular error caused high irritation among the judges.

The other inconclusive observation from this research concerns perceptual notions such as comprehensibility, irritation, grammaticality, and acceptability. The way that judges responded to different perception variables in the same studies suggest that these notions are relatively distinct constructs, yet quite a few studies also suggest that these indices tended to be related sometimes, while other times they appeared totally unrelated (Albrechtsen, Henriksen & Fræch, 1980; Chastain, 1980; Santos, 1988). Chastain (1980) concluded that many errors are unacceptable for reasons other than frustration or irritation and that errors are often comprehensible but unacceptable. Albrechtsen et al. (1980) failed to find connections between comprehensibility and linguistic correctness in their study on spoken communication strategies and grammaticality, and as previously reported, they argued that irritation is related to the number of errors in a unit, not the error types. Munro & Derwing (1995) failed to find relationships between intelligibility, comprehensibility and phonetic error counts though they did find that phonetic error counts correlated with the degree of foreign accentedness. Results from Albrechtsen et al.’s and Munro and Derwing’s studies tend to contradict those from Varonis and Gass (1982), who concluded that “for some speakers grammaticality has a large effect on comprehensibility, while for others, the effect is slight.” (p. 127)

2.2.4. Error gravity and native speakerhood

Some researchers have been interested in whether native speakers and non-native speakers judge errors differently. Galloway’s (1980) native speakers reacted differently than did his non-native speakers to learners’ Spanish tense confusion errors. This particular error gave more communication trouble to the non-native speakers than to the native speakers. This result does not seem to contradict that of Santos (1988), who found
that her "non-native speaking professors were more severe in their judgements than native speakers" (p. 69). However, Birdsong and Kassen (1988) found that their native speaking teachers of French were the most severe in error evaluation, followed by non-native speaking teachers, more proficient learners and less proficient learners, and in that order. Hence, the most we can say from these studies is that the speakerhood (a term suggested in James, 1998, p. 49) of the judges does seem to affect error perception.

2.2.5. Error gravity and judges’ age

Although findings on native speakerhood are not conclusive, there are congruent indications in the literature that error perception in general is related to a judge’s age. Magnan’s (1983) spoken error evaluation study found that her native speaking French adult judges were more bothered by verb and pronoun errors while her younger student judges found gender-related errors most irritating. Santos’ (1988) study discovered that older professors were less irritated by the language in an essay (as opposed to the content of the essay) than younger professors were. Delisle’s (1982) written German error perception study found that certain errors including gender error became less important with an increase in age among Delisle’s 10 to 17 year-old high school student judges. Also, in Vann, Meyer and Lorenz’s (1984) ESL written error study, when the 164 professor judges were categorized into three groups — (a) 55 and older group, (b) 45 to 54 year-old group, and (c) 35 to 44 year-old group — it was found that the 55 and older group were the most tolerant, followed by the 35 to 44 year old group, with the 45 to 54 year-old group being the least tolerant. Vann et al. did not suggest an explanation for this but concluded that age is a factor in predicting these responses. The results from all these studies tend to suggest that the background characteristics of judges such as age are important factors in error evaluation.
2.2.6. A survey of selected error evaluation studies

In sum, a review of this literature tends to suggest that there are more disparities than consensus in terms of what we now know about error evaluation. In a more recent, extensive review on error evaluation research, Rifkin and Roberts (1995) attributed the inconclusive results of this research primarily to a narrow definition of error and serious methodological flaws. They criticised that some error evaluation studies neglected context, did not account for the order effects of the tasks these studies used, and failed to use common, objective instruments. I find some of Rifkin and Roberts's criticism relevant but I wish to argue for an approach that takes the notion of context one step further, and I will begin with a review of some relevant, major empirical studies.

Table 2.2 gives the summaries of 18 error evaluation research studies presented in reverse chronological order. They were selected for this Table 2.2 based on two criteria. First, they all appeared in major second language education journals; and second, these were all empirical studies. Of these 18 studies, 10 examined spoken errors and 9 of them studied written errors (one study looked at both written and spoken errors). Table 2.2 shows that later error evaluation researchers began to pay greater attention to the linguistic context in which errors appear by abandoning the use of decontextualized sentences as learner language samples. For example, several earlier studies (Chastain, 1980; Delisle, 1982; Khalil, 1985; Piazza, 1980) used sentences as learner language samples, but more recent researchers (Haswell, 1988; Lennon, 1991; Santos, 1987, 1988; Vann, Lorenz, & Meyer, 1991) preferred the use of whole learner written compositions. However, some more recent error evaluation studies (e.g., Vann, Lorenz, & Meyer, 1991) were still using researcher-altered or researcher-edited texts, or even researcher-identified-then-inserted errors, rather than trying to account for errors in whole texts or other natural contexts for the occurrence of learners' errors.
Another weakness of many of the studies in Table 2.2 is that many of even the more recent studies put the judges in an unnatural, experimental setting to carry out the error evaluation, even though in a few of these studies the language samples they used were whole composition texts. For example, the judges in Vann, Lorenz and Meyer (1991) all responded to questionnaires, a setting that is basically artificial in terms of communication of errors. Even Santos' (1987) 40 biological and physical sciences professors marking two whole compositions was an experiment, given that these professors did not usually have to mark papers in this manner in their usual work. Hence, to this date, there has not been any error evaluation study examining errors and error gravity in a truly naturalistic context, using original, whole texts produced by learners and eliciting perception or evaluation of errors in a context where the errors are usually identified and assessed. I will discuss this important issue further.
Table 2.2. Summaries of some major empirical research on error evaluation

<table>
<thead>
<tr>
<th>Study and Year</th>
<th>Topic/Focus of study</th>
<th>Learner variety</th>
<th>Spoken/Written</th>
<th>Unit of error &amp; context</th>
<th>Source of learner errors</th>
<th>Key notions studied</th>
<th>Judges</th>
<th>Highlights of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Munro &amp; Derwing 1995</td>
<td>NSs* judged NNSs' speech samples</td>
<td>English</td>
<td>Spoken</td>
<td>36 speech excerpts</td>
<td>Extracted from NNSs-told stories prompted by cartoon</td>
<td>Perceived intelligibility, acceptedness, errors and comprehensibility</td>
<td>18 NSs</td>
<td>Accentedness correlates with errors, not comprehensibility and intelligibility.</td>
</tr>
<tr>
<td>Schairer 1992</td>
<td>NSs judged NNSs' speech samples: Phonetic analysis</td>
<td>Spanish</td>
<td>Spoken</td>
<td>18 taped speech samples</td>
<td>NSs read speech samples into tapes</td>
<td>Pronunciation, comprehensibility, agreeableness of voice, accent</td>
<td>28 NSs; 11 teachers and 17 non-teachers</td>
<td>No judges’ gender differences; teachers less tolerant than non-teachers; monolingual Spanish NSs more lenient.</td>
</tr>
<tr>
<td>Vann, Lorenz &amp; Meyer 1991</td>
<td>Evaluation of article, spelling and verb form errors.</td>
<td>English</td>
<td>Written</td>
<td>Two compositions accompanied by a questionnaire</td>
<td>Two researcher-altered texts containing inserted errors</td>
<td>Judges' background: e.g. academic discipline, age and gender of judges, hierarchy of errors</td>
<td>A total of 215 NSs in four groups of judges</td>
<td>Judges’ background (e.g. academic discipline) could predict evaluation; there exist no single hierarchy of errors.</td>
</tr>
<tr>
<td>Lennon 1991</td>
<td>Analysis of learner corpus of spoken English</td>
<td>English</td>
<td>Spoken</td>
<td>208 occurrences of doubtful acceptability</td>
<td>Narration by 20-24 year old German learners of English in UK</td>
<td>Discussed evaluation procedures and criteria</td>
<td>6 NSs</td>
<td>Proposed “extent” and “domain” and “infelicitous”; focused on advanced learner errors.</td>
</tr>
<tr>
<td>Haswell 1988</td>
<td>Frequency of written errors</td>
<td>English</td>
<td>Written</td>
<td>120 whole texts</td>
<td>Compositions from 1st, 2nd &amp; 3rd year students and workplace</td>
<td>Developmental contexts and errors</td>
<td>Researchers &amp; teachers</td>
<td>Post freshmen year students were experimenting with new language features.</td>
</tr>
<tr>
<td>Santos 1988</td>
<td>(1) Professors rated 2 compositions; (2) Professors rated underlined errors in 2 compositions</td>
<td>English</td>
<td>Written</td>
<td>Two whole texts</td>
<td>Compositions by intermediate ESL, students: natural unaltered texts</td>
<td>Content vs. Language; Comprehensibility, acceptability, irritation.</td>
<td>20 professors in (1) and 158 27 to 77 year-old students in (2)</td>
<td>Lexical errors &gt;***verb; irritation ***comprehensibility.</td>
</tr>
</tbody>
</table>

* NSs= native speakers; NNSs= non-native speakers
** "**" = "more serious than"
*** "+" = "not the same as".
Table 2.2: Summaries of some major empirical research on error evaluation (continued)

<table>
<thead>
<tr>
<th>Study and Year</th>
<th>Topic/Focus of study</th>
<th>Learner Variety</th>
<th>Spoken/Written</th>
<th>Unit of error &amp; Context</th>
<th>Source of learner errors</th>
<th>Key Notions Studied</th>
<th>Judges</th>
<th>Highlights of Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Santos 1987</td>
<td>Ranked error poorness &amp; indicated irritation level</td>
<td>English</td>
<td>Written</td>
<td>Two whole texts</td>
<td>Two compositions by intermediate ESL students; natural unaltered texts</td>
<td>Markedness: Unmarked to marked and Marked to unmarked</td>
<td>40 university science professors</td>
<td>Error evaluation is related to markedness.</td>
</tr>
<tr>
<td>Khalil 1985</td>
<td>NSs perceived Sentences (Context vs. No-context)</td>
<td>English</td>
<td>Written</td>
<td>30 sentences: 10 grammatically deviant &amp; 10 semantically deviant</td>
<td>Researcher reconstructed errors</td>
<td>Context and No-context; intelligibility and naturalness</td>
<td>440 university faculty members</td>
<td>Presenting context reduced judged intelligibility but not naturalness.</td>
</tr>
<tr>
<td>Vann, Meyer &amp; Lorenz 1984</td>
<td>Faculty members ranked erroneous sentences on 5-point scale.</td>
<td>English</td>
<td>Written</td>
<td>36 sentences containing errors of various types</td>
<td>Instructor-suggested, researcher constructed errors</td>
<td>Judges' background; error gravity</td>
<td>164 university faculty members</td>
<td>Age and academic discipline predicted tolerance of errors; Word order &gt;** It-deletion &gt;Tense&gt; Relative Cl. &gt; Word choice.</td>
</tr>
<tr>
<td>Magnan 1983</td>
<td>15 error types in 105 sentence pairs on tapes.</td>
<td>French</td>
<td>Spoken</td>
<td>105 pairs of sentences</td>
<td>Researchers constructed pairs of sentences on tapes</td>
<td>Gender related errors</td>
<td>352 French NSs; 104 adults plus 248 students</td>
<td>Adults were more bothered by verbs and pronouns while younger judges found gender-related errors more irritating.</td>
</tr>
<tr>
<td>Varonis &amp; Gass 1982</td>
<td>Four experiments: grammaticality, comprehension and pronunciation.</td>
<td>English</td>
<td>Spoken</td>
<td>Information requests at sentence level</td>
<td>Information request tasks by ESL students; sentences in previous study</td>
<td>Grammar; pronunciation; familiarity; comprehensibility</td>
<td>NSs (Background not given)</td>
<td>Grammaticality effect on perception only sometimes, but no effect on extremely good or extremely bad pronunciation.</td>
</tr>
<tr>
<td>Delisle 1982</td>
<td>Communicative-purposed-defined errors; repeated Politzer's 1978 but on written errors instead.</td>
<td>German</td>
<td>Written</td>
<td>60 pairs of sentences; learner indicated which one of the pair was worse</td>
<td>Errors by US learners of German (Politzer 1978)</td>
<td>Comparing perceptions of written errors vs. Politzer's spoken errors</td>
<td>19810 to 17 year-old high school students in W. Germany.</td>
<td>Some errors are more serious than others; gender-&gt;** verb morphology &gt;word order &gt; vocabulary &gt; case ending &gt; spelling; spelling, gender and word order errors become less important with an increase in judges' age; 10-12 olds perceived differently than older judges.</td>
</tr>
</tbody>
</table>
Table 2.2: Summaries of some major empirical research on error evaluation (continued)

<table>
<thead>
<tr>
<th>Study and Year</th>
<th>Topic/focus of study</th>
<th>Learner language</th>
<th>Spoken/written</th>
<th>Unit of error &amp; context</th>
<th>Source of learner errors</th>
<th>Key notions studied</th>
<th>Judges</th>
<th>Highlights of findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Galloway 1980</td>
<td>Evaluation of unrehearsed Spanish speech</td>
<td>Spanish</td>
<td>Spoken</td>
<td>Oral response to unseen randomly picked questions</td>
<td>10 second-semester university students in U.S.</td>
<td>Comprehensibility</td>
<td>Spanish NSs and NNSs, teachers and non-teachers</td>
<td>More intra-ling than inter-ling errors: NSs &amp; NNSs reacted differently to tense confusion errors.</td>
</tr>
<tr>
<td>Chastain 1980</td>
<td>Evaluation of U.S. university Spanish learner errors</td>
<td>Spanish</td>
<td>Written</td>
<td>Sentences</td>
<td>Instructor-identified errors</td>
<td>Comprehensibility</td>
<td>48 1st year college students in Spain</td>
<td>Word usage errors serious; Acceptability #*** comprehensibility</td>
</tr>
<tr>
<td>Albrehtsen, Henriksen &amp; Fierch 1980</td>
<td>Intelligibility and evaluation tests by NSs on collected ESL learner texts.</td>
<td>English</td>
<td>Spoken</td>
<td>Extracts of learner interviews conducted by NSs</td>
<td>Collected from 20 Grade-10 Danish students</td>
<td>Intelligibility; communication strategies; error frequency</td>
<td>120 adults NSs &amp; 180 16 to 17-year-old students</td>
<td>Significant difference between the two NS age groups; no connection between comprehensibility and correctness; Irritation related to error frequency, not type.</td>
</tr>
<tr>
<td>Piazza 1980</td>
<td>NSs' evaluations of U.S. learners' French errors</td>
<td>French</td>
<td>Spoken and written</td>
<td>100 sentences: 20 error types</td>
<td>Researcher-identified errors</td>
<td>Comprehensibility and irritation</td>
<td>264 French students aged 17-18</td>
<td>Verb form and pronouns &gt;** word order &gt; Tense usage</td>
</tr>
<tr>
<td>Guttermann 1978</td>
<td>Taped sentences</td>
<td>Spanish</td>
<td>Spoken</td>
<td>Sentences</td>
<td>Peace Corps volunteer' errors in training course</td>
<td>Comprehensibility and error frequency</td>
<td>70 college level El Salvador students</td>
<td>Article agreement &gt; article omissions</td>
</tr>
<tr>
<td>Politzer 1978</td>
<td>Matched guise; matched pairs</td>
<td>German</td>
<td>Spoken</td>
<td>Sentences: 60 pairs of sentences</td>
<td>Errors by U.S. learners of German</td>
<td>Judge's attributes Error gravity</td>
<td>146 teenagers</td>
<td>Vocabulary &gt; verb morphology &gt; word order &gt; phonology &gt; case ending, school experience influenced error perception</td>
</tr>
</tbody>
</table>
2.3. Proposing a constructivist approach to error evaluation

Although Rifkin and Roberts (1995) made the relevant remark that previous error evaluation research had neglected the context of errors, their reasoning behind this comment suggested a stance which, in my view, is of little help in making further progress in this line of research. Rifkin and Roberts criticized certain error gravity studies for not accounting for order effects in their experiments. They noted that many of the studies they had reviewed adopted subjective questions instead of the more desirable, objective assessments, making comparison of results from different studies impossible. They also suggested that the background of respondents in error perception experiments be more tightly controlled to avoid sampling skewedness, and they demanded a justification of the sample size in these studies.

2.3.1. The limitation of a conventional approach to error evaluation

Judging from their comments, I infer that Rifkin and Roberts have taken the view of a conventionalist (positivist), who assumes that a single reality exists independently from the observers (Guba & Lincoln, 1989)\(^2\). Adopting this conventionalist view, Rifkin and Roberts have come to the observation that greater progress in error evaluation

\(^2\) Habermas (1971) suggested three different knowledge-seeking sciences: the Empirical-analytic Sciences, the Historical-hermeneutic Sciences and the Critical Social Sciences. Popkewitz (1984) argued similarly that an educational researcher can be identified as a subscriber to one of the three basic paradigms according to his or her epistemological stance and the methodology adopted: the Empirical-Analytic, the Symbolic Interactionist, and the Praxis orientation. Guba and Lincoln (1989) called the Empirical-Analytic the positivists, and the Symbolic Interactionist and the Praxis orientation, the constructivists/post-positivists. Hitchcock & Hughes (1995) gave the label of interpretivists to Guba & Lincoln's (1989) constructivists, who in Popkewitz's terms can be divided into two camps: The Symbolic Interactionist and the Praxis orientation. In the present discussion, I will adopt the terminology of Popkewitz (1984) and Guba and Lincoln (1989) and use the term 'conventionalists' and 'the constructivists'. I will also maintain that there are basically two camps among the constructivists: the Symbolic Interactionist and the Praxis orientation, who both agree on Symbolic Interactionism with the latter also emphasizing emancipation in the process of research and as a result of research.
research can be made with more carefully designed quantitative, controlled experiments. While I do not in general object to an experimental or quasi-experimental research approach, I believe the premise expressed by Rifkin and Roberts is mistaken. Their static epistemological view erroneously posits that people's perception of errors is not an intersubjective, synthetic construction, but an objective, measurable entity that can be broken down, decontextualized and quantitatively assessed. The other weakness of this perspective is a narrow, deterministic experimental approach, not recognizing that error-related notions like the native speaker's "irritation", "comprehensibility", "acceptability" and "intelligibility" are meaningful only in a social-interactional, experiential, emic perspective.

Concluding their review, however, Rifkin and Roberts made the observant remark that future researchers should define what constitutes an error when at present error is understood only as L2 linguistic anomalies (p.532). I agree with this observation. In fact, the present thesis study aimed to realize this intent in my basic question for a phenomenological inquiry: What are the instructors' perceptions of language error in the context of teaching and assessing ESL technical writing? To a phenomenologist, this is equivalent to the question of perceiving not just 'errors in context' but also perceiving errors in a context in which 'errors in context' are perceived. That phenomenology must take into account the context of perception is a crucial starting point for the present inquiry. I will elaborate this point further in Chapter Three.

Up to this point, I have shown that the notion of error is a powerful one for the theory and the practice of L2 education. I have argued that the question of error perception and the constitution of error is a critical issue in the tradition of error evaluation research. I have also summarised the literature of error evaluation research and showed that our knowledge from this research is uncertain and obscure. Following Rifkin and Roberts (Ibid.), I will further argue that the obstacle to a better understanding in this
research is not just research methodology but is a narrow conceptualization of the concept of error.

2.3.2. Error and context

While the term *error* has been extensively used in teaching and in the second language education literature, not many writers have attempted seriously to define the concept of error. In the literature reviewed for this proposal, I have encountered a total of only five explicit definitions of error. Below I will subject them to an examination. The purpose of this exercise is twofold. First, through this review, I will show that the notion of error is in fact too complex to be captured in a definition of a few sentences. Second, my review supports Rifkin and Roberts’ (1995) observation that error so far has been conceptualized basically as a linguistic phenomenon. I argue that this is not enough.

Harmer (1983) in his book, *The Practice of English language teaching*, suggested the following definition for error:

Definition 1: An error is the result of incorrect rule learning; language has been stored in the brain incorrectly, in other words.
(p. 35)

Harmer’s definition follows that of Corder (1967). They see language learning as equivalent to learning the rules of linguistic systems. This reading has a base in cognitive psychology; however, it is not helpful functionally for identifying and evaluating errors. It does not make explicit who should judge correctness nor the basis on which correctness or incorrectness can be determined. Also, the conventional notion that language ability is something residing in a person’s brain is problematic. I will come back to this point later.
H. D. Brown in his 1980 version of *Principles of language learning and teaching* suggested the following definition for error, a definition which remained unchanged in the 1994 revised edition:

**Definition 2:** Put in another way, an error is a noticeable deviation from the adult grammar of a native speaker, reflecting the interlanguage competence of the learner. (1980, p. 165; 1994, p. 205)

Similar to that of Harmer, this definition tends to restrict the notion of error to grammatical competence, thus failing to account for sociolinguistic inappropriateness and pragmatic failures. However, it does suggest that the identification of errors requires a native speaker norm.

The third definition, given below, appears in Dulay, Burt and Krashen (1982):

**Definition 3:** We use *error* to refer to any deviation from a selected norm of language performance, no matter what the characteristics or causes of the deviation might be. (p. 139) [Italics in original]

The descriptive definition suggested by Dulay et al. (1982) can be seen as an improvement from Definitions 1 and 2 in that it points out the need for a norm to judge errors. However, it does not say who is to pass this judgement; nor does it explain what this norm of language performance is nor how this norm can be established.

Richards, Platt and Weber (1985) in their *Longman dictionary of applied linguistics* defined error as:
Definition 4: The use of a linguistic form (e.g. a word, a grammatical item, a speech act, etc) in a way which a fluent or native speaker of the language regards as showing faulty or incomplete learning. (p. 95)

This reading of error emphasizes rules of language use. It also has the advantage of recognizing the need of a native speaker to judge error. However, like Dulay et al.'s definition, it does not explain what "faulty" or "incomplete" means nor how to determine 'faultiness' or 'incompleteness'. It does not assume the existence of a norm and it does not consider error and its context.

Lennon (1991) in his study on spoken errors by four female advanced German learners of English proposed the following working definition:

Definition 5: A linguistic form or combination of forms which, in the same context and under similar conditions of production, would, in all likelihood, not be produced by the speakers' native speaker counterparts. (p. 182)

Of all the definitions suggested here, the one which Lennon (1991) proposed, endorsed by James (1998, p. 64), surpasses others in several aspects. In this definition, Lennon proposed to compare learner’s production with the language produced by native speakers in the same or similar context. An awareness of the context of production of learner language is evident in Lennon’s observation that “[M]ost erroneous forms are, in fact, in themselves not erroneous at all, but become erroneous only in the context of the larger linguistic units in which they occur” (p. 189). A similar point on the importance of the context of error production was also taken up strongly by Haswell (1988), who raised the importance of what he called ‘developmental context’. When comparing errors produced by two groups of college students’ writing at different years in their studies, Haswell observed that it is important to know the context of production, e.g., whether one
group of students were in fact experimenting in their writing with the more difficult features and hence produced more errors.

By adopting a comparative approach to error, advantages are gained. First, the expanded notion of errors now embraces not only syntactic or lexical anomalies but also violations of use rules such as Grice's maxims or Politeness principle (sociolinguistic inappropriateness and pragmatic failures). For example, in a formal context of meeting-new-friends, where native speakers may usually say “How do you do?” but the non-native speaker responded with “I am fine. Thank you,” grammatical analysis would not be telling; however, a comparative approach would show the non-native speaker’s pragmatic failure as this is not a usual native speaker’s reply in the similar context. Second, for a similar reason, a comparative approach can also reveal operationally learners’ covert errors (cf. Johansson, 1978; Schachter, 1974) and the possible use of avoidance strategies, a criticism many early error analysts received.

One study which has examined the use of socio-pragmatic competence of the second language speakers and compared it to that of the native speaker in the same context is Harlow (1990). In this study, significant differences were identified between those requests made by native speakers of French and those made by non-native speakers in three different scenarios: (a) requesting information and thanking a stranger for providing information, (b) requesting a service and (c) apologizing. Some of these differences, for example, were that the native speakers tended to use more polite indirect requests involving tu and modal auxiliaries such as pourriez-vous (or savez-vous) while the non-native speakers preferred using vous and direct requests such as Où est la pharmacie? Or Quelle est la route...? Although the data analysed in Harlow (1990) were realizations self-reported by 28 native speakers and 32 non-native speakers in questionnaires, the study demonstrates the value of a comparative approach in which the context of language
production is given due consideration.

2.3.3. The context of error production and the context of error perception

Lennon (1991) also proposed two useful constructs, *domain* and *extent*, to consider an error and its impact on the surrounding linguistic contexts. *Domain* refers to the rank of the linguistic unit (e.g., morpheme, word, phrase, sentence, paragraph, the whole text, extra-linguistic context) which must be taken as context in order for the error to become apparent. *Extent* refers to the rank of the linguistic unit (word, phrase, sentence, paragraph, the whole text, extra-linguistic context) from minimally the morpheme to maximally the sentence, paragraph or the whole text, which would have to be deleted, replaced, reordered, or supplied in order to repair production (p.191). Hence, for a sentence like “John is happy,” when in fact John is sad, the *domain* of the error is the extra-linguistic context” and the *extent* of error is a word, e.g., from “happy” to “unhappy”. Lennon’s examples for the use of these two constructs imply that the gravity of error may be assigned according to the level of fit between the form of the language and what exists outside the language and the text. To this date, however, there have not been any empirical studies verifying the usefulness of these two constructs.

The consideration of the context of production of learner errors and the larger linguistic milieu in the text where errors are identified are not all that have to be considered in error evaluation. What seems to have been neglected in all these definitions is also the context in which an error is identified and perceived. In fact, I believe that error gravity is not only related to the extent and domain of an error, and the linguistic context in which the error is produced, but also to the larger social context in which an error is produced.
perceived.

Consideration of the context of error evaluation, which includes both context of language production and context of error perception, in which error is perceived, is particularly crucial for error evaluation when learner language is considered at a discourse level. At this level, the genre (both speech and written), the role of the addresser and the addressee, the communicative purpose, the medium, the social formality, and the absence or presence of an explicit purpose of error evaluation are all relevant.

This review of definitions tends to suggest, first of all, that the notion of error cannot be adequately captured by a definition in a few sentences. In addition, the review also suggests three implications for further error evaluation studies:

(1) From Brown (1980, 1994), Dulay et al. (1982) and Harmer (1983) to Richards et al. (1985), and then to Lennon (1991), we can see error evaluation researchers' increasing awareness of the context in which error appears, though the larger context most of these writers have considered is basically linguistic.

(2) In spite of the consideration given to the context of error production, however, the role of the context of perception is still largely neglected in the most recent research.

(3) From Brown (1980, 1994) to Lennon (1991), we can also see the need for error evaluation to call on two important constructs: 'the native speaker' and a 'norm'.

The first two of these points will be further elaborated in Chapter Three when I

---

3 I note that the context of production may or may not be the same as the context of error perception. In face-to-face interactions, they are likely to be the same. In written communication, for example, where the writer and the reader are not co-present in time and location, they may be different.
discuss the methodology of the present study. For the moment, I wish first to examine the terms ‘the native speaker’ and ‘norm’. I will argue that in spite of their service to the discourse of error evaluation, the use of these two terms ‘native speaker’ and ‘norm’ has in fact taken too much for granted.

2.3.4. Expert speakers’ context-sensitive norm

Rampton (1995) proposed the use of the terms “expertise” (i.e., skill, proficiency, ability to operate with a language) and “allegiance” (i.e., identification with a language, with the values, meanings and identities it stands for) to decompose the term “native speaker” or “non-native speaker”. According to Rampton, the notion of ‘native speaker’ is problematic. It is too simplistic and cannot capture the complex realities of overlapping social and cultural identities in many multilingual communities and situations. Rampton posits that the stereotypical notion of native speaker implies a static, biological endowment of a language, but in reality particular languages are required in social settings, and people do not belong to only one social group, once and for all (p. 337). The notion of “the native speaker” is not only a dispreferred notion for sociolinguists; some language testers have made similar complaints. Davies (1995) concluded his article entitled “Proficiency or the native speaker: What are we trying to achieve in ELT?” with a preference for ‘proficiency’ over ‘the native speaker’:

The native speaker is a fine myth: we need it as a model, a goal, almost an inspiration. But it is useless as a measure; it will not help us define our goals. So in spite of my firm agreement with Birdsong and my conviction that there is a continuum between native speakers and non-native speakers, nevertheless, I recognize that for language teaching purposes what is crucial is the definability of partial proficiency. (p. 157)
Hence, to avoid the impression of suggesting language proficiency as a biological endowment in the native speaker, I will hereafter use an alternative term for my study: the expert speaker.

Rifkin and Roberts (1995) when commenting on the inadequacy of a narrow concept of error also pointed out the need to probe further into the notions of norm(s). They posit that:

... to better understand the process of L2 error evaluation, we need to investigate further the nature of the target language norm(s) against which an error is being evaluated. Error may not be absolute linguistic entities but rather flexible, norm-bounded constructs whose limits shift from judge to judge across speech communities. (p. 531)

The notion of norm as a fixed rigid standard by which errors are judged is problematic because any norm is basically context-sensitive and negotiable. The norms for different language media, for example, can be very different because of their different features, and people usually give more allowance to spoken variations and deviations than they do to written errors (Halliday, 1989; Ochs, 1979).

I believe that the norm people use to judge and evaluate language error is always influenced by the speaker/writer (learner) of the error, the genre of the language in question, the purpose of the evaluation activity, and the situation in which an error is identified. Take the speaker (learner) as a context as an example. For example, what people in general regard as a language error may yield a different impression when the speaker/writer possesses a certain social status or personal traits. Former US Secretary of State Alexander Haig speaking in congressional hearings in 1982 was perceived to be “strange” and not “ungrammatical”, attempting to achieve deliberate effects when he used
nouns as verbs and verbs as nouns (James, 1998, p. 75). In Chapter One, I reported Swain’s (1995, personal communication) observation that the audience would adjust their percepts of language deviations made by L2 speakers because of these L2 speakers’ special social status. I therefore hypothesise that in a similar fashion, in many teaching and assessment settings, the personal traits or the needs and motivations of an L2 language learner may also have an impact on the error perceptions of their teachers or raters.

By the same token, it can be reasonably predicted that the genre of the language, the purpose of the evaluation activity and the setting an error is identified in are all different contexts affecting a judge’s perception of a particular error. The present thesis study was launched with an aim to verify the impact of such contexts on peoples’ evaluations of language errors such as the writer/speaker of the language, the linguistic milieu (e.g. genre) in which a specific error arises, etc.

2.3.5. Establishing a constructivist perspective for error perception research

In recent years, language researchers’ ontological understanding of language has been less the abstract, objectified, universal entity, as Saussure and Chomsky saw it, and more the culturally relative tool of semiotic meaning as Humboldt, Sapir and Whorf once proposed. Gumperz and Levinson (1996) explained the ideas behind this perspective:

Linguistic relativity is a theory primarily about the nature of meaning, the classic view focusing on the lexical and grammatical coding of language-specific distinctions. In this theory, two languages may "code" the same state of affairs utilizing semantic concepts or distinctions peculiar to each language; as a result the two linguistic descriptions reflect different construals of the same bit of reality. (p. 7)
Gumperz and Levinson (Ibid.) elaborated this perspective further:

... [a] large part of the burden of interpretation is thus shifted from theories of context-free lexical and grammatical meaning to theories of use in context. Some important principles of the use of language may plausibly be argued to be universal (e.g. Grice's "maxims of conversation" or the turn-taking and repair systems of conversation, or even some principles of interactional politeness). Yet others seem much more clearly culture-specific. For example, the ethnography of speaking has shown how diverse can be the principles governing the production and interpretation of utterances in specific speech events — court proceedings, formal greetings, religious rituals, councils, and the like.

......(p.8)

With this understanding of what language is, I believe that like the 'meanings' participants make in interactions, the "norm(s)" and "criteria" that participants tend to adopt in many different settings are negotiated or are products of negotiations (Gumperz, 1982; Gumperz & Roberts, 1991). Gumperz and Roberts (1991) argued for the existence of different microscopic social systems which emerge from different interactions:

Ethnographers of communication have shown that speech events constitute miniature social systems that can be described in terms of associated beliefs and values, the social import of what is transacted, norms specifying who can participate and in what capacity, as well as expectations about appropriate themes and topics and about appropriate ways of speaking ... [e]vents themselves, as Goffman has shown, constitute separate social environments where participants respond to each other's moves and engage in types of reciprocity in accordance with situation specific rules of etiquette, norms and expectations. (p. 53)
Hence, given what we now know in sociolinguistics about how people communicate and act, the notion of norm(s) as a relatively rigid standard or set of rules, by which L2 errors are judged and which only the native speakers possess, seems to have neglected too much of what has to be accounted for when the interactional context is in fact an integral part of the error evaluation activity.

2.3.6. Error perception: the conventionalist and the constructivist

The conventionalist view that native speakers have the norms against which errors are judged is historically related to Chomsky’s (1965) notion of competence, an idealized ability that a native speaker has inherited and always possesses. Subsumed under this notion of “innate ability” is then the construct of norm(s), understood as a static set of rules residing within a native speaker who perceives and judges errors in a learner’s language. This set of rules is always, seen in this perspective, used in the same way every time by an idealized native speaker in every error evaluation activity. All of the conventional error evaluation studies I reviewed in this chapter tend to hold this assumption explicitly or implicitly. With this understanding of the norm, error evaluation is then depicted as a simplistic, straightforward, decontextualized, mechanistic process although, as I have shown in the review of the literature, some of the more recent researchers did begin to recognise the importance of an extended linguistic context of error (Lennon 1991; Santos 1987, 1988), and one study considered the production background of learner errors (Haswell 1988).

As I have argued, however, in recent years a conventionalist view of linguistic competence has come under attack from many fronts, noticeably by sociolinguists (Gumperz, 1982; Riley, 1996; Willing, 1992) and socio-cultural interactionists (Lantolf,
1983; Lantolf & Appel, 1994; Wells, 1981, 1985, 1986). These writers as well as the praxis researchers (Lather, 1986a, 1986b, 1993; Lemke, 1990) have argued for the significant role that context plays, the social-cultural aspects of language interaction, and the collaborative nature of context-sensitive, situated interaction. One implication of this constructivist perspective is that ‘norm(s)’ are a collaborative ‘contract’ (Wells, 1981) or a set of negotiated rules of systems agreed upon by the participants in the interactional context. Hence, what distinguishes a constructivist notion of “norm” from that of a conventionalist is that the norm is interactional, i.e. How the norm works in an error evaluation activity depends upon the nature of the interaction, which is a jointly constructed activity engaging both the expert speaker’s language and the L2 learners’ language. Such a perspective enables us to see the organic and dynamic relationships among error, the norm(s) and the context of the communication of errors.

Hence, the constructivist perspective is distinguished from the conventionalist perspective in that the norm(s) arise from the interaction itself. In other words, the expert speaker (as well as the L2 learner) may have some pre-conceptualization of what counts as an error or what not, but the actual judgement of L2 error is a function of the interaction and a host of other contextual elements, including the learner’s language, the communicative intents of the participants, the status and roles of the participants, the expert speaker, and so forth. In other words, this perspective assumes that the activity of perceiving an error is not only a matter of evaluating it in its linguistic context but is an activity affected by the larger context of the interaction.

Moreover, this constructivist perspective draws the difference between the context of production and the context of perception. For example, this distinction would allow an expert speaker to perceive a learner’s written language when the expert speaker and the learner are not co-present in time and space. An example of this would be an ESL instructor, the expert speaker, identifying errors in a language learner’s written essay for
assessment purposes. This context is different from that in which the very same instructor gives face-to-face error corrections on the same learner’s draft during a composition lesson.

The constructivist perspective seems superior to the conventionalist perspective in that:

1. It more accurately depicts the error perception activity. The dynamics in the constructivist perspective account for the different or changing norms people may adopt in different (social) contexts.

2. It substitutes the mechanistic notion of ‘competence of a native speaker’ with one which accounts for the context-sensitive and collaborative nature of interaction.

3. It takes into account not only the context of error and the context of interaction but also the context of perception to suggest that error evaluation is not only determined by linguistic rules, but also rules of use, the various contexts in which the learner’s language is situated, and the contexts in which the learner language is perceived.

2.4. Defining the research questions

2.4.1. Defining context and genre

Since the constructivist perspective I propose relies very much on the notion of context, I wish to elaborate it further and explicate some of its implications relevant to this study. Widdowson (1998) suggested defining context in a schematic way; he viewed context as a set of variables (constructs) enabling the understanding of meanings, when at the same time he also recognised that these variables can be modified in the communicative process (p. 11). Goodwin and Duranti (1992, p. 2) observed that the
notion of context "involves a fundamental juxtaposition of two entities: (1) a focal event; and (2) a field of action within which that event is embedded". Following Widdowson (1998) and Goodwin and Duranti (1992), I define a contextual element in this thesis study to be a variable in the field, acting upon an error in focus in an evaluation activity.

As I have argued earlier in this chapter, the evaluation of errors in classroom L2 assessment is to varying degrees influenced by the specific needs and strengths of individual students (Berwick, 1989), the purpose of the evaluation activity (Elley, 1989; Engeström, 1996), and the teacher’s own beliefs in and attitude towards language teaching and learning (Pennington, 1990). I also believe that the genre of a learner’s written document — one kind of “enabling constructs” (Widdowson, 1998, p. 17) and a form of “co-text” (Widdowson, Ibid.) — may affect the error evaluation, given what we now know about the relationships between a particular genre and the use of its corresponding syntax and lexis (Bhatia, 1993; Hatch, 1992; Hoey, 1983; G. James, 1994; Swales, 1990; Trimble, 1985). So in the event of an ESL instructor evaluating a written error in a student’s composition, major contextual elements should at least include the following:

1. Learner (writer) as context,
2. Teacher (judge) as context,
3. The purpose of evaluating the error as context, and
4. The genre of the learner’s written document as context.

Of all these various contexts, the notion of genre is perhaps one which is most complex and requires a detailed discussion. Some writers consider the term “genre” particularly applicable to speech communication (Bakhtin, 1986). Other writers (Berkenkotter & Huckin, 1995; Russell, 1997) have been more interested in the recognizable structure of written text types and their situatedness. On the other hand, some researchers (Wells, 1990) stress that the recognition of a genre by a specific
discourse community is indispensable while other researchers focus on the creativity and
dynamics of genre growth and death (Miller, 1984). Still others (Bhatia, 1993; G. James,
1994; Paltridge, 1997; Swales, 1990) have arrived at very systematic descriptions of
existing genres for linguistic or pedagogical purposes.

Because the data for the present study were collected in a course aiming to teach
written technical genres to ESL students, I found a definition given by Russell (1997, p.
515) most suitable to the present inquiry:

Put in simplest terms, a genre is the ongoing use of certain
material tools (marks, in the case of written genres) in
certain ways that worked once and might work again, a
typified tool-mediated response to conditions recognized by
participants as recurring (to paraphrase Miller, 1984).

Here Russell stressed the typicality of genre (description) and its recurring
applications (prescription), which are both important for the genre learning of many ESL
students, who may not have the equivalent genre literacy in their L1 nor have been
exposed authentically to the genre types they are learning. In other words, a definition of
genre may prove inadequate if it does not address the typicality of what is to be learned
and how this typicality and its related variations can effectively handle recurring situations.

Russell's definition also seems particularly appropriate because it downplays the
use of genre to create new knowledge. While this may appear to be a weakness, it is
pedagogically realistic. In most L2 classroom situations, the tasks are often simulations,
whereas in real world situations, the tasks requiring the use of a genre are often novice
attempts to create new knowledge. An example of this would be the laboratory reports
submitted by university students to their professors compared to the laboratory reports
produced by research scientists and engineers. The former are usually simulation work, as
students' answers to display questions, while the latter are mediated tools for scientists to
create new human knowledge, as answers to laboratory scientists' self-imposed referential questions. In the students' learning of laboratory reports, it is probably not as relevant to stress genre as a tool to create new knowledge, but useful to emphasize the effectiveness of its typicality in recurring situations.

Another reason for adopting this definition of genre as opposed to the others is that recent L2 testing research has witnessed a shift towards an interactionist perspective (Chapelle, 1998; Douglas 1998; McNamara, 1997). The notion of genre as typicality in recurring situations seems to fit best an interactionalist, performance-based assessment, which is based on what Chapelle (1998) called "a sample of performance in similar contexts" (p.43).

2.4.2. Research Questions

I have surveyed related literature to draw implications for the present study. In particular, I have proposed a constructivist perspective of error perception to highlight the relationships between context and error perception. The present study was undertaken to yield answers to the umbrella research question "what is error?", and to identify the relationships between context and error gravity. In light of the literature review I have presented in this chapter, I began my data collection in an ESL technical writing course at a Canadian university to find answers to a set of research questions.

I began with Research Question No.1, which specifies that the data for this thesis study were collected in a teaching context. This question aims to identify certain language instructors' percepts of the errors their ESL learners had made in their written work in a context in which the instructors were actively participating in certain assessment activities. Their percepts inform an understanding of the meaning of error in the context.
Chapter Two

Research Question No. 1: What is error in the perceptions of the four ESL instructors during their assessment activities?

Research Question No. 2 aims to verify if and in what ways contextual elements such as learners, the instructor, the linguistic context of the error, and the specific purpose of the assessment activity actually act upon the instructors’ understanding of error. In this research question, the contexts of perception refer to the contexts I have discussed, namely, (a) learner (writer) as context, (b) teacher (judge) as context, (c) the purpose of evaluating the error as context, and (d) the genre of the learner’s written document as context.

Research Question No. 2: What are the relationships between errors and the contexts of perception?

In order to elicit instructors’ assignment of gravity to their students’ errors, I also asked Research Question No. 3 to identify the conditions in which serious errors emerged and the criteria the instructors used to determine error gravity:

Research Question No. 3: What are serious errors? What are the conditions with which they are identified?

As I believe that error evaluation is a part of assessment activity in L2 education, which itself is an integral part of the larger activity called teaching, I wished to identify the relationship between error evaluation and the objectives and the nature of the technical writing course. Hence, I asked Research Question No.4.

- 45 -
Research Question No. 4:  What are the relationships between instructors’ percepts of errors and the pedagogic objectives and focuses of the course as they perceived them?

2.5. Summing up

To sum up this chapter, I have discussed the relevant literature to draw implications for a theoretical orientation and the research questions in this thesis study. I have pointed out that in spite of the crucial role error has played in language education practice and SLA theory development, our understanding of the nature of error and error gravity is obscure and uncertain. I have noted that there has been an increasing awareness of the important role context can play in recent error evaluation research, but this awareness has stopped at linguistic contexts (the context of error production), and the question of a naturalistic setting for the evaluation activity, the question of the context of error evaluation, has never been systematically addressed in previous studies.

I have also pointed out that the main cause of methodological weaknesses in previous research is a conventionalist (positivist) epistemology and a conventionalist ontology of language. These views inadequately metaphorise language as an objectified, universal endowment in a native speaker which implies that various human percepts of violations of these rules — such as the severity level of violations, acceptability, comprehensibility and irritations — can all be decontextualized and objectively measured. Against these conventionalist views, I have proposed that, based on a relativist ontology of language, an interactionist perspective for error evaluation would help to make greater progress in error gravity research. Seen from these perspectives, I predict that error perception evaluations are activities subjected to the influence of contextual elements such as the participants in the activities, the purposes of the activities and the contexts in which
the errors were perceived or evaluated. In addition, I have also argued that a naturalistic inquiry into error evaluation can help expand our understanding of what error is.
CHAPTER THREE

APPROACH AND DESIGN OF THE STUDY

In this chapter, I explain why I adopted a phenomenological approach for this study. I present the rationale for the design of this thesis study; I also report the methods I used to collect and analyse the data.

3.1. The phenomenological approach

3.1.1. What is a phenomenological approach?

Phenomenology is an important school in the humanistic tradition (Lincoln, 1992). It shares a set of commonalities with other constructivist research approaches. Moustakas (1994) in his book *Phenomenological research methods* suggested the following list of common qualities of all humanistic science research, under which the phenomenological approach is subsumed:

1. Recognizing the value of qualitative designs and the methodologies and studies of human experiences that are not approachable through quantitative approaches.
2. Focussing on the wholeness of experiences rather than solely on its objects or parts.
4. Obtaining descriptions of experience through first-hand accounts in informal and formal conversations and interviews.
5. Regarding the data of experience as imperative in understanding human behaviour and as evidence of scientific investigations.
6. Formulating questions and problems that reflect the interest, involvement and personal commitment of the researcher.
7. Viewing experience and behaviour as an integrated and inseparable
relationship of subject and object and of parts and whole. (p. 21)

Phenomenological curriculum inquiry is an approach based on the assumption that human knowledge springs first and foremost from perception and human lived experience (Moustakas, 1994; Sullivan, 1990; Van Manen, 1990; Willis, 1991). Willis (Ibid.) defends this assumption:

Human experience, of course, inevitably includes perceiving, thinking and acting inextricably connected in certain ways, but it all begins with perceptions, and in this sense phenomenological inquiry is inquiry into what is primary in human experience. (p. 173)

Willis further explains the role of consciousness and its profound implications for human knowledge:

Human beings can have perceptions of the external world without becoming conscious of them, but our consciousness of external perceptions — our life-worlds — is where each of us lives in the most deeply personal sense. Not only do we feel our perceptions more acutely when we become consciously aware of them within our life-worlds, but we begin autonomously to consider what we can do about them. (Ibid., p. 175)

A phenomenological approach begins with intentionality (human consciousness). Central to the notion of intentionality are noema and noesis. Noesis refers to the interpretive act of perceiving, feeling, thinking, remembering, or judging an object. Noema refers to the object noesis is directed to, the object of our intentionality. The consideration of the co-existence of noema and noesis signifies the important notion of context. The ultimate aim of a phenomenological study is to acquire an interpretation of the structure of the object in scrutiny and identify the implications and outcomes of this
3.1.2. Justifying a phenomenological approach for this study

I adopted a phenomenological approach for the present study for four main reasons. First of all, historically a phenomenological approach, often initiated by a question in the form of "What is X?", has been considered one of the best tools to probe into the basic structure of a phenomenon. Here are some examples. Van Kaam (1959) asked: "What is really feeling understood?" Moustakas (1961, 1972, 1975) asked the question: "What is loneliness?" M. Heidegger (1971) asked: "What is language?" Van Manen (1982) asked: "What is pedagogy?" Giorgi (1985) asked: "What is learning to drive?". Holmes (1993) asked: "What is the experience of homelessness?". The present study, initiated by the basic question "What is language error?", was undertaken in a similar spirit.

Secondly, I have reported in the previous chapter a growing awareness in recent error evaluation research of the context of error, and I have argued for the importance of considering also the context of perception. A phenomenological approach is suitable to answer to these calls because it resents separating context from the subject under scrutiny and prioritizes the wholeness of human experience and perception. Holmes (1975) in his explication of Husserl’s theory of the noema pointed out the importance of accounting for context in studying lived experience. He suggested three essential components of the noema, (1) the intentional object, (2) the noematic senses (percepts), and finally (3) the particular manner of their being given — the context.

Thirdly, recent evaluation research tends to relate error perception and gravity to notions such as interpretability, comprehensibility, irritation, and acceptability (Rifkins & Roberts, 1995). However, these notions are fluid and very difficult if not impossible to
define in conventional research paradigms; the complexity of their meanings can best be revealed in the emic, experiential perspective of a phenomenologist. Furthermore, the intrinsic concept involved in these notions suggest the need for social selves and an interactional perspective. This is congruent with that of the phenomenologists, who hold that the unit of analysis in phenomenology is never a person (an ‘I’), but a person in his or her social relation to others, a dialogue (I-thou) (Sullivan, 1990, p. 29).

Finally, a phenomenological approach is congruent with the interactionist view on error perception. Inquirers from both phenomenology and symbolic interactionism hold that reality is socially constructed, consensual, intersubjective, and yet can have as many phases as the number of viewers or participants (Giorgi, 1992; Prus, 1996). These two approaches both seek answers to meaning questions (Heap 1995) — the meanings of people and things for people in the context in which they are situated; both approaches aim at understanding human behaviours by understanding the language people use to interpret people and things in the situations (Blumer, 1969; Mead, 1934). For all these reasons, I carried out the present inquiry with a phenomenological approach.

3.1.3. Phenomenological research methods

Contemporary phenomenologists all refer to the works of the same group of philosophers: Husserl, Heidegger, Schutz, Marcel, Gadamer, Merleau-Ponty, and Dilthey. Different phenomenologists, however, although they share some common core characteristics, have emphasized different aspects of the philosophy. This is understandable because, after all, phenomenologists are constructivists who believe in what Lincoln and Guba (1989) called multiple realities. Van Manen (1990) even suggested that a phenomenological inquirer should invent an approach for a particular inquiry (p. 173), a point echoed in Willis (1991, p. 184).
My task in this section of the thesis is to report the considerations I had when choosing a set of research methods. Tables 3.1, 3.2, 3.3 and 3.4 highlight the major emphases of four different schools of phenomenology and the different methods each has adopted. These four schools are *Empirical phenomenological research, Hermeneutics research, Heuristic research* and *Transcendental phenomenological research*.

As can be seen in Tables 3.1 to 3.4, empirical phenomenological research tends to use the experiences of other people to ask and answer a meaning question. Hermeneutics emphasizes historical and textual analyses. Heuristic research usually requires the collection of experiences from more than one situation or context; it also prefers the use of stories and focuses its descriptions and interpretations on individual participants. Transcendental Phenomenology emphasizes Epoche, the elimination of suppositions and empirical knowledge in the pursuit of the essence of a phenomenon; it also tends to rely on the researcher's self-reflections to discover the meaning of things.

Although I recognized that in the present thesis study I would derive the meaning of error from the interviews the instructors gave me, I did not undertake this study using Hermeneutics because this approach tends to focus on historical or textual analyses, and I did not think that this was particularly relevant to the issue of language error and language error perception among the dynamics of teacher-student interaction.
Table 3.1. Empirical phenomenological research

<table>
<thead>
<tr>
<th>Methodological approach</th>
<th>Characteristics and methods</th>
</tr>
</thead>
</table>
| Empirical phenomenological research (Van Kaam, 1959, 1966; Giorgi, 1979, 1985; Van Manen, 1990) | 1. To obtain comprehensive descriptions that provide the basis for a reflective structural analysis.  
2. To disclose and elucidate the phenomena of behaviour as they manifest themselves in their perceived immediacy.  
3. Method 1 (Giorgi, 1979; Van Manen, 1990):  
  (1) Reads the entire description of the learning situation straight through to get a sense of the whole.  
  (2) Reads the same description more slowly and delineates each time that a transition in meaning is perceived with respect to the intention of discovering the meaning.  
  (3) Eliminates redundancies, clarifies or elaborates to oneself the meaning of the units one has just constituted by relating them to each other and to the sense of the whole.  
  (4) Reflects on the given units, still expressed essentially in the concrete language of the subject, and comes up with the essence of that situation for the subject. Each unit is systematically interrogated for what it reveals. Transform each unit, when relevant, into the language of psychological science.  
  (5) Synthesize and integrate the insight achieved into a consistent description of the structure of learning.  
  (1) The research delineates a focus of investigation and formulates a question in such a way that it is understandable to others.  
  (2) Start with descriptive narrative provided by subjects who are viewed as co-researchers; query the person and/or engage in dialogue.  
  (3) Read and scrutinize data so as to reveal their structure, meaning, configuration, coherence, and the circumstances of their occurrence; emphasis is on the study of configuration of meaning involving both the meaning and how it is created. |
Table 3.2. Hermeneutics

<table>
<thead>
<tr>
<th>Methodological approach</th>
<th>Characteristics and methods</th>
</tr>
</thead>
</table>
| Hermeneutics (Gadamer, 1976; Willis, 1991; Smith, 1991; Moustakas, 1994) | 1. All experience is given validity by our consciousness; undistorted reality exists for us in the facts of consciousness given by inner experience.  
2. To understand this experience requires a study of history and the experience in its historical grounding.  
3. Phenomena are amoral but interpretations are moral and interpretations unmask what is hidden behind the objective phenomena.  
4. Interpretation is not isolated; it is the basic structure of experience.  
5. The dialectic hermeneutic circle: our prejudgments are corrected in view of the text, the understanding of which leads to new prejudgments; the prejudgments that lead to pre-understanding are constantly at stake, their surrender could also be called a transformation. (Moustakas, 1994, p.10)  
6. The text, or interview protocol, provides an important description of conscious experience.  
7. Reflective, creative interpretation of text is needed to achieve a fuller, more meaningful understanding, to bring 'before me something that otherwise happens "behind my back"'. Reflection gives not only a description of the experience but also an analysis of the interpretation of the underlying conditions that account for the experience.  
8. Methods: 1. a fixation on meaning 2. Dissociation at some point from the mental intention of the subject (author of text) 3. the necessity to interpret the protocol as a whole, a gestalt of interconnected meanings 4. Their potentiality for multiple interpretations. (Moustakas, 1994, p. 11) |

Aim: to describe and understand human experience by examining the meaning and intention of experience (especially texts) in the historical context in which it is situated.
Table 3.3. Heuristic research

<table>
<thead>
<tr>
<th>Methodological approach</th>
<th>Characteristics and methods</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heuristic research (Moustakas, 1990)</td>
<td>1. The self of the researcher is present throughout the process, while understanding the phenomenon with increasing depth, the researcher also experiences growing self-awareness and self-knowledge. Hence, heuristic processes incorporate creative self-processes and self-discoveries.</td>
</tr>
<tr>
<td>Aim: to initiate a process of internal search through which one discovers the nature and meaning of experience and develops methods and procedures for further investigation and analysis.</td>
<td>2. Six phases of heuristic research: (1) initial engagement, (2) immersion into the topic and question, (3) incubation, (4) illumination, (5) explication and (6) culmination of the research in a creative synthesis.</td>
</tr>
<tr>
<td></td>
<td>3. Heuristic research (HR) shares many commonalities with Empirical Phenomenological research (EPR) but according to Moustakas (1994) they are different in four aspects: (1) the Duquesne studies (EPR) focus on a situation in which the experience investigated occurs; heuristic research, on the other hand, is a wide open investigation in which typically the research participant widely and deeply explores the phenomenon. Hence, heuristic studies usually involve more than one example or one situation. (2) EP researchers seek descriptions of experience. In HR studies, the researchers seek to obtain self-dialogues, stories, poems, artwork, journals and diaries and other documents that depict the experience. (3) EP studies construct structures of the experience; HR aims toward composite depictions that remain close to the individual stories rather than elucidating situational structural dynamics. Heuristic studies culminate in a creative synthesis while empirical phenomenological studies end with a general structural description. (4) In EP studies, the researcher disappears in the process of interpretation in structural analysis. The Heuristic research participants remain visible in the examination of data and especially in the individual portraits, and continue to be portrayed as whole persons.</td>
</tr>
</tbody>
</table>
Table 3.4. Transcendental Phenomenology

<table>
<thead>
<tr>
<th>Methodological approach</th>
<th>Characteristics and methods</th>
</tr>
</thead>
</table>
| Transcendental phenomenology Moustakas (1994, p. 34)  
Aim: A return to the self to discover the meaning of things as they appear and in their essence. | The approach is characterized by four major processes: |
|  | 1. Epoche: The elimination of suppositions and empirical knowledge. All everyday understandings, judgements and knowings are set aside, and phenomena are revisited, freshly and naively from a vantage point of a pure or transcendental ego. |
|  | 2. Reduction: Each experience is considered in its singularity, in and for itself and described textually in its totality, in a fresh and open way. |
|  | 3. Imaginative Variation: Utilize imagination and varied frames of reference to approach the phenomenon from divergent perspectives, different positions, roles or functions. A theme analysis in this step will help consider the countless possibilities that are intimately connected with the meanings of the phenomenon. |
|  | 4. Synthesis: The final step is the intuitive integration of the fundamental textual and structural descriptions into a unified statement of the essences of the experience of the phenomenon as a whole. |
I did not choose Heuristics, which usually requires the active participation of the researcher in the process and focuses upon the individuality of the participants. At the end of the day, I wanted this study to reveal some new meanings of error and inform language educators about error gravity, rather than focus on maintaining instructors’ individual traits and their specific contexts in the experiences they told. Nevertheless, I appreciate that Heuristics emphasizes the discovery of new meanings of human phenomena via intense observation and introspection. I present my interpretative account of each of these instructors’ percepts in Chapter Five when I discuss the findings of this thesis study.

I did not choose Transcendental phenomenology because it requires epoche, with the researcher holding no presupposition, preoccupation, beliefs and biases. Although in the process of carrying out this study, I exercised “bracketing” on my own preconceptions and resisted leading the participants in certain pre-determined directions, I found it impossible to begin my data analyses with epoche. I realized that I had come to the study with my own subjectivity, teaching and learning experiences, and an understanding of the related literature. As the data themselves unfolded, the themes emerging from the transcriptions often turned out to be key points of discussions in the literature on error-related research.

I adopted an approach which basically followed the Empirical phenomenological school (Giorgi, 1979, 1983, 1985, 1992; Van Manen, 1982, 1990). The reason that I chose this school is that, first of all, empirical phenomenology emphasizes the lived experiences of other people (Van Manen, 1990):

The point of phenomenological research is to “borrow” other people’s experiences and their reflections on their experiences in order to better be able to come to an understanding of the deeper meaning or significance of an aspect of human experience, in the context of the whole of human experiences. (p. 62)
At the time of the study, I was not teaching and hence my study had to rely upon data from other teachers as a starting point for my own understanding of the issues. This was the first reason. The second reason for my choice was that Empirical phenomenology is a research tradition with lucid, explicit research methodology and methods (Giorgi, 1979, 1983, 1985, 1992; Marton, 1986; 1994; Van Manen, 1990), and I agree with Van Manen (1990) that there is a need to demystify phenomenology. Empirical phenomenology is especially capable of this task. Van Manen explicated the idea of demystifying phenomenology:

"Phenomenology is the study of essences," said Merleau-Ponty (1962, p. vii). But the words "essence" should not be mystified. By essence we do not mean some kind of mysterious entity or discovery, nor some ultimate core or residue of meaning. Rather, the term "essence" may be understood as a linguistic construction, a description of a phenomenon. (p. 39)

Thirdly, according to Richardson (1999, p. 67), empirical phenomenology advocates the use of discursive practices — using participants' language constructed interactively and conversationally to arrive at descriptions and interpretations; this fits logically and coherently well with the basic views of social constructionism (Blumer, 1969) and, consequently, with the constructivist perspective of error perception I proposed in Chapter Two. Like a phenomenologist, a constructivist recognises that the relevant entities (e.g., "irritation", "acceptability", "seriousness") appearing in the discursive practices (the conversational constructed interviews) are all locally created artifacts, and hence there is no need to prove their evidential status (Richardson, Ibid., p. 67), thus saving the researcher from the ontological debate on realism versus idealism and dualism versus nondualism.
In his discussion on contextualizing and implementing a phenomenological study, Van Manen (1990) advocated a researcher “inventing an approach” (p. 173) and he argued for adopting “approaches and procedures uniquely suited to this particular project and this individual researcher” (p. 163). Although the approach I adopted was primarily based on empirical phenomenology, this approach also includes my interpretation of teachers’ percepts as heuristics of my own. I have already argued, based on Willis (1991) and Van Manen (Ibid.), that others’ perceptions are a valuable starting point for my understanding, and in this thesis study the perceptions the instructors provided me with served well as the major sources of my findings. However, at the end of the day, this thesis study also requires that I offer my own interpretation of these findings; ultimately it was I who was seeking the meaning of the phenomenon under study.

There is a practical reason for this study to arrive at an eventual interpretation of my own. The contributions that the participating instructors could give to this research were limited, and they had already contributed a lot by giving me the time they had agreed to. Because individual instructors only verified the descriptions of their own perceptions and did not have time to read those from the other instructors, they were not able to see what I could see as a researcher, and hence it was not possible to seek their interpretations of all the percepts I have gathered. Moreover, of all the participants including myself, I was probably the only person in a position to have carried out an extensive review of the literature. Consequently, I came to the data at the end of my analyses with this understanding, informed also by discussions in the literature.
3.2. Methodological issues

3.2.1. Data collection methods

One major kind of data in this thesis study was collected via the method of stimulated recalls (Tuckwell, 1980). Because such a method involves a participant reporting verbally in introspection or retrospection, it is also known as protocol analysis (Ericsson & Simon, 1993), a method widely used in many qualitative studies including some phenomenological and phenomenographic research. For example, in some studies by phenomenographers, Marton and his associates (e.g., Marton, 1986; Marton & Booth, 1996, 1997), gave students articles to read and asked them afterwards to report the gist of what they understood; the aim of their research was to define a hierarchy of learning processes and outcomes.

In a similar vein, the primary method for collecting data in this thesis study was to ask each of the four participating instructors in interviews to read again a student’s written report the instructor had previously marked, and then ask the instructor to report in retrospection his or her percepts at an earlier time when he or she encountered errors or mistakes in the assessing process. The student’s text and the instructor’s own markings on the student’s report served as stimuli for the instructor’s recalls.

The general descriptions of instructors’ beliefs, assumptions and knowledge of language teaching and the course they were teaching were also an important part of the

---

1. Phenomenography can be considered a subfield of phenomenology, combining phenomenology with ethnography. Simplistically put, a phenomenological study investigates into a local culture or an institution in which the researcher is situated (e.g., Marton, 1986, 1994).
data for this study. I collected them in one of the three interviews each instructor gave me; these interviews were carried out as conversations.

To understand the context in which my data were collected and the context in which the instructors’ percepts were initiated, I also carried out classroom observations and textual analyses to gain different channels to the phenomenon under examination. I went into the participating instructors’ classrooms, the larger contexts from which their perceptions of error (the noema) emerged, and observed. I also studied students’ previously assessed written reports and performed analyses on them. I report how I carried out these data collection procedures and data analyses in detail in Sections 3.3 and 3.4 of this Chapter.

3.2.2. Reliability and phenomenological research methods

Phenomenologists do recognize the paramount importance of reliability in their methodology since for them the question of reliability, the differentiation of illusions from stable perceptions, is an integral part of their validity issue. It has been argued that reliability in phenomenology is not the same as in the discourse of the psychometric tradition (Wertz, 1986). The discussion on phenomenologists’ conceptions of reliability starts with the primacy of perception but the argument rounds itself in dialectics. Wertz (Ibid.) maintains that “the multiplicity of perspectives entailed in various sense modalities does make perception vulnerable to illusion”, but “the problem of the perspectivity of perception contains within itself its own solution in the end, since it is only by the assumption of a new perspective related to the first as “better” that a former can be disclosed as illusory” (p. 190). Wertz develops further his argument dialectically:
What is most important about perception for our problem is that inconsistency, variance, relativity, and difference, far from being a threat to the establishment of a stable reality, is the very precondition and guarantee that I can look again and again and see the same thing, that I can touch and see the same thing; it is the way in which the meaning of "sameness" and "thing" is established. (p. 190)

In other words, the solution to the possible unreliability of perception is continued perceptual consciousness and the variation of our focus of awareness. This view gives a crucial hint to explain why Moustakas (1994) advocates the phenomenological method of Imaginative Variation (cf. Table 3.4) and Van Manen (1990) the method of Varying the Examples (p. 121). By Imaginative Variation, Moustakas means "seeking possible meanings through the utilization of imagination", "varying the frames of reference" and "approaching the phenomenon from divergent perspectives" (p. 97). By varying ways to look at descriptions (examples), Van Manen (1990) means bringing the 'invariant' aspects of the phenomenon itself to come into view (p. 122).

3.2.3. Validity and phenomenological research methods

Many writers have argued that researchers in the humanistic tradition do not face the same validity issue as the conventionalists or natural scientists do (Bogdan & Biklen, 1992; Eisner, 1991; Hitchcock & Hughes, 1995; Lather, 1986a, 1986b, 1993) because of the differences in their respective ontologies and epistemologies. The conventionalists take the realist view that a single reality exists independently from the observer; hence the issue of capturing this single reality — the validity and especially construct validity — is of paramount importance (Guba & Lincoln, 1989). The constructivists, however, believe that there are multiple realities, which are socially co-constructed by people in the local context
(Ibid.). For this reason, some writers, Wolcott (1990) for one, have even argued that the pursuit of validity might be meaningless and that what should be considered of supreme importance is the understanding of the complex layers and webs of meanings.

Wolcott’s (1990) view that understanding is most crucial finds echo in a piece of earlier writing by the phenomenologist Shapiro. In his paper entitled “Verification: Validity or understanding”, Shapiro (1986) drew upon the works of French phenomenologist Maurice Merleau-Ponty (1962) and Michael Polanyi’s (1958) Personal knowledge, criticized the positivist notion of validity, and argued for self-verification that emphasizes understanding as a criterion of truth. He concluded his argument by pointing out that although understanding is a personal process, its practice need not be asocial because the history people live in and the language people use to formulate this knowledge are shared.

While I agree with Wolcott and Shapiro that validity cannot serve in the same way as a key criterion or goal for a constructivist inquiry such as this thesis study, I interpret validity as a verification of the standard (quality) of an inquiry, and agree with many constructivist researchers (e.g., Bogdan & Biklen, 1992; Eisner, 1991) that qualitative methodology is not “a licence for anything goes”. Certain precautions and procedures must be exercised to ensure that an inquiry can meet the standards set forth in the research community. The consideration of these procedures for the proposed study will be discussed in the following section, but before that, I wish to conclude this section by discussing validity from a phenomenological perspective.

The examination of validity of a phenomenological inquiry also depends upon the congruence of the phenomenologist’s inquiry process to the reality his or her claims seek to represent. The two core procedures in a phenomenological inquiry are seeking descriptions of the phenomenon and arriving at an interpretation of it. The distinction
between interpretation and description in phenomenology is a subtle one. According to Giorgi (1992), interpretation is a “development of a plausible but contingent line of meaning attribution to account for a phenomenon” (p. 122). Here Giorgi emphasizes the term *contingent* and that in interpretation one argues only for the plausibility of an account, not for its exclusivity. For Giorgi, description is “the clarification of the meaning of the objects of experience precisely as experienced” (Ibid.); hence, the difference between the two is “the plausibility of a meaning” (interpretation) and “the precision associated with the fulfilment of a demand or requirement” (description) (Ibid.).

At the level of method, this distinction helps to see the different strategies phenomenological inquirers should adopt at different stages of an inquiry. The stage of collecting co-researchers’ descriptions of the phenomenon requires that the primary inquirers should persist in seeking clarity and detailedness in a co-researcher's (participant's) reports. The emphasis is on *precision*. Later at the stage of analyses, the researcher, collaborating with a co-researcher (participant) in constructing a consensual interpretation, should persist in seeking openness and variations in the interpretation. The emphasis is on *variation* and *consensus*.

Eisner (1991) endorsed constructivist approaches and suggested three ways for them to ensure validity and reliability. They are *structural corroboration* (i.e., triangulation of data type and the consideration of disconfirming evidence and contradictory interpretations), *consensual validation* (triangulation of viewpoints of observers or co-researchers) and *referential adequacy* (the capability of providing references to observations, inferences and conclusions — references which exist in the consensual, experiential worlds of the co-researchers) (p. 110).

I obtained different data types in this thesis study, and following Eisner (Ibid.), I performed structural corroboration on the data I collected in the study. What I obtained
from the instructors in the interviews regarding their beliefs and teaching practice was checked against my observation in their classes, and against their recalls of specific errors. Also, about 10% of the transcriptions in the study was verified for accuracy by a second transcriber. This will be reported in detail in a later section.

The consensual validation was carried out at two levels, one at the level of description, and the other at the level of interpretation. Individual instructors were all given the transcription of their interviews to verify and respond to. They were also given an interpretation of mine, to which they also responded. This will also be reported in detail later in the section describing my data treatment and analyses.

3.2.4. The form of the study

The present study was carried out in a form Patton (1990, p.191) called a mixed-form paradigm, using naturalistic inquiry, qualitative data, theme analysis and some simple statistical tallies. I had no hesitation to use statistics as a tool in a context-rich study. In an earlier unpublished paper (Tong, 1996) evaluating a language collocational study that I completed for a course at OISE/UT, I unfolded my own understanding of this mixed-form paradigm and argued for the inclusion of quantitative measures in context-rich, qualitative studies from a constructivist perspective:

My position now is that we must look at the issue from a new angle: the experimentalists and quasi-experimentalists need not have to actually put on a positivist hat. The experimentalists and quasi-experimentlists have a long and solid tradition and have established a set of readily accepted methodological procedures. Of course, I see it a myth to think that following this set of procedures faithfully, we will have objective truth. But if only we recognize that this methodological tradition is no more objective or subjective than any other acceptable methodologies, we should not be afraid to use it as one of our many tools. The attraction of statistical analyses over other context-rich methods is that it comes with a set of established
criteria and practice; hence, evaluation, interpretation and judgement about the quality of a study are always accessible and relatively easier.

A more recent example of this type of mixed-form inquiry in the symbolic interactionist tradition can be found in Wertsch (1998). In his narratives-as-a-cultural-tool study of U.S. college students' accounts of the origins of their country, Wertsch used a mixture of qualitative and quantitative measures to examine the construction of main characters in the students' narratives of the birth of their nation. These measures were (1) frequency of mention; (2) patterns of agency (e.g., the ‘who” in a “who did what” statement); and (3) patterns of presupposed presence (p. 91). The results of Wertsch’s study were also presented both qualitatively and quantitatively.

3.3. The context of data collection

The bulk of the data of this study were obtained between January 1998 and July 1998 from four ESL language instructors and their students in one of the two degree-compulsory ESL technical writing courses at the Faculty of Engineering in a university. However, the very last follow-up response from one of the instructor participants ended in July 1999. This particular university is situated in a major Canadian city with a high percentage of immigrants in its population. At the time, a high percentage of immigrant students attended the Faculty of Engineering of this university. A policy stipulated that all entering students who had not obtained a score of 85% or above in English in their provincial high school academic course had to take a written English placement test.

2. According to a faculty newsletter published in September, 1999, a total of 900 students entered their nine undergraduate programs that year. Out of these 900 students, a total of 572 (who had not obtained 85% or higher in English in their provincial high school academic course) wrote an English placement test and at least 220 students in the test were determined to be needing ESL instruction.
administered by the faculty. From this placement test students were divided into three groups. The high ranking group from the placement test attended a technical writing course taught by graduate teaching assistants in the English Department of the university. The lowest ranking group, considered relatively weak even in general English academic writing, had to first take a term-long ESL English academic writing course before entering the ESL technical writing course. The middle group from the placement test, considered only proficient enough to benefit from a technical writing with an ESL focus, went directly to this ESL technical writing course, which in the Spring Term of 1988 was taught by five ESL instructors, four of whom were participants in this thesis study.

3.3.1. Justifying the selected context of the study

While the present context of the study was selected primarily because of its availability and the likelihood of voluntary participation by the instructors and students, it was also my belief that the constructivist perspective I proposed in Section 2.3.6 for error evaluation would be verified and exemplified in a variety of contexts, including the one in which the data of this study were collected.

I note that interactions such as error evaluations in classrooms can be very different from those outside classrooms, and students’ writing in a technical writing course is intrinsically different from engineers’ writing in their real working world. I take it that language error evaluation outside classrooms, where authentic communication is taking place, would make very interesting studies. Nevertheless, using a technical writing course such as the one in the present study is an exploration of a legitimate error perception phenomenon, given that classroom teaching and learning, like many other human endeavours, can be seen itself as a genuine, social activity, authentic in its own way. Put differently, it is a reality that in L2 assessment (a context this study explored), an ESL
instructor has to identify and evaluate errors and to determine error gravity in order to assign a score to a piece of student’s written work. Hence the acts of instructors perceiving ESL errors in assessment and in giving feedback are authentic, social-interactional acts in their own right.

The nature of the course I selected for the study also afforded a probe into the assessment activities in a technical written communication course, which is somewhat different from that in many English academic writing courses. For instance, there have been various studies exploring the use of an analytic assessment scheme for ESL academic writing (Carlisle & McKenna, 1991; Hamp-Lyons, 1991; Vann, Lorenz & Meyer, 1991; Vaughan, 1991); to date, however, very little research has been carried out to explore the activities of L2 assessment of technical writing, a kind of English for Specific Purposes (Hutchinson & Waters, 1987) or English for Science and Technology (Trimble, 1985). The course in which the data of the present study were collected provided a valuable opportunity to reveal the relationship between technical writing assessment and error gravity.

3.3.2. The participants

At the time of this study, the eight sections (each consisting of about 16 students) of this technical writing course for non-native speakers were taught by five instructors. In December 1997, I began to communicate with all five instructors teaching the course. I explained to them that the key participants in my study would be the instructors themselves, although after obtaining their consent, I would still have to secure consent from their students as well. In my explanation, I also stated explicitly my ethical obligations in the data collection procedures, and I requested their voluntary participation. At this point, I suggested that this would include eight 2-hour sessions of classroom
observations, six one-hour instructor interviews, four hours of their reading and validating my descriptions of their percepts, and five hours of their reading my final analysis.

*Teacher participants.* In January 1998, after the instructors had discussed my invitations, the course co-ordinator gave me administrative consent to conduct the research, however, with a reduction on the number of hours the instructors could participate in the study. Table 3.5 summarizes what this collective, common consent included. In addition to what I specified in this summary, three of the four consenting instructors upon my request also agreed to let me audio-tape their teaching sessions. The fifth instructor, whom I talked to only once over the telephone and have never met, declined my invitation for personal reasons.

Table 3.5. Summary of the teacher participants’ consent to participate in the study

<table>
<thead>
<tr>
<th>Data Collection Procedures</th>
<th>Purpose</th>
<th>Total Time (Estimated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom observations</td>
<td>4 - 8 sessions of classroom observation</td>
<td>8 hours</td>
</tr>
<tr>
<td>Instructors’ Interviews</td>
<td>3 interviews, each lasting no more than an hour.</td>
<td>3 hours</td>
</tr>
<tr>
<td>Validating descriptions</td>
<td>Reading and validating descriptions and reflections of teaching and assessment activities</td>
<td>3 hours</td>
</tr>
<tr>
<td>Follow-up feedback</td>
<td>Reading and giving feedback to description</td>
<td>1 hour</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>15 hours</td>
</tr>
</tbody>
</table>
The four consenting instructors signed back a copy of the consent letter I had given them for their consideration. I present a copy of this letter in Appendix A. Prior to their individual consent, I had already obtained administrative consent from the coordinator of the course, who was also one of the four consenting instructors. Appendix B displays a sample of this letter. Table 3.6 gives a profile of the four teacher participants in the study, whom I have called Anna, Edward, Kelly and Marge. All teacher participants were native speakers of English. Anna was born in Europe and emigrated to Canada when she was about 8 years old. Edward emigrated to Canada in his teens. Kelly was born and grew up in Canada. Marge grew up in an English speaking country and moved to Canada in the 70's.

Table 3.6. Profile of the teacher participants in the study

<table>
<thead>
<tr>
<th>Pseudonym</th>
<th>Gender</th>
<th>Age</th>
<th>Highest education level</th>
<th>Years of language teaching</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>Female</td>
<td>Mid 30s</td>
<td>Ph.D. in Second Language Education (SLE)</td>
<td>&gt; 10 years</td>
</tr>
<tr>
<td>Edward</td>
<td>Male</td>
<td>Early 40s</td>
<td>Doing an M.A course in SLE</td>
<td>&gt; 4 years</td>
</tr>
<tr>
<td>Kelly</td>
<td>Male</td>
<td>Early 40s</td>
<td>M.A. in SLE</td>
<td>&gt; 8 years</td>
</tr>
<tr>
<td>Marge</td>
<td>Female</td>
<td>Early 60s</td>
<td>Ph.D. in Applied Linguistics</td>
<td>&gt; 30 years</td>
</tr>
</tbody>
</table>

Student participants. Student participants in this thesis study were either immigrant students or visa students. The vast majority of the participants were native speakers of Cantonese or Mandarin/Putonghua, and they came from either Hong Kong, China, Taiwan, Malaysia or Indonesia. In every section there were also often three to four
students from Korea, Japan, or Sri Lanka, and one or two from Eastern European countries such as Poland and Romania.

During the first and second week of the Spring Term of 1998, I visited each of the classes of the consenting instructors, explained to their students the purpose of my study and my being there. I gave each of the students a letter of invitation (See Appendix C). The consent I requested from a student included his or her allowing me to sit in and audio-tape or take note of the class on condition that I would not interrupt the lesson, and granting me the right to have access to all the written work he or she submitted to their instructors for assessment and feedback. The students indicated their consent or disagreement in a reply slip attached to my invitation letter. All the students expressed consent to my sitting in class and audio-taping the lesson. A total of four students from different classes denied my request to have access to their written compositions or exercises. I noted the names of these students and I excluded their written work from my data set; I did not make copies of their written work and I did not use them for either the instructor interviews or for my own analyses.

3.3.3. Data types

My data collection began with classroom observation during the second week of the Spring Term, 1998. I collected four types of data in this study. The first type of data were classroom observations. These data appeared as audio-tapes in the cases of Anna, Edward and Kelly, who had all agreed to my audio-taping. I also sat in on two of Marge’s sessions and took field notes; she did not agree to my taping her classes. These data provided me with direct percepts of the context of teaching, the instructors’ teaching agenda, their concerns, the climate of learning of these students and their needs and weaknesses. These data were used to crosscheck with the data I obtained from the
instructor interviews, especially with their recalls of specific serious errors. I discuss this data type further in Section 3.4.1.

The second type of data was the instructor-marked, written assignments or practice activities written by the students. These data were passed on to me by the instructors. Appendix D presents a summary of the students' text types I collected via the instructors. Selected texts from these data were used as stimuli in the stimulated recalls. I often referred to these data when I was transcribing the instructor interviews. These data also provided a basis for certain textual analyses, which I did to inform my interpretation of the instructors' percepts. I discuss these data in Section 3.4.2.

The third type of data consisted of instructors' stimulated recalls of specific errors in the learners' written samples. These data were actually instructors' recalls of their error evaluation experiences. I elicited them in the first two of the three 1-hour interviews each of the instructors gave me, and they provided me with the primary evidential bases for my descriptions and interpretations. These data included individual instructors' percepts of errors, their error evaluation procedures during their assessment procedures, and specifications of the kinds of errors they considered as serious, minor, or irritating. I later imported these data into a computer software package for qualitative research, QSR NUD*IST Vivo (NVivo) 1.0 for coding. Based on the language the instructors used in their recalls of specific errors, and the four major categories in the assessment scheme the instructors were all using (See Appendix E for a copy of the assessment scheme the instructors were using), I generated a list of error labels in the NVivo 1.0 software program to code the errors each of the four instructors commented upon. I discuss this in Section 3.4.3.

The fourth type of data was the instructors' general descriptions of their own beliefs, assumptions and knowledge of language teaching and of the course they were
teaching, the kind of students they had and their general views on language teaching and learning. These data allowed me to interpret the instructors' beliefs, assumptions and knowledge (their BAK systems\textsuperscript{3}) and to relate them to their percepts of errors and error gravity. These data were later also imported into NVivo 1.0 and then coded in a theme analysis. I discuss this in Section 3.4.4.

Table 3.7 gives a summary of the four data types, the methods of collection, data treatment, the analyses performed and the uses of the data in relation to my research questions.

---

3. \textsuperscript{3}Woods (1996) coined and defended the use of the acronym 'BAKs' to refer to teachers' beliefs, assumptions and knowledge in relation to their teaching.
Table 3.7. Summary of data types, methods of collection, treatment, analysis, and their uses

<table>
<thead>
<tr>
<th>Data type</th>
<th>Method of collection</th>
<th>Data treatment</th>
<th>Data analysis</th>
<th>Data use</th>
</tr>
</thead>
<tbody>
<tr>
<td>(A) Classroom observations</td>
<td>Audio-taped or summarized as field notes</td>
<td>Reviewed during transcribing and used to cross-check with other data types</td>
<td>Used for cross-checking data types and arriving at an understanding of the teaching agenda of the instructors</td>
<td>Provided the contexts of learners’ error production and instructors’ teaching and error corrections</td>
</tr>
<tr>
<td>(B) Marked written samples from learners</td>
<td>Collected from the instructors</td>
<td>Used as stimuli in the stimulated recalls to obtain instructors’ recalls of specific errors</td>
<td>Referred to during transcription to identify the error an instructor commented upon</td>
<td>Allowed me to examine the genre and the linguistic contexts of errors to compare with instructors’ error perceptions</td>
</tr>
<tr>
<td>(C) Instructors’ recall of specific errors in the learners’ written samples</td>
<td>Instructor interviews: Stimulated recalls</td>
<td>After being transcribed, these were verified by the instructors</td>
<td>Coded in NVivo 1.0 to generate statistical counts of serious errors and their types: content, organization, vocabulary or accuracy</td>
<td>Provided answers to the question of error gravity; was used to check against the general description data</td>
</tr>
<tr>
<td>(D) Instructors’ general descriptions of own beliefs and practice</td>
<td>Instructor interviews: General descriptions</td>
<td>After being transcribed, these were verified by the instructors</td>
<td>Imported into NVivo 1.0 and coded in a theme analysis to provide descriptions and interpretations of individual instructors</td>
<td>Instructors’ beliefs, assumptions and knowledge in language teaching (BAKs) provided bases for my interpreting instructors’ recalls of specific errors</td>
</tr>
</tbody>
</table>
3.4. Data collection procedures and data analyses

The data collection in general progressed uneventfully, though one difficulty I experienced in the process was time pressure, described below. In the following I report the data collection procedures according to the data type.

3.4.1. Classroom observations

Each of the consenting instructors suggested some possible dates for my classroom observations and asked me to choose some possible dates. However, the difficulty I had was that the two 2-hour meetings per week of all the classes of this course were scheduled to take place between 12 p.m. and 2 p.m. on the same days. This meant that on any day I could visit at most only one instructor's class. The other constraint was that there were assignment sessions in the course and during these sessions, the students did nothing but write and the instructor did not do any teaching; hence, making classroom observations on these days would be of little use to this study. Because of these constraints and because sometimes the sessions did not last the full 2 hours, I was not able to obtain a full 8 hours of observation time from Kelly and Marge. Of the four instructors, I had more observation hours in Edward's classes than I had in the classes of the other instructors because he kindly granted me the permission to visit any of his classes at any time and on any day provided that of course that I caused no interruptions. Because Edward had granted this permission, I was able to complete a total of over 33 hours of class observation time, though about half of this time was in Edward's class. The specific hours I observed classes for this thesis study are given in Table 3.8.
Table 3.8. Summary of classroom observations by instructor

<table>
<thead>
<tr>
<th>Instructor</th>
<th>Duration of observation time</th>
<th>Format of data</th>
<th>Total duration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>7 hours and 40 minutes</td>
<td>Audio-taped</td>
<td></td>
</tr>
<tr>
<td>Edward</td>
<td>17 hours and 55 minutes</td>
<td>Audio-taped</td>
<td></td>
</tr>
<tr>
<td>Kelly</td>
<td>4 hours and 9 minutes</td>
<td>Audio-taped</td>
<td></td>
</tr>
<tr>
<td>Marge</td>
<td>3 hours and 30 minutes</td>
<td>Field notes</td>
<td>33 hours and 14 minutes</td>
</tr>
</tbody>
</table>

These data helped me gain a solid feel of what was going on in the course and a good understanding of the teaching agenda, the kind of learners and the learning climate, and the contexts in which learners produced their language errors.

3.4.2. Marked written work

I collected a total of 331 pieces of students' written compositions via the four instructors. A total of 285 pieces of these compositions were assignments. An assignment was a piece of student's written work which was awarded a mark that would be counted towards the final course grade. Appendix F gives a sample of a prompt for an assignment task and Appendix G shows a page of student's work for an assignment. The assessed written work which I collected via Edward and Marge included 46 practice activities. A practice activity was a student's written work which was marked and corrected by the instructor but was not given a mark, or even if it was, the mark was not counted towards the final course grade. The use of assignment tasks and practice tasks provided my study with two different conditions for the context which I have called "the purpose of the
(evaluation) activity”. When eliciting Edward’s recalls of specific errors, I used samples from both types of tasks for him to comment on, to see if there were any differences in his comments. The only other instructor I collected practice tasks from were ten argumentative articles from Marge, which were actually a diagnostic test Marge gave her students on commencement of the course. I did not use these articles for recalls of specific errors because they were not technical writing tasks.

Since the course aimed to teach technical writing, the syllabus of the course contained several text-types the students had to learn to write. The course began with a diagnostic test of students’ ability to write an argumentative essay, a text-type most students were familiar with prior to attending this course. The course then moved on to the teaching of several text-types engineers are supposed to produce on their jobs: Incident report, technical instructions, inspection report, instructions manual, progress report, investigation report, and evaluation report. The tasks students were given usually consisted of a one-to-three page prompt containing not only all the contextual clues and information necessary for the writing of the required report but also some irrelevant or unnecessary information to train students’ ability to discriminate the relevancy of information.

In addition to accurate and appropriate use of language, the success of the writing tasks therefore required that the written compositions appear in a suitable genre presenting all the relevant textual features required of this genre. For example, in the case of a set of technical instructions for the installation of a piece of equipment, the required documents should consist of (a) an “Introduction” or an introductory paragraph, stating the primary use or function of this particular piece of equipment, (b) a “Materials and tools” section, listing all the needed materials and tools for installation procedures, and finally (c) a section called “Instructions”, giving all the numbered steps one by one to complete the installation. In the case of an incident report, the expected structural elements, as the
coursebook suggests (Blicq & Moreno, 1998, p. 86), should include: (a) summary, (b) background (c) facts and events and (d) outcome.

I note also that while the student compositions were supposed to appear in the appropriate, required text-type, the prompt they had been given for the writing task often appeared as a narrative, often containing extracts of conversation. A sample of the prompt for students' tasks is displayed in Appendix F.

It was only right that the instructors wanted to return to their students the assessed written work as early as possible; however, this sometimes left me little time to gain access to the marked scripts, particularly when marking had been done only shortly before they were returned to the students. Hence, in general, it was really the availability of marked scripts from the instructors that determined which written work I had access to. Nevertheless, I should note that it was not necessary for me to have collected all the students' scripts since I had more students' written samples than the instructors could cover in their two 1-hour stimulated recalls.

3.4.3. Recalls of specific serious errors

The major parts of my data came from the instructor interviews. As can be seen in Table 3.7, the two types of data I obtained via the instructor interviews were instructors' recalls of specific errors and instructors' beliefs, assumptions and knowledge of language teaching and the course they were teaching. The instructors' recalls of specific serious errors always took place in their first two interviews. The four instructors' beliefs, assumptions and knowledge of language teaching and the course they were teaching were all elicited in the last of their three 1-hour interviews.
According to Tuckwell (1980), instructors’ recalls of specific phenomena should be elicited as soon as possible after the instructors’ assessment procedures. On the other hand, I also respected instructors’ intentions to return their assessed student work to the students as soon as possible after doing the assessments. This meant that I had to rely on the instructors to make adjustment to their schedules so that these interviews could take place shortly after the assessment and right before returning them to the prospective students. Because of great efforts on the part of the instructors, I was able to elicit this part of the data successfully. The instructors provided me with the students’ assessed written work or the photo-copies of these marked scripts, which had already been returned to the students. The time between instructors’ actual assessments and the elicitations of recalls usually ranged from 20 minutes to 2 days. During their recalls of specific errors, I did not detect cases in which the instructors experienced any difficulties in making their recalls. I present in Table 3.9 some of the cue questions I used in this part of the data collection process.

### Table 3.9. Samples of cue questions eliciting instructors’ recalls of specific errors

**Sample 1: A cue question I asked Anna:**

In that order, [I passed the photocopies of Anna’s marked scripts to her: Scripts: A-4-1 to A-4-23] I hope, if you could primarily talk about your perceptions of the time, your judgements and specially when evaluating the seriousness of errors. You could relate them to specific students.

**Sample 2: A cue question I asked Edward:**

Now uh, basically I- what I would like you to do is to recall your perceptions at the time when you encountered errors in essays, and not- probably not much on a change of perception… I mean, if you see things that you did not see before, that’s fine too. But that wouldn’t be my focus, and also I sort of I never- there is not what you would call a “hidden agenda” in these things. It’s just- I just randomized the order of these reports a little bit. There are about 20, and it would be good, if we could go through these 20 scripts.
Sample 3: A cue question I asked Kelly:

Okay. Let me remind you that this is basically a recall interview. That is, I would like to ask you to recall your perception and feelings at the time when you encountered errors in the texts. By perception I mean the feelings, judgements, evaluation, your decisions, why you made that sort of decision, why you made certain responses, all right. And no names will be mentioned, only the only the uh only the numbers? So could I ask you to look at K-1-001? Uh, and please tell me the kind of-

Sample 4: A cue question given to Marge:

Now actually I would like you to recall although I don't mind you re-evaluating, you know, the errors now. But I would like you to recall and say say your feelings and your judgement, your attitude, your mood when you saw these errors. I particularly would like you to talk about grave errors, serious errors, and errors that disturbed you or irritated you.

It was my intention to elicit from each of the instructors a similar number of specific errors from their students. However, after my first interview with Anna, who was also the first instructor I interviewed in the data collection process, I realized that the number of specific errors Anna had covered was relatively small. I tried to raise this number during the second interview, but overall my attempt was not too successful. This may have had to do with Anna's personal conversational style; she not only spoke eloquently but also went into length and depth on every issue and provided vivid examples and stories. However, in the end, the number of specific errors she was able to cover in her two 1-hour interviews was still relatively lower than that covered by the other instructors. In Table 3.10 I present a summary of the number of assessed scripts and the number of specific errors each of the instructors covered in their interviews.
Table 3.10. Frequency of assessed scripts and errors recalled in the instructor interviews

<table>
<thead>
<tr>
<th>Instructor</th>
<th>No. of assessed scripts covered in the instructor's recalls</th>
<th>No. of specific errors covered in the recalls</th>
<th>Duration of specific error recalls</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>4</td>
<td>36</td>
<td>95 minutes</td>
</tr>
<tr>
<td>Edward</td>
<td>33</td>
<td>153</td>
<td>155 minutes</td>
</tr>
<tr>
<td>Kelly</td>
<td>24</td>
<td>100</td>
<td>135 minutes</td>
</tr>
<tr>
<td>Marge</td>
<td>13</td>
<td>94</td>
<td>110 minutes</td>
</tr>
<tr>
<td>Total</td>
<td>74</td>
<td>383</td>
<td>8 hours and 15 minutes</td>
</tr>
</tbody>
</table>

3.4.4. General descriptions of instructors' beliefs, assumptions and knowledge

I elicited from individual instructors information about their personal backgrounds and education, their own language learning experiences, their beliefs in language teaching and learning, and their general views on errors and error gravity. Table 3.11 consists of verbatim transcriptions of sample questions I used during the interview to elicit these data. In general, I attempted to elicit clarity and detail in participants' descriptions, as advised by phenomenological researchers (Giorgi, 1992; Van Manen, 1990). To this end, the instructors were given considerable freedom in their interviews to talk about whatever and in the manner they wished, although I had prepared a set of common questions to refer to in all the instructor interviews. During these semi-structured interviews, I looked at my list of questions from time to time to make sure that I had covered all that I had wanted to ask.
Table 3.1 Sample questions used to elicit descriptions of instructors' beliefs, assumptions and knowledge about language learning and their teaching

A. Instructor's educational background and other languages they spoke

- Can you say something about your training, your education and what you have been doing uh since graduation?
- Do you speak other languages, other than French?
- Do you speak German?

B. Instructor's teaching experiences and beliefs

- Before teaching this ESL course, did you teach other courses?
- Do you enjoy teaching in this program?
- You taught other ESL courses before. How do you compare this one with the other ones you had?

C. Instructor's own language learning experiences and beliefs:

- So what is your view on language learning? Like, what are some of the more effective ways of learning a language, in general? Do you want to talk about it a little bit?
- Some factual questions like- I understand that you are teaching linguistics and at the same time also teaching pedagogy, but now I am interested whether other languages, what the other languages you have acquired or learned? What are these languages?
- In your view, what makes good teaching?

D. Instructor's understanding of the objectives of the present course and his or her students

- Can I ask you to give a general description of the program you are teaching, like what kind of students do you have?
- So in this regard, how important do you see that you should understand your students, individual students in the teaching?
- When you mentioned that you want your students to understand the language, what does that mean? Can you elaborate further?
E. Relating language learning to language teaching

- Now if I ask you to connect your own learning experience of English with your own teaching, how would you make that connection, in terms of—do you see any link—obvious link or strong link between your own learning experience and your beliefs in language teaching?
- Do you think that your own learning experience of a foreign language such as French, has anything to do with your teaching philosophy now?
- Do you think the two [your teaching philosophy and your language experience] are related? Or is it more related to what you know about second language education?

F. Instructors’ percepts of error

- When you see errors in this category, for example, do you feel sometimes irritated?
- Maybe my last question is: what is error in your view?
- What role do you think in general errors play in the context of learning and teaching a language?

G. Instructor’s own evaluation policy or preferences

- I guess my second question, do you try to mark or identify every error you see in students’ scripts?
- Yeah, but, so so—when you when you’re marking for a score, do you find yourself trying to pick out every mistake?
- Do you also think that you don’t mark, you don’t mark or respond to article errors because they are not as serious? Would you say that? Is that a=

3.4.5. Data treatment

The instructor interviews were audio-taped and transcribed according to a standard set of conventions (shown in Appendix H). I exercised care during the transcription process and cross-checked these data with my field notes, the audio-tapes of the classroom observation, and copies of students’ written work to ensure accuracy at obscure points. To check the reliability of my own transcription from the audio-tapes, I passed two 1-hour tapes of the instructor interviews and a hard copy of the corresponding
transcription to an independent verifier. The verifier was a non-native speaker of English, formerly an ESL instructor but at the time a Ph.D. sociolinguistics student attending a US university in Washington, DC. Out of the 42-paged, single-lined transcription, the verifier identified a total of 15 minor discrepancies between my transcription and the audio-tapes she had listened to and interpreted. These were mostly typographical errors, articles, prepositions, and missing fillers. None of the discrepancies identified indicated major differences in understanding or serious distortions of meaning.

I then gave the transcription of individual instructors' interviews, together with a description I had produced of individual instructors' error percepts and some follow-up questions to individual instructor participants for their verification. Appendix I is a page of a sample of an instructor's feedback chart to my transcriptions. Based on the verified transcriptions, I formulated a description of each instructor's percepts of errors and some follow-up questions for each of the instructors to respond to. Participants' answers to the follow-up questions subsequently led to further amendments and supplements to the original descriptions and transcription. A sample of a description of an instructor's percepts and some follow-up questions I produced is given in Appendix J. A sample of this instructor's follow-up responses to my description and interpretation of their perceptions is given in Appendix K.

Based on the responses from the instructor participants, I revised the transcriptions and then imported them into QSR NUD*IST Vivo 1.0 (NVivo) (1999) (Licensed to me) for a computer-assisted theme analysis of the general descriptions of instructors' perceptions and the coding of specific errors commented upon by the instructors. The theme analysis assisted in organizing my data and formulating my interpretations of the instructors' percepts; the coding of specific errors the instructors commented on provided quantitative measures of serious error types responded to by individual instructors.
To ensure reliability, I performed triangulation procedures on the different data
types by comparing the data from different data sources. For example, from the recall of
specific errors, I detected that Kelly’s perception of error was very much structured by his
teaching advice and teaching points. In order to verify this observation, I looked at my
field notes and listened to the classroom observation tapes. I eventually found that Kelly’s
teaching indeed differed from the other instructors in that prior to an assignment he
particularly highlighted and gave advice against certain errors. When these errors came up,
he marked them in the recalls as serious errors. In my last interview with Kelly, I brought
up my observation that his assignments of error gravity was closely linked to the advice
and instructions he gave every time before the written assignment. Kelly confirmed this
observation.

What the instructors said in the general description was also used to cross check
with their recalls of specific errors. This use of data triangulation enabled me to explicate
instructors’ knowledge about themselves. In the interview, when asked about his marking
strategies, Edward expressed that he did not respond to article errors as a policy because
he felt that the English article system was too complicated to make any feedback useful. In
spite of his own perceptions of himself, to his surprise, the statistical counts of specific
errors indicated that he responded to article errors quite frequently. Edward explained that
he made exceptions to the rule when he detected that a student was relatively more
advanced and therefore might benefit from his marking this aspect of their English usage. I
discuss further some of the discrepancies existing between different data types in Chapters
Four and Five.

3.4.6. Analysis of descriptions of instructors’ beliefs, assumptions and knowledge

I subjected instructors’ general descriptions of themselves and their teaching to a
theme analysis assisted by the PC software program NVivo Revision 1.0, coding these as “free nodes”. The coding process was very much a ground-on-data procedure. During this process, I was all the time aware of the need to code the data as much as possible from the perspective of the individual instructors. This means that coding was done according to the language individual instructors used to describe and evaluate the error under discussion, although eventually the coding was my own interpretation of what they were actually saying. What was helpful was that the instructors not only explicitly identified the type of each of the errors they recalled according to the common assessment they all used, but also often explained clearly why they considered an error to be serious. I explain this in detail in Section 3.4.7. In the case where I suspected that there were slips, I went back to the instructors and asked for verification in the instructor interviews and in the follow-up questions; and these slips were corrected in their reading of the hard copies of the transcription.

In the theme analysis of the general descriptions, I performed all three methods suggested by Van Manen (1990, p. 94).

(1) Wholistic method: I read the whole text to look for themes at a macro level. Because the transcriptions contain conversational exchanges, this method allowed me to look at not only each conversational turn as a unit but also meanings conveyed over more than one turn.

---

4. A node in the NVivo Revision 1.0 program is an object in a project which is used to represent anything that the project users may wish to refer to, such as constructs, people being studied, concepts, places, mental states, features of the research project, etc. Some nodes are grouped, connected, or nested. For example, the node of “gender” can be connected to two other son nodes at a lower level “male” and “female”, which are themselves unconnected. A free node is a stand-alone node, which is not connected to another node at the top level; a free node may have son nodes at a lower level. See L. Richards (1999).
(2) Highlighting method: I read any phrases that stood out. In this reading, I found that, in particular, notions that have been discussed in the error evaluation review were coded, such as error gravity, comprehensibility, irritation, error type comparison.

(3) Line-by-line method: I read past each sentence and each sentence cluster line by line very carefully though I found that this method was not particularly helpful.

The NVivo Revision 1.0 software program allowed me to use all these three methods efficiently whereas in a previous trial using the NUDIST 4.0 program, doing so had involved very cumbersome operations. This is because unlike NUD*IST 4.0 (an old version of NVivo Revision 1.0), which requires the insertion of a hard return at every unit boundary in the texts before they are imported into NUD*IST 4.0, NVivo Revision 1.0 mercifully allows coding to take place at any time and at any point in the texts. The analyses of these data generated a group of themes, each of which collated relevant text extracts for easy, effective perusal. (See Appendix L for a project summary and a list of themes generated in the analysis.) These themes facilitated my interpretations of the instructors’ general perceptions data and reliability checks. One reason for this is that the NVivo Revision program enabled one to not only browse, code (or de-code) at any time and any location in the transcriptions but also to collate, group or re-group any nodes or any sets of nodes which had already been created.
3.4.7. Analyses of the instructors' recalls of specific errors

The purpose of coding the specific error recalls was to compute the relative frequencies of different serious error types perceived by individual instructors. These data provided evidence for me to see the serious error types and their different weights in individual instructors' percepts and to compare differing percepts of error gravity of the instructors. I also used this result to verify findings from the general description data.

When my data collection procedures began in January, 1998, I already learned from the course co-ordinator that all the instructors should be using the same assessment scheme she had developed for the course. This assessment scheme is given in Appendix E. As can be seen in Appendix E, this assessment scheme contained four major categories: content, organization, vocabulary and accuracy. When I coded the specific errors in the instructors’ stimulated recalls, I inevitably structured their responses around this assessment scheme.

When commenting on errors in students' writing, each of the instructors usually identified an error and categorized it according to the assessment scheme (in Appendix E), and then finished by commenting on its severity. An instructor’s comment on a specific error typically consisted of some or all of the following moves, not necessarily in this order:

1. Identify the location of the error in the text.
2. Categorize the error according to the assessment scheme.
3. Mention the evaluation of error gravity (the seriousness of the error).
4. Give reasons why the error was considered serious or not.
5. Estimate the possible cause of the error.

Here are some samples of specific error recalls by different instructors:
Sample 1

Researcher: Okay, this one is number 7. [K-M-7].
Kelly: Okay, this has a lot. His vocabulary is, a lot of mistakes. And there are words that if you look on the second line, the second word of the second line of the second paragraph. Uh I think the word is "flawns". I don't know, I mean, [reading aloud the sentence] "Trees were broken and flawns were damaged."? I mean, I I have no idea. It's uh, and uh, is there any more? Uh, spelling mistake like here: "Cottact" for "Contact". [Text: "Cottact Bell Canada to repair the telephone line." ] [K-M-7. Para.332]

Sample 2

Anna: And the other thing I think that, okay I am not on this whole page, one thing I put down on the bottom is "What about the furnace, air-ducts?". We need to mention what works as well as what doesn't. For me, I guess, that's the one of- it's important because they need to know that, we talked about that, but they have to talk about what's okay as well as not okay, like not just the problem. In any investigation inspection or progress report, we've said that you have to say what's done, and what's okay, and then also what's not done, and so that's not there. So as far as content goes, that needs to be there too. [A-4-17, Para.582]

Sample 3

Edward: "E", one, six. Okay, right. "E", one, six. [E-1-6] Yes, of course, we're not allowed to say names. [laugh] Okay, the first comment is obviously related to the content in a way. I made a comment regarding uh the fact that the student had included all the details, all the essential details but did not provide, did not provide any conclusions. I mean, I think in this particular exercise, I was expecting huh a kind of judgement. The exercise, I don't know if you know the exercise at all, but the exercise concerns an incident where somebody was coming back from huh a meeting, and when he opened up his briefcase huh to work on his computer, realized that he had taken the wrong computer case at some point of somebody. So, but the point here is that the computer that he got was far more expensive than the one he had lost, and then therefore it's highly unlikely the computer was stolen on purpose. In other words, it was more likely to have been displaced. I expected the student to make a statement regarding that. So the comment in this case was regarding the the missing logic in this case. [E-1-6, Para.108]

I coded each of the error recalls according to how the instructors categorised each of the errors they commented on. In a few cases where I doubted whether the specific errors were accurately recalled (I had detected slips of the tongue in some of the instructors' recalls), I asked the instructors to verify them again in the follow-up questions.
Not all the errors an instructor commented upon were serious errors although I always asked an instructor to identify them in each of the students' texts. Occasionally, the instructor reported first a minor error, or an error, and then added that it was a minor error. In these cases, the concerned error was coded as a minor error (non-serious error), as in the following examples:

Kelly: Organization, one sentence has two steps in it, but that's not that serious. [K-1-008, Para. 122]

Edward: Uh, nothing serious, it's the same again. I said "snap the Mini-Minder in" instead of "Mount it". I changed the testing step again. [K-11-011, Para. 705]

Marge: Okay, I also noted a minor Grammar error. This student makes very few Grammar errors, in fact, a Tense error uh which didn't change the meaning, but it's carelessness on both parts really and so I pointed it out to her. [M-M-1, Para. 19]

Because NVivo Revision 1.0 does not require pre-set unit boundaries for data entry, the data treatment procedures for this thesis study required that I decided on a working definition of a unit for the coding. When coding this part of the data, I treated the first mention of a new error as the beginning of a unit, and this unit ends immediately before the mention of the next error. There were also cases where between the mentioning of two consecutive errors, the instructor made other remarks, e.g., comments on language teaching and learning in general and on the specific needs and strengths of individual students, etc. I did not treat these as units of analyses for this data type, but I did use these comments as part of the data for understanding individual teachers' BAKs. I used 16 content codes, 10 organization codes, 20 vocabulary codes and 25 grammatical accuracy codes, based on the language individual instructors used to evaluate these errors in their recalls, to code a total of 323 errors in 74 pieces of students' written documents.
I coded the stimulated recalls of specific errors and generated a list of error types (which is actually a list of tree nodes, as opposite to free nodes) in the NVivo Revision 1.0 program. I present this list in Table 3.12 and wish to explain it with an example. In Table 3.12, the first error type label under the category of content reads:

82 (1 1 12) /Error Type/Content/Ambiguous expression.

This expression Error Type/Content/Ambiguous expression indicates that there is a node called Ambiguous expression. This particular node, however, is located under another parent node called Content. While Content is the parent node of Ambiguous expression, it is itself the son of another parent node called Error type. Hence the error type labels (error names) are in fact located at the lowest level of a node tree, which has a top node called Error type. The numbers in the bracket (e.g., (1 1 12), (1 3 23)) in each label form a position indicator; it is just another way to indicate the parent-son relationships of the nodes. For example, the third number “23” in (1 3 23) for an error type called Adverbial means that this error type is the 23rd error type under the node Grammatical accuracy, which itself is the third node under the parent node Error type. This number (e.g., (1 3 23)) indicates therefore the relative position of an error label in the whole node tree. These position indicators are not in ascending or descending order in Table 3.12 because the list is in fact ordered alphabetically by node (label) name.

When coding, every time I identified an error type in an instructor’s stimulated recalls, I coded it using one of these labels in Table 3.12 based on how the error was described, identified and evaluated by the instructor. When I carried out my data analyses, I counted the number of errors I coded under each label (node). I then tallied the resulting frequencies.
To check the reliability of the coding I performed, I gave to a second coder the transcription of Edward's second 1-hour interview (which was selected randomly) together with the list of coding labels I had generated from my own coding. The transcription of this interview contained a total of 88 specific serious errors. The second coder was at the time a practising ESL instructor in Canada, a native speaker of English who holds an M.A. degree in second language education. Without prior discussion or negotiation, out of the total 88 errors, a total of 75 of the errors were identically coded by both this ESL instructor and myself, achieving an agreement rate of 83%.

Table 3.12. Error type labels generated based on the comments by the instructors on specific errors

<table>
<thead>
<tr>
<th>Content</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>81 (1 1) Error Type/Content</td>
<td>98 (1 3) Error Type/Grammatical Accuracy</td>
</tr>
<tr>
<td>82 (1 1 12) Error Type/Content/Ambiguous expression</td>
<td>99 (1 3 23) Error Type/Grammatical Accuracy/Adverbial</td>
</tr>
<tr>
<td>83 (1 1 3) Error Type/Content/Contradicting statements within text</td>
<td>100 (1 3 1) Error Type/Grammatical Accuracy/Article</td>
</tr>
<tr>
<td>84 (1 1 13) Error Type/Content/Conveying message that may cause undesirable reader behaviour</td>
<td>101 (1 3 1 1) Error Type/Grammatical Accuracy/Article/Definite article and relative clause</td>
</tr>
<tr>
<td>85 (1 1 8) Error Type/Content/Copying from prompt</td>
<td>102 (1 3 8) Error Type/Grammatical Accuracy/Comma splice</td>
</tr>
<tr>
<td>86 (1 1 6) Error Type/Content/Did not consider real world constraints</td>
<td>103 (1 3 21) Error Type/Grammatical Accuracy/Concord agreement</td>
</tr>
<tr>
<td>87 (1 1 14) Error Type/Content/Fact mentioned in one section requires mentioning in another section</td>
<td>104 (1 3 10) Error Type/Grammatical Accuracy/Connector</td>
</tr>
<tr>
<td>88 (1 1 9) Error Type/Content/Failure to present info and argument logically and schematically</td>
<td>105 (1 3 16) Error Type/Grammatical Accuracy/Coordination</td>
</tr>
<tr>
<td>89 (1 1 7) Error Type/Content/Inaccurate information</td>
<td>106 (1 3 11) Error Type/Grammatical Accuracy/Count Non-count noun</td>
</tr>
<tr>
<td>90 (1 1 4) Error Type/Content/Information contradicting the prompt</td>
<td>107 (1 3 15) Error Type/Grammatical Accuracy/Fragment</td>
</tr>
<tr>
<td>91 (1 1 2) Error Type/Content/Irrelevant or unnecessary information</td>
<td>108 (1 3 18) Error Type/Grammatical Accuracy/Gerund</td>
</tr>
<tr>
<td>92 (1 1 11) Error Type/Content/Lack of necessary emphasis</td>
<td></td>
</tr>
<tr>
<td>93 (1 1 5) Error Type/Content/Missing elaboration</td>
<td></td>
</tr>
<tr>
<td>94 (1 1 1) Error Type/Content/Missing information</td>
<td></td>
</tr>
<tr>
<td>95 (1 1 1 1) Error Type/Content/Missing information/Background</td>
<td></td>
</tr>
<tr>
<td>96 (1 1 10) Error Type/Content/Wrong focus</td>
<td></td>
</tr>
<tr>
<td>Chapter Three</td>
<td></td>
</tr>
<tr>
<td>---------------</td>
<td></td>
</tr>
<tr>
<td>109</td>
<td>(1 3 13) /Error Type/Grammatical Accuracy/Modal</td>
</tr>
<tr>
<td>110</td>
<td>(1 3 2) /Error Type/Grammatical Accuracy/Preposition</td>
</tr>
<tr>
<td>111</td>
<td>(1 3 9) /Error Type/Grammatical Accuracy/Pronoun</td>
</tr>
<tr>
<td>112</td>
<td>(1 3 19) /Error Type/Grammatical Accuracy/Relative Clause</td>
</tr>
<tr>
<td>113</td>
<td>(1 3 20) /Error Type/Grammatical Accuracy/Run On</td>
</tr>
<tr>
<td>114</td>
<td>(1 3 3) /Error Type/Grammatical Accuracy/Sentence structure</td>
</tr>
<tr>
<td>115</td>
<td>(1 3 12) /Error Type/Grammatical Accuracy/Subject of the sentence</td>
</tr>
<tr>
<td>116</td>
<td>(1 3 17) /Error Type/Grammatical Accuracy/Subordination</td>
</tr>
<tr>
<td>117</td>
<td>(1 3 4) /Error Type/Grammatical Accuracy/Tense</td>
</tr>
<tr>
<td>118</td>
<td>(1 3 14) /Error Type/Grammatical Accuracy/Topic-Comment vs. Theme-Rheme</td>
</tr>
<tr>
<td>119</td>
<td>(1 3 7) /Error Type/Grammatical Accuracy/Verb</td>
</tr>
<tr>
<td>120</td>
<td>(1 3 5) /Error Type/Grammatical Accuracy/Verb-Tense agreement</td>
</tr>
<tr>
<td>121</td>
<td>(1 3 6) /Error Type/Grammatical Accuracy/Verb-Voice</td>
</tr>
<tr>
<td>122</td>
<td>(1 3 22) /Error Type/Grammatical Accuracy/Wrong part of speech</td>
</tr>
</tbody>
</table>

**Organization**

| 123 | (1 2) /Error Type/Organization |
| 124 | (1 2 8) /Error Type/Organization/Failure to present info or argument schematically |
| 125 | (1 2 4) /Error Type/Organization/Info in a wrong organizational unit |
| 126 | (1 2 3) /Error Type/Organization/Missing organizational unit |
| 127 | (1 2 5) /Error Type/Organization/Not following the organization of a document type |
| 128 | (1 2 1) /Error Type/Organization/Organizational unit too long |
| 129 | (1 2 2) /Error Type/Organization/Too much info in an organizational unit |
| 130 | (1 2 9) /Error Type/Organization/Unnecessary organizational unit |
| 131 | (1 2 6) /Error Type/Organization/Use the organization of a different document |
| 132 | (1 2 7) /Error Type/Organization/Wrong ordering of information |

**Vocabulary**

| 133 | (1 4) /Error Type/Vocabulary |
| 134 | (1 4 1) /Error Type/Vocabulary/Ambiguous expression |
| 135 | (1 4 1 12) /Error Type/Vocabulary/Ambiguous expression/Ambiguous word |
| 136 | (1 4 18) /Error Type/Vocabulary/Awkward expression |
| 137 | (1 4 11) /Error Type/Vocabulary/Copying from prompt |
| 138 | (1 4 16) /Error Type/Vocabulary/Lack of cohesion |
| 139 | (1 4 10) /Error Type/Vocabulary/Lack of parallelism when necessary |
| 140 | (1 4 5) /Error Type/Vocabulary/Low information content word |
| 141 | (1 4 17) /Error Type/Vocabulary/Over-using a word |
| 142 | (1 4 4) /Error Type/Vocabulary/Plural marker |
| 143 | (1 4 13) /Error Type/Vocabulary/Register and formality |
| 144 | (1 4 3) /Error Type/Vocabulary/Spelling |
| 145 | (1 4 12) /Error Type/Vocabulary/Unnecessary politeness marker |
| 146 | (1 4 9) /Error Type/Vocabulary/Use of abbreviation |
| 147 | (1 4 2) /Error Type/Vocabulary/Using expression unknown to reader |
| 148 | (1 4 7) /Error Type/Vocabulary/Word choice |
| 149 | (1 4 6) /Error Type/Vocabulary/Word ending |
| 150 | (1 4 8) /Error Type/Vocabulary/Word formation |
| 151 | (1 4 14) /Error Type/Vocabulary/Word order |
| 152 | (1 4 15) /Error Type/Vocabulary/Wrong choice of expression |

- 93 -
I should note that it was not the purpose of this study to construct an objective, reliable labelling system for these “serious” ESL errors. The present coding scheme was generated from an emic perspective. It is therefore only natural that this labelling system contains some overlapping labels. For example, an instructor at one instance considered an item a vocabulary error because it was an ambiguous expression, but at another instance, the same instructor called another language item also “an ambiguous expression” but then categorised it as a content error.

3.4.8. Reliability and validity of the theme analysis

During the data analysing process, I carried out consensual validation with the instructors to revise and improve the transcriptions of the interviews and my descriptions of the instructors’ percepts of errors. Marge’s and Kelly’s follow-up responses to my descriptions were sent back to me by post and by email. (See Appendix J for a sample of my initial interpretation of Marge’s error percepts and the follow-up questions I posed. See Appendix K for a sample of Marge’s response to my follow-up questions.) The follow-up questions with Edward and Anna were primarily carried out in two independent, 45-minute meetings; these data were also audio-taped and transcribed. In general, the follow-up responses from the instructors did not show that I had formed misinterpretations of instructors’ percepts at even the initial stage. They helped achieve the following:

1. Clarified obscure points in the interview transcriptions.
2. Extended the instructor interviews by allowing me to ask further questions.
3. Provided different focuses and variations on my interpretation.
4. Brought into light instructors’ own tacit knowledge about themselves and their teaching.
5. Encouraged the instructors to reflect on their teaching and assessment strategies.

3.5. Summing up

To sum up the present chapter, I adopted a naturalistic, emic and context-rich approach to elicit data to answer the research questions. Although I experienced some difficulties during the process of my data collection, I exercised appropriate procedures to ensure the quality of this study. These procedures, the details of which I have reported, included structural corroboration (triangulation of data types), consensual validation (triangulation of observers/participants), and referential adequacy (quality documentation of observation, inferences and conclusions).
CHAPTER FOUR

FINDINGS

In Chapter Three I described the context of the present thesis study. The data were collected from and via four language instructors teaching an ESL, degree-compulsory, technical writing course to engineering students at a university in a major Canadian city. In Chapter Two I introduced the four research questions guiding this thesis study. These questions are:

Research Question No. 1: What is error in the perceptions of the four ESL instructors during their assessment activities?
Research Question No. 2: What are the relationships between errors and the contexts of perception?
Research Question No. 3: What are serious errors? What are the conditions with which they are identified?
Research Question No. 4: What are the relationships between instructors’ percepts of the errors and the pedagogic objectives and focuses of the course as they perceived them?

In this chapter, I present my findings to answer these four research questions in Sections 4.1, 4.2, 4.3 and 4.4 respectively.

Section 4.1 basically shows that the definitions of errors three of the instructors provided to my question “What is error?” were much narrower than the notions of error they actually utilized in their assessment activities. These narrow definitions conformed to one suggested by error analysts in earlier SLA research. They were, however, to varying degrees, amended by the instructors themselves in their later responses to my follow-up questions.
Findings in Section 4.2 demonstrate that while the assessment scheme prescribed for the course seems to have imposed a structure on the four instructors' percepts of errors in their students' written work, a host of contextual elements also influenced their error evaluation procedures. These contextual elements included the specific assessment purposes, the genre of the document in which errors were identified, and the instructors' repeated or prolonged engagement with the students' written texts.

Responding to Research Question 3, findings in Section 4.3 show that the criteria that these instructors consciously used in their error evaluation activities included comprehensibility, error frequency, and irritation. In the process of an error evaluation, however, the determination of error gravity was often affected by factors such as perceptions of student’s learning needs and strengths, or the genre of the document containing the error(s) under focus. A tally of the kinds of serious errors (e.g., grammatical errors, content anomalies, or discourse infelicities) individual instructors responded to shows that there were basically two different patterns, dividing up the four instructors equally into two groups according to the serious error types they responded to the most. Two of the instructors, Edward and Marge, responded more to grammatical errors than content errors, whereas the other two instructors, Anna and Kelly, responded more to content anomalies than grammatical errors.

My analyses performed to answer Research Question 4 in Section 4.4. suggest that the two different response patterns of the two groups of instructors were due to individual instructors’ beliefs, assumptions and knowledge (the BAK system) about their teaching. Put differently, the different response patterns of the two groups of instructors were related to what they considered as important objectives and teaching points in the context of the course and how they viewed error as a social phenomenon.
4.1 Research question No. 1: What is error in the perceptions of the four ESL instructors during their assessment activities?

This section reports individual instructors' views in relation to this research question then provides an overall summing-up of what error is in the perceptions of these instructors.

4.1.1. Edward

When the first interview started, Edward did not object to my using the term “error”. He had been using the term “error” himself 67 times during the recalls of specific serious errors (and had only four times used the term “mistake”) prior to my question on the definition of error, mostly referring to grammatical anomalies but also occasionally missing details, content anomalies, and missing organizational elements in compositions such as background or conclusion. However, when I asked him the very first time towards the end of the first interview to define error, he gave the surprising comment that he did not like to use the term “error” much:

Researcher:  Okay, my last questions, probably reveal more about my- what I am really interested in is In your view, what is error, what is error, as a- as a working notion in the context of what you teach and what you assess? Can you say something about what it is, in your view?

Edward:   Error, uh. I guess sometimes I am reluctant to use the term errors, simply because it's a value judgement. I tend to say things like- I would rather say "you have problem with this", "you have problem with that", or "you're weak in this area", "you're weak in that area". [Para. 373]

Edward gave an explanation for preferring not to use the term “error” though at the same time he felt that maybe he needed to:

Edward:   It's not uh, you know, "computer was missed" is not an error as, well it's an error, but I don't call it an error, I don't call it an error. Somehow it doesn't seem like it. It's just an imperfect expression. It's just using the wrong thing. I mean it's an error in a sense, but uh I mean, the fact is that he may be taking something from his native language, and just translating it to English. [Para. 379]
Later when I asked him once again in the second interview to define errors, Edward expressed that errors are violations of rules:

Edward: Yeah yeah yeah. Uh, what is an error? Well, simplistically, it'd be just incorrect application of rule or something like ignorance, ignorance of a particular rule I guess that in the end, is uh, that is what error is. If you want me to, that's what I would define it. [Para. 970]

In terms of interpreting the term “error”, Edward was different from all the other instructors in that he attached a religious connotation to the term “error” and explained that this may have to do with his Catholic background:

Edward: Well- error has a religious connotation. I think if you brought it in there, I am not a Catholic any more. But if you are brought up as a Catholic, an error is an error. I mean error is, you know, biting the apple in the Garden of Eden. That's an error. [Para. 388]

So Edward preferred to limit the notion of error to incorrect linguistic forms, and he had the opinion that content and organization anomalies should not be called errors:

Edward: I mean, I guess, to me errors are such that if I use the term "error", then I tend to regard its formal features as an error. So when I say error, then to me that would be almost like language error. But errors, I tend to associate more with incorrect form, I guess this is how I would- an error is what I would correct- would call an incorrect form. So that there might be- there might be difficulty in organization, there might be missing details, there might be missing details, but I don't see them as errors in that way. So errors. [Para. 374]

When I asked Edward if the term “error” was not preferred, how should anomalies be conceptualized, Edward suggested using the term “mistakes” or “imperfections”:

Edward: Imperfections, because I think right and wrong is- Error, I don't like the term, because to me error implies maybe it's my Catholic background, but an error is an error, as a sin, right? Then error is something that is wrong. So I don't like telling students that you commit fifteen errors, or something because- but to me errors, if you ask me what error is, I would say error to me is a linguistic thing, to me is an incorrect form. But I mean that I might be picking at details, and it's
certainly not the only thing I paid attention to, but error as such to me is that. [Para. 377]

Edward: An error is something that uh that is something you are supposed to repent or something. To me it's a much too value laden a term, that's why I don't want to use it for even grammar, I guess, "mistake" would be a better term in a way than an "error", but uh but that's my feeling. [Para. 389]

4.1.2. Kelly

Kelly's discussion on the conceptualization of errors revealed three key points.

1. Error is related to declarative knowledge and declarative knowledge only.
2. His students in general did not commit errors, but mistakes, which he felt are more serious than errors.
3. Errors and mistakes are different. Serious mistakes refer to particularly students' violations of rules or knowledge the teacher, Kelly himself, had previously taught.

As had happened in Edward's interview, I started Kelly's first interview with a cue question containing the term "error". All throughout most of the one-hour interview, Kelly did not seem to have any queries regarding the use of the term. Although tallies of words used in the interview transcription later did show that Kelly himself used the term "mistake" a lot more frequently (30 times) than "error" (9 times), he did not seem to intend a difference between the two terms, as shown in the following examples:

Kelly: O.k., serious errors. The person made basically one big serious error for the whole thing ... he put two instructions in one step- uh two steps in one instruction [Para. 18]

Kelly: Irritating, uh. No, this one is quite good. The only other serious error, in fact, is uh, a lot of people made it too. In the instructions, they say "take colour picture and make a copy". ......what the person needs to do the, what it's called? In the the equipment list, they forgot that colour picture. [Para. 34]

Kelly: Uh the other mistake the person did was, a lot of vocabulary errors. [Para. 46]

Kelly: Okay. This person made uh a lot of content errors. Uh, ... again, having uh too much information for each instruction. [Para. 99]
Similar to those examples I gave above for Edward’s recalls, these extracts show that Kelly was using the term “error” to refer to not only grammar anomalies and vocabulary anomalies, but also organizational and content anomalies. I was therefore surprised when towards the end of the same first interview Kelly replied this way to my next question:

Researcher: Can I ask you uh something more general in nature? Suppose someone comes up and says to you as a teacher “what is error?”, so how would you respond to that?
Kelly: An error or a mistake?
Researcher: Do you make a distinction? Do you distinguish between the two?
Kelly: Oh, yeah. An error is when someone does something consistently, or does something wrong. I guess, consistently. Mistake is they don’t do it right, like the person who got 60% of the articles right and 40% wrong, that’s a mistake. [Para. 190]

At this point, Kelly actually suggested a criterion to differentiate errors from mistakes:

Kelly: If the person missed every article, or missed every article or missed 89% of the articles, then I would say that’s an error, and they should, you know, let’s do- I’ve done it, I’ve done one-to-one work on errors. [Kelly, Para. 193]

These conceptualizations of errors and mistakes suggest that Kelly considered mistakes in the realm of procedural knowledge (performance deviations) and errors the results of ignorance or lack of proper declarative knowledge. Consequently, Kelly felt that his students did not make a lot of errors, only mistakes, as he expressed in the following:

Kelly: The thing is that most of my students don’t make errors, they make a lot of mistakes, but not that many errors. So if I do find errors, I do do one-to-one. Luckily, this is a high-level class, and I don’t find that many errors. [Para. 219]

In a verification response to my description of his perceptions, Kelly made a point that his distinction between errors and mistakes was not so much based on the declarative knowledge of the students but rather on whether he, Kelly himself, had already taught the specific feature. In other words, if an anomaly was something Kelly had already covered in the course, he would consider it a mistake and penalized the student more heavily. This is actually an observation he had already confirmed once during his third interview.
Second, he objected to my initial observation that he did not seem to react to mistakes and errors differently; he wrote: "I penalize mistakes much more than I do errors. Mistakes made me sad while I understand the cause of errors." He added: "Thus, I am most critical when a student makes a mistake on something that we have either discussed in class, or that I had repeatedly written down on his previous papers."

I therefore interpret this to be what Kelly was saying: The term "error" is reserved for improper declarative knowledge of rules, grammatical, organizational and content rules; "mistake" is a term for a procedural deviation. As an instructor teaching the course, he was primarily concerned with "mistakes", which to him were more serious than "errors".

4.1.3. Marge

In contrast with the way Kelly and Edward responded to the cue question, Marge showed her sensitivity to the scope and meaning of the term "error" right at the very beginning of the first interview, but like the other instructors, there were discrepancies in what she declared as the definition of error, and how she applied the term "error" in her language.

As soon as I asked Marge to begin commenting on the serious errors in her students' texts, she asked me to clarify whether I wanted her to comment on just grammatical errors, or any problems in the organization and the content categories:

Researcher:  Now actually I would like you to recall although I don't mind you re-evaluating, you know, the errors now. But I do like you to recall and say say your feelings and your judgement, your attitude, your mood when you saw these errors. I particularly would like you to talk about grave errors, serious errors, and errors that disturbed you or irritated you.

Marge:  Are you- do you want me to talk about the language, not talking about the content and the organization.

Researcher:  Well, I- actually mmm, I would like you to talk about anything that is not right.

Marge:  Well, okay, fair enough. So it could be any of those things, okay. [Para. 12]
However, there were contradictions in how she used the term. In the following, I see that she avoided calling a content anomaly an error:

Marge: Uh, yeah, the first thing, okay, the first thing I reacted to really was the fact the subject was not complete. This is not a grave error. This is a good student. I want to help her to improve in very precise ways. So uh, by reacting to the first so called error, it's not really an error, it's not a grammatical mistake. It's the fact that she is using too much detail in a situation, where conciseness is needed. [Para. 16]

In another instance, however, I found again that Marge kept using the term "mistake" to refer to an organizational anomaly, as in the following segment:

Researcher: Yeah, for example, the cue text, the prompt is obviously a narrative, whereas this is a more a functional report, afterwards [after event] report?

Marge: They are- because the textbook teaches it specifically, and directly, so yes, they are aware of it. That's in fact where the major thing is that we are testing. So in terms of error, if it is deviating from that form, they are in fact making an error at this point. Before they have been taught, this is the format of an inspection report, this is the format of an incident report. And even though they know that there are variations within that format, if they deviate overall from the format, it's a mistake. [Para 95]...

Here, I can see Marge in one way intending that error is a violation of a rule (a use rule or a morpho-syntactic rule) which the students had been taught and therefore should know but which they failed to execute. Later, in a follow-up verification, Marge confirmed that she used the two terms "error" and "mistake" interchangeably, and that the word "slips" to her would be related to carelessness. In her verification of my description of her perceptions, Marge also reaffirmed her view that errors are violations of linguistic (morpho-syntactic) rules, which are applied regardless of context, whereas genre and content anomalies are context dependent. This is a very interesting observation, and I will take it up further in Chapter Five. In her most recent response to my follow-up questions, Marge observed that there are language rules and writing rules, and she wrote:

I think my answer is that breaking the linguistic rules marks you as a non-native speaker, which may be more or less acceptable depending on the toleration of the host culture. The consequence of breaking the second set of rules results in a less effective or inappropriate piece of writing. It is a writing, not a language problem, and certainly not confined to
non-native speakers of English. [Marge, Follow-up response]

I interpret this message to mean that there are language errors and there are writing errors although in the case of a violation of a writing rule, Marge might prefer to say "an inappropriacy" rather than "an error".

4.1.4. Anna

During the interviews, Anna used both "error" and "mistake" to refer to all kinds of anomalies. She used the term "mistake" herself much more frequently (35 times) than "error" (17 times) in the three interviews. Of these 35 counts of "mistakes" and 17 counts of "errors", 25 uses of "mistakes" and 5 uses of "error" occurred before I asked the cue question regarding her definition of error. It is obvious in the specific recalls and in the general descriptions of her BAK system that Anna was using the two terms interchangeably.

When asked what error is, Anna put forward a dialectic view of error. For her, error serves a positive function but it is also something to be eliminated in the process of achieving a learning goal. This is a view that, as I interpret it, actually sums up what we have learned from SLA, a view that sums up from not only error analysis research (which basically views that errors are undesirable and to be eliminated in the learning process), but also interlanguage studies (which views that errors come from a learner's own developing rule-based system, which is approximating and evolving towards that of a proficient speaker) and learner process and strategy research (which holds that errors are results of and evidence for learners' uses of learning strategies and internal processes):

Anna: I think it's a lot of things, just depending on who you are or which way you are looking at it, what hat you're wearing, you know.... You're [the learner] working something out here. It's- it's a mistake. It's something you're doing in a way, it's the wrong, it's a rule you don't have, it has stopped as a rule- being an English rule, you know. Huh so you know, it's obviously a developmental
thing and you have to work through it. On the other hand, it's something in
some way does not fit what the native speaker does. It's always that dilemma
of saying "yeah you have to make mistakes in the process to get it to this
point, and at the same time you want to eliminate them", you know and yet the
question is you're never really gonna be eliminating them. [Para. 179]

In a follow-up verification interview, Anna posited that errors are violations of
rules that native speakers go by, and there are not just syntax rules and lexical rules, but
also discourse and genre rules (use rules), and the notion of error covers any anomalies in
any one of the four categories in the assessment scheme they were using: content,
organization, vocabulary and grammar. Like Marge, she viewed a slip as being due to
carelessness.

4.1.5. Summing-up

Findings to answer this first research question displayed the four instructors' 
diverse, personal understandings of "errors" and "mistakes". At times, there seemed to be
discrepancies between the instructors' own declarative knowledge on these issues and
their actual behaviours in their recalls of specific errors. For example, Edward, Marge and
Kelly all said they did not like to use the term "error" for the content and organization
anomalies their students made, but they had been using the term "error" to refer to content
or organization anomalies during their interviews, until the point at which I specifically
asked for a definition of "error". My questions on the definitions of errors and related
terms seem to have brought into light their tacit knowledge in a number of error-related
issues, but also the discrepancies between their tacit knowledge and their actual
behaviours.

On the issue of whether error is reserved only for linguistic anomalies, the four
instructors tended to be divided into two camps. Anna and Kelly both applied the term
"mistakes" or "errors" to all kinds of anomalies, although Kelly maintained that error was an indicator of improper declarative knowledge and was not relevant to his teaching context. Edward and Marge, on the other hand, considered only linguistic deviations as errors. Table 4.1 summarizes the instructors' different views on the nature and scope of errors, mistakes and slips, and the amendments some of them implied in their responses to my follow-up questions.

The title of my thesis study is "A phenomenological study of ESL instructors' perceptions of written errors among adult ESL learners". I have chosen a broad notion of error and intended that this thesis study would look into all kinds of errors this group of ESL instructors perceived, whether they are ESL errors, linguistic inaccuracies, discourse anomalies or writing errors. Therefore, although the definitions that three of the instructors' initially declared in their interviews were restricted to only linguistic (morpho-syntactic) anomalies, and although some of instructors tended to differentiate ESL errors from writing errors (i.e., content or organization imperfections or mistakes as they called them), I continue to adopt an inclusive notion of "error" for my discussion and use the term as such.

In order not to misrepresent the instructors' views, nevertheless, I have tried to exercise care in this regard when reporting the language the instructors used, and to make clear the scope of the terms the instructors were referring to each time. As a general rule in this chapter, I will use the term "anomaly" to refer to any deviation from an instructor's expectations, be they morpho-syntactic, content, organization, lexical or discourse (language use) deviations. On the other hand, the term "error" in the remaining portions of this chapter always refers to a violation of a morpho-syntactic rule.
Table 4.1. Summary of views of the instructor participants towards scope of errors, mistakes, slips and content and organization anomalies

<table>
<thead>
<tr>
<th>Instructor</th>
<th>Scope of errors</th>
<th>Mistakes</th>
<th>Content and organization anomalies</th>
<th>Amendment in the follow-up responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>“Error” applies to a violation of rule; these rules include discourse and genre rules.</td>
<td>Same as errors.</td>
<td>These are also considered errors.</td>
<td>No amendments</td>
</tr>
<tr>
<td>Edward</td>
<td>He disapproves of the term “error” and prefers to use “mistake” or “imperfection”. If he had to use the term, “error” is for linguistic incorrectness.</td>
<td>Same as errors.</td>
<td>He preferred to call all anomalies including those in content and organization “mistakes” or “imperfections”.</td>
<td>There are organizational rules and violations of such rules can be considered errors.</td>
</tr>
<tr>
<td>Kelly</td>
<td>“Error” only applies to a violation of a rule due to lack of declarative knowledge of the rule, which he had taught his students already.</td>
<td>Most of his students made mistakes because they are careless and lack editing skills.</td>
<td>He preferred to call all anomalies including those in content and organization “mistakes”.</td>
<td>In assessments, he dealt with mistakes, which include anomalies of all kinds. Mistakes result from improper procedural knowledge.</td>
</tr>
<tr>
<td>Marge</td>
<td>“Error” refers to a violation of a linguistic rule only; Error rules are decontextualized; genre and discourse are context dependent.</td>
<td>Same as errors.</td>
<td>They are not called errors.</td>
<td>There are writing rules and therefore writing errors. Writing rules vary with the context but grammar rules do not.</td>
</tr>
</tbody>
</table>
4.2. Research question 2: What are the relationships between errors and the contexts of perception?

In spite of the four instructors’ diverse views on the scope and nature of errors, mistakes, and content and organizational anomalies, one common invariant element present in the context of the instructors’ assessment activities was the use of the marking scheme. This assessment scheme provided conceptual categories to structure instructors’ perceptions of language anomalies.

4.2.1. The marking scheme

Although the instructors in the study provided different definitions and scopes for the concepts of errors, slips, or mistakes, it is perhaps even more important to see how they reacted to them at an operational level. As I have reported in Chapter Three, the instructors’ recalls of specific anomalies in their students’ writing were very much structured by the assessment scheme they were all using.

Marge, the coordinator of the course, explained this marking practice in very direct terms when I asked her why they had to evaluate content anomalies:

Marge: And sometimes I- next to the parentheses, when there is irrelevant detail, I'll write "not relevant here", if it's something that can be used somewhere else.

Researcher: Why- well, just a side question, maybe.

Marge: Yeah.

Researcher: Why why do you think you should respond to content?

Marge: Because of our marking scheme. Have you seen the marking scheme? [Para.44]

Anna responded that she would always evaluate anomalies according to the four categories in the scheme when I asked her to look at specific anomalies for the first time:
Anna: Okay, so I don't know what you want me to, just talk? Talk aloud?
Researcher: Primarily just talk about some of the errors you you think that are more serious.
Anna: The errors that I think are more serious?
Researcher: Yeah.
Anna: What kind, do I divide-. I mean, I tend to think of this always in terms of these categories [Anna points at the marking scheme.] I don't, I mean, I don't, I mean I don't- I wouldn't say, you know, put content over grammar or something like. So you know, if I think about=
Researcher: That's fine. Whatever you think is whatever I want to get. [Para. 74]

I reported in Chapter Three that I had coded all the recalls of specific anomalies according to the language the instructors used and the marking scheme they were using. This was possible because this was how the instructors categorized each of the anomalies that had caught their attention. For example, in nearly all of Anna's students' marked assignments, I could see marking and responses to anomalies in different categories with different colour pens, just as she explained:

Anna: I always mark the grammar and the comments differently, so I'll do grammar in one colour, and then comments about organization and content in others because I'll find then people would get it back, and then often the comments on organization might be, you know, "clearly organized", "good intro [introduction]", "clear topic sentence", and stuff like that, making them separate out from what they are doing with organization and content and the grammar. [Para. 287]

The instructors' recalls on specific anomalies also often progressed from one category in the marking scheme to another. In the following extracts, Kelly commented particularly on the two of the four categories in the assessment scheme:

Researcher: Next one, "K", one, zero, one, one. [K-1-011].
Kelly: Okay. Uh, [sigh] Did I add this up right? I'm sorry. [Kelly checks his scoring.] Two six, two four. Yeah. Okay, uh almost the same thing, again it's content and vocabulary is what the person has to work on. Uh, vocabulary, ambiguous sentences, uh, using again "ccw", instead of counter clockwise. Uh, like in step No. 3, uh he's used the word- uh number "4" (screws) instead of spelling it out. [Text: Use the Philips No. 4 screwdriver to remove the 4 screws at each corner of the backplate.] This person lost a lot of marks relative to the rest of the class in vocabulary. And content, uh, he has a couple- one one sentence with two steps in it. Uh, little minor things, forgetting the colour picture. Nothing serious. Vocabulary
is the most serious. [Para. 128]

Edward in the following extract commented upon three of the four categories in the assessment scheme.

Researcher:      Which one is this?
Edward:          "E", one, five. [E-1-5] This one. Okay, obviously I gave high marks for both vocabulary and language, and I said it was well-organized and well-written. I made reference to a missing logical point, again; in other words, in this case he did not say anything about whether the computer has been recovered or not. The prompt made it very clear that it's unlikely that they would find the computer because it's hand luggage, not check-in luggage. Because there was only limited information available to them huh from the briefcase they have got. So it's not very likely that they would find it. And I also made a thing about not following the memo format. Okay, let me see now. So here, for example, I I I corrected the article error again, "the Washington National Airport". It's almost the only error he committed. So as I said, when I deal with a student that is otherwise competent in every respect, then I then I do article, because there was not much else they need to do in terms of improving their language so. [Para. 226]

In Edward’s recall here, the missing logical point was a content error; the memo format was an organization error; the article error belonged to the grammatical accuracy category.

Hence, all through this part of the interviews in which the instructors recalled the specific anomalies they identified, corrected or commented upon, they spoke of errors and anomalies in terms of the assessment scheme. I discuss this assessment scheme in Section 4.3. in greater detail.

4.2.2. The specific purpose of assessment and error perceptions

The data also show that the specific purpose of assessment structured at least two of the four instructors’ responses to errors at an operational level, in terms of which categories to emphasize and how detailed the corrections or marking should be.

As I have explained in Chapter Three, there were basically two types of student
reports for the instructors to assess: (1) assignments, which carried marks counted towards the final course grade and (2) practices, which were writing activities that were not counted towards the final course grade. When Anna and Edward were marking a practice task, they placed emphasis on content and organisation and displayed lower sensitivity to language and vocabulary anomalies.

Edward told how he dealt with the practice tasks differently than he did with assignments:

Edward: Uh probably, uh probably in terms of the practice ones, uh I may be more haphazard, I might have been correcting things more in random than the assignments because- in terms of the language itself. I think in the assignments I tend to focus more on the meaning and how they do the organisation simply because in a sense that is the course is orienteci towards that. [E1-5-6]

Anna was one of the two instructors who reported that she marked practice tasks differently than she marked assignment tasks. Anna suggested a reason for being more careful and more comprehensive in identifying errors in an assignment: In the assignment tasks, the scores the instructor gave would be counted towards the course grade; the scores given to the practice tasks would not. The instructor as a rater therefore was held accountable for what he or she has put down as a mark; and in order to feel confident enough to assign a mark to an assignment, the instructor had to be able to see what had been put down as corrections and comments.

Anna: So huh although for this, there is always the dilemma, because with the assignment, I feel like I have to put enough on it for to justify the mark. Huh, on the other hand, they would have to be useful.[A1-6-15]

Earlier than this, Anna explained in detail why correcting and marking anomalies or errors in an assignment is important.

Anna: Once they [the students] actually do it for evaluation purpose, so when they do an assignment today, there I want to, there I feel the need to be even more accountable for
what I write on that for feedback, probably because this is gonna be a marked assigned
now to it. So I want to give that same feedback, but even more so, I have to make clear,
you know, if you gonna give someone a 60 or a 50, they need to know why they have a 50
or the 60. And I can't just put a number without putting enough comments on them, so
they can see it. And it's also for myself, so that I don't put the grade. I mean, my
assessment too as I worked through and writing is to get a sense of what everybody's
problems are and where these students stand. So before I can put a number on it, I need to
be able to see the kind of comments I put on. [Para. 26]

At this point I should add that although Anna and Edward were not as comprehensive in
their responses to the practice tasks as they were to the assignments, the two instructors often
got over typical errors and anomalies in their classes and discussed them with their students, as
Edward put it:

Edward: So the quality of the feedback may not be the same for an informal assignment [a practice task].
And often what I do also with informal assignment is, although I don't make too many
corrections on the paper, I do take them up in classes as you have seen, and I often give
them a model or something like that, instead of correcting errors. I tend to make
comments on the more general level. So I think that it is probably in terms of individual
feedback is not of the same quality, but at the same time, I try to compensate for that by
taking them up in class or giving them a model for the solutions. So I think that I treat
them seriously because there is no point in giving them- but uh it's a different kind of
approach. [Para. 93]

I asked both Kelly and Marge if they had responded to students' anomalies
differently in the two conditions, the assignment and the practice task; they both replied
that they would not behave differently. I have not been able to check their own
observations against other data sources because I do not have any of Kelly's student
practice tasks, and the only student practice tasks I had obtained from Marge was not a
technical writing piece, but a piece of academic writing which she had asked the students
to do in the very first session in the term, primarily for diagnostic purposes; I therefore did
not use it in the recalls of specific errors. However, I examined the feedback in this
diagnostic test and found that Marge basically responded to all the obvious grammatical
errors in the students' texts.
4.2.3. Genre as context acting on error perception

As Marge suggested, if the term "error" applies only to grammatical accuracies and that error is often decontextualized, then it seems only reasonable that contexts cannot influence the perception of an error, regardless of the context type. In other words, a violation of a morpho-syntactic rule is called an error regardless of where it has been identified, in whatever condition, whoever committed it and whoever identified it. There is, however, one instance in Marge's data to show that the application of a morpho-syntactic rule is not always relevant to any genre. This was indicated in Marge's recall:

Marge: Uh, he's missing out a whole verb, and I am not sure about this point. I meant to ask him that, but I didn't. I am not sure whether he thinks he is writing note form. Because this would be perfectly good note form language. For example, he has written "Major damage found at this mark." [The student had provided a sketch map accompanying his report. The mark was given on the sketch map.] You and I might write that when walking along there. [Marge said this because the prompt was a narrative of an investigator reporting what could be seen as he or she was walking along the paths in the resort site.] He hasn't written "Major damage was found at this mark." or "is found at this mark." He has missed out the "be" form. They are told clearly that they are not allowed to write in note form. But in note form this would not be an error. In here, grammatically it is an error. [Para. 183]

One interesting side question arising in this example is: Was this an instance of the student not knowing the syntactic rule (the "be" form in passive constructions) or was this an instance of this student not knowing the required use of this syntactic rule in this genre? Or was this just a careless slip? There is little doubt here Marge concluded that the student was committing an error, a violation of a syntactic rule, meaning that the student had known he should use a complete passive sentence but failed in executing it successfully. Personally I would guess retrospectively that the possibility of the student making a careless slip is very high, given that he or she had made several passive constructions flawlessly in the same text, as in the following:

"The inspection was held on Sept. 19 by myself."
"The resort has been closed for six years."
"Now some mining work will be needed in that area."
*"The road is rough and fallen trees are found in the interval."
**"The bridge acrosses a stream is made of unpainted timbers."
"Some parts seem to be rotting, and they have to be replaced."
"Fallen trees and subsidence should be cleared up."
**"The 120m major damaged section should be rebuilt." [M-M-5]

But when Marge said, "They are told clearly that they are not allowed to write in note form", she was actually referring to a use rule, a regulation on whether to use a set of morpho-syntactic rules, not about whether applying a certain morpho-syntactic rule correcty, i.e., not whether rules such as "NP Movement Rule" or "Structure Preserving Principle" (cf. Radford, 1988, p. 420-445) have been applied correctly, but whether these rules were applicable in the first place.

Marge’s suspicion at one point that this student might have thought that he was writing in note form was not totally unsupported. In this particular student’s Investigation Report, there was another instance where the “be” form was missing in a lead-in for a list of items:

“Conditions found along the road:

i) 0-1.8 Km The road is rough and fallen trees are found in the interval.” [M-M-5]

As can be seen here, the very first sentence following the erroneous passive lead-in is a correctly constructed passive sentence although this sentence does contain a word choice error, also pointed out by Marge’s written comment. This suggests that I cannot in absolute terms exclude the possibility that (if the original anomaly was not a careless slip as I would guess) the student might have been mistaken in deciding whether to use a passive construction rather than the student had intended a passive construction but failed in his execution.

Hence, this particular example seems to suggest the following:

(1) While the linear formation of this particular sentence must comply with morpho-syntactic rules, there exists at least a rule at the genre level
regulating whether one should apply these morpho-syntactic rules in the first place.

(2) Marge might think that error applies to violations of grammatical rules only; in practice, it seems that she was treating violations of use rules as errors as well.

In any case, however, the point that is most relevant here is that the norm Marge used to judge errors, in Marge's view, has to be related to the genre in discussion.

Marge's observation was about the interaction between the genre and the determination of an error proper. Kelly spoke about the interaction between an organizational anomaly and the genre of technical instructions manual:

Kelly: Uh, in the manual, I explained to the class that there should be only one step for each instruction, so one imperative verb for each instruction. And that's important because the user has to be able to read one simple line, do a step, read another line, do a step. If the two steps are in one line, the reader has more difficulty following the manual, and one two three four five times in this one, he put two instructions in one step or two steps in one instruction. So like "Unplug the photo copier and turn it around.". There's the "unplug" and "turn". "Remove the bulb and replace it". There's "remove" and "replace". They should be broken up into two different steps. And for technical manuals, that's really important. So probably that's where the person lost quite a few of his marks, most of his marks. I think I deducted like, [coughs] 32% of the organizational marks got taken away for doing that, and doing it so persistently. [Para. 20]

It is obvious that this one-step-one-verb rule (and I notice that in fact all the instructors referred to this rule when marking this document type) applies only to writing technical instructions, and not to the writing of many other genre types. Kelly's sensitivity to this interaction between genre and specific error types was manifested when he suggested why a particular error seemed particularly prevalent in the assignment of a user's manual:

Kelly: Okay. This person did well. This is a good student, probably one of the best in class. Let me just- yeah. Again, it's vocabulary and content, but it's really the type of exercise that creates this. [Para. 155]

Here, by vocabulary, Kelly meant the erroneous use of more than one verb in a
procedural step.

Hence, in these two examples, respectively from Marge and Kelly, I can see that sometimes genre as a context can exert great influence on instructors’ error perceptions, to the extent that it can determine whether an error or an anomaly actually exists.

4.2.4. Change of percepts resulting from repeated engagement with scripts

There are instances in the data which show that errors and anomalies may exist but they did not always necessarily catch the eyes of the judges. Raters or instructors missing errors in their reading is part of the reality in language evaluation and assessment, however uninteresting and trivial this phenomenon is, as in the following:

Edward: So I should have corrected "aboarded" to "boarded". I think the original prompt probably said "aboard this fight" and uh I think he sort of formed the wrong word. But I think my my my my assessment of the language is fine. [Para. 125]

Kelly: Yeah, there're some, there're some vocabulary mistakes. uh, like "number" over here should be abbreviated. "N" "O" period. uh I didn't I didn't see that.

Researcher: Yeah, is it because you think that it's, I mean=

Kelly: No, that's a mistake that I'm gonna correct right now. I corrected on everybody else's paper. I must have- it's hard to... [Para. 26]

Marge: Oh, god, you know, I missed that last thing, look at that at the very end. "It's essential to improve your company competitive." And I am just going to add the word "edge", if you want to add it to yours, uh just to give her new= [Para. 497]

While missing errors may be due to raters’ fatigue, carelessness or other factors, it seems that prolonged or repeated engagement with a script may result in the instructors’ change of percepts. During their recalls of specific errors, on several occasions, the instructors experienced changing percepts in their evaluations of the errors in the students’ report, resulting in an evaluation of their initial assessment. The following extract records that Anna experienced a change of percept after she had gone through a few scripts during
recalls of specific errors:

Anna: Probably not. Yeah, I think I'm gonna change this mark anyway because I thought I was looking-. actually I didn't have time because he was- there were about five or six or more that I set aside. I said I would, by the time I've gone through it earlier, I said that these are the ones I should re-look at because I think that the marks are low. I didn't have time to do it, so I'll just do it later, either look at them again, or make adjustment later when I talk to them individually and go over them. But uh, yeah, because it wasn't meant to be finished, the whole lot, and then you go back to the ones you did way back, and even though I am looking at all of them as I go, you still end up marking low for some reason, and then by the time you reach the end, you realize you still have to adjust that. [Para. 218]

This extract shows that Anna was aware that she had to keep readjusting the norm she was adopting in light of her percepts of the performance of other students. While Anna's change of percept here was on the performance of a whole report, Edward reported in the following a change of percept on the seriousness of a specific error:

Edward: I guess I should have made the correction on the article usage because uh somehow in front of proper noun, somehow it looks worse. [Laughs]. Sometimes it's "the Washington National Airport". I mean it's a proper noun and "You go to Pearson Airport", not "the Pearson Airport". So the thing, uh [Reading to himself in an inaudible voice.] [Para. 167]

In a similar situation, Marge re-evaluated the seriousness of errors and anomalies in one of her students' writing:

Marge: Uh, we've got to the last page or so, "in order to operate an power factory tentatively". I am glad you brought this to my attention, because the language work is what really matters. Okay, so, the recommendation, what is she recommending? Researcher: Yeah, it doesn't have a label, and it doesn't have a section.
Marge: No, no. I mean, another thing is, is she recommending that they do or they don't? It's not clear, as to if they are supposed to hire somebody or not hire somebody. And then she immediately jumps "the sooner technical problem get solved, the sooner the power plant gets better". Boy, I am I am inclined to lower her, you know. [Para. 560]

In sum, the instructors' changes of percepts during the interviews suggest that when assessing they often had to readjust the norm they had been using, resulting in an evaluation of their previous evaluation, implying that the reliability of assessment in fact
varies with prolonged, repeated engagement with the scripts. Examples given in the present section also show how complex the assessment procedures can be and that percepts of errors can change when evaluation is carried out in a different context.

4.2.5. Summing-up

In answering Research Question No.2, I report that several contextual elements influenced the instructors’ perception of error. In spite of the instructors’ diverse, personal declarative knowledge of error and other related notions, they all used the same marking scheme. This marking scheme provided important categories for the instructors to make sense of their error evaluation and their assessment activities. There was also some evidence to suggest that in some situations genre as context and assessment purpose as context changed the instructors’ responses to errors, in turn resulting in a different percept of an assessment. Repeated or prolonged engagements with the students’ anomalies also tended to result in a perception of a different assessment.

4.3. Research Question No. 3: What are serious errors? What are the conditions with which they are identified?

In this section I organize the findings in three parts to answer this question. First, I discuss the types of errors and anomalies which were considered serious by the instructors. Second, I report the conditions in which the instructors identified serious errors and serious anomalies. Finally, I show how these conditions affected error gravity.

4.3.1. What kind of errors did the instructors consider serious?

In coding the errors the instructors recalled in their interviews, I used a total of 16 content codes, 10 organization codes, 20 vocabulary codes and 25 grammatical accuracy
codes, based on the language individual instructors used to evaluate these anomalies in their recalls, to code a total of 323 errors in a total of 74 pieces of students' written documents. As I reported in Chapter Three, the instructors remarked some of these anomalies they regarded as minor (non-serious) anomalies. So in the end, there were a total of 293 coded serious anomalies. The frequencies of these serious anomaly types they attended to in the recalls by category are given in Table 4.2.

Table 4.2 shows that two anomaly categories seem to have attracted a lot of the instructors' attention: content and grammatical accuracy. The second rightmost column in Table 4.2 shows that on the average 38.8% of the errors the instructors considered serious were content anomalies, and slightly above one fourth, 27.34%, of the errors they commented on were grammatical errors. These figures have to be considered approximations only because one of the instructors, Anna, had a very high response rate in the content category, causing a high average in the content category, particularly compared to Marge, who made few comments about content but many about grammatical accuracy.

Table 4.2. Frequencies and percentages of anomaly types considered serious by the instructors by anomaly category

<table>
<thead>
<tr>
<th>Category and the weight each category carries in the Marking scheme</th>
<th>Anna</th>
<th>Edward</th>
<th>Kelly</th>
<th>Marge</th>
<th>Averages of percentages</th>
<th>All instructors</th>
</tr>
</thead>
<tbody>
<tr>
<td>Content (22%)</td>
<td>23(63.9%)</td>
<td>36(30%)</td>
<td>35(43.8%)</td>
<td>10(17.5%)</td>
<td>38.8%</td>
<td>104</td>
</tr>
<tr>
<td>Organization (28%)</td>
<td>4 (11.1%)</td>
<td>27(22.5%)</td>
<td>20(25.0%)</td>
<td>11(19.3%)</td>
<td>19.48%</td>
<td>62</td>
</tr>
<tr>
<td>Grammatical Accuracy (28%)</td>
<td>7(19.4%)</td>
<td>43(35.8%)</td>
<td>4(5.0%)</td>
<td>28(49.15%)</td>
<td>27.34%</td>
<td>82</td>
</tr>
<tr>
<td>Vocabulary (22%)</td>
<td>2(5.6%)</td>
<td>14(11.7%)</td>
<td>21(26.3%)</td>
<td>8(14.0%)</td>
<td>14.40%</td>
<td>45</td>
</tr>
<tr>
<td>Total (100%)</td>
<td>36</td>
<td>120</td>
<td>80</td>
<td>57</td>
<td>100%</td>
<td>293</td>
</tr>
</tbody>
</table>
Table 4.3 shows a ranked order to represent the extent of attention individual instructors tended to pay attention to each anomaly category.

Table 4.3. Ranked order of serious anomaly categories recalled by the instructors

<table>
<thead>
<tr>
<th>Instructor</th>
<th>Ranked 1st</th>
<th>Ranked 2nd</th>
<th>Ranked 3rd</th>
<th>Ranked 4th</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>Content</td>
<td>Grammar</td>
<td>Organization</td>
<td>Vocabulary</td>
</tr>
<tr>
<td>Edward</td>
<td>Grammar</td>
<td>Content</td>
<td>Organization</td>
<td>Vocabulary</td>
</tr>
<tr>
<td>Kelly</td>
<td>Content</td>
<td>Vocabulary</td>
<td>Organization</td>
<td>Grammar</td>
</tr>
<tr>
<td>Marge</td>
<td>Grammar</td>
<td>Organization</td>
<td>Content</td>
<td>Vocabulary</td>
</tr>
</tbody>
</table>

In Table 4.3 it is shown that both content anomalies and serious accuracy errors ranked higher than the other two categories when the instructors recalled serious anomalies in their interviews. This suggests that while linguistic errors were two of the four instructors’ one primary concern, these instructors also paid great attention to the content anomalies in their students’ writing. I will take up the issue of content deviations and grammatical errors as two serious anomaly types in the next section when I answer Research Question No. 4. There I also discuss why content anomalies were two of the instructors’ primary concerns. At this point, I move to report the instructors’ conscious use of certain criteria in their error evaluation activities.

There are ample instances in the data to show that the instructors’ percepts of serious errors was a function of context. That is, their determination of serious errors was often affected by contextual elements such as the learner, the purpose of evaluation activity and the document type. To see what “serious errors” really meant to the
instructors, it is essential to see what criteria the instructors actually used in determining error gravity. In Sections 4.3.2 to 4.3.4 below, I discuss three criteria. They are comprehensibility, irritation, and error frequency. I then move to discuss in Sections 4.3.5 to 4.3.6 how genre as context and learner as context interacted with error gravity. I do not discuss instructor as context in a section of its own because I deal with this when I report on individual instructors under other section headings. I note the instructors’ individual differences and interpret that instructor (teacher or judge) as a context is probably the most determining element of all.

4.3.2. Comprehensibility as a criterion for evaluating error gravity

In many previous studies and writings on error analysis, two frequently studied criteria for error evaluation are comprehensibility (Galloway, 1980; Gutterman, 1978; Khalil, 1985; Lennon, 1991; Munro & Derwing, 1995; Schairler, 1992; Varonis & Gass, 1982) and irritation (Albrechtsen, Henriksen & Færch, 1980; Piazza, 1980; Santos, 1988). Comprehensibility as a criterion referring to whether an error is serious or not is determined by whether the error impedes comprehension, hinders communication, or changes the intended meaning.

In Chapter Two, I showed in a review of the related literature that previous studies have studied constructs such as “serious errors”, “comprehensibility”, “irritation” and “acceptability” (Albrechtsen, Henriksen & Færch, 1980; Galloway, 1980; Gutterman, 1978; Khalil, 1985; Munro & Derwing, 1995; Piazza, 1980; Santos, 1988). In this section, I will show that:

1. Comprehensibility was an important criterion to determine serious errors.
2. Even when an error was comprehensible, it does not mean that it was not serious, however, because there were situations in which other pedagogic considerations came into play.

In the interviews, although I myself used the term “serious errors” and “irritating
errors” in my cue questions, I did not actually have to mention terms such as “comprehensibility” or “comprehension” or “communicability”. Rather, the notion of comprehensibility arose itself in the general descriptions of the instructors’ BAK systems and in the recalls of specific anomalies. Three of the four instructors thought that prepositions and articles in their given specific contexts were minor errors because they did not hinder communication or comprehension. Here are some examples taken from the recalls of three of the instructors. The first extract is taken from Edward’s interview:

Researcher: Yeah, well, let me go back and pick up two terms you have used. One I think I introduced it, that is ‘serious’, serious errors. The other is ‘disturbing’ you mentioned some time ago in the interview.
Edward: Yeah.
Researcher: Do you tend to equate those two types of errors? I mean, like errors, are errors- serious errors usually irritating or=
Edward: Disturbing. I’d say disturbing, I am trying to think I take the perspective of not a language instructor, but of an ordinary person with whom I am communicating. That’s why I say disturbing. They are not serious because because as I said, just because you don’t put the ‘third person singular’ ending to a verb doesn’t mean that I don’t understand anything you are saying. Of course I understand you perfectly. So they are not serious errors, but they are disturbing errors, to I think a non-language instructor. If it is not a language instructor, they will think that you know this student is not- it seems so basic to be able to put a “s” at the end of it. [E1-4-18]

Edward’s comment here once again shows that comprehensibility was a major criterion he used for determining serious errors, but at the same time, he claimed that minor errors are usually disturbing to a lay person although they may not impede meaning and communication.

Kelly also commented on article errors, using comprehensibility as a criterion for error evaluation:

Kelly: Uh most of the mistakes he makes are minor, basically articles missing or prepositions, usually missing or the wrong one. So you know “on the boat house” instead of “at the boat house”. But none of his grammar mistakes impeded comprehension. And and they were mainly minor... [K2-3-4]
Marge commented on a tense error and considered it minor, suggesting that she was using comprehensibility as the criterion for error gravity though she also estimated that the error was due to carelessness:

Marge: Okay, I also noted a minor grammar error. This student makes very few grammar errors, in fact, a tense error uh which didn't change the meaning, but it's carelessness on both parts really and so I pointed it out to her. [M1-1-31]

Why is comprehensibility a criterion to determine error gravity at all? Marge's example below gives a reason: Errors that distort meaning, hinder communication, or do not present the required message can result in undesirable grave consequences:

Marge: The way I treated errors, the way I talked to students about errors, is the consequences of the errors, that is, in this situation, make this kind of errors, this is going to be the consequence. Not that error is, not that accuracy is is important for its sake, which I think is the French attitude. It's- if it's going to offend somebody, you're going to make a huge mistake here. somebody is going to misunderstand you, the bridge is gonna fall down, you know, that there are consequences to making that kind of mistakes, like the "s" and the "ed" endings. [Para. 909]

But meaning or communicative effectiveness as a criterion alone was sometimes not helpful enough in terms of determining error gravity. In the following Edward seemed to be, on the one hand, hesitant to call this particular error serious because it did not impede communication, but on the other hand, he found such an error rudimentary.

Edward: "E" zero eight. [E-0-8] Okay, I just said "well-written report". The only persistent problem appears to be verb-tenses." [Reading the text in an inaudible voice.] Yes, verb-tenses. First, tense, "We had met at my manager office ten days ago." So, you know, use of present perfect with a clear past. Well, in fact, I made the comment because twice in successive sentences he has committed the same error: "You met me ten days ago." "You had given me air sample at that time, so I could measure the pollution rate twice a day."

Researcher: What did you think about this type of error?
Edward: Uh it doesn’t impede meaning, but uh but it’s uh but it’s uh uh uh (pause)
Researcher: Rudimentary, would you say?
Edward: I think it’s one of those tense errors that, is well rudimentary, I think it’s one of those tense errors again maybe native speakers would not commit. [Para. 252]
I draw attention to Edward’s hesitation and his five-second long pause in this segment; I interpret that what seems to have bothered Edward here was not that comprehensibility as a criterion for determining error gravity is problematic, but that it would be a problem if he were to insist on using comprehensibility as the only criterion for error evaluation; this is because somehow the gravity of an error has to be judged also based on the relative position of where the error is located on an instructor-perceived language developmental scale or an L2 acquisition continuum.

The fourth instructor, Anna, did not speak directly about comprehensibility as a criterion for serious errors. As a matter of fact, she gave each of her students an error record sheet, which she called “Revising your writing” (displayed in Appendix M). This sheet contained a list of what Anna considered major grammatical errors and minor grammatical errors. The major errors included sentence structure, unclear message, relative clause structure, word order, coherence, conjunction and transition words, fragment, run-on, comma splice, verb tense, verb form, preposition, word form, wrong word or expression. The minor errors included spelling along with other error types such as agreement, article, punctuation, capital letters, plural/singular, pronoun reference or reference agreement. In both my first interview and also the follow-up verification interview with Anna, I asked her the question of how she determined her major errors and her minor errors. On both occasions, she answered that she could not remember exactly the source, but that she had probably adopted this from a book for instructors of English writing and had revised it a few times over the years. She inferred that the original source was probably not an ESL book, because some of the items in the list tend to be errors committed by native speakers, such as run-ons and fragments.

4.3.3. Irritation as a criterion for evaluating minor errors

As I said in Section 4.3.1, during the interviews, I did not mention in my cue
questions the notion of comprehensibility, which eventually arose in each instructor’s recalls. I did, however, in my cue question mention irritation and irritating errors. I did this because I wanted the instructors to report not just error gravity, but also their emotive feelings, which I feared they might not be inclined to report. Two of the instructors expressed that they were often irritated or disturbed, not by serious errors, but by certain types of minor errors which are violations of trivial, simple rules:

Edward: Prepositional errors, I tend to correct them because again they are things that kind of look funny. Disturbing, you could say, not serious. I mean you know. [E1-10-31]

Edward on another occasion explained what he meant by “disturbing errors”, using the third-person-single error as an example:

Edward: Uh I said the only rationale is that now I tried to mark the ones that’s are either are are are recurring and uh and then like, I have to say it this [way], errors that I find disturbing as a reader.
Researcher: Uh hum.
Edward: That that may may not mean that writing is non-communicative. It just means that it gives a poor impression. In other words, it looks sub-standard, for example, uh not putting the third person singular “s” at the end of the verb or something. [Para. 52]

Anna commented that “all errors bug me”, but like Edward she also suggested equating minor errors with irritating errors:

Anna: The minor [errors] does not mean that they are not important. It’s because they are important too. The idea is that in relation to these [She points at the major errors], you want to clear them up. And I emphasize that the minor errors are the errors that are irritating ... [A1-9-15]

She explained further why and how these minor errors are irritating.

Anna: Well, I think. I think it’s even for native speakers. If you have handed in something, let’s say, a native speaker’s report is full of spelling errors. Huh, it’s like, you know, wearing the wrong dress to the whatever you know performance. It’s people get an impression. And impression isn’t so much about what it’s about your language, it’s about carelessness, it’s about how
you do your work. And because it distracts them, not in terms of meanings. Say with agreement, because I think for a lot of native speakers it’s kind of like, [inhale] “God, ask her again? Get it? Don’t they get it? Yeah, doesn’t he get it?” I can ask again! You know, what I mean, and then it’s like in his whole mind is preoccupied with this because it’s like you have someone’s, like I tell them, kicking their head against the wall, like someone, you know, knocking every time. At the end, you [the instructor or the rater] are not paying attention to content, even if the person has great ideas and good organization, everything has been so distracted by this. Then you give them a “D”. I mean, I don’t do it, but I am saying, if you go to a credited course with a native professor... [Para. 117]

The scene that Anna depicted in this long extract is interestingly vivid. She used phrases like “whole mind is preoccupied”, “kicking their head against the wall”[sic], “knocking every time” to describe a rater’s provoked emotion. The gist of this message is that when a rater or reader has been irritated and distracted by this type of error, comprehensibility as a criterion for error evaluation no longer applies, and all that the rater is left with is irritation.

4.3.4. Frequency as a criterion for evaluating anomaly gravity

Comprehensibility and irritation were both criteria used to evaluate one single error; but error frequency was a notion that the instructors applied to more than one anomaly. Table 4.4. consists of three samples from each instructor, showing how quantification of errors played a role in their anomaly evaluation.

<table>
<thead>
<tr>
<th>Instructor</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>I mean I guess, yeah because sometimes people presume a lot of spelling mistakes. If I put Spelling all over it, they [the students] kind of just think “Oh, my god.” You know, woo. They don’t put it in perspective. [Para. 123] You know, a lot of the errors, you know. And I do point those out. [Para. 166] Yeah. I thought it was- his vocabulary was good, he doesn’t have a lot of mistakes. [Para.</td>
</tr>
</tbody>
</table>
Edward  In this case it was probably he just omitted a lot of details so. Yeah, uh, a lot of details are missing, factual details. [Para. 185]

Otherwise, the language is fine. It's just like reading a telegram. Again, fair enough. I probably wouldn't do anything different. And again as I said, I reacted more to the fact that there are too many missing details. [Para. 194]

Yeah. Because he's using it correctly many times, but he's also using it at the times when he shouldn't be using it. [Para. 265]

Kelly  Uh the other mistake the person did was, a lot of vocabulary errors. [Para. 46]

It just could be better. Uh, the other mistake this person made is- there're a lot of vocabulary errors. [Para. 56]

Okay. This person made uh a lot of content errors. Uh, (pause) again, having uh too much information for each instruction. [Para. 99]

Marge  I mean he's using complex grammar, when you look at Mike's sentence structure, it's complex and it's fairly accurate. It's just his individual- and the trouble is there are a lot of individual errors. [Para. 338]

For this particular person, she is not making too many grammar errors, she's still making a fair number of preposition errors, so uh I do like to point out. She's like many of the advanced students, mostly preposition and article errors. [Para. 28]

Yeah, he he's got a lot of the word- the uh word ending problems. which of course are various grammar rules, but in most cases that I've mentioned so far, he's got problems whether to assume it's a count noun or not. [Para. 151]

In the extract below, Kelly related quantification of content anomalies to gravity:

Kelly:  Uh, there're other ambiguous sentences too. Like "Unscrew the four screws on each corner of the back plate." It sounds like there are four screws on each corner. What the person really wants to say is at each corner, there's a screw. So there's only four screws in total. In this case, he makes it sound like there are 16 screws. So that's ambiguous. So a lot of mistakes like that, which probably is serious. [Para. 61]

In the second extract, he again explicitly related the frequency of grammatical
errors (of various types) to a mark reduction in the grammatical accuracy category of the assessment scheme:

Researcher: So are you saying this person didn’t choose his word wisely from an audience analysis?
Kelly: Yes. One of the things I try to have the students do is to know who the readers are, and use vocabulary appropriate to them. But I think the other area the person probably dragged down the mark is his grammar. A lot of little grammatical mistakes. [Para. 84]

Edward in a similar fashion explained why frequent errors should be given priority in instructors’ feedback:

Edward: You know, and and I find that the distinction between meaning and and form sometimes artificial, so I do focus on grammar too, but I try to more to focus on recurring problems, problems that uh- problems that may have either mmm a bad effect, because I say this in class, there are some problems that are not serious, but they happen so frequently that they should be corrected simply because of the frequency. [Para. 30]

I have reported that Anna labeled some of the grammatical errors as major and minor in the sheet “Revising your errors” she gave her students. In this error tally sheet, Anna categorized spelling errors as minor, but at one point during the interview she also noted that a repeated minor error is not minor and is worthy of special attention, as she said in the following:

Anna: If a person, for example, keeps spelling the word “panel” wrong throughout the whole thing, then I am really, and I am gonna say at some point, you know, right now it's not that important, as you go on, this is the word you’re probably gonna use, so just make a note of it and practice it at home. You may know how to spell it in your head, but it's not getting down to your hand, and it needs to be automatized. You know, write it up. [Para. 34]

4.3.5. Teaching advice as a criterion for anomaly gravity

One criterion which seems to have been used by two of the four instructors is to consider an anomaly serious when it went against the teaching advice this instructor had
previously given. In his verification response to my description of his perceptions, Kelly reported explicitly the use of this criterion:

Kelly: Thus, I am most critical when a student makes a mistake on something that we have either discussed in class, or that I had repeatedly written down on his previous papers. For example, before we review "numbers" (i.e., symbols or spelled out), I do not penalize a student if they have written out a number instead of spelling it out. I would comment on it so that the student can think about it. However, after we review number use, I expect the student not to make these mistakes. ... [Kelly, Follow-up response]

After reading Kelly's follow-up response, I went back to Kelly's interview transcription and found an example Kelly mentioned in his follow-up response. In this example, a student made the mistake of writing the word "number" instead of "No." I asked Kelly in a recall if it was a serious mistake. He said:

Kelly: Well, no, as in this scheme of things, it's not very important; however, about two or three weeks ago, I taught it to the students, and what I teach the student, I hope that they remember and use it in their writing. So, in that way, it is serious. Uh, just because uh I taught it, I hope that they use what I teach. [Para. 32]

Still on another occasion during another interview, Kelly explained explicitly his policy on evaluating errors when commenting on another student's vocabulary mistake:

Kelly: Vocabulary, uh, what he does is, uh, mistakes that we talked about in class, and I am pretty hard. Mistakes that we haven't discussed yet, like the person before with the conjunctions, I'm I'm not that hard on him although I explained what I mean, and tell him we're gonna have a lesson on it. But here, uh, when he wrote like "Approximately 7", we talked it about in class. We don't use "approximately" and we give exact numbers. So I think I was hard on him with vocabulary. [Para. 313]

Another vocabulary mistake this particular student made was the use of "low information content words" (Blicq & Moretto, 1998, p. 390), a mistake that went against what Kelly had just taught the class:

Kelly: I told the class, we did exercises, not to, you know, use vocabulary like that. It just makes the reader have to read slower, less confidently. So, uh a lot of his mistakes- he doesn't have a lot of mistakes, but a lot of mistakes we've dealt with in class. And in engineering, they are heavy counts. [Para. 316]
The other instructor who related serious errors to prior teaching advice is Edward. When he was commenting on the use of a definite article in a defining relative clause, he emphasized that this particular serious error was a point he had made in prior teaching activities.

Edward: Yeah. The same kind of correction "the air sampler which you gave me". I think, I don't remember. I think I actually raised this point in class. Well, uh I am sure I did. The fact that you cannot use a definitive reference on this unless you are qualifying it. Because in this case, it's known necessarily for the reader and the writer at the same time. [E-0-13]

In Section 4.3.6, I discuss this particular error in the work of Edward's students in detail when I exemplify the relationship between serious error type and genre as context.

4.3.6. Learner as context and error gravity

I have given in Sections 4.3.2 to 4.3.5 evidence to show that the instructors were using certain criteria in their evaluations of error. In this section, I wish to show that perception of the learner also plays a role as a contextual element in the evaluation process. When this happened, the instructor was usually ready to abandon or downplay another criterion. Three of the four instructors reported that their perceptions of a learner's individual strengths and weaknesses can play a part in shaping the instructors' error evaluation behaviours.

Anna

In the error record sheet "Revising your writing" that Anna's students used (Appendix M), spelling was categorised as a minor error type. I have already reported that, to Anna, a repeated minor error did not remain minor; it was worthy of special treatment. Another exception to Anna's general division of major and minor errors appeared in a story of one of her engineering students who kept misspelling the word
"machine". This story exemplified, in her perception, an interactive relationship between error gravity and learner as a context:

Anna: For example, one student, I think it was a lot of the comments that I wrote. I circled the word and put in the column the word "machine". And the next thing she did for me, it was the same mistake. I thought, well, here I have to just point out, you know, as an engineer, if you don't know how to spell this word, huh, let's say it's not important, you know, I know you know how to spell it. I am not saying you are stupid, and you don't know how to. I am saying, it just means going home. ... but it's just repetition. It's just your hand doesn't know what your head is telling it to do, and it's writing it out so that it's automatic. You know, and and that's one that I sort of think I am- not necessary- but huh sometimes- so I guess in some way it's something serious. [Para. 143]

The message here is clear. To the instructor, repeatedly misspelling a word such as "machine" is perhaps not much of an error in its ordinary sense; it could be just a simple bad habit. But considering that the student is an engineering student, misspelling a word like "machine" is certainly not something that the instructor thought could be taken lightly. Other instructors’ recalls also exemplified similar evidence, which I will show in the following.

Edward

Edward reported that in general he would not respond to article errors:

Edward: I don't correct articles simply because uh because it's much too complicated. And I think students will learn it they learn it by themselves
Researcher: You don't correct articles at all? Is that a policy?
Edward: Generally, I don't anymore. No. [Para. 57]

The justification for this policy according to Edward, is that the English article system is too complicated to teach and learn.

Researcher: Do you also think that you don't mark, you don't mark or respond to article errors because they are not as serious? Would you say that? Is that a=

- 131 -
Edward: I think so. Well, not as serious, I mean uh I think that what I what I think. And as a
language teacher, I know how complicated the article system is. A person who is not
a language teacher may still be offended by the error, but having having and I used
to correct articles when I began, but again, having realized how complicated it is to
learn, never mind to teach them to learn, I have given up because there's no point
correcting them without going into detailed explanations of why and the rules are
too many. I don't think they are that serious, and ... [Para. 68]

However, in his recalls of specific errors, I was able to identify a total of 16 corrections on
or responses to erroneous article usages or erroneous usages of article before a defining relative
clause. Edward explained these exceptions to the rule:

Edward: The only case I can see myself correcting articles deliberately is when I get a
student who is pretty well, you know, pretty well competent in every respect and
then an article is something that he or she is not very good at, then I say fine. I mean
that's- I guess I could say sometimes I do tailor my marking sometimes to the level
of the student. I mean uh a student with the problems at all levels, I don't both with
Articles or even the third person singular. [Para. 73]

At one point during the interview, when Edward discovered that he had responded to an
article error again, he gave the following explanation:

Edward: So here, for example, I corrected the article error again, "the Washington National
Airport". It's almost the only error he committed. So as I said, when I deal with a student that
is otherwise competent in every respect, then I then I do article, because there was
not much else they need to do in terms of improving their language, so. [Para. 231]

In other words, an exception to Edward's policy of not correcting articles was his
perception of learners' individual needs.

Marge

I have reported how Anna and Edward adjusted their percepts of error gravity due
to individual students' roles or their specific needs. This was also the case with Marge. I
reported in Chapter Three that there were a common set of steps that all the instructors
took in their recalls of specific errors. While the other instructors' recalls basically

- 132 -
progressed from one error category to another and one error to another and then commented on the individual student’s characteristics only occasionally, Marge’s recalls of specific errors were different in that they were often organized around her knowledge of her students, demonstrating strongly an understanding of individual students’ needs and strengths. In some of her recalls, Marge would give the name of the student whose text she was about to read, then she made comments on the specific strengths and needs of the student in relation to the error or errors concerned. Below are some of these samples:

Marge: It [Laughs] I don’t know if you remember, when I was talking to him [the student], but uh, he spelt “block” as B-R-O-C-K-E-D [brocked], so he was carrying a pronunciation error into his writing. I changed it into “block”, and I did=. [Para. 73]

Marge: Tony, yeah, not a lot of grammar mistakes, Tony. He is quite good. He’s got about 76. I mean my highest mark is 78, you have to realize that. [Para. 406]

Marge: Yeah, yeah. Okay, okay, so we’re starting with Mike [M-4-009]. Okay, Mike. Mike is interesting from the point of view that he is one of the best responders in class. That his idea is about the report are always very good, one of the best in the class actually. [Para. 295]

Marge: John is again a bit of a mix like Eddie. Sometimes he seems to do really poor work, and the other time he comes up with really good vocabulary, his vocabulary is quite sophisticated, and unusually so. I mean I haven’t seen- actually he showed me a couple of poems he has written. It’s quite sophisticated vocabulary. So he seems from the vocabulary point view, he is good. [Para. 393]

Marge: Okay. Should we go on to Lilian [M-4-004]?
Researcher: Yeah.
Marge: Uh, she works so hard; she tries so hard; and she still got her grammar mistakes. I don’t think Lilian has really improved very much. [Para. 429]

Marge: Okay, now Julie [M-4-005] is quite interesting. She’s not one of my favorite students. She is one of my best students. But she tends to be, because she is good, she gets impatient with what I am doing with the rest of the class. [Para. 511]

Since Marge had, all the time, kept in the back of her mind the needs and strengths of each of her students, it is no surprise that Marge adjusted her reactions to serious errors according to the learner, as in the following:

Researcher: Is this like a policy of yours that you respond to prepositions, or=
Marge: No, depends on the student’s frequent errors. I mean if- prepositions for this group are a frequent error for most of them. For this particular person, she is not making
too many grammar errors, she's still making a fair number of preposition errors, so uh I do like to point out. She's like many of the advanced students, mostly preposition and article errors. So I just indicate them and then I expect them to try work it out themselves. [Para. 29]

As a matter of fact, because Marge knew her students and their needs and strengths very well, she was ready to overlook even frequent errors in a student's work and see errors as an unavoidable consequence of a learner's risk-taking experiments with newly learned features. This is how she commented on the relatively high number of errors a student made:

Marge: My initial reaction to this student, is no- nothing to do with errors. This is that he made major real effort to use his own language, so consequently, he has made more errors than the others. He doesn't- uh uh I'm very careful not to penalize too much in a situation like this, when they are really taking a risk, and they are really trying to write their own grammar, own vocabulary. Uh, but I do you know note his grammar so that he can work it out and work together. And he did today actually. [Para. 107]

4.3.6. Genre as context acting on errors

I have reported in my answers to Research Question No. 2 that in some cases, genre as context can actually determine whether "an error is indeed an error", i.e., whether certain morpho-syntactic rules are still applicable to a particular genre. In one case, genre restricted the use of a morpho-syntactic rule; in another case, it relaxed the use of a morpho-syntactic rule. What I show in this section is that the determination of serious anomalies, in the perceptions of the instructors, was also influenced by document type as context, with document type being one of the main teaching objectives in the course.

There was a close link between the genre of an assignment and what was expected of the students in terms of their morpho-syntactic production. Before the students began a writing task, they were told in the prompt what their writing was supposed to achieve; they were also given in the prompt all the information needed to complete the writing task.
The prompt was usually given in one genre while the required document would be in another. For example, in one assignment (displayed in Appendix F), the prompt the students were given was a narrative consisting of a conversation, in which one person explained the technical procedures to another person. Having read this prompt, the students were expected to produce a user’s manual consisting of some technical instructions. In another task requiring the students to produce an investigation report, students were given an eye-witness account from the investigator, as if he or she was walking along the paths inspecting the site. While events and things in the prompt were presented in a narrative, the students were expected to organize the content differently in the Investigation Report. In both examples, the students were required to transform the information in one genre to another, and of course in these tasks, students’ ability to use a genre and its related language was in fact one of the focuses in the instructors’ assessments.

Because of a close relationship between the genre of a piece of writing and the frequent use of certain types of morpho-syntactic structures, the instructors expected each time in students’ assignments or practice tasks not only a correct presentation of the organization and content, but also a correct use of the certain genre-related morpho-syntactic structures. These expectations may have been part of their tacit knowledge, but the questions that I asked in the interviews regarding such a link definitely forced the instructors to come to an awareness of this. In the following extracts, we can see Edward and Kelly observing this genre-error relationship:

Edward: Well as I said, they were just- the problem is that with this kind of exercises, a lot of the steps that you see in the prompt will be reproduced exactly because that’s exactly how it is- that there is nothing else you can do. You know, the prompt is very restrictive in a sense when you are using it for Technical Instructions. The prompt more or less tells you exactly- It’s an information transfer thing and you are quite restricted in what kind of vocabulary you can use. [Para. 624]

Kelly: Okay. This person did well. This is a good student, probably one of the best in class. Let me just- yeah. Again, it’s vocabulary and content, but it’s really
Kelly: Well, it's the type of exercise that creates- the grammar is very basically easier in this. It's all imperative verbs. And I was really pleased to see the students got that. I think only one instruction with one student, they didn't use the imperative verb. [Para. 157]

In her interview, Anna also spoke on the relationships between the kind of errors she expected and the specific document type the students were writing in:

Anna: I mean I'm trying to think of- like we are back on reports now. I mean. We are actually, uh we are not on the technical instructions. Uh I think so. Some they make very few mistakes. I mean I am looking just at whether it is comma splices for those, or fragments or whatever else. Now that they had- but still in this one, they had more; it's always different. In an incident report, they may be struggling with the perfect tense, you know, or some other tense like that, whereas in this one, there were a lot more passives; it seems that they were putting in the time so that they don't have to show up the problem. [Para. 222]

In my classroom observation data, I also noticed how a genre highlighted the role of certain morpho-syntactic structures. The instructors teaching the course had all agreed readily on a list of document types they were to cover in the teaching scheme. There was no agreement at all on what grammatical items to teach. A lot of the in-class time on the course was devoted to discussion about the organisation and function of specific genres, but the instructors also spent time on grammatical items as they saw fit, and a lot of this discussion involved responses to errors in students’ recent work. I was therefore not surprised when, within the first three weeks in February 1998, I found Edward and Anna both discussing restrictive and non-restrictive relative clauses in their classes because that was the time the instructors had just finished correcting incident reports, their first assignment.
From a functional perspective on grammar, English relative clauses are often used to identify things, people or events, or to give extra information. Halliday observed that a defining (restrictive) relative clause functioning "as an expansion is essentially to define, delimit or specify" (Halliday, 1994, p. 243). It is easy to see why such a function has to be performed in an incident report, especially in the beginning of the report when people or things involved in the incident have to be identified from a more general class so that what has happened to them can be talked about. A sample of an assignment of an incident report, which the instructors used and I am about to discuss, is given in Appendix N. The example was taken from the course book, Blicq and Moretto (1998, p. 117-118). In this task, the writer and his or her colleagues were travelling in a company van on a field assignment to video-tape a construction site. They stopped at a motel for the night. The next morning they found that their van had been broken into and the video camera they were going to use had been stolen. The writer wrote to report the incident and ask for a reshipment of the video camera. In the following I list the relative clauses Anna used in her class discussion, all produced with errors by her students for this assignment.

1. **"The company was broken into last night on the parking lot of the motel that we stayed yesterday."**
2. **"There will be one bus a day passes through Freehampton, and I will pick it up at Freehampton."**
3. **"The serial number of the video camera that was stolen is Nabuchi TX200, 217848B where I rented from Meadow Electronics in Toronto."**
4. **"I would like to pick it up on Jan 31, 98 at Freehampton's Greyhound station, which the bus stops only on request."**
5. **"Yesterday, we stopped for the night in the Clock Inn where is 3.5 km west of Clearwater village."**

A follow-up, correction activity for what the instructor considered to be a serious error type is of course a familiar, logical act from many instructors. The point I make here is that what the instructor was doing about this particular syntactic form was a consequence of her teaching this particular document type.
Still another example illustrates how error gravity interacted with genre. I have reported that Edward said he adopted a policy of not correcting articles in general but I have also reported that according to the results of my coding, he in fact responded to erroneous article usage for a total of 16 times, treating them every time as grave errors. As Edward explained, exceptions to his policy were often due to learners’ individual needs. On the other hand, an examination into Edward’s corrections on these article errors in a total of 34 incident reports he marked also suggests that many of these erroneous article usages were in fact definite article uses related to a restrictive relative clause, a salient feature in this type of report. I observe the salience of the restrictive relative clauses in incident reports in the following episodes.

One of Edward’s students [E-0-9] in the very first mention of an air sampling device, called “air sampler” in his incident report, wrote this sentence:

*“I’m writing to inform you that unfortunately [sic], the air sampler broken at Cormorant Dam accidently.[sic]”

Edward in his written feedback suggested deleting “I’m writing to inform you that”, inserting “was” before “broken” and inserting “which you gave me” after “sampler”. In the interview, he recalled this erroneous sentence as follows:

Edward: Uh mm there was a missing auxiliary in which I put in there for the passive, and when you say the air sampler, you cannot really say the air sampler like that, you cannot make a definite ref- this is the case where the article is in fact correct, I mean, but if you are going to use an definite article, then you must put some kind of a qualification to put it in context, so I gave- but I did not explain to him why. [E-0-9]

Having gone through two more student reports, Edward once again in this part of the interview came to another student report which contained the same erroneous form.
"The air sampler has been broken accidentally in Conmorant Dam on 23 December, 1997."

This time Edward again corrected the sentence by inserting "which you gave me" after "sampler". So the sentence now became one that contains a restrictive relative clause:

*"The air sampler which you gave me has been broken accidentally in Conmorant Dam on 23 December, 1997."

In his recall of this error, Edward gave an explanation for what he did:

Edward: Again, I did the same thing as before. He said "the air sampler" and I said "the air sampler which you gave me", because otherwise the reader will not probably understand- what do you mean "the air sampler"? [E-0-15]

Grammatically speaking, as Edward put it when he commented on the previous report [E-0-9], "this is a case where the article is in fact correct". but then it would not be the meaning which should be intended, and hence it was in fact a discourse error! Or in Lennon's (1991) terminology, this is an error which has an extra-linguistic context as its "domain" (See also my discussion in Section 2.3.3, Chapter Two.)

The context is that prior to writing this report, a senior staff member in another department in the company had given the writer an air sampler to take measurements of the air when he or she was on a construction site. Later, unfortunately, an accident happened while the writer was taking samples, and the air sampler was broken. The task of the writer was to report this incident to this senior staff member and request a reshipment of another air sampler. As any experienced writer would know, the writing of this kind of report requires that in the beginning of a text, certain relevant things and people must be identified first (and in this case, it would have to be the air sampler). It is this requirement that determines the nature of error or anomaly here and how it should be evaluated and corrected.
As the interview continued, Edward found still once more he had corrected the very same mistake in the same sentence structure in a third student's report [E-0-13]. This student had written:

*"I would like to tell you that the air sampler had been broken in an accident."

Edward again added "which you gave me" after the word "sampler". This is how he explained this in this specific recall:

Edward: Yeah. The same kind of correction "the air sampler which you gave me". I think, I don't remember. I think I actually raised this point in class. Well, uh I am sure I did. The fact that you cannot use a definitive reference on this unless you are qualifying it. Because in this case, it's known necessarily for the reader and the writer at the same time. [E-0-13]

In an examination of the students' reports, I looked through this batch of written reports, including the ones that Edward did not have the chance to cover in the interview. I conducted a simple frequency count of students' uses of this particular syntactic structure in the case when they used it, the alternative way they handled the transmission of this message in the case they did not use this structure, and the responses Edward gave. I present the results of these figures in Table 4.5.

There were a total of 34 reports collected from the two sections of students which Edward was teaching. As Table 4.5 shows, a total of only 7 out of 34 students managed to do what they were supposed to without using the restrictive relative clause, and a total of 27 (nearly 80%) of the students attempted to use this particular structure. Table 4.5 also shows that 17 of all the students (50%) failed to use this construction correctly, resulting in Edward's remedial work and re-teaching of this structure in class.
Table 4.5. Frequency count of the use of restrictive relative clauses by Edward’s students

<table>
<thead>
<tr>
<th>Type of student strategy</th>
<th>Evaluation of student’s usage of this syntactic form in their reports</th>
<th>Number (%) in the category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Student used a simple sentence to mention the air sampler for the first time and then reported about the air sampler using a definite article</td>
<td>Grammatical</td>
<td>7(20.5%)</td>
</tr>
<tr>
<td>Student used relative clause with a definite article correctly</td>
<td>Grammatical</td>
<td>10(29.4%)</td>
</tr>
<tr>
<td>Student attempted to use relative clause with a definite article but used incorrectly</td>
<td>Ungrammatical</td>
<td>17(50%)</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>34</td>
</tr>
</tbody>
</table>

The purpose of this tally is to show the salience of this particular syntactic structure in this particular assignment, at least as the students (the writers of this task) perceived it.

In his recalls, Edward explained that he had talked about relative clauses in class, and according to my classroom observations, he did this at least twice. Because I wanted to see whether he would share my observation on the link between grave errors and genre, I asked him why he had talked about relative clauses, restrictive and non-restrictive, towards the end of our second meeting on February 11, 1998. I reminded Edward that it was about incident reports, but he did not seem to remember. During the third interview in May, 1998, however, I brought up the question again because by then I was already quite certain about this link. This time he answered that he introduced relative clauses because his students perceived it as a punctuation issue. When later I reminded him that it came out of the students’ writing in the first assignment, he began recalling the situation more clearly.

Researcher: I think, I believe, it was also- it came out of the students' writing that you asked them to do for initial assessment.
Edward: Well, that's probably true. Yeah, that's how I tend to decide what to cover in terms of grammar. I tend to look at the diagnostics and then at that point, make a decision as to what, what points to cover. I mean they tend to be more or less the same throughout, but it's not always easy to relate it to the topic. I mean I try to relate to the topic in some way, but- [Para. 953]

I had little doubt that the link between genre and the seriousness of certain morpho-syntactic forms is part of the instructor's tacit knowledge. What I was interested in knowing was when Edward corrected the erroneous definite article to a restrictive relative clause, and made students do all the follow-up activities in class, was he aware of the salience of this structure in the document type? I therefore pressed on:

Researcher: What I like to bring to your attention, is to see whether you are conscious of this, is that the use of relative clause is often related to identifying something or people, that function is uh salient in a sense uh in certain document type?

Edward: Yeah. Certainly. Sure. When you write a report, that's exactly what they are. They identify or they provide further information about somebody that is already known. Sure, like I say, that's why you introduce probably in the beginning, because incident report I treated them first before I go to anything else. [Para. 955]

I am inclined to believe that these two experienced instructors took up the teaching and practice of this syntactic form as a conscious reaction to students' specific grammatical needs. These responses can be expected especially from experienced instructors such as the ones in this thesis study. I also believe that instructors' prolonged and repeated engagement with errors in general would eventually reveal to them that certain genres demand more frequent and proficient use of certain morpho-syntactic forms and hence they should adjust their expectation of errors accordingly.

4.3.7. Summing-up

I have in this section reported that among the serious anomalies the instructors identified, the two most highly ranked anomaly types were content anomalies and grammatical accuracies. I have also reported how comprehensibility, irritation and
frequency were used by the instructors as criteria for evaluating anomaly gravity. It seemed that although comprehensibility was an important criterion, it was not the only criterion, and there were other principles at work in the process of anomaly evaluation, notions such as irritation and frequency counts of errors, which were possibly pedagogic considerations rather than defined by communication needs or communicability. In this section, I have also reported anomaly gravity due to learner as context and genre as context and showed that these two contexts gave rise to changes of perceived gravity among the instructors. They also prompted certain teacher reaction to specific anomaly types in anomaly treatments. Finally, I have given evidence to show that percepts of the gravity of certain syntactic errors can be a function of the genre they appear in.

4.4. Research Question No. 4: What is the relationship between instructors’ perceptions of errors and the pedagogic objectives of the course as they perceived them?

To answer this research question, I provide data that explain why in their treatments of all anomaly types, two instructors reacted more often to content anomalies, reflecting how they interpreted the nature of the course. I also explain why two of the instructors responded more to grammatical accuracy than to other error types in their recalls of specific errors. A discussion of these two observations together illustrates the relationship between individual instructors’ perceptions of anomalies and the way they understood the pedagogic objectives and the focuses of their instruction in the course.

4.4.1. Content anomalies

That two of the instructors reacted much more frequently to content anomalies than to other error types (See Table 4.2) is particularly interesting given that the content category accounted for only 22% of the whole 100 marks for any particular assignment.
(The marking scheme in Appendix E shows that content and vocabulary both carried 22% out of 100% of a student report; on the other hand, grammatical accuracy and organization each carried 28%).

In the parts of interviews where I elicited instructors’ general descriptions, I asked them to justify a content category in their marking scheme. The way they responded to this question tended to give clues to why there were individual differences in their responses to different anomaly types.

Edward

In the follow-up verification interview Edward gave me, one of the follow-up questions I pursued was whether he thought content was an important category and how he justified the inclusion of content in his marking scheme. This was his reply:

Edward: So what was I saying? Uh, content. Yeah, I think content is important, I mean these are university students, they are given a task, and uh you know, I mean, you are supposed to be able to use their brain and then think about the task itself. And and I guess is that if a piece is well-written, then I am willing to overlook a fair amount in terms of content, not organization, but content. I am willing to overlook a fair amount, as long as I don’t get the impression that this is some kind of memorized memorized piece when the writer is trying to uh pretend they know more than they really do. ... as long as- as long as I feel that piece is not a memorized piece, where the content is zero because the person is just trying to make it useful in a particular context, I don’t care, but I think that content is important. Well, you know, you are trying to communicate something, uh I mean, it’s uh, you know, there has to be some kind of logic to it, partly because if there is no logic, the writing- the language will probably become more prominent. [E4-3-3]

What Edward seemed to be implying here is that content is an important vehicle for language proficiency. It gives language its substance and meaning, and instructors’ responses to content anomalies are a way for instructors and learners to relate meaning to language.

When asked how to assess the successfulness of his teaching, Edward was modest in his response. However, he expressed that he enjoyed doing this course more when
compared with an English academic writing course he was also teaching. He stressed that he enjoyed teaching this course more because the course placed emphasis on content:

Edward: I don't know if I have succeeded, because most of my students are fairly pathetic in class, but try to try to uh, try to get them involved than just get them to talk in the language, I guess, as much as anything else through content-based instruction, in the way. That's why I probably enjoy teaching the Technical Writing course more than the other first year course, the non-credit course because at least in Technical Writing, you are supposed to communicate with some kind of purpose in mind. So I guess that's why. [Para. 900]

Marge

When I asked Marge how she would justify instructors' responses to content, she referred to the marking scheme:

<table>
<thead>
<tr>
<th>Researcher:</th>
<th>Why why do you think you should respond to content?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marge:</td>
<td>Because of our marking scheme. Have you seen the marking scheme?</td>
</tr>
<tr>
<td>Researcher:</td>
<td>Yes, I have.</td>
</tr>
<tr>
<td>Marge:</td>
<td>Okay. The marking scheme gives 22 marks to content. You know, again, it it- again, I am trying to re- I am trying I try to stick to the marking scheme if possible. [Para. 47]</td>
</tr>
</tbody>
</table>

On hearing this, I pressed for a further justification for the marking scheme:

<table>
<thead>
<tr>
<th>Researcher:</th>
<th>Yeah, yeah. But how do you- how do you justify uh uh looking into the content of these written reports, for example?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Marge:</td>
<td>Because for these students, particularly, the content is probably more important than anything else. I mean, if they are writing a report for their supervisor, it doesn't matter how beautifully written it is in terms of grammar and vocabulary. If it doesn't make sense, in terms of what the job is, and what has to be done about the problem, then they as engineers are gonna be penalized, something they have to. We have to work very much on content. This is not a sort of an ordinary ESL class, where we are just practising writing. We teach them to write engineering reports. So for me, the content is probably the most important aspect of the whole thing. And organization, the second most important aspect, language is really- comes third. [Para. 52]</td>
</tr>
</tbody>
</table>

My interpretation of the reason Marge gave here is that she saw a link between the predictive validity of the present assessment scheme in the future needs of technical writing in students' professional work. Because the reports the students would write in the future
would be a tool to solve real world problems, therefore, the content of the reports they wrote now was the most important category of all four in the marking scheme. In the following, I will contrast Edward’s and Marge’s perception on this issue with that of Anna and Kelly.

Kelly

Kelly was the one of the two instructors who responded most frequently to content anomalies. (Of all Anna’s responses to anomalies, 63.9% belonged to the content category; 43.8% of Kelly’s responses to anomalies were content anomalies.) The reason he gave for responses to this anomaly category again reveals why he did it in the way he did. But Kelly’s response in this respect is rich and interesting, so I would like to give it a more thorough discussion.

One of the key notions that Kelly used to explain why a mistake is considered a mistake is “know your reader” (or “not knowing your reader” being a learner mistake). This is a notion that he used to explain the gravity of three of the four anomaly types (content, vocabulary and organisation) though mostly on vocabulary and content anomalies. I went through all the errors he recalled and did some tallies.

\textit{Organization mistakes.} Kelly used the know-your-reader rationale to explain the seriousness of two organization mistakes in his recalls. An example he gave is as follows:

Kelly: ... what I try to teach the students is that to think about what they’re writing to group things in a logical order. When I wrote up, I wrote this test [the prompt], so when I wrote up the thing, I purposefully had the person [in the prompt] walking around, from the cottage to a shed, back to the cottage, to the boat-house. It kept on bouncing back, but I expected the person [the writer] to put, if he’s gonna to discuss the cottage, he’d discuss it all at once, not just followed the organization I wrote up and walked around. But this person just wrote it as he walked around, so it makes no sense to the reader. [Para. 210]
Chapter Four

**Vocabulary mistakes.** He used the know-your-reader rationale to explain the seriousness of 11 vocabulary mistakes. An example Kelly used on a vocabulary mistake was as follows:

Kelly: Uh, and on the last page, uh, he was one of the few, there is something in textbook I reviewed called uh low information content words, kinds of words that the reader looks at and uh it's just adds words to it that the reader has to figure out, and so words like "aforesaid", "upon performing the aforesaid renovations". I told the class, we did exercises, not to, you know, use vocabulary like that. It just makes the reader have to read slower, less confidently. [Para. 316]

**Content mistakes.** Kelly used the know-your-reader notion to explain the seriousness of 12 content mistakes. For example:

Kelly: It's uh, uh, he didn't write how he did the inspection. I mean, and that's something and that's something I go over with the students quite a bit that they should explain. So I wrote: "How did you do it? Did you talk the police? Did you walk around? Did you go swimming in the lake?" I mean, there's different ways of doing the inspection. Also, he didn't even say what he inspected in the background. So the reader doesn't know what's being inspected until they get to the main findings. So that's, you know, things that are missing. [Para. 222]

Kelly: Okay, still staying in the background, which is still only two sentences long. Uh, he explained uh, it's knowing the reader, uh he is writing to a Board of Director that used the cottage as a summer place. So he is writing to a group of people who go to the cottage all the time. Yet he writes exactly where the cottage is, on Tenor Lake, located approximately 50 kilometers north of Ottawa. So he is giving too much information for the reader. You might need to know where the cottage is because you don't know. But the Board of Director, everybody on the Board knows exactly where the cottage is, so they don't need that details. [Para. 224]

Throughout Kelly’s recalls of specific errors, I counted Kelly using the know-your-reader notion a total of 25 times to explain anomaly gravity: 12 times for content anomalies, 2 times for organisation, and 11 times for vocabulary anomalies. To put this in a perspective where individual instructors’ behaviours can be better understood, I have also counted the number of uses of “know your reader” as an explanation for all serious anomaly categories by all four instructors, as shown in Table 4.6.
Table 4.6. Frequency count and percentages of the uses of the know-your-reader notion to explain anomaly gravity by instructor

<table>
<thead>
<tr>
<th>Instructor</th>
<th>No. of times the know-your-reader notion was mentioned in the instructor recalls</th>
<th>No. of distinct anomalies &amp; percentages the notion was used to explain anomaly gravity</th>
<th>Total No. of all anomalies commented on</th>
</tr>
</thead>
<tbody>
<tr>
<td>Anna</td>
<td>11</td>
<td>10 (25.6%)</td>
<td>39</td>
</tr>
<tr>
<td>Edward</td>
<td>3</td>
<td>3 (2.4%)</td>
<td>125</td>
</tr>
<tr>
<td>Kelly</td>
<td>45</td>
<td>25 (26%)</td>
<td>96</td>
</tr>
<tr>
<td>Marge</td>
<td>3</td>
<td>3 (4.76%)</td>
<td>63</td>
</tr>
</tbody>
</table>

Table 4.6 demonstrates the frequencies and percentages of “know-your-reader” being mentioned in individual instructors’ evaluations of serious anomalies. The figures suggest that in nearly one out of every four serious anomalies Anna (25.6%) and Kelly (26%) commented upon, they explained the anomaly gravity by relating it to the notion of “know-your-reader”. The transcriptions of other two instructors, Edward (2.4%) and Marge (4.76%), however, did not reveal this strong tendency. My explanation for Kelly’s and Anna’s relatively higher percentages of responses in the content or vocabulary categories was their individual emphases on communicating to their students the know-your-reader notion, and it was through responses to content and vocabulary errors that Kelly and Anna were able to engage with their students in a dialogue which could persuade them of the importance of an audience-oriented approach to writing.

Anna

I have already pointed out Anna’s tendency to use the know-your-reader notion to explain anomaly gravity. In the following, I will give samples to show how Anna responded to content anomalies. Moreover, I will also show that one important reason for
her doing so was that she perceived that (a) there were important concepts these students would be required to know in the business culture they would eventually go into, and (b) it was only by responding to content anomalies in the students’ writing that she was able to engage her students in conscious talk about the necessity and power of these concepts. But first I will set the scene.

This particular set of examples came from a writing task designed by Anna but was also used by the other instructors. The task required the students to write an investigation report as an Engineering technician who was hired by a consultancy firm to go into the bungalow of the Parsenon family (See Appendix O for a copy of the prompt text.). The Parsenon family bought this house new in 1985 and since then had been complaining about the high gas consumption to the construction company, who had developed a total of five bungalows in this area on Gregory Avenue including Parsenon’s. The construction company, who had found nothing wrong with the construction of the house, now finally decided to authorise an outside engineering consultancy to investigate the legitimacy of the complaint. Hence the purpose of the investigation was to go into Parsenon’s house, identify the causes of the problem and make recommendations. In the following I will give some examples to show how Anna responded to content anomalies in this task.

One content anomaly that many of Anna’s students made in this assignment appeared in the very beginning of their reports.

Anna: Okay, well as far as content on this page, one of the things that I think is important, is that the topic. This is in the title and in the first sentence because this person says: “I investigated the building and the heating system of nine Gregory Avenue Bungalows”. In fact, the purpose was not= Researcher: This is the one that= Anna: A number of students did that and the purpose was not to talk about nine Gregory Avenue Bungalows, it was about one problem. And that would really throw someone reading it as a reader I think. So the topic is uh off there although they managed to get uh lead into the summary a bit, so. [Para. 576]

What one can see on the first page in this student’s text [A-4-17] is, first of all, a report title. It reads: “Investigation report of gas consumption in nine Gregory Avenue
Chapter Four

Bungalows”. And the very first sentence in the very first paragraph reads: “I have investigated the building and the heating system of nine Gregory Avenue Bungalows.” There appear to be two problems there, one in the report title, and the other is in the statement of purpose; the source of the two problems was likely to be the same. I coded this anomaly as “lack of proper emphasis”, under content.

Still another example that Anna gave as a content anomaly had to do with what facts to put down to serve the primary function of the report, i.e., to make a point at a global level such as a stance or a position on an issue. Information in the prompt indicated that the insulation index of the concerned bungalow “does not meet present government standards” but “it met government standards in 1985,” the time when the Parsenons moved in. One student mentioned the former but not the latter. Anna commented that what the student put down was a fact (i.e., that the insulation index was not meeting present standards) but it was a misuse of a fact.

Anna: And also I added to him on the bottom of the page where he talks about the insulation, I said you didn’t tell me, when I read it the second time before I gave him another mark, you didn’t say that. You said that this would not meet the government standard today, but you didn’t say that in 1985 it did. So with that information is false, it sounds like- so that was another thing that we pointed it out. And then we=

Researcher: That’s also a matter of what you mean, I mean it’s not just=
Anna: What that means to say: It’s okay now but it wasn’t then. Yeah, that’s about building requirements.
Researcher: It’s not a matter of whether this is factual or not. I mean what he has put down is a true statement, but it’s not the whole truth.
Anna: Right, right. And it’s like you need to know something about the building code [Para. 673]

I scrutinized Anna’s recalls of anomalies in this batch of reports. From her explanations of the gravity of content anomalies, I sorted out three distinct though related features which, I think, she perceived to be important in a business culture. These three distinct but related notions are (a) know your reader, (b) show awareness of real world constraints, and (c) know who you are in the business world.
Know your reader. The prompt of the task made it explicit that the Marsland construction company wanted the report. It also made clear that Marsland wanted the Parsenons to read the report too because this time they wanted to get the Parsenons off their back. Based on the data and information in the prompt, one would come to the conclusion that the construction company should not be held responsible for the high fuel consumption. Hence one difficulty in this report is the appropriate presentation of the facts the investigator had identified, particularly in answering the question of “whose fault?”

Anna commented on her students’ performance in this respect:

Anna: For someone, like imagine you never think about that, like and that’s why I think for some of them- some of them came right out and said it’s so and so’s fault and not so and so’s fault. Some of them avoided it totally. So I mean you don’t want to come out, if the Parsenons gonna be reading it, saying it’s their fault, and they may have to fix it because it’s their problem. [Para. 677]

Hence the dilemma was knowing that the Parsenons would be reading the report too, having to say the fact that the construction company should not be responsible, but in the meantime not offending the Parsenons, who had been all this time adhering to a false assumption.

Be aware of real world constraints. Another content difficulty in this report is the writing of the recommendations. Parsenons’ complaint was real and had to be dealt with, even if it was true that the construction company had not been at fault. The prompt implicitly suggested some possible causes of high fuel consumption, higher than all the other houses on Gregory Avenue. These possible causes are: The humidifier was not well-maintained; additional insulation would make a difference though the present insulation standard in the house met the government standard in 1985; the Parsenons always kept their home warmer than people in the other houses did. Several students made the mistake in the recommendation of only suggesting that the Parsenons insulate the house, a suggestion which could cost up to one thousand dollars. Anna thought that this was a big mistake; her comment was that the Parsenons should be “given options”. I coded it under content as “not considering real world constraints”:

- 151 -
Anna: I know, it's like, but then he has one section, he was sort of- he just went straight forward, and said they [the Parsenons] should have you know put a thousand dollars on insulation into it. Given this, I really think you should... you can either say, well, insulate this if you want to insulate that too, there is- here’s a degree of cost. But to- maybe for most of them, it was just- they were just thinking and that’s really where the difficulty is, you know, what you say: Is it better than that one? They were just thinking, as an engineer, and while you do it, you do the best thing. It doesn't matter how much it costs, that's what you should do. [Para. 299]

Since there were so many possible causes of high fuel consumption in this house, according to Anna, each of the possible causes should be discussed and given as an option, especially when some of the options (such as asking the gas company to replace the corroded humidifier plates in the furnace) would not cost anything to the house owner. Instead, this particular student, as well as several others, just recommended one solution, the best solution, which is also the most expensive, not considering that the Parsenons may not have that kind of money available, or may not want to spend that kind of money. As many working adults can tell, these sorts of mistakes are not uncommon from students who have no experience in the working world. Perhaps they are particularly common from students who tend to metaphorize solutions to human problems with fixing a machine in the best but most expensive way, ignoring many of the real world constraints that surround the problems. Anna continued:

Anna: You know, uh, yeah, you got to fix this bridge, it costs you three million, and that's what you do. Uh, the fact that we don't have three million is irrelevant, well, it is relevant! Because that's where you are writing it for [Laughs]. So, you know, again, it's the culture of the business, and they are just starting to learn their science, let alone the politics of it. Right, I mean, you know, it's more than just the fact here. [Para. 299]

Here Anna hit on the notion of the culture of business, a world where many factors have to be calculated and weighed before a decision can be reached.
Chapter Four

Know who you are in the business world  In another student’s report [A-4-18], the question of who-is-the-writer also came up. The sentence Anna commented upon in the student text reads like this: *“I was authorised by my supervisor, Marg Brock, to evaluate the gas consumption in Mr Parsenon.”*[sic.] The name of the company that the writer and his or her supervisor Marg Brock were working for is called H. L. Winman & Associates, an engineering consultancy firm:

Anna:  We talked a bit about this when I started with him about who is the reader [Anna talked to this student in class shortly before the interview.] because I said to him, they [the readers] don’t really want to know who is your supervisor, Marg Brock or not. I said, but what you want to say is you represent Winman now. So you are “we”, you are Winman, and Winman was authorized by the client, Marsland Construction Company, you know, who cares about Marg Brock? Uh, and then he organized this part, which is good, and I said good, but ... [Para. 652]

There were two content anomalies in this sentence. The first one was the unnecessary mention of the writer’s supervisor. The other anomaly, which I wish to focus on now, was the use of “I” in the sentence. At this point in the interview, Anna wanted to continue to move on to other anomalies in the text, but I almost lost track of her because I was still wondering if the student had really understood his own error. I asked Anna to go back a little bit:

Researcher:  Can I go back a little bit? Just now when you talked to him again [in class], did he ask you about these?
Anna:  I actually pointed it out because I went through with him and it took me a little while because he still did not get it right away, and I kept saying, okay like when you go out, I gave him another example, I said, okay, let’s say you are doing a project for MacDonald’s, a report for MacDonald’s, and you work for Winman. When you go out there, you are not John Wong, you are Winman. And who asked you to do it? It’s not Marg the supervisor, it’s McDonald’s who asked you to write this report. It’s Winman, you know, so who you are changes. Uh, if you are in your job, and your supervisor asks you to write a report for another department in your company, and then you know, you are John in this department, so that that whole thing is a culture thing, I think, you know, who am I, what am I doing, what am I writing this for in the context of it and then all that. [Para. 654]

From this segment, once again I see Anna identifying a “culture thing”. I also see that Anna had really made great efforts to get the message across. This was not a question of misusing a pronoun, “we” or “I”. This was about an understanding required of the
writer about the workings in the business culture, and about who one really is in the various business contexts. What I can also see here is that Anna had in this segment of recalls shifted from telling why she was giving feedback on content anomalies in a writing task with an ESL focus to exemplifying the kind of cultural knowledge that a writer has to bring to the task for effective communication. This example once again pointed to an observation I made earlier: Anna, (or perhaps even Kelly and even Edward or Marge also) felt the need to respond to content anomalies because it is through responding to these anomalies that she could impart a bit of the culture of technical and business communication, and all that social understanding and world knowledge has to be put to use in the writing. It was through responding to content anomalies that the instructor was able to engage the students to talk consciously about ESL literacy. Such a need is pertinent because, as Marge put it, this was no ordinary ESL course, but a preparation for communication needs in a different community with a culture unfamiliar to the students.

In my answer to Research Question No. 3, I reported that two of the instructors, Kelly and Anna, demonstrated a bias towards content anomalies in their recalls, a behaviour that distinguished them from the other two instructors, Marge and Edward. My answer to Research Question No. 4 up to now has explained why this was the case by establishing that:

1. The ways Kelly and Anna perceived serious errors have to be understood as the ways in which these two instructors perceived the pedagogic objectives and the main focuses of the course. An umbrella concept that seems to be at work here for Kelly and Anna’s perceptions is their understandings of how the notion of know-your-reader should be understood by the students.

2. Anna’s attention to content errors was also very likely to be motivated by her understanding of the importance of imparting business culture to her ESL engineering students.
4.4.2. Grammatical errors

In this section, I show that Edward and Marge’s percepts of the importance of grammatical accuracy are also in similar fashion related to certain distinct features in their understanding of the pedagogic objective of the course.

Marge

As I have reported, in the general description of instructors’ BAK systems, Marge cast her opinion that the primary objective of the course was content, and that organization ranked second and language ranked third. How should this perception be understood when in the recalls of specific errors Marge responded more to grammatical errors? I brought up this apparent contradiction when I sent Marge a description of her perceptions. I asked her about this in a follow-up question:

Follow-up Question No. 2: “Do you see a contradiction between your highest grave error rate in the accuracy category and at the same time your recognizing that the course objectives are primarily content and organization? What are some of the possible explanations, in your view, of this incongruence?” [Marge, Description]

This is how she answered me in an electronic message from overseas:

This is a very important question as it relates to conclusions you have already drawn. Be very careful to consider all the feedback we provided in your analysis. For example, the feedback provided to students in the marked reports would always concentrate on language errors as individual students made different errors, and as I explain in the transcript, I deliberately concentrate on those errors at their request.

I dealt with content and organization in a different way. Because the report format was prescribed by the textbook, I used OH [Overhead Projector] models to demonstrate effective and non-effective reports. I would then relate it to their marks in terms of what was a superior as opposed to an adequate attempt. I would assume the other instructors also focused on content and organization in classroom feedback and grammar mistakes in tutorial feedback. [Marge, Follow-up]

Marge’s answer in her e-mail was that although in her marking she seemed to have focussed on grammatical errors, this did not mean she did not respond to other anomaly
types. She usually provided in class her feedback on content and organization anomalies, often by representing in class good and bad models. I went through my field notes, which I had taken during the limited hours I was granted to sit in her class, and I found my notes verifying Marge’s explanation. On the other hand, I also saw that in fact all the instructors teaching the course gave in-class feedback and corrections on organization and content anomalies. I therefore find that Marge’s email explanation may not account for the incongruence I had asked her to explain.

However, I do hypothesize that Marge’s bias towards grammar in the specific recalls could have also been the result of an observer effect, namely that Marge focussed on grammatical errors because she conceived that morpho-syntactic errors were the kind of things that I was looking for. I have some evidence for this hypothesis. I have reported in Section 4.1.3 Marge’s interpretation of the term “error” earlier in our first interview. Marge later defended again this narrower notion of error in her answer to another of my follow-up questions, in which I asked if discourse anomalies can be considered errors and if not, how are they different from grammatical errors. This is how my question read:

Follow-up question No. 3: Are syntactic rules and pragmatic/discourse rules different? In what ways are they different so that a violation of the former is called an error but a violation of the latter is not? [Marge. Description]

In her follow-up response in writing, Marge, for the first time since I began a dialogue with her on errors, agreed that there are in fact two kinds of errors: linguistic errors and writing (discourse) errors. Marge’s answer was that morpho-syntactic errors and writing (discourse) errors are not the same because the determination of discourse errors varies with the context in which the language is situated, while linguistic errors are context-independent. Furthermore, she observed that discourse errors are in fact writing errors, not language errors. As she put it:

Marge: I am not sure how to answer this question [to Follow-up Question No. 3] because I do not know how you define rules of discourse. For example, there are rules of
discourse which a native speaker would follow in any situation (unless he made a slip) but they tend to be related to syntax (e.g. deixis). However, there are rules which apply to a particular genre (i.e. a technical report) which change when the context changes. I think my answer is that breaking the linguistic rules marks you as a non-native speaker, which may be more or less acceptable depending on the tolerance of the host culture. The consequence of breaking the second set of rules results in a less effective or inappropriate piece of writing. It is a writing, not a language problem, and certainly not confined to non-native speakers of English. I hope this helps. [Marge, Follow-up]

I find these views interesting and revealing, especially the part about language errors being context-independent, and I will discuss this issue further in the next chapter. But when Marge cast the opinion that language errors are not the same as writing errors, she seemed to be suggesting that although organizational and content anomalies are also errors, they are writing errors but not ESL errors, which is more of the error types she had been more concerned with. I believe that when she meant language errors, she intended a notion such as ESL errors, or errors resulting from weak linguistic competence in a second or foreign language. And all the time, she did not think that I was looking for any writing errors.

In sum, as she suggested, perhaps Marge’s bias towards grammatical accuracy could have been more a representation of her feedback strategy than a balanced picture of what she would consider as serious anomalies. There is also the possibility that, as a conjecture, her bias towards grammar was a result of how she had pre-conceptualized errors for the purposes of my study.

Edward

I have reported how Edward responded when I asked him to justify looking into the content of students’ writing. In general, he felt that the content was there to serve and contextualize language and to give language use the logic of things. This understanding of the meanings in marking content anomalies would probably explain why Edward was not in general as keen on this type of anomalies as Kelly and Anna were.
Edward’s distribution of comments on specific errors demonstrated a high percentage of grammatical errors. How should this behaviour be understood in the light of his understanding of his own teaching mission in this course?

One interesting idea that emerged from Edward’s general descriptions, which all the instructors had spoken on, but about which Edward talked in much greater depth, is the social stigma that is attached to linguistic inaccuracy.

Edward: I mean accuracy I worry about in another sense, also simply they are stigmatized, that it doesn’t matter if they write the letter- it doesn’t matter that I can understand it because if the spelling is incorrect, and the punctuation is incorrect, I still think of them somehow as being substandard in terms of their intelligence or whatever. So they have to be accurate not just because it is important but because of the ways they penalize themselves socially, you know, by making difference in getting the job and not getting the job, getting the proper care or not. So I think I worry accuracy in that sense, trying to get it across as they are being important, not because they are rules you have to obey or something.

Para. 928

In this segment, Edward told why linguistic errors are a great concern. He took them seriously because he perceived that society as a whole penalizes those who commit linguistic errors and consider them substandard, and it is to the learners’ advantage to reduce their errors so that they would not be penalized.

When I asked Edward to compare ESL errors and native speakers’ errors, he mentioned linguistic errors as social stigma again, but this time his focus was on errors committed by native speakers.

Researcher: ... How is uh how do you compare native speaker's errors or mistakes, and second language learners' errors?
Edward: Uh, I haven't spent much time studying native speakers' error- but uh but uh or mistakes, whatever, uh.
Researcher: I’d like to seek your intuitive feeling.
Edward: I said I tend to hold a sociolinguistic view on a lot of these things, and so I find it stigmatizing. Uh, I probably have a value judgement. I mean if I find my native speaker consistently make errors, then I'll say, you know, uh, either you are lacking in intelligence, or you have a poor education, or
Edward’s understanding that linguistic errors are stigmatized may have made him think that one of his teaching goals was to rid the learners from giving the impression of being a non-native speaker. This is suggested in the segment immediately below. While he overtly explained that he had responded to this particular student’s article errors because this was an “A minus” student, Edward also implied that attaining the proficiency (at least the impression) of a native speaker was his teaching goal, as is shown in the following extract:

Edward: So you know I think you can start focusing on things that make him [this A minus student] different from the native speakers. I am conscious of this, or maybe this is because I was coming from anthropology and sociolinguistics, but I mean I know perfectly well that a person might be perfectly fluent in English and all he has to do is to screw up the article system and people would start thinking of him as substandard. That’s how a native speaker might think. So so that’s so I am conscious of that, I mean, for some people it's not worth bothering with errors like that, because there are far more serious errors. This is not a serious error, but again this is the kind of errors that tell somebody who is reading this that he is not a native speaker, so I think you know he is okay otherwise.

When commenting on irritating errors, Edward also said that he thought that violations of trivial rules like third person singular “gives a poor impression,” and in his words, “it looks sub-standard”. From these examples, I infer that Edward’s understanding of how society as a whole perceived grammatical errors is constitutive of his own error evaluation principle.

Another explanation for Edward’s higher response rate to grammatical mistakes than to other anomaly types could be that Edward preferred attending to recurring problems. This is a marking or teaching strategy perhaps only relevant to grammatical errors, as teaching in the course tended to progress from one document type to another, and each task had a different content and context. The only recurring anomaly type in Edward’s data which could be easily identified were of course grammatical errors, either
in the same task or from task to task.

4.4.3. Summing-up

In Section 4.4, I show that the four instructors could be divided into two groups according to the patterns through which they each responded to serious anomaly types. Edward and Marge tended to respond more frequently to grammatical errors than other anomaly types, and Anna and Kelly responded more to content anomalies than to grammatical errors. These two response patterns to serious anomalies, I infer, tend to be related to how the instructors understood the main objectives and focuses of their teaching.

Anna and Kelly were keen to impart to their students an audience-oriented approach; Anna also emphasized the importance of learning the culture of business, which she believed would be essential for effective technical writing. Edward and Marge did not see the importance of content in the same way as Anna and Kelly did. Edward understood content to be important because it provided a context for the use of language. Marge believed that content was important as a predictive validity issue; to her, content was important primarily because it would be the most important thing for the writing of practising engineers working in their professional world.

On the other hand, Edward and Marge attached greater importance to grammatical errors primarily because they saw the undesirable social consequences that grammatical errors could bring to ESL writers or speakers, and to their readers or listeners. In Edward’s view, grammatical errors are stigmatized, and lay people would even consider certain grammatical errors indicators of “substandard” language. It is against these undesirable social consequences that Edward believed his learners should rid their writing of errors.
CHAPTER FIVE
DISCUSSION:
AN INTERPRETATION OF THE INSTRUCTORS' PERCEPTIONS

I reported in Chapter Four the findings from this thesis study. In this chapter I wish to offer an interpretation of these findings, but before doing that I will begin with a summary of Chapter Four. I do this to provide a platform for my interpretation of these findings. I also find it appropriate to amalgamate my interpretation of instructors’ perceptions with a discussion of the relevant literature. It is the primary task of this chapter to give a coherent interpretation of the instructors’ percepts, but at the same time, I also interpret what I know from related literature, since what I know from the literature informs my interpretation of findings in this thesis study.

5.1. Summary of instructors’ perceptions of anomalies

Findings in my thesis study confirm the value of taking a constructivist approach to error perception. In this perspective, error evaluation is understood as a dynamic, context-sensitive, interactive process. This thesis study in a constructivist approach showed that the assignments of gravity to grammatical errors and other anomalies were often under the influence of the following contexts: (a) learner as context, (b) the assessment purpose as context, (c) the genre in which the error appeared in, and most importantly (d) teacher as context.

5.1.1. Learner as context

One of the contextual elements affecting the instructors’ error perception was the learner. The instructors reported that based on the needs and strengths of a particular
learner, the instructors adjusted the gravity they assigned to the error identified in a way different than they usually did. For example, Marge gave more allowance to a student who she believed was experimenting with newly learned language items. Similarly, Anna considered the misspelling of the word “machine” to be very serious, primarily because she felt that this was a word an engineering student should not have misspelled. Edward made exceptions to his do-not-correct-article-error policy when he came across an advanced student who might benefit from such corrections. In sum, these reassignments of error gravity were out of certain pedagogical considerations based on the perceived needs of the specific learner being considered.

5.1.2. Assessment purpose as context

Instructors’ assessment activities were influenced by the specific purpose of the assessment. At least two of the instructors changed their response patterns to errors depending on whether a mark had to be given to the students’ work. When marking for grades was not necessary, one instructor in general paid attention to only content and organization anomalies but not to grammatical errors. This suggests that the error evaluation procedure is likely to be structured by the purpose of the evaluation activity, as individual teachers perceived them to be.

5.1.3. Genre as context

There were instances in which genre considerations (e.g., discourse rules) determined the instructors’ perceived need to apply certain morpho-syntactic rules. These were instances in which the determination of grammatical errors was actually dependent on the genre in which the errors appeared. These instances demonstrate the great extent genre as a context can exert on error evaluation. On the issue of assigning gravity to errors, the genre in which errors appear also played a role. In Chapter Four, I identified a
close relationship between genre and certain morpho-syntactic forms, namely, reliance on the use of relative clauses with a definite article to identify things, people and events in incident or investigation reports. In the example I showed there, a total of 27 (about 80%) of Edwards' 34 students attempted to use a restrictive relative clause with a definite article to identify for the first time a device called "air sampler" in the writing of their incident reports; out of these 27 students, a total of 17 failed to use a restrictive relative clause with a definite article correctly, resulting in Edward's remedial teaching or reteaching of this syntactic structure. There I also presented samples to show that Anna carried out a similar reteaching of the same syntactic form in her lessons.

5.1.4. Teacher as context

In Chapter Four, I identified certain differences among the instructors' conceptualizations of errors. Some of these differences were in the definitions of errors, mistakes, and slips instructors initially offered in their interviews and in the instructors' response patterns to serious anomalies. During the interviews, two of the instructors, Marge and Edward, used the term "error" to refer to violations of grammatical rules and this kind of violation only. While "error" to the third instructor, Kelly, was purely determined by declarative knowledge, the fourth instructor, Anna, preferred to call any anomaly an "error". However, prior to my question asking explicitly for a definition of error, in their stimulated recalls when the instructors commented one by one on the serious errors in each of their students' written reports, they all used the terms "errors" and "mistakes" interchangeably to refer to anomalies in either content, organization, vocabulary or grammatical accuracy. It was only later in their follow-up responses that the instructors amended these definitions to a broader notion that covered all sorts of anomalies; they recognized that in their assessment practices that they all did not differentiate mistakes from errors. For example, Marge in her follow-up response explained that content and organizational anomalies were writing errors while grammatical
errors were ESL errors. She also maintained that breaking linguistic rules marks the writer as a non-native speaker, while violations of discourse and genre rules are considered writing or literacy errors.

Patterns emerged in the instructors’ responses to anomaly types. Two of the instructors, Edward and Marge, commented more on grammatical errors than other anomaly types whereas, the other two instructors, Anna and Kelly, responded more to content anomalies than other anomaly types. In comparing the language and the serious error types the instructors responded to, I can see that on the one hand, although Edward and Marge followed closely the assessment scheme they were using, the “point(s)” they were making in their comments on errors displayed a tendency towards ESL language proficiency or grammatical competence and did not really echo the objectives of a genre-based technical writing course. On the other hand, in Anna’s and Kelly’s responses to serious error types and the teaching points they were making, what they were trying to get across to their students did seem to echo the perceived objectives of the course, placing emphasis on the development of ESL literacy.

Differences between these two groups of instructors were also apparent in their two typically different understandings of their teaching focuses. Edward and Marge considered it important to rid their learners of grammatical errors, which they believed were stigmatized (Edward’s view) and may lead to serious social consequences (Marge’s view). Anna and Kelly, though they also discussed in general the consequences errors could bring to the reader and the writer, emphasized in their teaching an audience-oriented approach to technical writing (Anna’s and Kelly’s view) and an understanding of the culture of the business world (Anna’s view). These different understandings of the teaching “points” (in the sense of teachers’ specific motivations to teach and learn) between these two groups of instructors seem to relate to the difference in their response patterns to specific serious errors.
Differences in the instructors' evaluation behaviour also suggest that of all the contexts being examined, instructor as context is likely to be the most overriding one. Individual instructors applied unique perceptions of students' errors. This is probably because it was the instructors, the participants in this thesis study, who determined and took the lead in the teaching activities, of which the evaluation was a part; it was the instructors who were judges of the particular evaluation activities; and it was the instructors who considered what, when and how contexts came into consideration in the evaluation of a specific error.

In spite of the individual differences among the instructors in their evaluation, my analyses revealed certain commonalities among the instructors. One commonality was the use of the same criteria in evaluating serious errors and minor errors. The criteria the instructors used included at least comprehensibility, irritation and error quantification. My data show that the instructors were not always aware of how various contextual elements acted upon error evaluation at the time of the evaluation activities, though they all tended to agree about this mechanism in their follow-up responses. On the other hand, however, they all reported their overt, conscious uses of comprehensibility as a single, major criterion for determining serious errors, along with the use of irritation and error quantification as evaluation criteria. Two of the instructors', Edward's and Marge's, uses of error evaluation criteria indicated that they were particularly conscious of how society as a whole perceives people who make grammatical errors and other anomalies as well as the social stigma laypeople may attach to errors.

On the question of what error is, the majority of the instructors, Edward, Kelly and Marge, tended to focus on offering me a nominal definition of error. Three of the instructors' percepts of errors stopped at the point where errors are considered violations of native speakers' language rules; they did not discuss at all what error meant to them. However, in her answer to my question "What is error?", the fourth instructor, Anna, reported a very adequate, summative synopsis, in my view, of the notion of error from an
SLA perspective. This perspective sums up not only the notion of error in error analysis but also in interlanguage studies and learning strategy research. This understanding of error indicates the many faces error can have to different people at different times and in different contexts. I therefore think it worth repeating the excerpt projecting this view:

Anna: I think it's a lot of things, just depending on who you are or which way you are looking at it, what hat you're wearing, you know. ... You're [the learner] working something out here. It's- it's a mistake. It's something you're doing in a way, it's the wrong, it's a rule you don't have, it has stopped as a rule- being an English rule, you know. Uh so you know, it's obviously a developmental thing and you have to work through it. On the other hand, it's something in some way does not fit what the native speaker does. It's always that dilemma of saying "yeah you have to make mistakes in the process to get it to this point, and at the same time you want to eliminate them", you know and yet the question is you're never really gonna be eliminating them. [Para. 179]

In sum, certain contextual elements acted upon the activity of error evaluation. The instructor as context was the most important of all the various contexts identified. The other factors that came into instructors’ considerations in the error evaluation processes were the genre the error appeared in, the specific needs and strength of a particular learner, the specific assessment purpose, and how seriously the error hindered comprehension. In answering the question of what error is, there was also the suggestion from one of the instructors that error has different meanings to different people in different situations.

5.2. Re-defining the scope of error

Having summarized the key findings in this thesis study, I will subject these findings to a discussion, aiming to provide a coherent interpretation of the instructors’ perceptions of errors. I wish first of all to redefine the term “error”. I side with Anna, who proposed in her interviews that the term error covered all kinds of instructor-perceived deviations in their students’ written work. I suggested to the other three instructors that their behaviours in the recalls of specific errors demonstrated a much wider notion of the term error than
the ones they had suggested. This resulted in Kelly's, Marge's and Edward's later amendments to their initial definitions and the acceptance of writing errors and genre errors as well as the interchangeability of "error" and "mistake". Hence, from this point on, I use the term "error" to refer to all sorts of anomalies appearing in these ESL students' written language production. These include discourse infelicities and any kinds of anomalies in the domains of content, organization, vocabulary, and grammar.

The definitions that Kelly, Edward and Marge initially offered in the general descriptions of their BAK systems were basically the ones Corder (1967, 1981) suggested. The scope of this definition, very much restricted to a violation of a morpho-syntactic rule at the sentence level, is obviously much narrower than the one discussed in the first language learning (L1) literature and in the more recent literature of SLA. In the teaching of L1 writing, error is a construct which has a wide connotation (e.g., Haswell, 1988; Nystrand, 1982; Shaughnessy, 1977). Shaughnessy (1977), for example, considered errors which are beyond the sentence level, although her discussion mostly referred to syntactic deviations. These errors beyond the sentence level included the use of rhetorical functions and organization, style and register and, topic development. Nystrand (1982) in his article "An analysis of errors in written communication" considered five types of constraints; they are graphic, syntactic, semantic, textual and contextual constraints. In more recent L2 literature, Lennon's (1991) proposal to consider extra-linguistic context as a unit for examining error gravity is also a broadened conceptualization of error. C. James (1998), in his ESL error description system considered substance errors (e.g., misspellings), syntax errors, and text errors; under text errors he included not only lexical errors and grammar errors, but also discourse errors, the violations of lexico-grammar, for example, developed by Widdowson (1979) and Halliday (1985, 1994). I therefore find it justified to use the term "error" to denote not only grammatical errors but also discourse errors or writing errors, as Marge suggested in her follow-up response. However, the distinction between grammatical errors and writing errors warrants further discussion.
While I eventually gained endorsements from all the instructors on a broadened notion of errors, I note that the discrepancy between what the instructors commented on as (serious) errors and the definitions of errors the instructors initially offered requires an explanation. It was clear in my cue questions in the recalls of specific errors, which in every interview happened before my question on the definition of error, that I was asking the instructors to identify “errors”. Except for Marge, who asked me on the spot whether I was referring only to grammatical accuracies or including content and organization, all three other instructors started commenting on all kinds of anomalies and did not query the scope of the term I intended. As a matter of fact, during the interviews, all four instructors and I myself were using the term “mistakes” and “errors” interchangeably to refer to grammatical inaccuracies and to content and organization anomalies.

Hence my explanation for the discrepancy in what the instructors said and how the instructors behaved is this: The definitions of error that Edward, Kelly and Marge suggested in the interviews all originated directly or indirectly from the error analysis literature these instructors had been exposed to in their education and professional reading. I infer that these definitions were all recalls of a kind of knowledge which Dewey (1938) called compartmentalized knowledge and Whitehead (1929/57) called inert ideas; it did not come from the instructors’ own practical experience (Schön, 1983). Whitehead defined “inert ideas” as “ideas that are merely received into the mind without being utilized, or tested, or thrown into fresh combinations” (p. 1); Dewey (1938, p. 48) also considered that this kind of knowledge is transmitted as compartmentalized abstractions; it can be elicited as declarative knowledge but is often disconnected from a learner’s actual life experience. In sum, I interpret that the definitions which the instructors initially offered, a definition much narrower than the working notion of errors they had been using, is the kind that is seldom applied and realized in the instructors’ everyday educational activities (Johnson, 1999, p. 8).

In my later communications with Edward, Kelly and Marge in July 1999, all three
instructors agreed to the existence of writing errors. In the case of Kelly, although he had maintained that his students made a lot of mistakes but not errors, in his follow-up response, Kelly suggested that in practice he was not able to distinguish “mistakes” and “errors”. In the case of Marge, in the specific recalls she had refused to consider organizational and content anomalies errors, but later when having gone through the transcriptions of her interviews and reflected on my follow-up questions, she proposed that there are not only grammatical rules and grammatical errors but also writing rules and discourse rules. From this I interpret that she would agree to my calling them “writing errors” although “errors” may not be a term she herself would prefer using. Edward also proposed in a later discussion with me that there is the need to tell students that there are genre rules and writing rules, and conceptualizing them as rules might help ESL students improve in their written content and organization, although Edward also thought that grammatical errors are harder to teach and correct than are organizational and content problems.

I appreciate the instructors’ amendments, with which I agree. I also observe in the data that all the instructors in their recalls of errors responded to mistakes and errors alike (but not to obviously careless slips), suggesting that the difference between mistakes and errors in the actual assessment practices of all the instructors in this thesis study was non-existent. This respond-to-all policy they adopted was very likely to be a direct result of the practical difficulty Kelly mentioned, of it being too hard to tell written errors from written mistakes, but then it was also a policy that had its pedagogic value. This policy finds support in Johnson (1996, p. 123), who argued that previous researchers and teachers following Corder’s (1981) distinction between errors and mistakes had devoted a lot of their energy to errors but should now start paying attention to mistakes, which are signs of improper procedural knowledge development, given that procedural knowledge is also an important aspect for language learning.
5.3. Error and genre as context

In her follow-up response, Marge also made the interesting remark that while there are grammatical errors and writing errors, grammar rules are different from writing rules because grammar rules apply universally regardless of the context, whereas writing rules and discourse rules vary with the context. This is seemingly a view derived from early versions of Chomskyan linguistics, which held that syntax is autonomous, and in this regard, I have three points to make.

First of all, Marge’s observation that grammar rules apply universally was challenged by Kelly’s and her own recalls of specific errors. In Kelly’s recall, the genre of technical instructions imposed externally a morpho-syntactic restriction on a sentence, namely that in this specific genre, there should be only one imperative verb in one sentence. In Marge’s recall, she observed that the genre of note form would actually relax the morpho-syntactic rule of English passives and allow the omission of the auxiliary verb in a clause. Both these two instances are counterexamples showing that a use rule imposed by a genre may change the relevancy of applying a specific grammar rule, exemplifying the maximum extent that genre as context can exert on error evaluation.

Indeed the effects of meaning and speech acts on grammar rules is an issue that Lakoff (1987) discussed thoroughly. In Lakoff and his colleague’s recent work, Lakoff and Johnson (1999, p. 483) repeated an argument from Lakoff (1987) to show that the notion of “autonomous syntax” in Chomskyan linguistics is in truth a metaphor, not a fact (in the sense of a theory of correspondence as truth), and that Chomskyan linguistics attempted to complete its logic by admitting that part of syntax that agrees with their pure formalist agenda and excluding from their syntax those syntactic regularities that do not. The examples that Lakoff used to show this is a group of sentence constructions that Chomskyan linguists once called “main-clause” constructions (e.g., “Here comes the bus!”, “Who on earth can stop Jordan?”); early syntacticians took interest in these
constructions because they believed that these constructions could stand alone and hence had a purely syntactic constraint. Lakoff (1987), however, showed that these clause constructions can in fact exist in some adverbial subordinate clauses (e.g., “I’m leaving because here come my bus.”, “The Bulls are going to win because no one can stop Jordan.”), but not others (e.g., *“I’m leaving if here comes my bus.”, *“The Bulls are going to win if who on earth can stop Jordan?!”). Lakoff furthermore showed that the governing rules of these occurrences and non-occurrences are not syntactic rules, but speech acts and semantics. The instances I identified in this thesis study regarding genre as context make the same point that a use rule associated with a genre can determine the relevancy of a morpho-syntactic rule; the only difference is that in Lakoff’s examples the use rules were speech acts realized by sentences; the use rules identified in this thesis study were at a genre level.

Second, I wish to argue for the view that whereas there are grammatical errors (since there are morpho-syntactic rules), there are also writing errors because there are writing rules (genre rules or discourse rules). I admit that the term “error” may be indeed an undesirable label to some teachers — and in this study to Edward and possibly Marge as well. However, that I consider writing (genre) deviations as errors is a perspective based on the works of Halliday (1984, 1985, 1994). In this perspective, language is described in a systemic theory, following a European functional tradition and particularly the Prague school. This functional paradigm of language in the writing of Halliday and other systemic linguists (e.g., Berry 1975, 1977; Martin 1992) considers that language at work develops into a “system network”. According to Halliday, this system network is “a theory about language as a resource for making meaning” (Halliday, 1994, p. xxvi). Unlike Chomskyan linguists’ interests in formal, syntagmatic, sentential grammar, systemic linguists aim to identify paradigmatic relationships among specific lexico-grammatical codes and relationships between specific lexico-grammatical codes and specific social acts (genres) in the culture in which the users are situated. A systemic linguist’s mission is then the description of syntagmatic relationships because, in Halliday’s words, “there is no
difference between describing something and relating it to everything else” (Ibid., p. xxvii); therefore, the task of a description of a feature in a language is to reveal its relationships to all the others, as the use of a restrictive relative clause with a definite article is related to the function of identifying an object or a person in an incident report. It is in this sense that we see how the lexical or syntactic choice of a sentence can be related to a paragraph, and then to the whole essay or a document as a tool for a social act.

As many genre researchers have pointed out, the syntagmatic relationships between a genre\(^1\) and the language that goes with it are particularly stable in certain discourse communities such as classrooms (Lemke 1990), professional domains or academic circles (Bhatia, 1993; McKenna 1997); these relationships are still even more stable and identifiable in written texts (Beaugrande & Dressler, 1981; Hoey, 1983; Halliday & Hassan 1976; Swales, 1990) than those in speech. As a matter of fact, I see that the recognition of these relationships is particularly fundamental for the teaching of technical writing because it is the recognition of these relationships that justifies the substance we teach and the way we organize this substance. Seen from this perspective, then these writing (discourse or genre) rules are certainly fundamental for the learning purposes in order that ESL learners can learn to perform effectively a socially situated, communicative act; consequently, the violations of writing rules or discourse rules can be just as serious as, if not more serious than, the violations of syntagmatic morpho-syntactic rules.

---

\(^{1}\) Although ‘genre’ is defined differently by different writers, it is a term containing some commonly agreed notions, e.g., it has a social purpose, it has recurring properties and it is reusable in similar contexts. I note that some writers (Biber, 1988; Paltridge, 1996) differentiate “genre” and “text type”. Biber’s (1988) “genre” refers to types of social activities such as prayers, sermons, songs and poems, and his “text type” refers to groupings of texts sharing similar co-occurrence of linguistic patterns. Using both ‘genre’ and ‘text type’, however, can be very confusing. For example, what Paltridge (1996) called different text types (e.g., procedure, description, exposition, and recount) would overlap with what Trimble (1985) called rhetorical organizations or rhetorical functions. I use a definition of ‘genre’ suggested in Russell (1997), which is inclusive of text types. Russell (1997, p. 515) defines genre as “the ongoing use of certain material tools (marks, in the case of written genres) in certain ways that worked once and might work again, a typified tool-mediated response to conditions recognized by participants as recurring”. (See Section 2.4 for a discussion in this thesis.)
Third, although I gave examples to show that writing rules override grammar rules and I consider that there are grammatical errors and writing errors (genre errors or discourse errors), I also hold that there are differences between them. Some studies have already established a distinction between writing ability and language proficiency from a cognitive perspective (e.g., Cumming 1989). My focus on the distinction of the two here is one that is socially oriented. The distinction becomes more lucid in light of Leech’s (1983) use of “constitutive rules” and “regulative rules”, two notions coined by Searle (1969, 1995, 1998) to differentiate what Leech called pragmatic (discourse) rules and language (grammar) rules. According to Searle (1995), rules such as “drive on the right-hand side of the road” is a regulative rule while rules of chess are constitutive rules. Breaking a traffic rule to others means bad driving, but it is still driving, and driving can exist prior to the existence of that rule. In a chess game, if a rule of chess is violated, people no longer consider it a game of chess. Hence the rules of chess define the game, but traffic rules do not define driving. Leech (Ibid., p.8) observed that use rules in language are regulative rules while morpho-syntactic rules are constitutive rules. Leech explained:

Hence if one tells a lie in English, one breaks one of Grice’s maxims (a Maxim of Quality); but this does not mean that one fails in any way to speak the English language... On the other hand, if one breaks the rule for tag-questions by saying *We’ve met before, isn’t it?* rather than *We’ve met before, haven’t we?* One thereby fails in some particular way to speak the English language.

It is basically an institutionally assigned social convention that people call “a violation of a traffic rule” “bad driving” and not “non-driving”, but that people understand “a violation of a chess rule” to mean “not really playing the game” and not “bad chess play”. Hence errors are what Searle called “social facts”, and this is true with either “grammatical errors” or “discourse errors” (discourse infelicities). Violations of grammatical rules can result in granting a reader or a hearer the right to say that the
speaker or writer has failed to use the language. Violations of discourse rules (writing rules or genre rules) can result in granting a reader or a hearer the right to say that something is bad writing or bad speech. In either case, however, the notion of error is by nature a social construct, and what Searle would call an institutional reality (Searle 1998, p. 124).

5.4 Error evaluation and the contexts of teaching

Some previous error gravity studies (Delisle, 1982; James 1977; Politzer, 1978) have suggested the existence of certain error hierarchies according to their severity. For example, C. James (1977) found that of the ten ESL error types he compared, the most serious type were transformations and tense, followed by concord agreement, case, negation, articles, and word order. James (1998, p. 232) also summarized Politzer’s (1978) study into a chain of inequalities:

MOST SEVERE: LEAST SEVERE

Lexis > verb morphology > word order > gender > phonology > case marking

The problem with many such suggested universal error hierarchies is that they seem to be meaningful only under the quantification of when other things are equal, but as Archibald (1994, p. 79) put it, in reality and in the context of things as they happen, “[A]ll things are, however, rarely equal and when comparing two errors it is usually the case that several factors are involved in each.”

My findings in this thesis study uphold a constructivist approach to error perception and yield the implication that because error evaluation is basically a context-sensitive activity, the existence of universal error hierarchies, such as the one James (Ibid.) summarized above, is neither feasible nor useful — because of its lack of descriptive and
predictive power. The present findings suggest that the instructor, the learner, the genre, and the evaluation purpose all have a role to play in the assignment of gravity to errors.

An interactive view of error evaluation is also supported by a scrutiny of the criteria these instructors consciously used in their error evaluation activities. The instructors often considered that comprehensibility was an important, single criterion for error evaluation. Since comprehensibility is a meaning criterion, it is difficult to see the validity of an error hierarchy such as the one mentioned above (e.g., Politzer, 1978), in which the level of seriousness is determined by grammatical categories. After all, comprehensibility can be blocked by a lexical item, by the word order, or by a case marker, depending on the meaning an error carries and the context the error appears in. The observation that comprehensibility has little to do with grammatical errors tends to concur with a finding by Munro and Derwing (1995). They found in their study of second language learners’ speech that phonemic errors, phonetic errors and grammatical errors were more related to accentedness than to perceived comprehensibility; furthermore, these errors are still much less related to intelligibility, which they believed was a more direct measure of what the listeners in their research study actually understood than perceived as comprehensibility.

This thesis study has also identified instances in which comprehensibility gave way to other criteria and these criteria were all of a pedagogic nature. The use of comprehensibility to evaluate errors, especially grammatical errors, is understandable since the grammar category of the assessment scheme the instructors were using actually specified meaning as a criterion, (e.g., “meaning is not obscured by the errors”, “meaning obscured or confused” — See Appendix E). It is difficult to pinpoint in my study exactly why the instructors made comprehensibility a criterion for assessing grammatical errors. The specification in the assessment scheme that comprehensibility was a criterion for evaluating grammatical errors may have come as a subject requirement, as in the
assessment of mathematics, where accuracy is always a requirement and as in the assessment of physical education, where speed and stamina are often requirements; in the subject of language or communication, comprehensibility may be a fundamental requirement. Comprehensibility as a criterion for writing assessment and error evaluation could be a predictive validity issue too, because the students would have to write comprehensible texts for their future jobs. Comprehensibility as a criterion may also have come from the influence of a communicative language teaching approach, in which case it is basically also a pedagogical consideration. In any case, I interpret that the use of comprehensibility as a criterion for evaluating grammatical error to a great extent must have been pedagogically motivated. This means that comprehensibility was primarily used as a criterion for grammatical error evaluation to serve teaching.

I identified many instances to show that learner as context and assessment purpose as context had effects on error evaluation. Some of these instances were, for example: Kelly determined error gravity based on prior instructional advice; Edward made exceptions to his do-not-mark-article-error policy when he identified a more advanced student; Anna considered the misspelling of the word “machine” serious because it was from an engineering student; Marge adjusted her percepts (allowance) for the errors in a piece of writing when detecting the student’s attempt to experiment with newly learned items. These phenomena together suggest that error evaluation was subject to the activity of assessment, which in turn served the activity of teaching.

Individual instructors’ different response patterns to errors also suggest that instructor as context was perhaps the single, most important contextual element of the various contexts in the error evaluation activity. This result tends to concur with findings in many previous experimental studies (Albrechtsen, Henrilsen & Færch, 1980; Delistle, 1982; Magna, 1983; Munro & Derwing, 1995; Schairer 1992), which identified that judges’ personal traits and characteristics such as age, gender and native-speakerhood are related to error evaluation. A teaching activity in its general sense, be it an assessment
activity or an error evaluation procedure, is in fact a conscious action resulting from a teacher's own interpretive processes (Woods, 1996). After all, in this study as well as many other teaching contexts, it was an instructor who, based on his or her BAK system, decided whether, which, what and when contextual element(s) should be considered in a specific error evaluation activity. Hence, it is very likely that findings from many experimental error evaluation studies I have cited here and in this thesis study are making the same observation, namely that the BAK system(s) of the judge or judges, in relation to the purpose of the error evaluation activities, exert a crucial effect on the results of the error evaluation procedure.

The key, umbrella interpretation this study provides to the issue of error evaluation is that error evaluation is context sensitive. The idea that evaluation should be tailored to the context is not new in language curriculum evaluation (e.g., Elley, 1989). In most recent L2 writing and error evaluation literature, as I reported in Chapter Two, we can see an increasing emphasis on the context of error evaluation (e.g., Haswell, 1988; Lennon, 1991; Taylor, 1986). The present thesis study exemplifies the observation that contextual elements actually acted on the error evaluation activity and influenced its results.

In Chapter Two of this thesis, I argued for the primary premise of phenomenology, namely that the pursuit of the essence of error as a phenomenon should begin with human perceptions. In this study, I interviewed four ESL writing instructors and elicited their perceptions of error and error gravity in their assessment activities, a context in which error evaluation is naturally carried out. In Chapter Four I reported specific percepts from the instructors to show that use rules overrode grammatical rules. These examples actually validate Lennon's (1991, p. 189) observation that "[M]ost erroneous forms are, in fact, in themselves not erroneous at all, but become erroneous only in the context of the larger linguistic units in which they occur."

Lennon (Ibid.) also proposed the use of two constructs, *domain* and *extent*, to
measure error gravity. *Domain* refers to the rank of the linguistic unit (e.g., morpheme, word, extra-linguistic context) which must be taken as context in order for the error to become apparent; *extent* refers to the rank of the linguistic unit which would have to be deleted, replaced, reordered, or supplied in order to repair production. In this thesis study, that all of the instructors considered content errors as one of the four important error categories and that two of the instructors responded to this error type more often than they did to organizational errors, grammatical errors and vocabulary errors both suggest that Lennon’s *domain* of the errors could explain well error gravity since a content anomaly can be made apparent only when the extra-linguistic context, the largest of a *domain* unit, is considered as well.

In Chapter Two of this thesis, I followed Haswell (1988), Lennon (1991) and Rifkin and Roberts (1995) and argued that previous understanding of language error was too narrow, and in order to better understand error, we need to broaden our conceptualization of error. Three of the instructors in my study did not say much other than offering some nominal definitions of errors. While I appreciate Anna’s reflection that error has different meanings to different people and at different times, the one meaning I wish to add to the understanding of error, in light of findings in this study, is the explanation of error by Activity Theory and the concept of a mediating tool in sociocultural theory. I understand that this meaning of error has not been endorsed by the instructors because they were no longer available for consultations. Nevertheless, I will argue for this perspective all the same.

### 5.5. The activities of assessment and error evaluation

I began this study from a phenomenological standpoint. As I interpret its findings I have come to a sociocultural orientation to explain the results. I see no conflicts in doing so, and I defended the compatibility of the epistemological outlooks of these two
approaches in Chapter Two of this thesis. To justify the use of Activity Theory to explain my findings, I will now begin with the unchallenged assumption that the instructors were using their assessment scheme to do something, namely to provide teacher feedback on a student’s written work, and perhaps assign a score to it whenever required. One outcome of this assessment was the instructors’ identification of errors (especially serious errors), and their comments or marks on these errors, some serious and some not. It seems also an indisputable assumption that the instructors’ assessments were purposeful and the ultimate goal for doing so was teaching, the enhancement of students’ learning, though I note that there may have been other intermediate goals in the process, such as assigning students’ grades. The assessment, of which the error evaluation was a part, was an integral part of a larger activity called teaching. The question I now ask is: What is the meaning of error in all this?

The Activity Theory I use to reveal the meaning of error was first proposed by Leont’ev (1981) and has been further developed by Engeström, (1992, 1996) and used by other writers (Russell 1997; Wells, 1996) in educational practices. According to Wertsch (1985), Activity Theory aims to answer the question of “What is the individual or group doing in a particular setting?” (Wertsch, Ibid. p. 211; cf. Lantolf & Appel, 1994, p. 17). A justification for applying Leont’ev’s (1981) Activity Theory to a setting of language teaching is given by Lantolf and Appel (1994, p. 17):

The highest level of analysis within the theory, the level of activity, is defined as the social institutionally determined setting or context based on a set of assumptions about the appropriate roles, goals, and means to be used by the participants in that setting. Setting, in Leont’ev’s framework, does not mean the physical or perceptual context in which human function; rather, it refers to the sociocultural interpretation or creation that is imposed on the context by the participants (Wertsch, 1985b, pp. 203, 212). Examples of activity settings include play, work, education, worship and leisure time.
Lantolf and Appel's list of examples here could have well included the teaching and assessment of writing. With this understanding of Activity Theory, I therefore see that the purpose of my analysis here is to explain how the instructors participating in my research interpreted the contextual elements in their error evaluation activities, which were actually situated in the larger context of teaching, which itself in turn is another activity.

The simplest activity system in sociocultural theory consists of two elements only, as argued for in Wertsch (1998), who preferred to call this system "the irreducible tension between the agent and his mediational tool" (p. 25). Wertsch (1998) posited that "it makes sense to give the relationship between agent and instrument a privileged position" (p. 24) and argued for "a focus on the agent-instrument dialectic" (p. 24). In his example of the pole vault as a tool for pole vaulting, Wertsch reported the development of athletes' uses of pole vaults, and concluded that "it is futile, if not ridiculous, to try to understand the action of pole vaulting in terms of the mediational means—the pole—or the agent in isolation" (p. 27). What is relevant about Wertsch's observation to the present discussion is that it would not make sense to try to understand the meaning of error without looking at how error was acted upon by the instructors.

An activity system which I use to interpret the assessment and error evaluation activities originated from Engeström (1987, 1993). A full form of this system actually consists of six elements, but for my present analysis, a simpler system consisting of only three main components will do the job just as well. So for simplicity sake, the system I use is composed of: (a) Subject(s), (b) the Mediation tool, and (c) Object/Motive and outcome. The third component (c) in this system actually breaks down into two or three related elements. To complete my interpretation in the present discussion, I also call on two other notions associated with the use of a mediational tool, the affordances and constraints of a tool (Wertsch, 1998, p. 38). Affordances refer to the power and uses a particular mediational tool render to its user, the abilities which the user would not have
possessed without this particular tool (Ibid., p. 38). Constraints, on the other hand, are the limitations or impossibilities a particular tool imposes on its user (Ibid., p. 40). A Vygotskian perspective maintains that what one can see and do has a lot to do with what is used as a tool in seeing and doing. While a particular mediational tool enables its user to look at things in certain ways and do certain things, it also at the same time restricts or stops the same user from looking at things in certain other ways and therefore imposes on the user certain other limitations.

To apply this activity system to the assessment activity and the error evaluation activity I have researched, I present a diagram of two connected triangles in Figure 5.1, depicting these two related activity systems. The triangle on the upper left represents the activity of instructors’ assessments and the one on the lower right depicts instructors’ error evaluation activities.

In the system of assessment, which is the triangle on the upper left, the subjects in the system are an instructor and his or her students. I should note that because of the power relationship in a classroom, the instructor was the primary subject (agent) while the students were the secondary subjects (agents) though both parties participated in the assessment activity. The instrument in this activity was of course the assessment scheme the instructor used. The object of the activity was students’ written work. Attached to the object of this activity system is the motive and the outcome. The motive of this type of classroom assessment had a strong pedagogical mission though in general, there could be multiple motives. In this particular case, the assessment was carried out with the motive to serve both the immediate teaching objectives and the end-of-course evaluation. One important constituent element of this activity is of course the error evaluation activity as an outcome.
Figure 5.1: The activity systems of assessment and error evaluation
My application of the activity system, however, does not stop here. I see a use also in applying this activity system to the error evaluation procedure, which is the triangle on the lower right in Figure 5.1. In other words, the activity system of assessment is connected directly to the error evaluation system because the outcome in the former system is now the tool in the latter, and this level of activity system explains how the notion of error was utilized further by the instructors. In other words, in this system of error evaluation, the instrument an instructor and the students actually used here is the notion of error. While the object of the subjects’ attention is the examination of the adherence and non-adherence to writing and grammar rules, one primary motive (at least as seen from the instructor) was to provide evidential basis for the assessment and the scoring; the outcome was an array of what I would call “potential points of engagement”. The term “potential points of engagement” perhaps requires some further explanation. In reality, these points could appear in many different forms. Sometimes they were points that the instructor covered one by one in a class feedback session; sometimes they appeared as points of discussion in one-to-one instructor-student tutorials; other times, they appeared as mere written marks and comments on paper that students had to interpret and incorporate into their rewrites. But the substance of these points of engagement were in fact the “points” or focuses which the instructors deemed important for students’ rewrites.

5.6. The affordances of the notion of error

While both the assessment scheme and the notion of error were used as mediational means, there are differences in these two means. The first difference is that the use of the assessment scheme may last, so to speak, throughout the whole assessment process although it may only be referred to occasionally, but attention to errors and the evaluation of errors, embedded in the assessment process, would happen only some of the
time since in the assessment process, the instructor would also be doing many other things, such as interpreting meaning and envisioning the situation of writing (Cumming, 1990a).

The second difference is that while the notion of error and the assessment scheme were both psychological tools (Lantolf & Appel, 1994, p. 8), the assessment scheme also appeared as a document. This implies that it had a "materiality" aspect (Wertsch, 1998, p. 30), and as Wertsch observed, the material aspect of this tool, i.e., it being a text, could render a kind of permanency and stability, which the notion of error did not seem to possess. On the other hand, in contrast, the construct of error working at the background of the assessment activity but realized only in its outcomes is a much more fluid and flexible notion than the assessment scheme. It was the use of this fluid, flexible notion of error that we saw prompt two of the instructors, Anna and Kelly, part from the assessment schemes and focus on content anomalies in their attempts to engage their students’ attention to what they wished to profess. In sum, I interpret that it was the notion of error that gave Anna and Kelly the freedom to engage their students in talks about texts in the way they intended, and this is what I refer to as the affordance of the notion of error as a mediational tool.

The affordance of the notion of error is in great contrast with the constraints of the assessment scheme the instructors were using, and it is in the constraints which the assessment scheme tended to impose on its users that we see better the power of the notion of error. In the following, I will discuss the limitation this particular assessment scheme imposed on its users. In addition, I will also propose that in order to implement a more coherent curriculum, the instructors could consider either contextualizing this assessment scheme or supplementing it with an error checklist that is based on an ESL literacy model.
5.7. The constraints of the assessment scheme

The marking scheme that the instructors were using had its value. It was basically an assessment scheme for academic writing, one that many of the students on the course had been previously exposed to and therefore should be quite familiar with. This specific assessment scheme, according to Edward, was adapted from Jacobs, Zinkgraf, Wormuth, Hartiel & Hughey (1981) and other unpublished marking assessments. Like many other analytic academic writing assessment schemes, this scheme adopted a multiple-trait approach, based on a concept of proficiency defined in traits such as content, syntax, organization, vocabulary, and mechanics. However, it is easy to see that the scheme did not explicitly require that a proficient writer, through the appropriate use of shared and unshared information and choice of language and vocabulary, demonstrate a sensitivity to the genre of the required text, and to the complex web of relationships among the genre, the assumed role of the writer, and the assumed role of the reader. For example, under the vocabulary category, the scheme specifies meaning as a criterion to evaluate lexical choice but says nothing about formality, register, or choice of words and expressions appropriate to the assumed reader/writer role, the task or the genre. Some of these were facets, as I showed in Chapter Four, considered important by the instructors.

The assessment scheme does not explicitly state that a writer is required to demonstrate adequate shared knowledge, but the data from the instructors’ recalls of specific errors show that this is a key concept the instructors (at least two of the four instructors) demanded their students to demonstrate in their writing. This suggests that the assessment scheme is not serving well the instructors for the purpose.

Because the notion of “reader’s and writer’s shared or unshared knowledge” is not explicitly stated in the scheme as a key trait, there was confusion in the instructors’ scoring in relation to this requirement. For example, the error of “reader’s and writer’s
shared/unshared knowledge was sometimes called a content error and other times considered a vocabulary error; hence, there were instances in which a student was possibly doubly penalized for such an error. On the other hand, the instructor as a rater was inconsistent in determining the category of an error related to "shared/unshared knowledge".

The other problem with such an assessment scheme is that it missed the instructors' stated main objectives of the course: the teaching of technical genres. One implication from my analyses is that the seriousness of errors could change with the genre under consideration. This implication in turn suggests that an assessment scheme decontextualized from its genre may not serve effectively as a tool because it would miss the interactivity between error types and genre types. For example, one frequent, serious error type that many of Kelly's and Edward's students made was the violation of an organizational rule: One imperative word for each numbered step in a set of technical instructions. This rule however applies only to this particular genre. If one looks into the assessment scheme, one reads the descriptors under organization as follows:

**ORGANIZATION**

28-22 Excellent: very good: ideas clearly stated and supported
well organized, using elements of intended form (e.g. short report, technical instructions, etc.); logical sequencing; cohesive.

21-14 Good-average: although main ideas are clear, it is somewhat choppy.
May be missing (or misplaced) one or two elements required in the form; logical but incomplete sequencing.

13-7 Fair-poor. The ideas are confused or disconnected. It is non-fluent; several elements of the form are missing, unclear or out-of-place; lacks of logical sequencing and development.

6-1 Very poor: Does not communicate the necessary information. No organization; or not enough to evaluate.

This category of the assessment scheme provides only a very vague, implicit remark on the link between the required organization and the name of the document type. The language here is abstract and hard to interpret, and it is doubtful whether the students
reading this assessment scheme can yield relevant interpretations of the writing requirements, even though the scheme might have worked fine for the instructors.

The same is true with, for example, the vocabulary category:

**VOCABULARY**

22-17 Excellent- very good: Sophisticated range/appropriate to the task. Effective word/idiom choice and usage: shows word from mastery.

16-11 Good- average: adequate range. Occasional errors of word/idiom form, choice, usage but meaning is not obscured.

10-6 Fair-good: Meaning is sometimes obscured and confused. Limited range: frequent errors of word/idiom choice, usage.

6-1 Very poor. Essentially a translation. Little knowledge of English vocabulary, idioms word from; or, not enough to evaluate.

Once again I see in this category some general phrases in the descriptors, such as “sophisticated range/appropriate to the task” and “effective word choice and usage”. What I cannot find is the kind of substantive, contextualized language the instructors were using to describe specific serious vocabulary errors in their recalls, such as “not to use low information content word”, “should not have used unnecessary politeness marker here”, “too much copying from prompt”, “do not use abbreviations of technical terms”, or “should consider the readers and not use expressions that they don’t know”. The students on the course were all given a copy of this assessment scheme on commencement of the course in the expectation that the students would use the scheme together with instructor feedback to improve in their rewrites. But I doubt that this good intention could be effectively realized with such incongruence between what the instructor wanted the students to pay attention to and what this assessment scheme provided in its verbal descriptions.

I have in this section subjected this assessment scheme to an examination using Activity Theory. I have shown that the notion of error actually played a role in the assessment by allowing two of the teachers to deviate from the constraints the assessment
scheme tended to impose to their rating. While Marge and Edward utilized the assessment scheme faithfully, Anna, by responding to error types she favoured, initiated her students into a culture of business and a discussion of literacy knowledge. Kelly, on the other hand, also deviated from the assessment scheme and centered his feedback on asserting the importance of an audience-oriented approach in technical writing. In sum, the notion of error was utilized by these instructors as a tool, just as their common assessment scheme was.

5.8. Theoretical relevance of the study

This thesis study makes its contribution to error evaluation research by exemplifying a constructivist approach to error perception. The outlook of this approach is based on a social constructivist view on the nature of language and people’s discursive practices. The advantage of a social constructivist agenda on the learning of technical writing is that it avoids declaring an autonomous, absolute notion of literacy by defining literacy as social practices (Street, 1984, 1999). Just as the four instructors were trying to do here in the teaching of this particular course, acquisition of effective technical communication relies on not only talking about texts to learn the lexis and syntax of technical English and the different technical text types, it is primarily a gradual socialization into socioliteracies, the learning of “communicative competence, knowing when and how to use resources from different channels, that affects abilities to operate in different domains” (Street, 1999, p. 3). Kelly’s and Anna’s attempts to communicate to their students the seriousness and causes of content errors these students had committed exemplify very well this notion of socioliteracies.

Findings from this thesis study have implications for other error evaluation studies which may take place outside classroom and educational settings. For example, findings from this thesis study suggest that error evaluation activities in workplaces and in everyday social activities may also exemplify the same relationships between the purposes.
of a particular social activity (e.g., job application, business transactions) and the purposes of the error evaluation subsumed under the social activity (e.g., errors in a job application letter; the outcome of business transactions due to errors). This thesis study suggests that in these social settings, error gravity may be closely related to the purposes of the activity of which the error evaluation activity is a part, the judges, and the other contextual elements.

This thesis study offers an understanding of error by applying Activity Theory to the notion of error. In the light of this application, one meaning of error for the instructors in this study lies in its being a tool for them, the primary user (a teacher), to engage in talks about texts (Olson, 1997; Wells, 1990) with their students, the secondary users. This use of error by these users may be unique to the context of teaching, but language error used as a tool can also happen in a non-teaching or non-pedagogical setting or activity. For example, people identify ethnicities, social identities or statuses by recognising the phonological errors in speech (Giles, 1973; Giles, Scherer & Taylor, 1979; Giles, Coupland & Coupland, 1991). The meanings of error in these settings are all interesting issues (Munro & Derwing, 1995) that future error evaluation studies could even examine.

Recent discussions in L2 assessment have shown an interest in context-based language testing (Douglas, 1998) and an invitation of an inclusion of a Vygotskian or an interactionist perspective (Chapelle 1998; McNamara, 1997). In such perspectives, performance assessment is seen as a situated, interactive activity which must be understood with a focus on its social dimension since evaluation itself is a social act (McNamara, 1997, p. 458). McNamara (Ibid.) called for a renewed understanding of assessment as a social fact. The present thesis makes its contribution to L2 assessment by revealing a socio-cognitive aspect of error evaluation as an integral part of the assessment activity in the locality of university ESL classrooms. It also notes the relevancy and power of contextualizing an assessment scheme to create desirable backwash effects. The view that error is a context-sensitive notion is consistent with recent development in
interlanguage studies (Selinker, 1995) and indigenous assessments (Douglas, 2000).

To date very little published research seems to have looked into the assessment of technical writing. L2 testing studies looking into the uses and effects of assessment schemes have displayed a biased interest in academic writing (Hamp-Lyons 1991; Horowitz 1991; Vaughan 1990). In English for Specific Purposes (ESP) and English for Science and Technology (EST), research interest has centred on the descriptions of different genres and their language (Bhatia, 1993; G. James, 1994; Swales, 1990; Trimble, 1985) but not the assessment of the learning and writing of technical genres and their language. This study makes its contribution by looking into assessing ESL technical writing in a classroom setting and has initiated a discussion on technical writing as literacy versus language proficiency.

The present study also demonstrates the value of using Activity Theory to analyze an assessment scheme, yielding some insightful observations about assessment practices. The Vygotskian perspective that a (semiotic) tool (e.g., an assessment scheme, a checklist, etc.) has affordances and constraints finds support in a study by Allan (1995), who showed that the instrument he used to elicit Hong Kong's university-level ESL learners' self-reports of reading strategies in fact shaped the responses these learners made. When an item, which had previously been self-reported to be the most popular reading strategy, was deleted from a checklist for a new round of self-report, the same learners tended to choose the second most popular choice item on the list, even though they were given the opportunity to report any strategies not listed.

In L1, there have been studies comparing the impacts of different writing assessment rubrics (Gearhart, Herman, Novak, & Wolf, 1995; cited in Wolf & Davinroy, 1998). In the light of findings in this thesis study, potential explorations in L2 testing could include the examination of assessment schemes as mediational tools to look at how individual assessment schemes can yield different backwash effects, and how they can
structure raters’ evaluation and feedback and affect its validity. L2 assessment research could also look into the use of assessment schemes contextualized to the task to improve validity, instead of persisting only in the pursuit of a universally applied, one-for-all assessment scheme.

There was little doubt in the percepts of the teachers participating in this thesis study that an error signifies a violation of a certain expert speakers’ rule, but there are situations where different rules should be called upon, and these different rules may interact. In this study, I present cases where discourse (genre) rules override morphosyntactic rules.

This thesis study suggests a notion of error much wider than the one defined in earlier error analysis research. Error has many phases and has different meanings for different people. Seen from the eyes of a learner, an error is evidence for how a learner relies on his or her interlanguage rules; it is also something to get rid of for the purpose of improvement. Seen from the perspective of a teacher, an error is a violation of a certain rule that expert speakers use; it is also a tool for a teacher to create potential points of engagement with his or her students. Furthermore, I also argue that error is a tool a teacher utilizes to do his or her own things. Woods (1993, p.1) reported that when he was in a project evaluating different teachers teaching, he observed the apparent uniformity of teachers giving the same lessons, using the same materials and books, and sometimes even the same pages and exercises, but was struck by the fact that there were important differences in what the students in these different classes were actually doing, and therefore a difference in what the teachers were actually doing. I have observed a similar phenomenon here in the use of the same assessment scheme by these instructors; ultimately, it was the percepts and uses of errors that explain the differences in what the teachers were trying to do, even though they were all using the same assessment scheme.
5.9. Practical implications

In view of the constraints that the concerned assessment scheme imposed on its users (i.e., instructors and students), and for the sake of promoting a more coherent curriculum (Johnson, 1989, p. xiii), there is a need, in my view, for the instructors in this context, and perhaps in situations like it, to adopt a feedback system that coheres more tightly to the perceived teaching objectives. I propose two possible ways to improve the feedback system in practice.

Method one would be to contextualize the existing assessment at the level of genre. Doing so would give a strong focus on the teaching of genre. Moreover, in terms of improving grammatical competency and reducing grammatical errors, this method can reveal better the relationship between, for example, the usage of certain morpho-syntactic structures and the genre. The contextualization to genre will also give hints as to which grammatical structures to emphasize in the teaching of different document types (Paltridge, 1996, 1997; Swales 1990).

It has been argued that contextualizing an assessment scheme to a specific task, as in the primary trait scoring method, is often too costly and time-consuming for most evaluation purposes (Hamp-Lyons 1991, p. 246). However, I would argue that although Hamp-Lyons' criticism may apply to some large-scale, high-stake language proficiency assessments for admissions or placement, the contextualization of an assessment scheme to a specific task I suggest here would be less costly and time-consuming. This would be, after all, a small-scale, classroom-based assessment in an ESL genre-based technical writing course. Also, unlike the kind of large-scale, high-stake assessments to which Hamp-Lyons' criticism applies, the purpose of the assessments in the present context are for teaching and course evaluation. Finally, using genre as a primary trait would no doubt give stronger coherence to the implementation of the curriculum.
Method Two would involve developing a generic assessment scheme or a writing checklist that is based on an ESL literacy model, inclusive of but not restricted to the construct of ESL language proficiency the assessment scheme was based upon. The concept of language proficiency and literacy often overlap, but literacy is often inclusive of language proficiency. Cumming (1990b) discussed a number of issues in ESL literacy and suggested that the acquisition of ESL literacy includes (a) literate knowledge and expertise, (b) proficiency in a second language, (c) different cultural practices, knowledge and values, and (d) new ways of interacting with people (p. 42). Put in the present context, this means that these students of technical writing need to develop, in addition to their ESL language proficiency, literate knowledge of and expertise in reading and producing technical texts, an understanding of the practice and values of business and technical personnel, and new ways of interacting with these people. All these seem to be in accordance with the objectives of teaching technical writing as genre (Berkenkotter & Huckin, 1995; Bhatia, 1993; Russell, 1997; Swales 1990), the key notions discussed in many technical writing texts (Blicq & Moretto, 1998; Huckin & Olsen, 1991), and the points of engagement intended by at least two of the instructors, Anna and Kelly.

It may be true that the students in the course needed a lot more help in their language proficiency than they did in literacy for many of the writing tasks they faced in university, but the solution to this dilemma exists outside the course and rests at an administration level. I myself had the experience of teaching a technical writing course very similar to the one in which the data of this thesis study were collected, in terms of context and the kinds of student and teacher participants. Having taught the course myself for a year after the completion of the data collection for this thesis, I have observed that it is not realistic to expect that these students’ L1 literacy can help them learn technical writing. What I have observed is that most of these students did not come to the university with strong L1 literacy. Many of the students I taught on the course in Fall 1998 and Spring 1999, I found out, came to Canada in their teens, so they were actually “nonnative
speakers of English, lacking required literacy in their own culture" (McKay, 1993, p.3). In any case, an expectation of an immediate transfer from their L1 literacy into their L2 writing is not realistic because of the cultural specificity of technical writing. Hence, though it might be an accurate observation that language proficiency is an area that demands direct instruction, it might be a mistake to just emphasize English proficiency and expect these students to pick up other components of literacy in the process.

The development of an adequate assessment scheme based on an ESL literacy model for a technical writing course might be an arduous task: moreover, there is not any ESL literacy model readily available for adoption. An alternative to this is a writing checklist based on a suitable literacy model. Such a checklist could be used during the planning stage of students' writing or as an aid to students' post-writing editing. An adequate blueprint for such a checklist can be found in Johns' (1997, p. 37) conceptualization of L1 socioliteracies within a situated cognition perspective. A diagram of this model is given in Figure 5.2. The model Johns suggested is built around the idea of a text and its contexts. The strength of this model is its emphasis on the knowledge required of the reader or the writer, and how this knowledge interacts with other facets such as the communicative purposes, the document type (name), roles of the reader and the writer, context and features in the textual form, register and content.
Elsewhere, Cumming (1995, p. 387) has suggested the explicit use of what he called "procedural facilitation prompts" to facilitate the teaching and learning of ESL writing. The five prompts he suggested are (1) setting and monitoring goals to accomplish specific writing purposes and to accommodate readers of their writing, (2) selecting appropriate words and expressions, (3) comparing equivalent expressions in L1 and L2, (4) using relevant grammar and spelling rules and (5) assessing the coherence between
parts of the compositions. Cumming further suggested the use of these prompts to organize teaching, to structure the teacher-student discussions and to structure the students’ composing processes. The functions of the checklist I am proposing are similar to the facilitation prompts Cumming suggested, perhaps with the only exception that these prompts should be geared towards technical document types.

In addition to the prompt types Cumming (Ibid.) proposed, I consider that these prompts should at least include the following: (a) document type and communicative purposes, (b) the writer’s and reader’s roles and relationships and (c) shared knowledge and unshared knowledge. Facilitation prompts, checklists or assessment schemes organized in these categories should serve better the perceived teaching objectives of the course. They would highlight the use of knowledge required for effective technical writing and empower the instructors and students alike, enabling them to engage in talk about an array of facets of literacy.

To sum up this section, I suggest that because of the constraints the tool (the assessment scheme used by the instructors in this thesis study) tended to impose on its users, the instructors teaching in this course should consider the adoption of an alternative assessment scheme. I argue that a basis for such an assessment scheme can be founded on the socioliteracies model suggested in Johns (1997). Such a model suits better the assessment of technical writing than one that rests on a general notion of language or ESL proficiency. If the development of a different assessment scheme were not feasible, then the instructors of technical writing should consider either contextualizing the assessment scheme they were using each time to the specific genre under discussion, or adopt a set of facilitation prompts or a checklist based on the socioliteracies model in Johns (1997) so that a more explicit, comprehensive feedback system can be available to the students and the instructors.
5.10. Limitations of the study

The phenomenological approach I adopt for this thesis study has provided me with a set of rich and insightful data for my analyses. I have exercised care in my analyses and generated some useful observations. However, I note the constraints I have faced in the process of my data collection and my data analyses of this study.

How much people understand about themselves and their perceptions of a phenomenon very much depends on the time and opportunity they have had to observe themselves and reflect upon these observations. As Willis (1991) observed, perception is an excellent starting point for human knowledge and understandings; however, in the pursuit of the real essence and meaning of a phenomenon, the reliability of humans' percepts of a phenomenon — the ability to tell incoherent illusions from reliable percepts — as many phenomenologists (Shapiro, 1986; Wertz, 1986; Van Manen, 1990) have suggested, depends a lot on prolonged engagement with the phenomenon. Time was one of the constraints that I faced in this study. Three of the participants in the study could only avail themselves for no more than three 1-hour interviews, a verification of the transcription of their own interviews, and their responses to the description of their own perceptions and some follow-up questions. I have reported some of the discrepancies between what individual instructors told me about themselves and how they actually behaved. The time that the participants had been able to devote to an examination of their own percepts was limited, resulting in difficulties in my asking them to give these sometimes crude and incongruent percepts, elicited by different methods, a coherent interpretation. This is one limitation this study was not able to break through.

If time is not too great a constraint and arrangements can be made, similar phenomenological studies in the future should consider the use of focus groups as an alternative to gain individual and collective percepts of a phenomenon. By this I mean
involving all the participants in a group discussion focussing on their percepts of the same phenomenon or phenomena. This technique has the advantage of letting the participants expose their individual percepts to others for sensitizing each other’s perceptions to improve participants’ internal reliability. For example, when I asked the question “What is error?” three of the instructors restricted their responses to a nominal definition. Anna took a different path and reported her understanding of its many meanings. In a focus group discussion, Anna’s report could have initiated or extended other participants’ percepts. In addition, focus group discussions can also help form more coherent percepts from both the group (collective percepts) and its individuals (individual percepts). The one-to-one, interviewer-interviewee data collection procedure I used in this thesis study certainly did not enjoy these advantages.

There is a methodological issue related to the coding of specific errors the instructors commented upon in the stimulated recalls. I coded these specific errors and performed member checks only in cases when I felt that there were ambiguities in the instructors’ comments. Alternatively, to be consistent with a phenomenological approach, I might have had individual instructors go through the transcription of their own recalls and code the errors themselves.

Another limitation of this thesis study is in the descriptive statistical analyses I performed on instructors’ recalls of specific errors. Although I believe that I use the statistics properly and I draw valid inferences from these statistics, I recognize that the total number of specific errors one of the instructors, Anna, recalled was much smaller than the total numbers of errors the other teachers recalled. I reduced this problem by computing the relative frequencies of the response rate of each of the instructors, but I understand that this part of the statistics and the inferences drawn have to be treated with caution.
Other possible contexts might have been worthy of further investigation. Freire (1997) studied the process of producing and using electronic messages in a workplace and depicted a vivid scene of the process of this type of technical writing. It will be very interesting and useful if future error evaluation researchers can investigate professional engineers playing the role of judges or assessors of language errors in a workplace, since the perspectives professional engineers take regarding language errors might be very different from that of language instructors. Another perspective that researchers could take is to look into the process of error evaluation constructed orally between learners and the instructor to see how they build up their conceptualization of error in those contexts.

Finally, this exploratory study was carried out as a case study in one context. It examined error evaluation in the context of an ESL technical writing course at a Canadian university. It has generated some insightful, fundamental findings. However, findings from this study are highly interpretive and localized and there is no way of knowing whether or how applicable these interpretations would be in another ESL technical writing course in a different North American university, or in other naturalistic contexts, such as job applications, business transactions, etc. I have adopted an interactionist perspective, assuming that phenomena and percepts are meaningful only in the contexts they are situated. However, I am also attracted to the belief that we should seek generalizations from one context to help understand phenomena in other similar contexts. In order to do that, however, further research is needed to extend and verify the interpretations arising from this study in other contexts so that the validity and applicability of the interpretations I have arrived at in this exploratory study can be put to the test.
REFERENCES


- 210 -


Appendix A

Date:

Letter to a potential teacher participant for the study

Dear:

I am writing to ask if you would be willing to participate in my Ph. D. dissertation research on the responses of ESL teachers to language accuracy. I am conducting this project under the supervision of my supervisor, Dr. Alister Cumming, at the Ontario Institute for Studies in Education / University of Toronto.

The purpose of this study is to enhance our understanding of ESL teaching and learning. In particular, I wish to investigate teachers’ responses to the accuracy of ESL learner language. Your participation in my project would involve:

1. My being present in one of your ESL classes with an audio or video recorder to record classroom activities for a total of 8 sessions.
2. My access to some of the written work of your students, written essays which you have already marked.
3. My interviewing you and asking you to reflect on your classroom activities and assessment styles. In these interviews, each of which should last no more than an hour, you will be presented with students’ work or classroom episodes to help you recollect from memory.
4. Verifying my analyses in written form.

You are not under any obligation to participate. If you do choose to participate, your identity will be entirely confidential. Your name, the names of your students and the name and location of the teaching institute will all be identified only by pseudonyms. This restriction applies to any use of the data in all presentations, written documents and the research thesis. You may withdraw from the research at any time, for any reason, or you may refuse to participate in any component of the research, if you so wish. Although this study does not aim to evaluate teachers’ professional performance and students’ learning performances, please be alerted that the final analyses or thesis may contain some evaluative interpretations. However, all such analyses will be presented to you for feedback and endorsements.

I trust that you will find your participation in this project helpful. If you would like further information about the study before deciding, I would be happy to provide this. Please contact me at the number below. If you wish to verify the information above, you may contact Dr. Alister Cumming, my supervisor at OISE/UT (923-6641 Ext. 2538). If you are willing to participate, could you kindly complete the consent form below and return it to me. Please keep a copy of this letter and the schedule for your records. Thank you very much for attending to this request.

Yours sincerely,

Anthony K. K. Tong
Doctoral student
Modern Language Centre
OISE/UT
email: akktong@oise.utoronto.ca
Data Collection Schedule

The estimated time you may be involved in this project, including your usual teaching and assessment, is summarized as follows:

<table>
<thead>
<tr>
<th>Data Collection Procedure</th>
<th>Explanation</th>
<th>Total Time (Estimated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom observations</td>
<td>4-8 sessions</td>
<td>8 hours</td>
</tr>
<tr>
<td>Teachers’ interviews</td>
<td>3 interviews, each lasts no more than an hour.</td>
<td>3 hours</td>
</tr>
<tr>
<td>Validating descriptions</td>
<td>Reading and validating descriptions and reflections of teaching and assessment activities</td>
<td>3 hours</td>
</tr>
<tr>
<td>Follow-up feedback</td>
<td>Reading and giving feedback to description</td>
<td>1 hours</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>15 hours</strong></td>
</tr>
</tbody>
</table>

---

REPLY SLIP

Dear Anthony Tong:

I have read your letter of , describing your project on the teachers’ responses to ESL accuracy, and kept a copy of it for future reference.

I would / would not (circle one) like to participate.

Name: _____________________________________________________________

Telephone: ________________________________________________________

Date: _____________________________________________________________________________

Signature: ________________________________________________________________________
Appendix B

Letter to a potential administrator/co-ordinator of an ESL program

Dear [Name],

In your capacity as the administrator/co-ordinator of the ESL writing program at [Institution], I am writing to ask if you would give consent to my collecting data in [Course] of the ESL classes, units for which you are responsible. These data will be collected for my Ph.D. thesis project on the responses of ESL teachers to language accuracy. I am conducting this project under the supervision of my supervisor, Dr. Alister Cumming, at the Ontario Institute for Studies in Education / University of Toronto.

The purpose of this study is to enhance our understanding of ESL teaching and learning. The various activities I plan to carry out in the classes of this particular program are listed below:

1. With prior written consent from the prospective teachers and prior written consent from their students, I will attend their classes to take written notes of classroom activities.
2. I will interview the prospective teachers. I will have obtained their informed written consent prior to these interviews.
3. I may interview some of the students in the classes. I will have obtained informed prior written consent for these interviews.

The teachers and students are under no obligation to participate. A teacher or student may, for any reason and/or at any time, choose to withdraw from the study. Similarly, I will terminate my data collection process at any time when you, in your capacity as the program co-ordinator / the administrator, so request.

In cases when I am granted the consent, the identities of all the teachers and students in the data will be kept entirely confidential. This means that the names of the students and the teachers, the name and the location of this program and the institution in the data will be identified only by pseudonyms in all presentations and the research thesis. The same restriction will apply to any use of these data in all future publications and written documents.

Although this study does not aim to evaluate teachers' professional performance and students' learning performances, the teachers and students are reminded that my analyses may contain some evaluative interpretations. However, all such analyses will be presented to the prospective teachers for feedback and endorsements.

I would like to ensure you that during the data collection process, there will not be any interruptions to any teaching and learning activities; every effort will be made to minimize possible inconvenience to the classes. I trust that at the end of their participation, the teachers will find their experiences helpful.

If you would like further information about the study, please contact me at the number below. If you wish to give consent to my collecting data in this ESL program, please complete the consent form below and return it to me. Please keep a copy of this letter for your records. Thank you very much for attending to this request.

Yours sincerely,

Anthony K. K. Tong
Modern Language Centre, OISE/UT
Schedule

Below gives a breakdown of the estimated time a teacher is involved in the project:

<table>
<thead>
<tr>
<th>Data Collection Procedure</th>
<th>Explanation</th>
<th>Total Time (Estimated)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classroom observations</td>
<td>4 - 8 sessions</td>
<td>8 hours</td>
</tr>
<tr>
<td>Teachers' interviews</td>
<td>3 interviews, each lasts no more than an hour.</td>
<td>3 hours</td>
</tr>
<tr>
<td>Validating descriptions</td>
<td>Reading and validating descriptions and reflections of teaching and assessment activities.</td>
<td>3 hours</td>
</tr>
<tr>
<td>Follow-up feedback</td>
<td>Reading and giving feedback to description</td>
<td>1 hour</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>15 hours</td>
</tr>
</tbody>
</table>

---

REPLY SLIP

Dear Anthony Tong:

I have read your letter of _________ describing your project on the teachers' responses to ESL accuracy, and kept a copy of it for future reference. The tick (✓) in the appropriate box below indicates my response to your request.

☐ I give consent to your collecting the data in the classes listed below provided that

(1) the prospective teacher has agreed to participate in the study
(2) students in the prospective teacher's class have agreed to participate in the study
(3) normal class activities will not be interrupted or disturbed by your presence and your recording and
6. you will terminate your data collection procedure immediately upon my or a teacher's request.

List of participating classes and their teachers (Note: This part is to be completed by the researcher):

<table>
<thead>
<tr>
<th>Class</th>
<th>Teacher</th>
<th>Class</th>
<th>Teacher</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

☐ I do not give consent to your collecting the data in the ESL program in my administration.

Name: __________________________________________

Title: __________________________________________

Telephone: _______________________________________

Date: ___________________________________________

Signature: _______________________________________

-222-
Appendix C

Date:

Letter to a potential ESL Learner participating in the study

Dear student:

I am writing to ask you if you would be willing to participate in my Ph. D. dissertation project on language teaching styles. I am conducting this project under the supervision of my supervisor, Dr Alister Cumming, at the Ontario Institute for Studies in Education / University of Toronto.

The purpose of this study is to enhance our understanding of ESL teaching and learning. Your participation in this research involves my presence in the ESL class you are attending to observe your participation in classroom activities. You do not need to do anything additional other than what you usually do in your class. Your participation also includes letting me analyze some of your written assignments which the teachers have already marked.

Please note that you have no obligation to participate in this study. Whether you are willing to participate in this project or not will have no effect on your course grade or any one of the assessments of your performance on the course. In the case when you agree to my request, your identity in the data will be kept entirely confidential. Your name, the name of your teacher and the institute at which the class is being conducted will all be kept confidential; they will be identified by pseudonyms in all the written documents, publications and the research thesis. It is not the aim of this study to evaluate students’ learning performance though the analyses may contain some evaluative interpretation.

If you are willing to participate, please kindly complete the consent form below and return it to me. Please keep a copy of this letter for your records. If you decide that you do not want to participate, I will not record any of your participation, speech and performance. Thank you very much for attending to this request.

Yours sincerely,

Anthony K. K. Tong
Doctoral student
Modern Language Centre
OISE/UT

-223-
Dear Anthony Tong:

I have read your letter of ____________, describing your project on the teachers' responses to ESL accuracy, and kept a copy of it for future reference.

Pick tick √ one of the following boxes.

☐ I would like to participate.

☐ I would not like to participate. Please delete my participation in your transcription. Also do not study my marked essays.

Name (please print): ___________________________________________________________

Date: _______________________________________________________________________

Signature: ____________________________________________________________________

-224-
Appendix D

Summary of students’ written texts collected for this study*

<table>
<thead>
<tr>
<th>File No.</th>
<th>Teacher</th>
<th>Type of task</th>
<th>No. of scripts</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-4</td>
<td>Kelly</td>
<td>Evaluation report</td>
<td>16</td>
</tr>
<tr>
<td>K-M</td>
<td>Kelly</td>
<td>Investigation report</td>
<td>14</td>
</tr>
<tr>
<td>K-1</td>
<td>Kelly</td>
<td>Instruction manual</td>
<td>13</td>
</tr>
<tr>
<td>M-0</td>
<td>Marge</td>
<td>Argumentative</td>
<td>10</td>
</tr>
<tr>
<td>M-1</td>
<td>Marge</td>
<td>Incident report</td>
<td>14</td>
</tr>
<tr>
<td>M-11</td>
<td>Marge</td>
<td>Instruction Manual</td>
<td>11</td>
</tr>
<tr>
<td>M-M</td>
<td>Marge</td>
<td>Investigation report</td>
<td>14</td>
</tr>
<tr>
<td>M-4</td>
<td>Marge</td>
<td>Evaluation report</td>
<td>9</td>
</tr>
<tr>
<td>A-1</td>
<td>Anna</td>
<td>Incident report</td>
<td>33</td>
</tr>
<tr>
<td>A-01</td>
<td>Anna</td>
<td>User’s manual</td>
<td>15</td>
</tr>
<tr>
<td>A-001</td>
<td>Anna</td>
<td>Technical instructions</td>
<td>20</td>
</tr>
<tr>
<td>A-M</td>
<td>Anna</td>
<td>Progress report</td>
<td>9</td>
</tr>
<tr>
<td>A-4</td>
<td>Anna</td>
<td>Investigation report</td>
<td>23</td>
</tr>
<tr>
<td>E-0</td>
<td>Edward</td>
<td>Incident report</td>
<td>36</td>
</tr>
<tr>
<td>E-1</td>
<td>Edward</td>
<td>Incident report</td>
<td>34</td>
</tr>
<tr>
<td>E-001</td>
<td>Edward</td>
<td>Inspection report</td>
<td>12</td>
</tr>
<tr>
<td>E-11-001</td>
<td>Edward</td>
<td>User’s manual</td>
<td>34</td>
</tr>
<tr>
<td>E-002-1</td>
<td>Edward</td>
<td>Inspection report</td>
<td>15</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td></td>
<td><strong>331</strong></td>
</tr>
</tbody>
</table>

* This summary gives a variety of report types that were taught in the Technical Writing course in which the data for this thesis study were collected. See explanation for students’ written report types in Section 3.4.2. File Nos. are there for my own identification purposes.
Appendix E

ORGANIZATION (O)  The assessment scheme the four instructors were using

28 - 22 Excellent - very good: ideas clearly stated & supported.
   Well organized, using elements of intended form (e.g., short report, technical instructions, etc.); logical
   sequencing; cohesive.

21 - 14 Good - average: although main ideas are clear, it is somewhat choppy.
   May be missing (or misplaced) one or two elements required in the form; logical but incomplete
   sequencing.

13 - 7 Fair - poor: The ideas are confused or disconnected. It is non-fluent; several elements of the form are
   missing, unclear or out-of-place; lacks logical sequencing and development.

6 - 1 Very poor: Does not communicate the necessary information. No organization; or, not enough to
   evaluate.

CONTENT (C)

22 - 17 Excellent - very good: Substantive. Thorough and relevant completion of the assigned writing task;
   no 'information' gaps.

16 - 11 Good - average: Adequate coverage of information. Addresses main points/issues but may not
   adequately develop support; mostly relevant to assigned task, although general.

10 - 6 Fair - poor: Little substance. Missing important information/inadequate development; irrelevant to the
   assigned task, although relevant to the writer's purpose.

5 - 1 Very poor: Irrelevant to both assigned task and writer's purpose. Or, not enough to evaluate.

VOCABULARY (V)

22 - 17 Excellent - very good: Sophisticated range/ appropriate to the task. Effective word/idiom choice and
   usage; shows word form mastery.

16 - 11 Good - average: Adequate range. Occasional errors of word/idiom form, choice, usage but meaning
   is not obscured.

10 - 6 Fair - poor: Meaning is sometimes obscured or confused. Limited range; frequent errors of word/idiom
   choice, usage.

6 - 1 Very poor: Essentially a translation. Little knowledge of English vocabulary, idioms word form; or, nor
   enough to evaluate.

GRAMMAR (G):

28 - 22 Excellent - very good: Effective complex constructions appropriate to the task. Few errors of
   agreement, tense, number, word order, articles, pronouns and prepositions.

21 - 14 Good - average: Accurate constructions, but simple and/or inappropriate grammar. Minor problems in
   complex constructions, but the meaning is not obscured by the errors.

13 - 7 Fair - poor: Major problems in simple/complex constructions. Frequent errors; meaning obscured or
   confused.

6 - 1 Very poor: Virtually no mastery of sentence construction rules. Dominated by errors; does not
   communicate; or, not enough to evaluate.

- 226 -
Appendix F

A sample of a prompt for an assignment task

You are completing your Professional Experience Year at AV Electronics in Kingston, Ontario. One day your supervisor, Jane Finch, asks you to come to her office.

"We did it," she says, "this AV 307 Colour Kit could make us millions of dollars."

"What is it?" you ask.

"You know all those XK-307 black and white photocopiers in all those copy shops around town? Well this," she holds up a box with some things in it, "converts the black and white copier to a colour copier."

"Wow," you say.

"The beauty of the kit is how cheap it is. Most shopkeepers will make a 50% mark up even if they just charge 10 cents a coloured copy."

"Right," you say.

"Anyway, I know you took that writing course at U of... where they taught you how to write technical instructions. So I want you to write a set of instructions to shopkeepers so they can convert their old copiers to colour. OK?"

"Sure!" you say. "How does the kit work?"

"OK, come here." She leads you to a XK-307 photocopier in the corner of her office. "First you gotta turn the machine around. See those four screws on each corner of the back plate? You have to unscrew them with a Phillips Number 4 screwdriver. Like so. Oh, I forgot, before that you have to pull the plug out of the wall, if you touch a live wire you might wake up in the intensive care ward of the hospital. So, we now unscrew the back, and do you see that light bulb on the left side? It is held in place by this plastic basket. You have to take this special wrench we developed, it comes with the kit, and open the bracket, like this, by turning it counterclockwise. Then you just have to unscrew the bulb and replace it with our new Model 307 colour bulb. And then, using the wrench again, tighten up the bracket to hold the bulb in place. OK. Now on the lower left side, over here, is the dry ink container. You just unsnap the old one, like so, to remove it. Then you take the cap off the top of the new three coloured ink container that we include in the kit, and snap it into place. Oh, you have to be careful not to tip the container over because the dry ink could spill out and screw up the photocopier. Then you just screw on the back cover, and presto, you now have a colour photocopier. In our trials, we never had any problems with the installation; however, I suggest the shopkeeper find some coloured picture, you know, from a magazine or something, and test the copier. OK. I need technical instructions written up in about 45 minutes because I have to demonstrate the machine to some shopkeepers who are coming here at 10:00. OK?"

"OK boss," you say. You will now write the technical instructions.
A sample page of a student's written work

3. Full instructions relating to boiler cleaning need to be issued by the head office.

4. Flow meters' manufacturers needs to be called in to do a complete check and recalibrate meters.

5. The plant can implement new techniques for reducing operational costs.

Analysis

I consider the factors of cost and feasibility when examining the ideas, numbers 2 and 6.

1. For maintenance of pumps and vacuum equipment, engineers prefer to have the straight replacement method. However, it is not as costly than setting up a preventive plan.

2. Harris and Markham seem to prefer old techniques that are known to work but are slow. And they are not very interested in new techniques. Mr. Markham is a good company man and is ready and willing to be a chief engineer.

Conclusion

A preventive plan is better to be implemented since it is cheaper. Harris and Markham are both not suitable to carry out some changes in the techniques of the plant through Markham as he is experienced.

Recommendation

I recommend that a maintenance preventive plan is needed for pumps and vacuum equipment. Since Markham is experienced, I recommend him to be next chief engineer. However, I also recommend the company to employ a new engineer who is responsible for implementation of new techniques to the plant. (Who will be told?)

22/17/30 = 76 You have covered the main points well
Appendix H

The transcription convention used in this study

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>&quot; &quot;</td>
<td>This indicates the segment inscribed was read directly from a text.</td>
</tr>
<tr>
<td>[ ]</td>
<td>A pair of square brackets encloses either an explanation providing additional information (e.g., they [the students]) or an address in the data bank for ease of reference (e.g., [M-M-006], [Edward, Para. 169]).</td>
</tr>
<tr>
<td>[Text: ]</td>
<td>This indicates that the words following the colon can be seen in the text.</td>
</tr>
<tr>
<td>[Text by teacher:]</td>
<td>&quot;Text by teacher:&quot; enclosed in square brackets indicates that the words following the colon is the exact wording written by the teacher on the text.</td>
</tr>
<tr>
<td>[Pause]</td>
<td>Square brackets enclosing the word &quot;Pause&quot; indicate a pause of 5 seconds or more.</td>
</tr>
<tr>
<td>( )</td>
<td>Round brackets enclose an uncertain transcription</td>
</tr>
<tr>
<td>x</td>
<td>An “x” indicates incomprehensible item about one word long</td>
</tr>
<tr>
<td>xx</td>
<td>Incomprehensible item of phrase length</td>
</tr>
<tr>
<td>xxx</td>
<td>Incomprehensible item beyond phrase length</td>
</tr>
<tr>
<td>-</td>
<td>A hyphen at the end of a word indicates it is a false start or an incomplete word</td>
</tr>
<tr>
<td>=</td>
<td>An equal sign indicates an interrupted sentence, usually by another speaker</td>
</tr>
<tr>
<td>oh</td>
<td>This represents sounds such as oh, ah, oohhhhh, etc.</td>
</tr>
<tr>
<td>uh</td>
<td>This represents sounds such as uh, huh, hum, uhuh, etc.</td>
</tr>
<tr>
<td>mmm</td>
<td>This represents sounds such as mmm, mmmm, mmmmm, etc.</td>
</tr>
<tr>
<td>yeah</td>
<td>This represents any colloquial variations of &quot;yes&quot;, sounds such as ya, ye, yep, yeah, etc.</td>
</tr>
<tr>
<td>woo</td>
<td>This represents sounds such as woo, wow, woooo, etc.</td>
</tr>
</tbody>
</table>
## Appendix I

A teacher's feedback to the transcription of her interviews

### Feedback Chart

<table>
<thead>
<tr>
<th>Interview No.</th>
<th>Page &amp; Line Nos.</th>
<th>Original Text</th>
<th>Suggested change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>#3 line 35</td>
<td>&quot;passions&quot; — does not make sense</td>
<td>I probably said 'persons'</td>
</tr>
<tr>
<td>1</td>
<td>#9 line 23</td>
<td>&quot;toying my angrain&quot;</td>
<td>I probably said 'to an ingrained'</td>
</tr>
<tr>
<td>1</td>
<td>#10 line 11</td>
<td>'there is subsidence exits'</td>
<td>there is subsidence exists (see my comments)</td>
</tr>
<tr>
<td>1</td>
<td>#10 line 24</td>
<td>found at this [mock]</td>
<td>does not make sense</td>
</tr>
<tr>
<td>1</td>
<td>#12 line 1</td>
<td>you can have me</td>
<td>you can help me?</td>
</tr>
<tr>
<td>2</td>
<td>#5 line 30</td>
<td>it can be reversed or remedial</td>
<td>it can be pursued as remedial</td>
</tr>
<tr>
<td>2</td>
<td>#7 line 67</td>
<td>doesn't make sense</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>#16 line 35</td>
<td>I am a hawk?</td>
<td>doesn't make sense</td>
</tr>
</tbody>
</table>

-230-
Appendix J

Sample pages of my interpretation of a teacher’s percepts and my follow-up questions for the teacher to respond to

Error perception and the Context

1. Marge’s interviews differed from all the other teachers’ in that her comments on her students’ errors were organized around her knowledge of the individual student. This means that throughout her discussion of learner errors, her comments displayed a good understanding of the needs and the strengths of individual students although there was no strong evidence in the data to indicate that her perception of error or error gravity was influenced by the specific needs or characteristics of the individual concerned:

Ma: ... This is a good student. I want to help her to improve in very precise ways. So uh, by reaction to the first so called error, it’s not really an error, it’s not a grammar mistake. It’s the fact that she is using too much detail in a situation, where conciseness is needed. And she knows the conciseness is needed. So extra details were marked as being uh excessive as opposite to incorrect. Okay, I also noted a minor grammar error. This student makes very few grammar errors, in fact, a tense error uh which didn’t change the meaning, but it’s carelessness on both parts really, and so I pointed it out to her. [First interview, Page 1]

Ma: Mmm hum. Yes, she is. Yes, she is getting mixed up between- which suggests that she is at the lower level of development, say, than John even though she is a good hard-working. And she is obviously a clear logical thinker. She can organize ideas better than him and get a higher mark. So she does reasonably well on Organization. I gave her about organization this time, 22 [out of 28]. You see, she got a 22. Very few of them got a 22, but only Sanders got a 23. [Second interview, Page 13]

The other teacher Anna, when commenting on errors and grave errors also showed some awareness of the needs of individuals, but that tendency was not as apparent as Marge’s was...

2. Marge’s data supports the premise of this thesis project; it reveals a strong link between error perception and learner as a context of this perception. As indicated in No. 1, since Marge has a good understanding of learner behaviour and learning strategies, she was aware of the tradeoff and tension between risk-taking in language learning and errors. Because of this, she was sometimes ready to look over even frequent errors in a student’s work and see them (errors) as a means for improvement.

Ma: My initial reaction to this student, is no- nothing to do with errors. This is that he had major real effort to use his own language, so consequently, he has made more errors than the others. He doesn’t- uh uh I’m very careful not to penalize too much in a situation like this, when they are really...

[A total of six pages of description of Marge’s percepts of errors from this point on.]
Follow-up questions for Marge’s further responses:

Marge’s transcription provided some valuable and interesting data for this thesis project, and I certainly enjoyed very much talking with her in the interviews I was kindly given. I now have some further questions, which I hope Marge can respond to as a continuation of our previous discussions so that I can have more of her view for further analyses and deliberation:

(1) What is error? How should it be defined? How is an error different from a mistake and a slip? Is the distinction between them useful?

(2) Do you see a contradiction between your highest grave error rate in the accuracy category and at the same time your recognizing that the course objectives are primarily content and organization? What are some of the possible explanations, in your view, of this incongruence?

(3) Are syntactical rules and pragmatic/discourse rules different? In what ways are they different so that a violation of the former is called an error but a violation of the latter is not?
Appendix K

A sample page of a teacher’s responses to my interpretation of her percepts of errors and my follow-up questions

Marge’s verification responses to my follow-up questions in an email message were:

1. I use ‘error’ and ‘mistake’ generally and interchangeably. I differentiate between a) developmental errors where students take risks and try out new forms they are practising, not always correctly and consistently, and b) errors that are consistently repeated, even though rules have been taught and forms practised in exercises. I use ‘slip’ to denote a careless mistake where I already have evidence of correct use.

I use error/mistake to refer to non-native use of syntax and lexis, not content and organization. As far as CONTENT is concerned, I focus on what is relevant in that particular context. For example, I expect conciseness in a technical report so that excess detail would be inappropriate in that particular context; however, it might be very appropriate in a murder inquiry. Consequently, I avoid the use of ‘error’ which indicates to the student that it would always be wrong. I would point out how focus or emphasis was lost.

Similarly, if the student omitted an essential piece of information, I would indicate that the report was incomplete, rather than deem it an error. I think it is very important for them to understand that language operates within specific contexts, and what is ‘correct’ in one instance may not be in another.

This also describes my attitude to problems of organization. The ‘logical order’ expected in a technical report is culturally determined by a North American engineering profession. If you like, it is what is currently fashionable. A student has to learn to follow those rules, but it does not mean that the way he/she has presented the information is necessarily illogical. If ideas are not clearly connected, I might consider ‘errors’ here, but they tend to be syntactic ones.

2. This is a very important question as it relates to conclusions you have already drawn. Be very careful to consider all the feedback we provided in your analysis. For example, the feedback provided to students in the marked reports would always concentrate on language errors as individual students made different errors, and as I explain in the transcript, I deliberately concentrate on those errors at their request.

I dealt with content and organization in a different way. Because the report format was prescribed by the textbook, I used OH models to demonstrate effective and non-effective reports. I would then relate it to their marks in terms of what was a superior as opposed to an adequate attempt. I would assume the other instructors also focussed on content and organization in classroom feedback and grammar mistakes in tutorial feedback.

3. I am not sure how to answer this question because I do not know how you define rules of discourse. For example, ... I hope this helps.
Appendix L

Project summary and a list of themes used in the NVivo 1.0 program*

NVivo revision 1.0.118

Licensee: Anthony K. K. Tong

Project Error01User Akktong Date 9/28/99 - 84033 PM

PROJECT REPORT

Project Folder C\QSR Projects\Error01
Project Commenced 6/24/99 - 41152 PM
Modified 9/26/99 - 83424 PM
Description My Research Project Description

PROJECT TEAM
Administrator Akktong
Team Members1
Login access required? No
Team Member 1 Akktong
Affiliation Administrator
Date Joined 6/24/99 - 41224 PM

PROJECT STATISTICS
Documents 4
Memos 0
Non-memos 4
Nodes 163
Extracts 0
Non-extracts 163
Free Nodes 2
Tree Nodes 161
Case Type Nodes 0
Case Nodes 0
Sets 4
Document Sets 3
Node Sets 1
Attributes 0
Document Attributes 0
Node Attributes 0
Models 0

- 234 -
(D) //Document Annotations
(F) //Free Nodes
(F 20) //Free Nodes/Accountability in marking
(F 14) //Free Nodes/Activity purpose and error perception
(F 14 3) //Free Nodes/Activity purpose and error perception/Marking scheme
(F 14 2) //Free Nodes/Activity purpose and error perception/Marking strategy
(F 14 4) //Free Nodes/Activity purpose and error perception/Pedagogic advice and error
(F 14 5) //Free Nodes/Activity purpose and error perception/Pedagogic objective and error
(F 14 1) //Free Nodes/Activity purpose and error perception/Type of writing activity
(F 14 1 1) //Free Nodes/Activity purpose and error perception/Type of writing
(F 14 1 2) //Free Nodes/Activity purpose and error perception/Type of writing activity/Practice
(F 18) //Free Nodes/Comprehensibility
(F 26) //Free Nodes/Course objectives and nature
(F 24) //Free Nodes/Cultural factors
(F 24 1) //Free Nodes/Cultural factors/Business culture
(F 15) //Free Nodes/Definition of error
(F 17) //Free Nodes/Errors and mistakes
(F 16) //Free Nodes/Errors as realization of learning strategies
(F 2) //Free Nodes/Errors count
(F 6) //Free Nodes/Errors gravity
(F 6 3) //Free Nodes/Errors gravity/Error frequency
(F 6 2) //Free Nodes/Errors gravity/Minor error
(F 6 4) //Free Nodes/Errors gravity/Recurring error
(F 6 1) //Free Nodes/Errors gravity/Serious error
(F 7) //Free Nodes/Errors perception
(F 7 2) //Free Nodes/Errors perception/Comparing seriousness of error types
(F 7 1) //Free Nodes/Errors perception/Irritating or disturbing error
(F 7 3) //Free Nodes/Errors perception/Task type on error production
(F 4) //Free Nodes/Genre as context
(F 9) //Free Nodes/Inferring causes of error
(F 9 4) //Free Nodes/Inferring cause of error/Background knowledge
(F 9 1) //Free Nodes/Inferring cause of error/Carelessness in proof-reading
(F 9 1 1) //Free Nodes/Inferring cause of error/Carelessness in proof-reading/Editing skills
(F 9 7) //Free Nodes/Inferring cause of error/Did not present coherent congruent facts in different
sections
(F 9 10) //Free Nodes/Inferring cause of error/Information processing load
(F 9 9) //Free Nodes/Inferring cause of error/Lack of planning in writing
(F 9 8) //Free Nodes/Inferring cause of error/Lack of planning in writing
(F 9 5) //Free Nodes/Inferring cause of error/Language ability affecting other categories
(F 9 6) //Free Nodes/Inferring cause of error/Misunderstanding the prompt
(F 9 3) //Free Nodes/Inferring cause of error/Prompt induced error
(F 9 2) //Free Nodes/Inferring cause of error/Writer-Readership
(F 9 2 2) //Free Nodes/Inferring cause of error/Writer-Readership/Not considering reader's
need
(F 9 2 1) //Free Nodes/Inferring cause of error/Writer-Readership/Not knowing the reader
(F 13) //Free Nodes/Interactional norm
(F 11) //Free Nodes/Interview cue question
(F 28) //Free Nodes/Language rule change
(F 3) //Free Nodes/Learner as context
(F 3 3) //Free Nodes/Learner as context/Individualizing learning
(F 3 2) //Free Nodes/Learner as context/Learners as norm

- 235 -
(F 31) //Free Nodes/Learner as context/Recognizing the learner
(F 23) //Free Nodes/Learner dealing with error
(F 27) //Free Nodes/Learners' error-related expectations
(F 19) //Free Nodes/Marking strategies
(F 12) //Free Nodes/Native speaker
(F 8) //Free Nodes/On-line error perception
(F 8 1) //Free Nodes/On-line error perception/Change of perception during interview
(F 8 3) //Free Nodes/On-line error perception/Rater laughed because of error
(F 8 2) //Free Nodes/On-line error perception/Rater reacting emotionally to error
(F 1) //Free Nodes/Script count
(F 10) //Free Nodes/Social stigma
(F 10 2) //Free Nodes/Social stigma/Otherness in error gravity perception
(F 10 1) //Free Nodes/Social stigma/Substandard
(F 22) //Free Nodes/Spoken and written error
(F 25) //Free Nodes/Teaching and learning beliefs
(F 21) //Free Nodes/The role of feedback
(F 5) //Free Nodes/Variability of error type in a piece

* See Section 3.4.7 for explanations of the symbols of this list.
The error record sheet Anna's students all used

Revising your Writing
Grammatical Corrections

<table>
<thead>
<tr>
<th>Correction Symbol</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Errors</td>
<td>Errors that make it difficult for the reader to understand your meaning.</td>
</tr>
<tr>
<td>SS</td>
<td>Incorrect sentence structure</td>
</tr>
<tr>
<td>?</td>
<td>Unclear message</td>
</tr>
<tr>
<td>RC (relative clause)</td>
<td>Incorrect relative clause structure</td>
</tr>
<tr>
<td>NO (word order)</td>
<td>Incorrect or awkward word order</td>
</tr>
<tr>
<td>COH (coherence)</td>
<td>One idea does not lead to the next</td>
</tr>
<tr>
<td>CON/TRAN</td>
<td>Incorrect or missing conjunction or transition word</td>
</tr>
<tr>
<td>Frag (fragment)</td>
<td>Incomplete sentence</td>
</tr>
<tr>
<td>RO (run-on)</td>
<td>Two independent clauses joined with no punctuation or connectors</td>
</tr>
<tr>
<td>CS (comma splice)</td>
<td>Two independent clauses joined by a comma</td>
</tr>
<tr>
<td>VT Verb tense</td>
<td>Incorrect verb tense</td>
</tr>
<tr>
<td>VF Verb form</td>
<td>Verb incorrectly formed</td>
</tr>
<tr>
<td>Prep Preposition</td>
<td>Incorrect preposition</td>
</tr>
<tr>
<td>WF Word form</td>
<td>Wrong word form</td>
</tr>
<tr>
<td>WW or WE Wrong Word or Wrong Expression</td>
<td>Wrong word choice or wrong expression (not expressed this way in English)</td>
</tr>
<tr>
<td>Minor Errors</td>
<td>Errors that do not seriously affect meaning, but that do irritate the English reader</td>
</tr>
<tr>
<td>Sp Spelling</td>
<td>Spelling error</td>
</tr>
<tr>
<td>Agr Agreement</td>
<td>Incorrect subject-verb agreement</td>
</tr>
<tr>
<td>Art Article</td>
<td>Missing or incorrect article</td>
</tr>
<tr>
<td>P/Punc (Punctuation)</td>
<td>Incorrect or missing punctuation</td>
</tr>
<tr>
<td>Cap/Sa Capital Letters</td>
<td>Upper or lower case letters needed</td>
</tr>
<tr>
<td>Pl/Sg Plural/Singular</td>
<td>Problem with singular or plural of a noun</td>
</tr>
<tr>
<td>Pro Ref or Agr</td>
<td>Pronoun reference unclear or incorrect pronoun reference agreement</td>
</tr>
</tbody>
</table>
Revising your Writing
Grammatical Corrections - Tracking your Progress

<table>
<thead>
<tr>
<th>Correction</th>
<th>Assignment</th>
<th>Assignment</th>
<th>Assignment</th>
<th>Assignment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Major Errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>?</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RC</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coherence</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CON/TRAN</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frag</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BO</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>VV</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prep</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>WW or WE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Minor Errors</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sp</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Agr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Art</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P/Punc</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cap/Sa</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pl/Sg</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pro Ref/Agr</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note: Anna's students were supposed to record on this page the written errors they produced in the assignments.
Project 4: Theft at Whiteshell Lake

You are the team leader of a four-person inspection crew en route to a remote site 815 kilometres from your office, where construction of a nuclear power generating station is in progress. You are travelling in a panel van and after 580 kilometres you and the crew agree to stop for the night. At 8:05 p.m. you pull into the Clock Inn, a small motel beside the road that skirts around Whiteshell Lake.

The following morning you are having breakfast in the motel's tiny dining room when Fran Pedersen, one of the crew, goes out to the van to fetch the road map. She returns almost immediately and gasps, "The van has been broken into!"

The four of you scramble out to the parking lot and can see right away that the window on the front passenger's door has been smashed.

"They were after the radio," Shawn Mahler observes, pointing to a gaping hole in the dash.

"Check if anything else is missing," you suggest. Already you are expecting the worst, but to your surprise find that only two other items have been taken, one inconsequential and one important: about $6.00 from a tray in the dash (parking meter loonies and quarters), and a video camera and videotapes from a storage box in the rear of the van.

You try telephoning your office, but it is too early and no one answers. The motel has a fax machine, so you write a memo to your manager and send it by fax. In it you describe what has happened and ask for a replacement video camera to be sent to you. Here is some additional information you draw on to write your report:

- You are driving company panel van licence number HLW 279; it is a Ford.
- Your trip was authorized by Travel Order N-704, dated one week ago, and was signed by your manager.
- The power generating station is being constructed beside the Mooswa River, 27 kilometres north of the small town of Freehampton.
- The Clock Inn is 3.5 kilometres west of Clearwater Village, on highway A1136.
- The third member of your crew is Servi Dashi.
- The video camera is a Nabuchi TX200 "Portacam." You rented it from Meadows Electronics at 2120 Grassmere Road of your city. Its serial number is 21784B.
- Your manager's name is M.B. Corrigan.
- The purpose of the video camera is to record construction progress visually. The videotapes will be edited and then shown at the Power Authority Directors' Meeting scheduled for the 22nd of next month.
- You telephoned the RCMP detachment at Clearwater Village to report the break-in and theft. They ask you to drop in and make your report in person. You plan to do this at the start of your drive to the construction site (which will be after you have sent your fax).
- In your report you ask your manager to ship you a replacement video camera by Greyhound bus the day after tomorrow. One bus a day passes through Freehampton, but it stops only on request. (You will drive to Freehampton to meet the bus, and will telephone your manager tomorrow to check that the video camera will be on that particular bus.)
- You use today's date as the date of your report.

Part 1.

Write the incident report to M.B. Corrigan. Prepare it as a memorandum with a fax cover sheet.
Sample of prompt for an investigation report used in Anna's classes

You are an engineering technician working for H. K. Winman and Associates (Professional Consulting Engineers). Your supervisor, Marg Brock, asks you to handle one of their clients Marsland Construction Company. She shows you a letter dated March 12 in which Mr. Marsland requests that H. K. Winman carry out an independent study on high fuel consumption for one of their homeowners, the Parsenons. "They want us to find out what's going on and suggest any remedies", says Marg.

You call up the company and set up an appointment with Jack Marsland for March 20 at 10:00. This is your first big project and you feel nervous. The next morning, you meet with Mr Marsland in his office on Dundas Street. He tells you about Mr. and Mrs. Parsenon. It seems that the Parsenons have complained about the high gas consumption ever since the house was new and they first occupied it in 1988.

Mr. Marsland continues, "Over the years, we've done numerous checks and found little wrong with the house. But they just keep complaining and nagging us. We think it would be best if an independent consultant like yourself could carry out a study to identify the cause of the problem and recommend how it be resolved. Then we could present the results to the Parsenons to persuade them that we have done everything we can to rectify the problem. Whatever you discover is wrong, we'll fix it if it's our fault. Then I can use your report to get the Parsenons off my back!" Mr. Marsland concludes.

You get to work right away, setting up an appointment for the next day to visit the Parsenon home at 1464 Gregory Avenue. When you arrive, Mr. and Mrs. Parsenon greet you at the door. They are very friendly and show you around the home. Over coffee and cookies, Mr. Parsenon, who tells you to call him Jim, explains that they have to keep their thermostat set at 22.5 °C all the time so that they don't feel cold, especially because the floors are so cold. "If I don't keep the temperature up and wear these warm bunny slippers, my arthritis acts up" says Mrs. Parsenon. "I'm sure there is something wrong with your heating system" says Jim. "I think it's the insulation they put in when they built the house. I don't think it's up to standard" says Mrs. Parsenon.

You then inspect the house. These are your notes:

- gas furnace good
- hot air ducts good condition
- humidifier plates in furnace really corroded
- might cause lower humidity in home but not lower temperature. Low humidity can make people feel colder
- Gas company should be replacing these plates during regular maintenance
- gas flow meter is fine. Checked with Montrose Gas company. Said it has been replaced 2 times in the past 12 months at the Parsenon's request, but Montrose said that there was no problem with the metre both times
- ceiling has minimum of 180 mm of wood chips which equals R16 insulation factor
- walls have fiberglass insulation with R8 insulation factor
- basement unfinished and uninsulated.

After checking the insulation you realize that there is no way it would meet the government standards for a bungalow today (it would have to be R40 in the ceiling and R20 in the walls), but it does meet the insulation requirements that were in effect when the house was built in 1988.

Once your inspection of the home is complete, you realize that you have to do more than simply this. You couldn't find any real reason for high consumption in your inspection of the building and the heating system. You decide to compare the Parsenon home with similar bungalows in the neighbourhood.

You decide that the best comparison would be houses on the same street. So you find 4 identical homes built by Marsland and 4 other homes built at the same time but by other contractors. You ask each of these homeowners if it is okay to use their consumption records for 1987. You also ask them to tell you if they have installed any additional insulation since their bungalows were built, and their thermostat settings for day and night. You put all the information in a table (see attachment).

You spend some time talking to the owners down the street at 1410 Gregory Avenue who have an identical bungalow. It seems that they keep their place at 20.5°C. You ask them if their floors are cold as well. Mr. Shulman says, "We had that problem, but after we insulated the basement, our floors get nice and warm. We don't ever wear slippers now. I now keep the temperature at 20.5 instead of 22."

- 240 -
You analyze the information in the table and conclude that the problem lies with the homeowners and not the construction company. In 1997, the gas consumption at the Parsenon home was 38 MCF or 4.37% more than the other eight homes. The Parsenons keep their home warmer than the other eight. Even the other Marsland built homes are like the ones built by the other contractors. You even compare those with insulation to those that don't have it. You decide to include the table with the report and interpret the findings for the readers.

As you plan your report, you realize that the Parsenons could do several things to reduce gas consumption. They could lower the thermostat or insulate the basement and even the walls and ceilings. You check out the costs and use your comparison table to present options. You check with the insulation contractor and she tells you that it would cost $450 for the basement and $550 for the ceiling and walls. The exact cost though depends on the type of insulation used. And then you realize you forgot to mention the humidifier plates on their furnace. It's the gas company's job to do this every year.

In the end, no matter which way you look at it, Marsland Construction is not at fault. The problem with high fuel consumption lies with the Parsenons. You have to tell them something they don't really want to hear which is that Marsland Construction is blameless and that they have to fix the problem themselves. You fix yourself a big mug of coffee and start to write the report.
## ATTACHMENT
### COMPARISON OF GAS CONSUMPTION
IN NINE GREGORY AVENUE BUNGALOWS
FOR CALENDAR YEAR 19XX

<table>
<thead>
<tr>
<th>House No.</th>
<th>Year Built</th>
<th>Size (Sq ft)</th>
<th>Consumption (in MCF)*</th>
<th>Thermostat Setting (Day °C)</th>
<th>Night °C</th>
<th>Additional Insulation</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>R. M. Parsenon Residence</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1404</td>
<td>1986</td>
<td>1004</td>
<td>908.4</td>
<td>22.5</td>
<td>22.5</td>
<td>None</td>
</tr>
<tr>
<td><strong>Identical Houses Built by Marsland Construction Company</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1396</td>
<td>1985</td>
<td>1004</td>
<td>670.8</td>
<td>22.0</td>
<td>18.0</td>
<td>None</td>
</tr>
<tr>
<td>1399</td>
<td>1985</td>
<td>1004</td>
<td>894.4</td>
<td>21.0</td>
<td>21.0</td>
<td>None</td>
</tr>
<tr>
<td>1407</td>
<td>1985</td>
<td>1004</td>
<td>880.9</td>
<td>21.0</td>
<td>17.6</td>
<td>None</td>
</tr>
<tr>
<td>1410</td>
<td>1985</td>
<td>1004</td>
<td>841.3</td>
<td>20.6</td>
<td>20.6</td>
<td>Basement</td>
</tr>
<tr>
<td><strong>Group Average</strong></td>
<td></td>
<td></td>
<td><strong>971.9</strong></td>
<td><strong>21.1</strong></td>
<td><strong>18.2</strong></td>
<td></td>
</tr>
<tr>
<td><strong>Nonidentical Homes Built by Other Contractors</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1506</td>
<td>1983</td>
<td>966</td>
<td>889.9</td>
<td>21.5</td>
<td>16.5</td>
<td>None</td>
</tr>
<tr>
<td>1515</td>
<td>1983</td>
<td>980</td>
<td>813.3</td>
<td>21.0</td>
<td>16.5</td>
<td>Ceiling, Walls, Basement</td>
</tr>
<tr>
<td>1524</td>
<td>1983</td>
<td>980</td>
<td>900.3</td>
<td>22.0</td>
<td>22.0</td>
<td>None</td>
</tr>
<tr>
<td>1581</td>
<td>1983</td>
<td>966</td>
<td>977.6</td>
<td>21.0</td>
<td>18.0</td>
<td>None</td>
</tr>
<tr>
<td><strong>Group Average</strong></td>
<td></td>
<td></td>
<td><strong>998</strong></td>
<td><strong>865.0</strong></td>
<td><strong>21.4</strong></td>
<td><strong>18.2</strong></td>
</tr>
<tr>
<td><strong>AVERAGE OF 8 CONTROL HOMES</strong></td>
<td></td>
<td></td>
<td><strong>968.4</strong></td>
<td><strong>21.2</strong></td>
<td><strong>18.7</strong></td>
<td></td>
</tr>
</tbody>
</table>

* MCF = 1000 cubic feet