RESEARCH
on
ENGLISH COMPOSITION
Research
on
English Composition

CHING-KANG LIU

THE CRANE PUBLISHING CO., LTD.
For my wife

Yu-ling
## Contents

Preface

<table>
<thead>
<tr>
<th>Chapter 1</th>
<th>The Composing Process</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>The composing process</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>Studies on L2 composing processes</td>
<td>7</td>
</tr>
<tr>
<td>1.2.1</td>
<td>L1 and L2 composing processes are similar</td>
<td>7</td>
</tr>
<tr>
<td>1.2.2</td>
<td>L1 and L2 composing processes differ</td>
<td>12</td>
</tr>
<tr>
<td>1.3</td>
<td>Implications of recent studies</td>
<td>18</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 2</th>
<th>Between Speech and Writing</th>
<th>21</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>What have linguists found</td>
<td>22</td>
</tr>
<tr>
<td>2.2</td>
<td>A cognitive perspective</td>
<td>27</td>
</tr>
<tr>
<td>2.3</td>
<td>The cultural and social factors</td>
<td>29</td>
</tr>
<tr>
<td>2.4</td>
<td>The speech-writing issues for EFL students</td>
<td>32</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3</th>
<th>Between Reading and Writing</th>
<th>37</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>The directional hypothesis</td>
<td>40</td>
</tr>
<tr>
<td>3.2</td>
<td>The nondirectional hypothesis</td>
<td>44</td>
</tr>
<tr>
<td>3.3</td>
<td>The bidirectional hypothesis</td>
<td>48</td>
</tr>
<tr>
<td>3.4</td>
<td>What do beginner writers do in learning writing?</td>
<td>51</td>
</tr>
<tr>
<td>3.5</td>
<td>What should teachers do in writing instruction?</td>
<td>52</td>
</tr>
<tr>
<td>3.6</td>
<td>The reading-writing relationship in</td>
<td>105</td>
</tr>
</tbody>
</table>
This book focuses on five aspects related to writing in English: the composing process, the relationship between speech and writing, the relationship between reading and writing, errors in writing, and how writing is evaluated. These aspects will be discussed in five separate chapters, respectively, by comparing theories of English as the first language (L1) and those of English as a second language (L2). For instance, in the first chapter, theories and models of the composing process of L1 will be introduced and discussed. As there has been no specific model of the writing process for L2, a general survey of how L1 composing theories are applied to L2 will follow. At the end of the first chapter, the educational implications of these composing theories will be discussed.

These five perspectives are chosen over others because I am involved in and concerned with issues related to these areas. As an English composition instructor for several years (since 1993), I have realized that writing in English is usually a nightmare for most ESL or EFL students. As a matter of fact, a number of English native speakers also have a difficult time dealing with English composing “blocks” (Rose, 1984) once in a while and cannot figure out why that happens (both in
L1 and L2). This problem drove me to work seriously on how writing is “processed.” I started with of investigating the writing process that has been studied in hope that universal models can be found and applied to both L1 and L2. The studies of relationships between speech, reading and writing provide information concerning how a writer fosters his knowledge outside the writing process. Error analysis provides general types of errors made in writing, and how errors impede a writer’s composing and a reader’s comprehensibility. Different approaches to evaluation of English composition provide different means to evaluate the written products produced by students. Besides, each approach reveals its advantages and disadvantages in practice. Each of these five topics contains valuable information that will enhance the understanding of the issues of both English composition in a narrow sense and writing in a broad sense.

This book is written for people who teach English composition in Taiwan and other non-English speaking areas and want to understand more about issues related to English composition. In addition, students who learn English as a second or foreign language should also know how L1 theories can be applied to L2, or more specifically, how theories based on the English-speaking subjects can be applied to ESL or EFL students.

Practical as they might be, these five aspects do not include all components that make English composition possible. I hope this book will help stimulate both instructors as well as students who are interested in this field to do better in the research of English writing.

I would like to thank my advisor, Professor Priscilla A. Drum, School of Education, University of California at Santa Barbara, who helped and supported me when I was writing the draft of this book. I would also like to thank Professor Jerry Keating, my colleague in the English Department, the Chinese Culture University, who polished the language part of this book and raised quite a few helpful questions as well. Above all, my warmest thanks are given to my wife, Yuh-ling Wu, who has been taking care of our two children and the family and has made it possible for me to finish this book.
The Composing Process

1.1 The composing process

The composing processes are the major thinking or reasoning processes that a writer employs during the act of writing. Traditionally, unless writing instructors or researchers are interested in observing how students’ written products are completed, composing processes are usually not included in a writer’s (or a student’s) writing achievement. As long as the products meet the requirements of grammatical or rhetorical conventions and are interesting to read, it does not matter how these products (reports, compositions, reviews, etc.) are completed. Failure of student writing is generally ascribed to a student being incapable of absorbing or inadequately practicing the contents of instruction of compositions, which usually consists of “analyzing classic examples of good form, learning the rules that govern those classic examples, and practicing following the rules (either in exercises of limited scope or by imitating the classic models)” (Applebee, 1986, p. 95). In other words, the quality of the products of writing is, accordingly, measured by the ability to incorporate those rules into one’s writing.

Although the traditional instruction of composition has its indispensable role in making writing behavior complete and successful, this input-output-centered pedagogy is insufficient to account for the fact
that more and more students produce essays full of errors that are sup-posed to disappear in the earlier grades and of ideas so ill considered and poorly organized (Bizzel, 1986) so as to impede the reader’s compre-hension ability. The students’ or writers’ cognitive processes during the act of composing, therefore, attracted researchers and educators interest and theories of composing processes emerged in the 1960s.

The first model of composing is introduced by Rohman and Wlecke (1964). They proposed that successful college-level writers typically experience three stages in composing: pre-writing, writing, and editing. At the pre-writing stage occur idea-generating activities that provide essential preparation for drafting. Rohman and Wlecke believed that not only the activities at the stages of writing and editing could be taught, but the activities at the pre-writing stage could also be taught and suggested that these activities, such as meditation, become the actual content of the writing course.

Following the same direction, Emig (1971) provided evidence in a case study of eight twelfth-grade students, claiming that “prewriting is a far longer process in self-sponsored writing” (p.92) than in school-sponsored writing (e.g., assignments from school). However, Rohman & Wlecke (1964) and Emig (1971) did not provide the detailed evidence of how “activities” function at the pre-writing stage. They assumed that there exists a linear relation between these three stages of the writing process.

In the 1980s, some researchers working on composing processes (e.g., Hayes & Flower, 1980; Witte, 1985) were not satisfied with the simplicity and linearity of Rohman-Wlecke model, contending that the simplistic linear-stage models of composing served better as models of the emergence of the written product than as models of the real writing and thinking processes that writers experience in producing written texts. Hayes & Flower (1980, 1983) described the act of composing, not as a sequence of stages, but as a set of distinguishable processes that the writer must orchestrate in the act of writing. They also present a model, shown in Figure 1.1, specifying how these processes are organized.
Hayes & Flower’s model specifies three main elements that must interact:

1. **The task environment.** It includes everything outside the writer’s skin, including the assignment or exigency for writing, the audience, and eventually the written text itself. Hayes & Flower (1983) stressed that “no matter what rhetorical problem is actually presented to the writer, what finally matters will be the problem the writers represent to themselves in the act of composing” (p. 208).

![Figure 1.1 Structure of the writing model by Hayes & Flower (1980)](image)

2. **The writer’s long-term memory (LTM).** As shown in Figure 1.1, the writer’s LTM includes not only the writer’s knowledge of the topic but also the writer’s knowledge of writing plans, conventions, genre, rhetorical problems, and so forth. Hayes & Flower (1983), again, pointed out that what counts in writing is not what the writer may be informed of but what he is able, or chooses, to draw from his LTM during the act of composing.
(3) **The writing processes.** This element consists of four subprocesses: Planning, Translating, Reviewing, and Monitor, which are the major thinking processes that this chapter will focus on.

Planning is seen as a very broad activity that includes generating content, organizing it, and setting up goals, but also deciding on one’s meaning, deciding what part of that meaning to convey to the audience, and choosing rhetorical strategies (Hayes & Flower, 1983). It includes the whole range of thinking activities that are required before one can put words on paper. In addition, Hayes and Flower (1980) pointed out that (1) planning goes on throughout the composing act and (2) the plan may not be encoded in a fully articulated or even in a verbal form (plans may be nonverbal images).

Translating is the act of expressing the content of Planning in written English. Hayes & Flower noted that one can distinguish when writers move from Planning (producing notes and doodles) to Translating (producing prose), it does not mean that writers can have a fully formed meaning that they simply express in words. With further explanation, Hayes & Flower (1983) said, “Writers have some more or less developed representation encoded in one form. The act of translating this encoded representation to another form (i.e., written English) can add enormous new constraints and often force the writer to develop, clarify, and often revise that meaning. For that reason, Translating often sends writers back to planning. Often these processes alternate with each other from one minute to the next” (p. 209).

Reviewing is the act of evaluating either what has been written or what has been planned. Hayes & Flower (1980) postulated that when the evaluation of a text or a plan is negative, Reviewing often (but not invariably) leads to revision. Reviewing sometimes occurs self-consciously while the writer intends to evaluate the “translated” output. It sometimes occurs automatically when the writer senses the error or illogical expressions during the act of Translating.

Through the use of the Monitor, the writer is able to switch back and forth among the processes and to embed one process or subprocess within another. Hayes & Flower (1983) explained that “The monitor
The Composing Process

The composing process may function differently from writer to writer and from writing task to writing task. Some people move into translation as soon as possible in writing a paper, whereas others will not write a word until they are reasonably certain that planning is complete. Further, a writer undertaking an easy task, for example, a short letter, may do little or no planning before writing the first sentence, whereas the same writer undertaking a difficult task, for example, a philosophical treatise, may plan for months before writing a word” (p. 209).

During the act of composing, these subprocesses (e.g., Planning, Translating, and Reviewing), do not necessarily occur in a fixed sequence. Any of the three can be applied in any order the writing job demands. These subprocesses, therefore, show the recursive quality that, e.g., in Reviewing the writer can call in Planning and Translating any time necessary.

In addition, the evidence they presented clearly contradicts textbook approaches, which often suggest arbitrary, discrete steps in composing: formulate a thesis, develop an outline, and write (Hillocks, 1986). The research of Hayes and Flower (1980) indicates that generating ideas occurs prior to formulating a thesis or outlining and even during the Translating and editing processes, but the “stages” are not discrete. They are frequently interrupted by other processes.

Hayes & Flower’s general model undoubtedly provides researchers and educators a clear picture of how a writer’s mental processes function during the act of composing. But after this, there comes the problem of methods for studying composing processes. Hayes & Flower argued that the most superior way of studying composing processes is using thinking-aloud protocols. The reasons are: (1) they provide direct evidence about processes, (2) they yield rich data and thus promote exploration, and (3) they can detect processes that are invisible to other methods. Some researchers (e.g., Nisbett and Wilson, 1977) criticized that the thinking-aloud protocols are not valid on the basis of the following reasons: (1) people are not conscious of their cognitive processes and, hence, should not be able to report these processes, and (2) even when people report their conscious processes, these reports can be
verbally distorted or incomplete. Hayes & and Flower (1983) agreed with such criticism but insisted that, compared with other process-tracing methods (e.g., behavior observations, and retrospective reports), thinking aloud protocols still provide the most sufficient information. To support this, they made further arguments (1) that if the subject is reporting memory traces that are verbal, then reporting appears neither to change nor to slow down task performance, and (2) that if the subject is reporting nonverbal memory traces, then reporting may slow task performance but does not appear to change its course or structure. Only if the subject is asked to report information to which he or she would not normally attend while performing the task, then reporting may modify the usual sequence of mental processes involved in performing the task. This will not happen in the course of collecting thinking-aloud protocols.

After Emig’s (1971) pioneer study on writing processes and Hayes and Flower’s (1980) composing process model, there followed a flood of similar studies. Some researchers emphasized planning or prewriting stages (e.g., Matsuhashi, 1981; Pianko, 1979); some emphasized revising (e.g., Nold, 1981; Sommers, 1979; Scardamalia and Bereiter, 1983; Witte, 1985). Based on the multitude of specific results, several general findings are summarized as follows by Applebee (1986):

1. Writing involves a number of recursively operating subprocesses (for example, planning, monitoring, drafting, revising, editing) rather than a linear sequence (Hayes and Flower, 1980, 1983; Flower & Hayes, 1981a, 1981b).

2. Expert and novice writers differ in their use of those subprocesses. For example, novice writers have poorly developed skills in differentiating planning from text production (Scardamalia, 1982).

3. The processes vary depending upon the nature of the task. More recent studies have suggested that the processes also vary depending on the instructional context (Dyson, 1984), personal history (Florio, 1984), and knowledge the writer has about the topic (Langer, 1984).
These general findings, though derived from the results of different studies, can serve well as a summary of the L1 composing process theories. The studies also reveal the limitation of recent process-tracing research. For instance, based on Hayes & Flower’s argument, the best method of studying composing processes is using thinking-aloud protocols. But this method not only costs too much in practice but also cannot have a large sample involved. The results, therefore, can hardly be generalized before more similar and support researches are conducted.

1.2 Studies on L2 composing processes

In this section, studies concerning the L2 (or more specifically English as a second language) composing process will be divided into two categories: One consists of the studies that reflect the L1 composing theories or models, and the other, of the studies that draw the conclusion that L1 and L2 composing processes differ to a certain extent.

1.2.1 L1 and L2 composing processes are similar

Some researchers (Gaskill, 1986; Hall, 1987; Arndt, 1987) in their studies find that there are no differences between L1 and L2 composing strategies. Gaskill (1986) analyzed videotapes of four undergraduate subjects composing protocols and their written products, concluding that the writing and revising processes in English resembled those in Spanish. Hall (1987) analyzed videotapes of four subjects and their multiple drafts for generating data, interviewed them and had them fill out postwriting questionnaires, concluding that a single system was used to revise across languages (English and Spanish). Arndt (1987), based on six graduate-level, Chinese-speaking subjects, also had similar findings that “the composing strategies of each individual writer were found to remain consistent across languages” (p. 257).

Similar results are found by Zamel (1982), who interviewed eight university-level “proficient” L2 writers, asking for retrospective accounts of the writing processes, and asked each student to write multiple
drafts of an essay. Zamel found that the writing processes of her L2 subjects were like those of the subjects described in L1 studies, concluding that competence in the composing process is more important than linguistic competence in the ability to write proficiently in English. The combination of the conclusions made by Jones (1982) and Zamel (1982) clearly indicates that both poor and proficient L2 writers borrow their L1 composing competence or strategies while producing L2 compositions.

To directly compare L1 with L2 in composing behaviors, Jacobs (1982) collected her students' writing (six native and five nonnative speakers of English) and interviewed the subjects after they produced their compositions. She found one apparent inverse relationship between “integrative thinking” and “grammatical accuracy” among her subjects, but this relationship had something to do with a student’s development as a writer in general. No significant difference was found between L1 and L2 in her study.

Edelsky’s (1982) product-based, longitudinal study also supports the proposition that fundamental L1 composing processes are applied to L2 writing. Her subjects were 26 first, second, and third graders enrolled in a bilingual program. She collected as data four different writing samples from each of her subjects over the period of one school year, finding that the knowledge of L1 writing forms the basis of new hypotheses rather than interferes with writing in another language.

The following most recent studies provide more evidence supporting that L1 and L2 composing strategies are similar. Hilderbrand (1985) observed her Spanish-speaking subject write in class and found that the subject preferred creative, personal writing, and thereby hindered the writing process of the academic mode. She found that it was the discourse types, not the L2 writer’s linguistic competence, that hindered the student’s composing process, and concluded that her subject did not discriminate L1 and L2 composing strategies.

Diaz (1985, 1986) established a process-oriented classroom environment and then observed what happened to the students and their writing. Diaz (1986) found that L1 composing process strategies and
techniques could be recommended for ESL students. In addition, when these strategies and techniques were used in “student-centered” contexts, the benefits to these students could go beyond their development as writers.

In looking at the knowledge transforming model (Bereiter & Scardamalia’s term, 1987) rather than strategies and techniques, Brooks (1985) proposed that writing competence was separated from oral proficiency and that the lack of L1 cognitive academic development—a somewhat broader concept than writing competence—affected most L2 learners’ composing skills. Jones and Tetroe (1987), taking the same position, analyzed protocols to examine the L1 and L2 planning behaviors of six Spanish-speaking L2 writers, all of whom were planning for graduate study. Collecting data over a six-month period, they found a great variety among their subjects in the amount of native language use in L2 writing, as well as “some decrease in [writing] performance” when writing in the second language. They, therefore, proposed that composing in a second language employs “cognitive capacity” that would be used “for other tasks when writing in the native language” (p.53). This finding suggests that composing in L2 is using Bereiter and Scardamalia’s\(^1\) (1987) knowledge transforming model, rather than knowledge telling model. On the other hand, Jones and Tetroe also

---

\(^1\) In 1987, Bereiter and Scardamalia proposed that writing (or composing) had both natural and problematic models.

One [natural] model makes writing a fairly natural task. The task has its difficulties, but the model handles these in ways that make maximum use of already existing cognitive structures and that minimize the extent of novel problems that must be solved. The other [problematic] model makes writing a task that keeps growing in complexity to match that expanding competence of the writer. Thus, as skill increases, old difficulties tend to be replaced by new ones of a higher order. They also refer to the first model as knowledge telling, and the second, as knowledge transforming. (p. 5)

Bereiter and Scardamalia (1987) have several valuable findings. One of them is that novice writers bear a close resemblance to what appears on the page and expert writers have a great deal of activity revealed in the thinking-aloud protocols that is not directly represented in the text.
pointed out that a lack of L2 vocabulary may result in using first language strategies in L2 composing, and “that the quality, though not the quantity, of planning transfers from L1 to L2” (p. 56).

With respect to writing ability, some researchers draw conclusions that a lack of competence in writing in English (L2) results more from the lack of composing competence in general (e.g., in L1) than from the lack of linguistic competence of L2. Jones (1982) investigates the written products and writing processes of two L2 writers (one “poor”, a Turkish speaker, a graduate-level student, and the other “good”, a German speaker, a freshman-level writer). Johns analyzed the composing strategies by examining two composing behaviors: writing or generating text and reading the text already generated. He found that writing strategies affected the writers’ rhetorical structures. The poor writer was confined to the text so as to hinder the generating of ideas, whereas the good writer allowed her ideas to generate the text. Jones, accordingly, concluded that the poor writer had never learned how to compose. In other words, a general lack of competence in composing, rather than a specific lack of L2 linguistic competence, is the source of the poor writer’s difficulty in L2 writing.

Zamel (1983) conducted another study of six (orally) advanced L2 students (four skilled writers and two unskilled writers), using a case study approach, observing her subjects while they composed, interviewing them about their conclusion of their own writing, and collecting all of their written materials (including drafts) for the production of one essay each, which they had unlimited time to complete. Vocal protocols were not used because the researcher suspected verbalizing aloud might interfere with real composing situation. As a result, she found the skilled L2 writers revised more and spent more time on their essays than the unskilled writers. In general, they concerned themselves with ideas first, revised at the discourse level, inhibited recursiveness in their writing process, and saved editing until the end of the process--all writing strategies similar to those of skilled L1 writers, as described in L1 writing process studies (e.g., Pianko, 1979; Sommers, 1980). As for the unskilled L2 writers, they revised less and spent less time writing than the
skilled writers. They focused on small bits of the essay and edited from
the beginning to the end of the process, very like the unskilled writers in
Sommers’s (1980) report of her L1 writing process study. Zamel (1983)
also found that writing in a second language did not have a major im-
 pact on the composing process in general. The conclusion she made is
that the skilled writers in her study clearly understood what writing en-
tails whereas the unskilled writers did not.

Some researchers find that transferring L1 composing strategies to
L2 writing may result from the fact that some instructors apply instruc-
tion of L1 writing directly to L2 writing. Jones (1985) pointed out that
“It is worth noting that many of the proposals for improving first lan-
guage composing are also effective in helping second language learners
develop acquired linguistic competence” (p.114). Urzua (1987) pro-
vided some empirical evidence in her investigation of four children, two
fourth graders and two sixth graders, over six months. Her data con-
sisted of transcripts of peer response sessions, weekly compositions, and
twice-weekly dialogue journals. She observed that children acquired
three significant composing skills: “(a) a sense of audience, (b) a sense
of voice, and (c) a sense of power in language” (p. 279). This finding
strongly indicates that what has proved effective in L1 classrooms is
also effective in the L2 classroom (Krapels, 1990).

In summary, three statements can outline the general findings to
represent the first section (section 1.2.1) that L1 and L2 are similar in
the composing processes.

(1) The linguistic competence of L2 does not affect or impede L2
composing competence.

(2) The composing processes do not differ between L1 and L2 in that
composing involves a number of recursively operating subprocesses
(for example, planning, monitoring, drafting, revising, editing)
rather than a linear sequence. And

(3) Expert and novice writers differ in using those subprocesses.

1.2.2 L1 and L2 composing processes differ
Results found or conclusions made by recent researchers do not provide a whole picture of the L2 composing process model. What has been provided is no more than one or two subprocesses that are unique to L2 and that do not occur in the L1 composing process.

Of all the unique problems observed in L2 composing processes, translating is the one that first attracts my attention because my observation in class (English native speakers learning Chinese as L2) provides the same evidence that Chinese learners frequently try to “translate” English ideas into Chinese. Obviously, the term translation is different from the Translation in the Hayes-Flower model. In the Hayes-Flower model, the Translation process acts under the guidance of the writing plan to produce language corresponding to information in the writer’s memory. The translation from L1 to L2 refers to the process that the mentally formed text in L1 is translated into L2. In other words, this latter kind of translation process will not occur in the L1 composing process.

Quite a few researchers provide evidences concerning the translation theory. Zamel (1982) found one of her subjects in her study (a graduate student) incorporated translation into her L2 writing processes. According to Zamel, this subject was “the most proficient writer” in her study (p. 201).

In 1986, Martin-Betancourt examined protocols of college-level L2 students and found that the L2 writing process was similar to the L1 process, with the exception of two composing behaviors--using (1) more than one language and (2) translating. For the first, Martin-Betancourt found that her subjects’ writing processes involved solving linguistic problems and that use of the first language in L2 writing added to such linguistic problems, especially in vocabulary. For the second, Martin-Betancourt found some L2 writers incorporate translation into their second language writing processes (although this phenomenon was inconsistent among her subjects in that some relied on Spanish very little while others used Spanish frequently: so frequently that it could be viewed as translating an expression in Spanish into that in English).
Translation, however, does not always have a negative effect in L2 writing. For instance, Friedlander (1990) pointed out that translating does not constrain writers, either in time or quality, as they produce L2 texts. In addition, all studies mentioned above come up to the same conclusion that translation from one language to the other does not impede the composing process. However, this finding does not guarantee that nothing happens while L2 students translate ideas from L1 to L2 in writing, and we need further studies to investigate this specific field.

Certainly, there are also researchers who take the opposite position, believing that translation should not exist at all. For instance, some of Zamel’s (1982) other subjects did not depend on translation to write in English. Indeed, they scorned such a procedure, one saying, “It would be like being pulled by two brains” (p. 201). This student’s comment corresponds with Johnson’s (1985) findings. Johnson’s six subjects felt that the use of first language L2 writing was ill advised for advanced L2 learners. However, they inevitably tended to rely on their first language when they were generating ideas about a culture-bound topic. But verbalizing “ideas” seemed to be more like the Translation in the Hayes-Flower model rather than the translation referring to converting L1 to L2 directly.

Liu (1994, 1995) conducted a study trying to locate the act of the L1-L2 translating process in the course of EFL college students’ English writing in Taiwan. The findings indicate that all of the five subjects employed the act of L1-L2 translation during the course of their English composition. According to the protocols produced by these subjects, the locations of the L1-L2 translating process occurred most commonly in the following locations (See Figure 1.2):

1. Translating Chinese notes into English notes;
2. Translating Chinese propositions into English propositions;
3. Translating Chinese sentences into English sentences.

Figure 1.1 shows the first two of the locations listed above. The third location refers to a simple process that the subject finished every composing act in Chinese before a Chinese sentence was produced. Then the subject translated the Chinese sentence into an English sentence.
Long Term Memory

-Generating Chinese ideas in Chinese
-Organizing Chinese notes in Chinese
-Translating Chinese propositions into English propositions
-Evaluating in Chinese

Generating English ideas in English
Organizing English notes in Chinese
Translating English notes into English propositions
-Evaluating in English

L1-L2 translation

Succeed
Fail

Figure 1.1 The L1-L2 translating acts in the composing processes in Liu’s (1994, 1995) study
Cultural factors, undoubtedly, make the study of the L2 writing process more complicated. Raimes (1985a) reported that her L2 composition class might represent at least half a dozen strikingly different cultures, very different educational backgrounds, ages ranging from sixteen to sixty, and very different needs for being able to write in a foreign language. Galvan (1985) examined ten doctoral students of education, investigating the cultural and linguistic factors in the composing processes. He drew data from interviews, direct observations, and assessments of writing skills, finding that his subjects’ second language writing was generally influenced by both their first language thinking and culture and their second language thinking and culture. Moving between two languages and two cultures caused Galvan’s subjects’ composing processes to be full of pauses and doubts.

To examine how culture-bound topics affect writing, Lay (1982) analyzed the compositions and accompanying composing-aloud audio-tapes of four adult, Chinese-speaking L2 students. She also interviewed her subjects about their writing backgrounds and current attitudes toward writing. She found that although intending to produce L2 compositions, her subjects incorporated their first language into their L2 composing processes. Her study indicated that (1) “when there are more native language switches (compared to the same essay without native language switches), the essays in this study are of better quality in terms of ideas, organization and details” and (2) “certain topics induce more native language switches” (p. 406).

Burtoff’s (1983) product-based study provided the conclusion that the more culture-bound the topic, the more different the discourse structures are. Other studies of Lay (1982) and Johnson (1985) all indicated that certain topics elicit more first language use in second language composing than other topics do.

Friedlander (1990) also provided information about the role of first language (Chinese) use in generating content. His findings indicated that using the topic-related language to plan content results in better planning and better compositions. He concluded that “when writers planned in Chinese on the Chinese topic (Qingming) and in English on
the English topic (Carnegie Mellon University), their plans and texts were rated significantly better than when they planned in Chinese on the English topic and in English on the Chinese topics” (p.123).

Some studies do not specifically provide evidence in how L1 and L2 theories are different, but in how L1 and L2 interact. Cumming (1987) examined six Francophone Canadian adult subjects and found they tended to use their first language for generating content for the three writing tasks they were given, which were personal, expository, and academic. Cumming also found that the expert writers in his study did a lot of thinking in French, a finding contrary to Johnson’s (1985) finding that L2 students did not use their first language at all.

In addition to the issues of translation and topic, the monitor theory\(^2\) has also attracted some researchers. To examine how the monitor theory functions in the L2 composing process, Jones (1985) videotaped nine advanced L2 students while they wrote three different compositions; and each composition was of a different discourse types—personal experience, descriptions, and generalization (the topics were drawn from Pianko’s (1979) L1 writing process study). Subjects were given unlimited time to complete each composition, and were interviewed on their writing processes for each composition and on their opinions about how their L2 writing processes differed from their L1 processes. When applying Krashen’s monitor theory to analyze the writing behaviors of

---

\(^2\) The central hypothesis of Krashen’s (1978) monitor theory is that language ability is developed through acquisition (unconscious, internal, and systematic processes) and learning (conscious, explicit mastery of rules). Learned information, therefore, functions as a monitor of production from the acquired system. Krashen believes that the monitor theory can also be applied to the acquisition of writing skills. However, the use of the monitor theory has several constraints: (1) the learner must have time in which to cause the monitor to operate; (2) the learner must be “focusing on form”; and (3) the learner must have knowledge of the rules that apply in the particular situation in which the monitor comes into play. In addition, there exist individual variations in applying the monitor system. Some people make excellent use of the monitor, while others either over-use or underuse it. In short, the monitor theory predicts that given certain conditions, the learner can apply consciously learned grammar rules to alter and improve the accuracy of his or her written or spoken utterances.
two subjects in this study, Jones (1985) reported that “monitoring does not lead to improved writing” (p.112).

Rorschach (1986) pointed out the same phenomenon that monitoring may impede L2 writers from developing a better text. She asked three advanced L2 subjects to write one essay each and to revise these essays afterwards. She then collected the data of external raters’ evaluations of the students’ essays and interviewed the students about why they wrote and revised as they had done. The results showed that reader awareness led the writers to focus on correctness rather than content. She, accordingly, concluded that her study called into question composition teaching that concentrates on form, reflecting Jones’s (1985) speculation about the relationship between instruction and overusing the monitor.

Raimes (1985b) also observed that her subjects wrote more, exhibited more commitment to the writing task, produced more content, and paid less attention to errors. Raimes (1987) again found that L2 writers were different from L1 writers in that L2 writers “did not appear inhibited by attempts to edit and correct their work” (p.458). Based on these studies, the role of monitor seemed more complicated than simply directing when to switch processes or subprocesses (in the Hayes and Flower model) or guiding L2 writers with correct grammatical rules or rhetoric conventions. More studies are also needed for the understanding of the function of monitor in L2 writing.

In summary, the findings provided by the researchers who find that L1 and L2 composing theories are somewhat different can be outlined as follows.

1. Translation from L1 directly to L2 is one process used by L2 writers, which does not occur in L1 writing.
2. Certain writing tasks, apparently those related to culture-bound topics, elicit more first language use when writing in L2 than other tasks.
3. The monitor theory does not seem to help L2 writers write better both in quality and in quantity. And
in applying L1 composing theories, as Raimes (1987) suggested, adaptation is more appropriate than wholesale adoption.

1.3 Implications of recent studies

Studies on both L1 and L2 composing processes share several recurrent features in research designs. Most studies rely on case study as a research methodology, resulting in a small number of subjects in a single writing process study (e.g., Flower & Hayes, 1981; Lay, 1982; Zamel, 1983; Gaskill, 1986). The subjects of most published studies are usually “subjects of convenience,” who are sometimes the researcher’s own students and therefore not chosen randomly (e.g., Zamel, 1983; Pfingstag, 1984; Raimes, 1985a).

The writing tasks of studies vary in number, in discourse mode—narratives, descriptions, expository, and argumentative, in topic type—personal, nonpersonal, culture bound, and non-culture bound, and in time allotted for completing a task (Krapels, 1990). All these features make the studies, as Krapels stated, “lack of comparability” so as to affect their generalizability, which in turn decreases the significance of a study’s findings (p. 51). However, these features do suggest what we should do in the future. First, for both L1 and L2, “the research designs of future studies should replicate the best and most appropriate designs of prior studies so that findings can be compared” (Krapels, 1990, p. 51). Second, larger samples should be used so as to make the findings of each study more representative. And third, product- and process-based studies should be equally emphasized because “given knowledge of how planning and translating (in the Hayes-Flower model) enter into the composing process and the kinds of influences that act upon them; it is possible to make reasonable inferences about the writing processes from written products” (Witte & Cherry, 1986, p. 143).

For studies of L2 composing processes, Krapels (1990) suggests the use of ethnography to increase a researcher’s understanding of the writing processes. She points out that:
Ethnography takes the researcher out of the artificial environment of the laboratory and into the real space of writing, the collective consciousness of people making and then sharing meaning, so that both researcher and subject become participants in the research process (p. 51).

Raimes (1987) also suggests that “students perceive a writing task as having a real purpose and audience, more is needed than just surface wording to the topic” (p. 461). With cultural differences that may impede the L2 composing process, this can be even more true in that students in L2 writing classes are often asked to write something just for practice and not for real use, which usually makes it harder for L2 writers to generate ideas from something not necessarily familiar. That is why writing a review or a paper of an experiment enables L2 students to produce more satisfactory texts than most written assignments or topics or prompts assigned in L2 composition classes.

Is the composing process a “black box” in the writer’s head? Although we have learned a lot from recent studies, we still do not know why some people learn writing faster and more efficiently than others do. But, as Hayes and Flower (1983) pointed out, “research on writing is still in an early stage of development. We are now enjoying, or suffering through, a period of exploration” (p. 218), there is certainly much work left to be “explored.” I hope the future studies of the composing process will soon provide more solutions to these problems.
Writing, especially literary works, used to be seen as the true form of language and was held as primary; spoken language was viewed as nothing more than an imperfect or reflection of it. Spoken language was not studied by linguists until the nineteenth century, when Grimm (1785-1863) in Germany began to study speech and then Henry Sweet (1845-1912) in Britain started phonetics as a separate linguistic branch. Soon this trend of seeing speech as the true language flourished and has been dominating the whole field of linguistic study until today. Writing has been treated as a visual symbol system (Sapir, 1921), visible marks (Bloomfield, 1933), derivative of the face-to-face conversational norm (Fillmore, 1981), or simply as an artifact (Aronoff, 1985). In a word, writing is no longer primary in linguistic study.

Linguists are right from a historical viewpoint because speech developed much earlier than writing. Individually, they are also right because human beings normally develop their speech earlier than they learn how to write. In addition, many humans who are able to communicate orally never learn how to write. Judging from the value or function of the two forms of language, we cannot deny that speech is more widely used than writing in (1) that in this world there still exist some tribes in which only speech is used, (2) that, in the society where both speech and writing are used, not everyone who can communicate orally
can write, and (3) that even those who can write speak much more than they write (except, probably, for those professional writers or speech-disabled people).

On the other hand, however, writing or composition classes in academic schools have never stopped. People who read and write have generally been in more important positions or have been viewed as having higher status in a society. In addition, written records have always been the only documents to be trusted or legal (e.g., video- or audio-tapes still cannot fully serve as judicial evidence) at least in the United States. Written language, though it cannot compare with speech in historical length, in personal development, or in use as a means of daily communication, is loaded with heavy responsibility. One important feature that accounts for this is the stability of the written form of communication which enables written documents to be not affected by time and place. Ever since there were written records, the descendants of that specific culture have had the chance to understand, to doubt, to reevaluate, or to recreate those records. Without the stability of written records, a lot of knowledge concerning human civilization becomes impossible.

Understanding these facts, as Chafe (1992) concluded that “writing and speaking each has its own validity” (p. 257), more linguists have started comparing the linguistic features of the two forms of language. Educators have started observing how children develop their written language from speech. Some researchers even treated learning writing as learning a second language (Neilson, 1979; Horning, 1987).

2.1 What have linguists found?

After decades of investigating how speech and writing differ, linguists have done studies ranging from lexical density (e.g., Halliday, 1979), syntactic structures (e.g., O’Donnell et al., 1967; Halliday, 1979; Beaman, 1984) to situational features (e.g., Goody & Watt, 1963; Chafe, 1982; Cook-Gumperz & Gumperz, 1981). These studies provide overall linguistic characterizations of speech and writing. According to Biber (1988) in summarizing the results of previous studies in this field, writing is claimed to be:
1. more structurally complex and elaborate than speech, indicated by features such as longer sentences or T-units and a greater use of subordination (O’Donnell et al., 1967; Chafe, 1982; Tannen, 1982a, 1985; Gumperz et al., 1984);

2. more explicit than speech, in that it has complete idea units with all assumptions and logical relations encoded in the text (Olson, 1977; Chafe, 1986);

3. more decontextualized, or autonomous, than speech, so that it is less dependent on shared situation or background knowledge (Gumperz et al., 1984; Olson, 1977);

4. less personally involved than speech and more detached and abstract than speech (Blankenship, 1974; Chafe, 1982; Chafe & Danielewicz, 1986);

5. characterized by a higher concentration of new information than speech (Stubbs, 1980; Brown and Yule, 1983); and

6. more deliberately organized and planned than speech (Ochs, 1979; Rubin, 1980; Akinnaso, 1982; Brown & Yule, 1983; Gumperz et al., 1984) (p. 47).

In reality, any one of these characterizations may be criticized as not being able to generalize to all spoken or written genres because most of the studies cited are based on observing linguistic features of one or two situations. In other words, if the situations change, the results can be different or even contradictory to other studies. To avoid misunderstanding, Tannen (1982a, 1985), for example, notes that the characterization that writing is more decontextualized while speech is more contextualized is true only between conversation and expository prose, the two genres most frequently used to present speech and writing. It is not true of speech and writing in general.

However, not all linguists agree on the characterizations listed above. Some find that the elaboration and complexity of sentences in speech are higher than those of writing (Poole and Field, 1976; Halliday, 1979), which is contradictory to the characterizations listed previ-
ously. Beaman (1984) suggested that this contradiction results from their choice of samples. Because of the sample selection problem, Beaman pointed out, “what looks like differences between spoken and written discourse may really be differences in the register, purpose, formality, or amount of planning time of each task” (p. 51).

In addition, the definition of the variable, for instance “sentence,” can also be a factor to cause different results in similar studies. Blankenship (1962) found sentence length in speech and writing to be nearly the same, which is also contradictory to what Chafe (1982) or Tannen (1982a) discovered. Biber (1988) commented that “a major problem here concerns the definition of ‘sentence’ in speech (in English), and since most studies do not define the particular use of the term, there is no basis for comparison” (p.49).

Not satisfied with these contradictory findings and finding that pure quantitative studies have not addressed the important issues concerning speech and writing, Akinnaso (1982) and Gumperz et al. (1984) proposed to study thematic cohesion in the spoken and written texts, attempting to uncover the underlying differences between speech and writing. Biber (1988), furthermore, combined the quantitative methods used in most of the previous studies in the field of speech versus writing and the qualitative notions that examine the underlying dimensions drawn from explicit linguistic features. He developed a new method to compare multiple writing styles and multiple speaking styles with respect to a variety of features that might differentiate some, but not all, of the samples. Biber’s new study method has a strong hypothesis that informal conversations (representing speech) and academic/official documents (representing writing) are the two ends of a multi-featured and multi-functional continuum. He pointed out that “no absolute spoken/written distinction is identified; rather, the relations among spoken and written texts are complex and associated with a variety of different situational, functional, and processing considerations” (p. 24).

1. multichannel

<table>
<thead>
<tr>
<th>OC</th>
<th>AL</th>
<th>PL</th>
<th>AP</th>
</tr>
</thead>
<tbody>
<tr>
<td>+</td>
<td>+</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

unichannel
Furthermore, he added that “there is no single, absolute difference between speech and writing in English; rather they are more or less similar with respect to each dimension” (p. 199). With this hypothesis,
he presented a very simple picture, as shown in Figure 2.1, for how to visualize the relationship between the multiple linguistic features and multiple dimensions.

This figure (Figure 2.1) is of course a simplified one to let the reader have a quick understanding of this new approach. In Biber’s real studies, he used six dimensions (1. involved versus informational production; 2. narrative versus non-narrative concerns; 3. explicit versus situation-dependent reference; 4. overt expression of persuasion; 5. abstract versus non-abstract; and 6. on-line informational elaboration). He observed twelve types of texts (or genres) (from conversations to official documents) to see the quantitative data of the co-occurrence patterns of the linguistic features that mark underlying functional dimensions (these linguistic features, such as nouns, verbs, etc., are treated as variables in six factors in each dimension). Biber’s series of studies (1985, 1986) proved that this multi-feature/multi-dimension (MF/MD) approach is the most efficient one to date.

This approach has already been applied to the field of general composition research. Grabe and Biber (1987) used the model of textual relations developed in this MF/MD approach in a pilot study of the linguistic characteristics of good and poor essays written by native and non-native writers of English. That study found almost no difference between good and poor essays. The most striking result is that student essays are unlike any of the published genres of English; they use the surface forms of academic writing (e.g., passives), but they are relatively non-informational and involved, and they are extremely persuasive in form [based on Biber’s study, being non-informational and involved is the characteristic of face-to-face conversation, rather than academic composition]. This finding indicates that compositions do not have a well-defined discourse norm in English (Biber, 1988, p. 204).

It is true that students generally fail to distinguish one genre from the other when they write. What causes this may be ascribed to the writing instruction in class or to the students’ incapability of acquiring their
judgment from reading model essays (the relation between reading and writing will be discussed in the following chapter). However, if Biber’s MF/MD method can be applied to the study of students writing with elaborate designs, it undoubtedly may provide rich information about where student compositions stand between speech and official documents.

2.2 A cognitive perspective

In order to see the distinctions between speech and writing underlying the linguistic features, Chafe (1991) proposes that writing differs from speech in that writing experiences a process with a “displaced consciousness.” He established two basic modes: (1) Language and consciousness in the immediate mode; and (2) Language and consciousness in the displaced mode (See Figures 2.2 and 2.3). Writing basically belongs to the second displaced mode with a separate type of experiencing consciousness. In other words, writers do not experience what they write while they are writing. They write what they have experienced before (See Figure 2.4).

Chafe (1986), based on empirical evidences, also concluded that speech and writing are different in that, owing to the constraint of both the speaker and the hearer, each “intonation unit” in speech can hold only “one new concept at a time” (p. 25), whereas a punctuation unit in writing may consist of more than one new concept at a time. Apparently the “unit of consciousness” in speech (intonation units) and in
writing (punctuation units) apparently functions differently in processing. The consciousness between the speaker and the hearer is constrained by time and full involvement (or attention) whereas in writing; it is less constrained in both aspects.

![Diagram](image)

Figure 2.3 Displaced mode by Chafe (1991)

![Diagram](image)

Figure 2.4 The mode of written fiction by Chafe (1991)

Hildyard and Olson (1982), looking at the attention from the viewpoint of the hearer and the reader, also concluded that the hearer pays primary attention to the theme of the context, building a coherent representation of what is meant, and that readers, on the other hand, “are able to pay closer attention to the meaning of the sentences per se, recalling more incidental but mentioned details and being more accurate in their judgments of what was in fact stated in the text” (p. 32).

Most researchers (linguists or psychologists) who are interested in mental processes focus either on speech or on writing. Chafe may be the first linguist who thinks it is worthwhile investigating how conscious-
ness plays different roles in speech and in writing. But like other cognitive studies, it is always difficult to define the domain of the variable (e.g., the domain of consciousness) and how to measure it empirically. Obviously, this theory is expected to be accompanied with more well-designed studies to provide sufficient empirical evidence to indicate psychological reality.

2.3 The cultural and social factors

The linguistic findings we have had so far contribute a great deal to the understanding of the explicit characterizations of what a typical spoken discourse is (e.g., conversation) and how it differs from a typical written discourse (e.g., expository prose). These findings, however, do not help much in explicating why so many voluble children, as well as adults, have a hard time learning how to write. If a person cannot write because he never has a chance to learn, the reason may be simple: it is generally agreed that writing needs to be overtly taught (Martlew, 1986). But what about one who also goes to school and learns writing under the same instruction as his peers but still fails to learn to write well? Indeed, in the real world, there exist such students and the causes are still everyone’s guesses. One of the possible causes may be related to the different modes of communication in speech and in writing. Cook-Gumperz & Gumperz (1981) pointed out that

Children whose speech shows dialect features judged deviant by “standard” English speakers are not necessarily the worst at literacy skills. Test results moreover indicate that the gap in verbal ability between minority group children in many inner city schools is relatively low at the start of primary school but increases to alarming proportion by the fifth year (p. 90).

Cook-Gumperz & Gumperz then added that what impedes educational success (with literacy included) is not grammar (linguistic features), but the social or cultural gap between the oral community and the written community. They proposed that the introduction of a writing system has
great impact both on the individual cognitive process and on cultural practices in a society.³ For the individual cognitive process, the introduction of writing systems changes the basic character of the storage and transmission of knowledge. In preliterate cultures, one of the key ways that knowledge is transmitted is through such oral performances as the recitation of mythological folk narratives and oral genealogies. What is stable over time in these situations are story schemata, not details of content (Cook-Gumperz & Gumperz, 1981, p. 91). Later they cited Tripp’s (1977) study, saying that a seven-year-old girl, while writing her story on paper, kept using repeated lines or song refrains, which are common ways of cognitive processing (to enhance memory by repetition) in oral cultures but definitely not in written cultures. As for cultural practices in a society, the change from oral to written transmission brings about a shift from a view of knowledge as a constant state which can be learned through open and varied means of creative retelling; to one of knowledge as incremental, that is, where the initial learning process is repetitious in order to teach the store of knowledge available-to-date, but to which further new knowledge can be added, since the old store is on record.... Eventually even the daily life of people who can add to the essential store of knowledge may be further separated. New classes of literati arise who specialize in and earn their living through the preservation, editing and interpretation of written information. In doing so, over time, these groups develop strategies of processing information and conventions for dealing with language that are quite different from those used in every day interaction and which, as they grow more

³ This impact may also account for the common fact that some academic failures can be very successful in non-academic world (i.e., franchise business) where their competence in oral cultures is still highly capable among the majority. Typical literate cultures, however, do not solve the problem of hunger, but “Taco Bell” does.
complex, must be learned through special schooling (Cook-Gumperz & Gumperz, 1981, p. 91).

This difference between speech cultures and written cultures stood out more obviously after industrialization which made the writing in mass media serve as a social need for the majority, not for a small percentage of elite. The increase of population in learning how to write certainly includes more people who were, in earlier centuries, never expected to become literate at all.

The impact brought forth by the introduction of writing upon the individual cognitive process and upon the social practice in a society may apply to both children and adults. For children, Sulzby (1986) in her study found her subjects (24 kindergarten children) did not take writing as what their teachers did. One of her subjects “drew” his writing in nine pictures on nine separate pages. When the interviewer asked the boy to read the story to her, the boy did so. And then,

Adult: (Laughs) And left that neat prize there in the box. (Both laugh) That’s all right! Are you going to write your story to go with it, Doug?

Doug: What?

Adult: Are you going to write the story to go with it?

Doug: Yes.

When Doug came back, he put his name on the first page, “Da end” on the last page, and page numbers on each page. Doug’s behaviors showed that he understood what writing should be superficially--with the author’s name on the first page, page number on each page, the end on the last, and pictures on each page. But he ignored the “words” because these symbols were not part of his life yet. This also accounts for his first response--What?--to the interviewer’s request to “write” something conventional to go with the pictures, because Doug probably thought that he had already done it. And what he performed later was simply “editing.”

Knowing all these possible differences exist between the oral and written cultures, the best way to encourage people from oral cultures to move into a culture in which “language should be precision in usage,
decontextualization of information and careful weighing of words” (Cook-Gumperz & Gumperz, 1981, p. 108) is to let them realize the practical use or function of such [written] form of language. Cook-Gumperz & Gumperz suggested letting children experience the usefulness of the written form, for example, encouraging them to list down what they had gotten on Christmas.

While almost everyone is anxious about the decline of literacy in Western (or American) society, Lakoff (1982) argued that “loss of literacy is not the same as loss of culture” (p. 240). She treated the change from literate culture to oral culture as the need of modern society. She also doubted the common assumption that only a literate society emphasizing linear ways of reasoning and the skills of concise expressions can develop high technology, because there has not been sufficient evidence to say this. Furthermore, she added that the change from literate cultures to oral cultures is still too young to show any significance, assuming that the multitude prefer oral cultures because they try to regain what has been lost in the literate cultures—“immediacy and warmth” (p. 257).

2.4 The speech-writing issues for EFL students

The gap between spoken and written cultures is so obvious, from linguistic features to cognitive, social and cultural factors, that some researchers even take writing as a second language (Neilson, 1979; Horn-ing, 1987)4. But do these findings also apply in places where English is used or learned as a second or a foreign language? Few researchers seem to invest much time and energy in it for this specific issue. For English native speakers, the issue may be viewed as simple by assuming that all literates are from the oral culture. In other words, all people who learn to write already know how to speak English. For ESL or EFL students, the issue is definitely more complicated. Basically, the ESL or EFL students can be divided into at least four different types:

---

4 This may also account for the insignificant difference in the writing products between Chinese and English students found by Grabe and Biber.
Type A: Fluent in speech and literate in English (with a good command of speaking, reading, and writing)
Type B: Fluent in speech but not literate in English
Type C: Not fluent in speech but literate in English
Type D: Neither fluent in speech nor literate in English

Being fluent in speech is defined as being able to communicate with English native speakers without missing a punch line. People who have a chance to live in the English-speaking community may acquire fluent speech in English prior to their writing. Most of the others learn to read and write first and then try to transform written language into oral expressions. In the United States, the population of type A may be very small and that of types B and C may be large. Type D may be new immigrants. Types A, B, and C ESL students are usually the population from which the samples are drawn. Type D probably will be treated differently in a specially designed bilingual class or individual studies.

However, most EFL students in Taiwan may belong to types C or D. They are not likely to be orally fluent in English because most of them are still under the traditional teachers-preaching-on-the-platform teaching methods. They spend their first six years learning English to get ready for the entrance examinations, not for real application. Not until they go to college will they have the chance to go to the language lab to improve their oral communication skills. They do not seem to have problems with the gap between the oral and written cultures. What they are faced with is how to borrow materials from the reading to their writing. However, the EFL students in college, especially in the colleges in which the English language is emphasized, will improve oral skills to such a level that oral cultures start to interfere with written cultures. At

---

5 This is obviously a very difficult thing for an EFL student to accomplish. In the United States, for instance, type C ESL students are usually foreign students who come to study in universities in the United States. The fact that some highly educated ESL engineers, who learned English in a non-English speaking society before coming to the United States, have mastery in the “English” used in their specific fields but have a hard time chatting with American native speakers at parties adds further evidence to Cook-Gumperz and Gumperz’s hypothesis that spoken and written cultures differ. In other words, the spoken cul-
this moment, the EFL students are likely to use a great deal of oral expressions in their writing. Since they lack a native speaker’s intuition to distinguish spoken English from written English, they have little idea how written language should be treated differently from oral language. For instance, they probably do not know that written language is more structurally complex and elaborate than speech, indicated by features such as longer sentences or T-units and a greater use of subordination. Written language may be more explicit than speech, in that it has complete idea units with all assumptions and logical relations encoded in the text. Besides, writing is more decontextualized, or autonomous, than speech, so that it is less dependent on a shared situation or background knowledge. As a result, the writing produced by the EFL students in Taiwan may be a mixture of oral and literate cultures. And it is time that English instructors let these students know the differences between oral and literate cultures (as are discussed in the previous chapter).

Since learning literate cultures prior to spoken cultures is not uncommon in non-English speaking countries, in Taiwan for example, what is discussed above is no less than a cruel fact indicating that not only there exist barriers to account for students’ failure in acquiring speech competencies, but there also exist interactions between strategies of Chinese writing and English writing and the problems between oral cultures and literate cultures of the English language. It is therefore essential for composition instructors to ask their EFL or ESL students to familiarize themselves with the differences between oral and literate cultures in the English society.

On the other hand, what is mentioned above may also be a serious issue among the newly-arriving immigrants who are already literate in non-English speaking areas. By now, educators or researchers on ESL
programs have been focusing much more on the relationship between reading and writing. For instance, Krashen (1984) claims that the development of writing ability and of second language proficiency occurs in the same way: via comprehensible input, which implies large amounts of self-motivated reading for interest and/or pleasure. “It is reading that gives the writer the ‘feel’ for the look and texture of reader-based prose” (p.20). This reading/writing relationship will be discussed in the next chapter.
Although some researchers find that children can create symbols to represent language before they can read (Chomsky, 1975; Read, 1981), it is generally believed that being able to read is a prerequisite of the acquisition of writing. However, different studies emphasizing different aspects have different results. Some show a correlational relationship between reading and writing, others indicate directional relations: either writing affects reading or reading affects writing, and still others propose that reading and writing are in effect two sides of the same coin. Stotsky (1983) reviewed all the studies he could find concerning the relationship between reading and writing (earlier than 1983) and offered the following conclusions on the correlational studies:

1. There are correlations between reading achievement and writing ability. Better writers tend to be better readers. (See Table 3.1)

2. There are correlations between writing quality and reading experience as reported through questionnaires. Better writers read more than poor writers. (See Table 3.2)

3. There seem to be correlations between reading ability and measures of syntactic complexity in writing. Better readers tend to produce more syntactically mature writing than poor readers. (See Tables 3.3)
<table>
<thead>
<tr>
<th>Study</th>
<th>Subject</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Loban (1963; 1966)</td>
<td>Elementary students</td>
<td>High correlation between reading scores &amp; writing quality</td>
</tr>
<tr>
<td>Fishco (1966)</td>
<td>Seventh graders</td>
<td>Significant ( r ) between ratings creative writing samples &amp; reading comprehension scores</td>
</tr>
<tr>
<td>Maloney (1967)</td>
<td>Ninth graders</td>
<td>Superior writers scored significantly higher than poor writers in tests of reading comprehension &amp; vocabulary</td>
</tr>
<tr>
<td>Calhoun (1971)</td>
<td>College freshmen</td>
<td>Significant relationship between writing quality &amp; reading achievement</td>
</tr>
<tr>
<td>Campbell (1976)</td>
<td>College freshmen</td>
<td>High relationships between writing ability, a criterion referenced reading skills test, and a standardized reading test</td>
</tr>
<tr>
<td>Grobe &amp; Grobe (1977)</td>
<td>College freshmen</td>
<td>Good writers had significantly higher reading scores than average writers</td>
</tr>
<tr>
<td>Taylor (1981)</td>
<td>College students</td>
<td>Significant relationship between composition class final grade &amp; reading comprehension</td>
</tr>
</tbody>
</table>

Table 3.1  Studies correlating reading achievement with writing ability
### Table 3.2 Studies correlating experiences with writing quality

<table>
<thead>
<tr>
<th>Study</th>
<th>Subjects</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Donelson (1967)</td>
<td>Tenth graders</td>
<td>Effective writer was also apt to be a reader</td>
</tr>
<tr>
<td>Woodward &amp; Philips (1967)</td>
<td>College freshmen</td>
<td>Poor writers tend to have less reading experience than good writers</td>
</tr>
<tr>
<td>Maloney (1967)</td>
<td>Ninth graders</td>
<td>Superior writers read more frequently than poor writers &amp; tend to be female</td>
</tr>
<tr>
<td>Thomas (1976)</td>
<td>College freshmen</td>
<td>Positive correlation between writing achievement and amount and diversity of reading</td>
</tr>
<tr>
<td>Felland (1981)</td>
<td>High school students</td>
<td>Superior writers read more books than average writers</td>
</tr>
</tbody>
</table>

### Table 3.3 Studies correlating reading ability & syntactic complexity

<table>
<thead>
<tr>
<th>Studies</th>
<th>Subjects</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Zeman (1969)</td>
<td>2nd &amp; 3rd graders</td>
<td>Use of compound and complex sentences increased as level of reading comprehension increased</td>
</tr>
<tr>
<td>Heil (1976)</td>
<td>1st, 2nd &amp; 3rd graders</td>
<td>T-unit length correlated significantly with reading comprehension</td>
</tr>
<tr>
<td>Thomas (1976)</td>
<td>College students</td>
<td>Significant correlation between sentence maturity and reading achievement</td>
</tr>
<tr>
<td>Johnson (1981)</td>
<td>3rd, 4th, 5th, &amp; 6th graders</td>
<td>Words per T-unit, number of words per clause, &amp; number of clauses per T-unit all correlated significantly with one or more of the reading measures</td>
</tr>
</tbody>
</table>
These conclusions provide evidences that reading and writing are closely related. But since correlation does not provide causality, we still do not know whether reading affects writing or vice versa or reading and writing are in fact two interrelated processes. In this chapter, I will not focus on the correlation studies because it would be more valuable to know the direction of causality so as to help teachers decide whether they should teach reading prior to writing or otherwise. Most recently, Eisterhold (1990) did a survey of the relationship between reading and writing, dividing the most recent studies into three types, each type under one of the following hypotheses: (1) The directional hypothesis, (2) the nondirectional hypothesis, and (3) the bidirectional hypothesis. The following three sections will be developed on the basis of these three hypotheses.

3.1 The directional hypothesis

The first hypothesis is “directional” because “reading and writing share structural components such that the structure of whatever is acquired in one modality can then be applied in the other” (Eisterhold, 1990, p. 89). Under this hypothesis, there are two basic directions: (1) studies that examine the influence of reading upon writing, and (2) studies that examine the influence of writing upon reading. For the first direction, Stotsky (1983) categorized them into two subcategories: “those attempting to improve only reading, with effects on writing; and those attempting to improve writing through reading instruction, the use of literary models, or additional reading experiences” (p. 634). As a result, Stotsky (1983) concluded that “no studies can be found that seem to fall in the first subcategory” (p. 634). For example, Schneider (1971) found no overall significant differences between experimental and control groups on measures of reading and writing after teaching composition skills through reading instruction in a fifteen-week course for remedial students at the junior college level. Matt (1977) used an experimental discourse model for reading instruction to improve the writing skill of 40 high school seniors in a nine-week course. The experimental group had
no writing practice during the nine-week period. The results showed that although the experimental group made significant gains in reading on an experimenter-constructed test based on the discourse model used during the instruction, they made no significant gains in writing. Based on the studies Stotsky reviewed, the direction of using reading instruction to improve writing (in general, not in paraphrasing what is read) does not seem to be effective. Stotsky, therefore, concluded that “reading instruction is...probably best undertaken for the purpose for which it was designed” (p. 636).

Contrary to what Stotsky found in his review, Taylor and Beach’s (1984) study indicated that instruction in writing influences neither writing nor reading but that instruction in reading influences both reading and writing. Belanger (1987) also found that instruction in reading can be effective in improving writing only when it focuses on a common element. He also pointed out that transferring from general reading improvement skills to written composition should be guided, i.e., having some writing instruction or demonstration involved. The transfer will not occur automatically.

For those attempting to improve writing through reading instruction, the use of literary models, or additional reading experiences, Stotsky (1983) found some studies providing encouraging findings. Mills (1974) conducted a four-year longitudinal study, finding that fourth grade children who read or listened to and then discussed children’s literature as a “springboard” to writing score significantly higher in their free writing than a control group that did not use children’s literature in this manner. Writing is measured by a composition rating scale as well as by tests of capitalization, punctuation, and total language from the Iowa Test of Basic Skills. Andreach (1975) used another approach of developing organizational skills through the use of literary models in a semester-long course in Grade 10. He found that experimental students write significantly better organized compositions than control students. The results of these two studies suggest that reading may help students improve their writing in some ways.
Theoretically, Goodman & Goodman (1983) propose that children use in writing what they observe in reading. Smith (1983) claims that a person can become a good writer only if he reads like a writer. He argues that “writing requires an enormous fund of specialized knowledge which cannot be acquired from lectures, textbooks, drill, trial and error, or even from the exercise of writing itself” (p. 558). In other words, only through reading can a writer learn all the intangibles he should know. The knowledge of writing can only be acquired from a particular kind of writing, e.g., if one wants to write for a newspaper, he should read the newspaper as a “writer” for that specific purpose. He criticizes formal conventional writing instruction that usually provides definitions like:

- a sentence is “a complete thought”...a word is “a unit of meaning” or...a paragraph is organized around a single topic.
- How would anyone recognize a unit of meaning, a complete thought, or a topic in isolation? (p. 559)

These definitions are only definitions, not rules of application; they do not provide any constructive use for beginner writers.

What is more important to the beginner writer is to learn the knowledge of writing by reading, as a potential writer, through the existing texts of interest. By doing this, a writer can acquire (1) a vast number of conventions of writing used in a specific field, including spellings, punctuation, and other “rules” of grammar, and (2) the sense of “cohesion” which links sentences to each other and to the non-language context (Halliday and Hasan, 1976). How could any of this be reduced to prescriptions, formulas, or drills? (Spellings and grammar rules do not help people write at this point) (Smith, 1983).

As for the second direction: examining the influence of writing upon reading, Stotsky (1983) found a number of studies that recommended writing activities which were believed could effectively improve reading comprehension and retention of information. For example, Taylor and Berkowitz (1980) found that grade 6 students who wrote a one-sentence summary after a passage from a social studies textbook did better on measures of comprehension and memory than
students who used a study guide, students who answered questions after reading the passage, or students who simply read the passage. Glover et al. (1981) also found that college level students instructed to paraphrase or write “logical extensions” of an essay they were asked to read recalled significantly more ideas from the essay than students instructed to write only key words or nothing at all while reading the essay. Kulhavy et al. (1975) found that high school students who took notes on a textbook passage perform significantly better than those who underlined or merely read it. Collins (1979) and Walker-Lewis’s (1981) found that expressive writing practice combined with reading in a semester-long course for college freshmen improved their reading comprehension significantly more than did reading instruction alone for a control group. Smith (1983) theoretically proposes that reading like a writer needs not interfere with reading comprehension, in fact “it will promote comprehension because it is based upon prediction” (p.563).

These findings to suggest that only through writing down what is read can the act of reading help students enhance comprehending, memorizing, or recalling the read materials. It would be insufficient to say that “writing itself” promotes reading comprehension because there may be other factors involved. For example, in the study conducted by Kulhavy et al. (1975), the high school students who took notes comprehended better than those who did not; and those who wrote paraphrases recalled more than those who simply took notes. I will assume that the results of this study probably only prove one thing: a person who spends more time on the texts may perform better on achievement tests. As we mention in chapter 1, writing processes consist of planning, writing, editing, and evaluating and each of these processes requires the “writer” to “review” the material at least once. In other words, those who write paraphrases may involve more processes than those who simply take notes, and thus comprehend or recall better the materials that they have read. My assumption may not hold without empirical supports. It, however, suggests that further more sophisticated studies in this field are still needed before drawing a conclusion that writing will directly promote reading comprehension.
3.2 The nondirectional hypothesis

The nondirectional reading-writing relationship assumes that reading and writing “derive from a single underlying proficiency, the common link being that of the cognitive process of constructing meaning” (Eisterhold, 1990, p.90). Squire (1983), postulating that composing and comprehending are two sides of the same basic process, outlines six different ways to demonstrate his claim. These six different ways follow:

1. The skill in processing language is basic to both reading and writing. He says that “composing is critical to the thought process because it is a process which actively engages the learner in constructing meaning, in developing ideas, in relating ideas, and in expressing ideas. Comprehending is critical because it requires the learner to reconstruct the structure and meaning of ideas expressed by another writer. To process an idea that one is reading about requires competence in regenerating the idea, competence in learning how to write the ideas of another. Thus both comprehending and composing seem basic reflections of the same cognitive process” (p. 582). This is also true for children because both reading and writing are thinking processes (Nessel, Jones, & Dixon, 1989) and for most children language is the major vehicle through which thinking occurs (Britton, 1970). It is “through language that children learn to label ideas, to classify, to relate the new to the known, to construct ideas or compose, to reconstruct or comprehend” (Squire, 1983, p. 582).

2. Classroom strategies for regenerating ideas are essential to teaching comprehending. This implies that both composing and com-

---

6 I disagree with Squire’s view on this specific point because an artist or a musician probably classifies different types of music (without labeling them verbally) based on something other than language. In my personal experience, when I paint, I classify colors, not with labels “yellow or red,” but with their physical properties analogous to real things that are retrievable in my memory, e.g., the color of the leaf is like the color of the stone in my yard which I see many times a day.
prehending demand similar processes. Summarizing, rephrasing, and regenerating a text will certainly promote comprehending. He cited Lanier’s (1982) conclusion that “thinking about the process of comprehending, that is, consciously considering the reconstructions that one composes, can enhance the basic process itself” (p. 582).  

3. Because language learning and language processing involve cognitive processes basic to every discipline, application to the discipline is critical if children are to learn to think in the discipline. For this statement, I would confine “every discipline” to one specific type of text. Smith (1983) claims that one should read whatever one expects to write. Over several years, practice in writing in that subject area will contribute strongly to the performance in reading and thinking in the discipline. On the other hand, Squire (1983) gave one example that well supports this statement.

A daily log in a fifth grade records observations as a child watches a polliwog growing to a frog. Such activities are almost routine science. So is writing up a science experiment—a laboratory report—in eighth grade physical science. But as the children write science, they process the ideas of science. They use the language of science. They learn to think in science. They prepare themselves to understand their reading of science. By the time they enroll in tenth grade biology they are ready to comprehend with reasonable understanding textbooks far more difficult in science than in any other subject field...Contrast this condition with what presently obtains in social studies.... As Applebee (1981) has shown,

---

7 I do not agree with what he says that “Activity without language does not become experience” (p.582). It may be true that language may enhance the impression of an “experience,” but it does not prove that an “experience” without language is not an experience. Yet, we all might have some experiences that are beyond description and they are always fresh (e.g., the experience of first love). In general, this second point is somewhat less convincing.
much of the writing in social studies is note taking—a not
unimportant but severely limited kind of processing.
With insufficient experience in composing social stud-
ies, children do not learn the technical vocabulary of so-
cial studies, do not learn how language is used in the
discipline, do not learn how to construct and compose
and to reconstruct or comprehend (p. 584).
Squire (1983), furthermore, argues that the reason why many stu-
dents cannot read the word problems in math tests is that they have
never written any. They have not had the opportunity to learn how
language is used in mathematics.
4. Children and young people require instructional experience in all
important modes of rhetoric if they are to comprehend and com-
pose using these varied forms and functions. Squire (1983) used
Stein’s (1981, 1982) studies with story grammar to support his
statement that children who know how to read before they enter
schools have internalized the basic story structure and have pre-
pared themselves to understand the basic patterns (See also Beck,
1982). Writing, like reading, requires attention to the various
modes and functions of language. The example given is about
New York and Texas (1978) testing the proficiency of thirteen-
year-olds and seventeen-year-olds in writing business letters, re-
ports, and persuasive pieces of prose. Results show that all stu-
dents have no problem writing business letters because they have
been taught how to write business letters. But in writing persua-
sive prose, many of the most talented students fail, owing to “they
had had next to no instruction or instructional practice in writing
or in reading persuasive writing” (p. 585).
5. Instruction in comprehending and composing must concentrate on
coping with the total process of constructing and reconstructing
ideas. One does not learn to read or write by practicing isolated
reading skills. The instruction should be holistic and include the
following processes (p. 586):
   Before Writing: Securing ideas
Organizing ideas
Determining point of view
Considering audience

Before reading: Preparing to comprehend
Relating to prior experience
Establishing purpose
Looking for the author’s stance

During reading or Writing: Composing or comprehending
Actively engaged emotionally and intellectually

After Writing: Evaluating
Editing and revising
Applying outside standards of correction

After Reading: Evaluating
Studying parts in relation to whole
Analyzing how effects are achieved
Applying independent judgments (Preferences, ethics, aesthetics)

Squire (1964) demonstrated in his study of the responses of adolescents to short stories, that the complete processes of comprehending, like the complete process of composing, proceeds through predictable stages (Squire did not specify whether these processes in parallel are linear or recursive).

6. A critical factor in shaping the quality of both composing and comprehending is the prior knowledge the pupil brings to reading and writing. Squire pointed out three aspects of prior knowledge:
   a) Knowledge about rhetorical structures--the story grammars or patterns of expository prose discussed earlier in this paper.
b) The accumulation of knowledge and experience prerequisite to understanding or writing about an important concept or idea.

c) The strategies children must acquire to unlock the world knowledge that they have accumulated—to learn how to ask themselves those questions before writing or before reading that seem most likely to enhance composing and comprehending.

Kucer (1985), exploring a number of theoretical issues related to the parallel role of internal revision in reading and writing, also agreed that “when the individual has much of the necessary background knowledge and it is organized in an appropriate manner, reading or writing is fairly predictable and more easily managed” (p. 6). Hiebert et al. (1983) investigated the relationship between the recognition and production of different text structures (description, sequence, enumeration, and comparison and contrast) by college students, and found that students understood better if they were able to recognize the text structure. They therefore concluded that similar knowledge bases about text structures can be hypothesized to underlie reading and writing.

To conclude, as Shanahan (1984) suggested, “if reading and writing involve analogous cognitive structures and processes, it is possible that instruction in one would lead to increased ability in the other” (p. 467). But, as is commonly known, experiments concerning invisible mental processes demand extremely careful designs for tests and measurement and usually need large samples to make the study significant. Different designs of studies may very possibly lead to different results.

3.3 The bidirectional hypothesis

Eisterhold (1990) viewed this bidirectional hypothesis as the most complex model, “which includes the claim that reading and writing are interactive, [and the] claim that they are interdependent as well” (p. 92). Shanahan (1984) saw the reading-writing relationship as developmental and suggested that “what is learned at one stage of development can be
qualitatively different from what is learned at another stage of development” (p. 467). Goodman & Goodman (1983) outlined seven characterizations, showing that the interrelationships of reading and writing can be much more complicated than being regarded as directional or nondirectional.

1. The relationship between reading and writing is less observable (or harder to identify) than that between listening development and speaking development (I personally will leave this point open to more research).

2. Both reading and writing develop in relation to their specific functions and use. Again there is greater parity for functions and needs of listening and speaking than for reading and writing.

3. Most people need to read a lot more often in their daily lives than they need to write. That simply means they get a lot less practice in writing than reading.

4. Readers certainly must build a sense of the forms, conventions, styles, and cultural constraints of written texts as they become more proficient and flexible readers. But there is no assurance that this will carry over into writing unless they are motivated to produce themselves, as writers, similar types of texts.

5. Readers have some way of judging their effectiveness immediately. They know whether they are making sense of what they are reading. Writers must depend on feedback and response from potential readers which is often quite delayed. They may of course be their own readers, in fact it’s impossible to write without reading.

6. Readers need not write during reading. But writers must read and reread during writing, particularly as texts get longer and their purposes get more complex, because the writers want to make sure that all the conventional requirements are satisfied and their potential readers can receive the communicative purposes.

7. Reading and writing do have an impact on each other, but the relationships are not simple and isomorphic. The impact on development must be seen as involving the function of reading or writing
and the specific process in which reading and writing are used to perform those functions.

In short, Goodman and Goodman (1983) agree that people not only learn to read by reading and write by writing but they also learn to read by writing and write by reading. Furthermore, Shanahan and Lo-max (1986) found that in their data, the second graders showed more of the interactive relationship between reading and writing than the fifth graders. They believed that subjects generally received more reading than writing instruction in school, and that the influence of writing on reading begins to wane in upper grades, so that, as children grow in literacy, the reading-to-writing influential direction becomes superior to the writing-to-reading influential direction. Since they assumed that the best model of reading-writing relationship is the interactive model, they suggested school teachers balance the emphases of reading and writing instruction (and practice) so as to enable students to make knowledge of reading and knowledge of writing interact.

On the other hand, for young children (aged from 3 to 6), Read (1981) does not believe that reading and writing are interactive, or reading has a necessary impact on writing. For young children, “there can be a stage at which children have two distinct systems, one for reading standard spelling and another for writing in their own invented orthography.” And at this stage, “children created, rather than retrieved, the spellings of familiar words” (p.107). Based on Frith’s (1978) finding that poor spellers can also be good readers, Read further concludes that “children’s ability to read and to write standard spelling differed” (p.113). In addition, he finds this phenomenon has no difference in different dialect regions (From Virginia to Boston). However, he generally agrees that as children grow and acquire more literacy, reading and writing start to play interactive roles.

3.4 What do beginner writers do in learning writing?

---

8 My intuitive comments for this are that to recognize a symbol is often easier than to recall one.
This section will not emphasize a writer’s mental/cognitive processes in writing (which was discussed in Chapter 1). What I will focus on is: in the course of writing, how does a beginner writer deal with writing under the presumable impact from reading, from instruction (both of reading and writing), from pragmatic needs, and from linguistic and rhetorical requirements? How do children learn to write? Do they have anxiety in learning writing?

Smith (1983), based on Miller’s (1977) estimation of how infants add words to their vocabulary and his observation that children master speech without formal instruction, with very little evident trial and error, and with no deliberate diagnostic or remedial intervention at all, proposes that writing should be learned “without awareness of the learning that is taking place” (p. 561). Goodman and Goodman (1983) also find that children start learning writing by playing at what grown-ups do. They play at creating the aspects of their world which attract them. By doing this, one function of the written language—labeling—seems to be internalized unwittingly. Nessel, Jones, and Dixon (1989) propose that “many pre-school children play at writing as they play in general. Scribbles, rudimentary drawing, and other pseudowriting satisfy the young child’s interest in exploring the uses of pencils and paper as toys” (p. 177). “Children do not necessarily have goals, as adults assume, to convey anything,” says Read (1981). “They are far less interested in the product than we are, and more involved in the process of writing something down. A child will show it to you. Then, however, his interest in it is ended; it is rare that he will make any effort to keep it and even rarer that he will read it” (p. 108).

These observations suggest that children do learn, or acquire, writing at the very beginning without being overtly or elaborately instructed. Smith’s (1983) claim may be correct that children do learn

9 Miller (1977) estimated that infants must add words to their vocabulary at an average rate of one for every hour they are awake, a total of several thousands a year.
writing as they acquire speech. But how? This problem will be discussed in the following section.

### 3.5 What should teachers do in writing instruction?

As children grow older (in elementary school) and become more aware of the communicative function of written language, “their writing ‘play’ becomes more focused and more purposeful” (Nessel, Jones, and Dixon, 1989). They are required to spell correctly, to be aware of the conventional requirements of forms (i.e., capitalization, punctuation). Children start getting the pressure (either from the teacher or from self-expectations) of making errors. Writing at this stage is no longer “free creation.” For some students, writing even becomes a communicative form learned in school to satisfy teachers’ requirements rather than a form to serve as a useful means of communication in real life (one of my friends used to say that telephones and tape-recorders can now take the place of all the jobs needed to be done by pen and paper before). It is also undeniable that in real life adults do much more reading and talking than writing (Goodman & Goodman, 1983). Children in natural environments do not seem to have enough chance to acquire writing as they acquire speech. Smith (1983), noticing that children in general acquire speech without much effort, assumes that learning writing should also take the route of speech acquisition. He therefore illustrates four features in speech acquisition that may help us improve writing instruction.

He points out that (a) there is an exquisite selectivity in acquiring speech. “They [Children] do not learn to talk like everyone they hear speaking, even those they may hear most. They learn the language of the groups to which they belong (or expect to belong) and resist the language of the groups that they reject or from which they are rejected. They learn, I want to say, from the clubs to which they belong.” (561) Sometimes children do not learn about spoken language from everything they hear spoken because they do not understand or because they are not interested. Sometimes simply because they listen like a “lis-
tener,” not like a “talker,” they have no desire or expectation that they should come away talking like the speaker (p. 562).

(b) Learning with engagement. “It is not learning that takes place as a consequence of someone else doing something, but rather learning that occurs concurrently with the original act--provided it is our act too. The other person’s behavior is our own learning trial. We learn when the other person does something on our behalf, something which we would like to do, which we take for granted” (p. 561).

(c) Language is embedded in the situation. Adults have neither the time nor the expertise to teach spoken language to children. Instead, they act as a source of information for children and as unwitting collaborators (Smith, 1983). They are overheard as they talk to each other, and thereby show children why and how speech can be used. They demonstrate language being used for purposes which children would expect to accomplish themselves. Often the explanation of the language is embedded in the situation in which it is used--someone says “Pass the salt” and someone else passes it. Children do not need to be personally involved to learn to say what they would like to be able to say. They learn when others do the talking for a purpose they want or expect to share. The adults do not expect children to be experts in advance, nor do they anticipate failure. There are no admission requirements.

(d) Learners are “spontaneous apprentices.” “Children do not regard the language they learn from as something remote, an attribute of others, but rather something they themselves would want and expect to do. They become ‘spontaneous apprentices’ as Miller (1977) puts it, engaging in the enterprises of the adults or peers who are their unsuspecting surrogates for the trial and error of learning (and who since they are experienced tend to have a variety of trials and very few errors, a most efficient form of learning)” (p. 562).

In reviewing recent literature about writing instruction, educators do move forward to active involvement in writing by encouraging children to begin with private writing, such as journals, learning logs, and read-talk-write (Botel, 1977; Nessel, Jones, and Dixon, 1989). Goodman and Goodman (1983) also suggest using mundane functions of
writing like shopping lists in the classroom, assuming that they can be helpful because all members of a literate community participate widely in the pragmatic functions of reading and writing which Michael Halliday (1975) has called the “goods and services” function of language.

Private writing may be the best way to build up a beginner writers’ confidence because the intuitive products will not be evaluated by anyone in anyway if the writer does not want it. But as life goes on, children cannot always do private writing. The time when they need to produce public writing will eventually come. At this moment, the best motivation that drives children to learn public writing, such as essays, stories, letters, scripts, poems, and instructions, may be understanding the purposes and functions of these forms of writing. Since public writing has intended readers, the writing processes should always include the sense of acceptable writing conventions, readers’ expectations, and the completeness of the communicative purpose. Smith (1983) again points out that all these complexities of writing can not be learned from writing instruction. Children or adults can learn these only by “reading as a writer.”

In addition, Smith argues that “reading as a writer” can not be taught either. Children or adults acquire the technique naturally as long as the learner has a strong intention to learn. There is no way of helping a person to see himself as a writer if he himself is not interested. He suggests starting with being read to. He says, “teachers should not be afraid that a child who is read to will become dependent or lazy. Children able to read something they want to read will not have the patience to wait for someone else to read for them, any more than they will wait for someone to say something on their behalf if they can say it for themselves” (p.564). As children are able to read themselves, they will learn how to read like a writer gradually. For example, when we read something, it sometimes happen that we pause to reread a passage, not because we do not understand the passage, but because something in the passage is particularly well put, and because we respond to the craftsman’s touch (Smith, 1983). Similar experience will occur to children if
they have the strong intention to use public writing as one commu-
nicative means.

With respect to instruction, Smith (1983) suggests teachers doing
two things: (a) Teachers must show the advantages of reading like a
writer. And (b) teachers must ensure that children have access to read-
ing materials that are relevant to the kinds of writing they are interested
in; teachers must recruit the authors who will become the unwitting col-
laborators.

In addition to what is discussed, beginner writers should also be
aware of how writing serves as a communicative form. As a basic prior
knowledge of public writing, Goodman and Goodman (1983) provided
the general features of writing (derived from their list-making example)
that a writer should be aware of any time he writes.

1. Its purpose and function are usually related to communication over
time and/or distance.
2. It occurs in a specific context in which the purpose or function mo-
tivates specific writing events.
3. It has an intended audience of one or more readers.
4. Its content is relevant to the purpose and audience.
5. Its structure and format are suited to the purpose and function
which will be familiar to or expected by the reader (See also Frase,
6. It must be so composed that it serves both the writer and the
reader.

If children or adults are used to reading like a writer, it would not be too
hard for them to realize that all these qualities are what they want to at-
tain from written texts, and they also will realize that these are the basic
requirements any writer should bear in mind whenever he writes.

As for the feedback from the instructor, Smith (1983) complains
that usually insufficient feedback from either school or family is re-
ceived because not enough writing is done. In recent literature, the
feedback is given in two types: (a) correcting or modifying children’s
writing and (b) providing alternatives or interactions during children’s
writing. The former may help to polish writing skills, but cannot ac-
count for their acquisition in the first place (Smith, 1983). The latter will, on the other hand, stimulate children to elaborate on the topic of interest. The interaction may be in the very beginning of the writing by offering optional ideas of what children might want to write about (Nes-

sell, Jones, and Dixon, 1989). It may occur during writing by providing alternatives (by instructors) and letting the child decide which ones to use (Smith, 1983). It may occur in the end when the instructor thinks he should help children to complete the composition (Smith, 1983). The feedback may be expected even by very young writers in writing their journals. For example, Goodman and Goodman (1983) found one second grader wrote the following statement:

*I won’t write no mor til yuo write me back* (p.596)

Generally speaking, the best feedback from school or family may be direct interactions. The interaction can also extend to the structure of the text, how well the background knowledge (personal experiences) are expressed, and how smoothly the reading materials are integrated. Smith (1983) argues that teachers have to learn how to read like a writer before they can help students because teachers need to know how to lead their students to the threshold where the students can start learning how to read like a writer.

3.6 The reading-writing relationship in L2 acquisition

Most studies about reading-writing relations do not distinguish their subjects (who speak L1 from those who speak L2). Plenty of studies about literacy acquisition concern the diversity of languages and cultures, but few of them emphasize writing. Still, there are some studies centering on how reading affects writing, e.g., Campbell’s (1990) study that deals specifically with how college students synthesize reading materials into their academic papers (it will be discussed later).

Before reviewing the studies focusing on the reading-writing relationship in L2 acquisition, I would, first of all, mention a theory that is common in literacy acquisition and that, hopefully, can reflect a picture of the possible relationship between reading and writing in L2. Among
the theories of literacy acquisition, Cummins’ (1981) interdependence hypothesis is probably the strongest case for the transfer of language skills from one language to another. Cummins states:

To the extent that instruction in Lx is effective in promoting proficiency in Lx, transfer of this proficiency to Ly will occur provided there is adequate exposure to Ly (either in school or environment) and adequate motivation to learn Ly (p.29).

Cummins claims that a common underlying proficiency (CUP) across languages allows the transfer of literacy-related skills across languages. Thus, L2 literacy acquisition strategies may be affected by literacy capabilities in L1.

Here Cummins, like most other researchers, assumes that the relationship between literacy-related skills (i.e., reading and writing) in L1 and L2 are similar. For instance, in providing evidences for his hypothesis, Cummins reviews the results of research on bilingual programs (e.g., Rosier and Farella, 1976; Leyba, 1978) to conclude that “minority children’s L1 can be promoted in school at no cost to the development of proficiency in the majority language” (p. 28). On the other hand, he investigates the relation between the arrival age and L2 language proficiency, finding that “older Japanese immigrant students, whose L1 literacy skills were better developed, acquired English proficiency significantly faster than younger immigrant students” (Cummins et al., 1981, p. 31). At this point, reading and writing are not distinguished and their relationship here is not of interest. But this CUP theory more or less accounts for the fact that the research trend concerning reading-writing relationship in L2 is not much different from that in L1.

Most researchers dealing with the reading-writing relationship concerning L2 (ESL or EFL) students accept the notion that reading instruction promotes both reading comprehending and writing abilities (Spack, 1985; Tudor & Hafiz, 1989; Horowitz, 1990; Costello, 1990). They do not seem to be interested in distinguishing the directional models (the reading-to-writing model or the writing-to-reading model) from the interactive model. What they want to ascertain is what content (e.g.,
fiction or nonfiction) provides better information to improve ESL or EFL students’ language proficiency (including reading and writing). For instance, Costello (1990) proposes that teaching short stories and autobiographical essays is the most productive way to promote writing ability not only because these two types of literary works help students acquire their writing conventions, but also because they provide knowledge of cultural background. Horowitz (1990) also agrees that the use of literature (fiction) in ESL classes has its own value, but he contends that it is necessary to specify the goal of a composition class before selecting the reading texts because different texts function differently in fitting in with students’ needs. For instance, reading nonfictional texts (e.g., journal articles) fits right in with the need of a student who intends to write such articles. And this is the knowledge of writing that fictional texts can never provide. Spack (1988) also argues that there is no evidence that the skills learned in writing personal essays provide students with the skills they need to produce the academic writing required in academic courses.

Campbell (1990), furthermore, substantiated the influence of reading upon writing by comparing the processes by which less proficient ESL writers, more proficient ESL writers, and English L1 writers incorporated background information (from reading) into written text (academic papers). She found that all groups used the background information most in the final paragraphs. However, ESL students used significantly more background information in the first paragraph than native writers. In addition, ESL writers were more likely to acknowledge the source of material than native writers. As expected, native writers received higher holistic scores, and ESL students’ use of background knowledge text did not appear as integrated into compositions because the surrounding language was less academic in style and tone.

Campbell’s study may not reflect the true issue because she does not group her subjects with writing performance, but on the basis of SAT scores (She arranged the groups on the basis of their SAT scores (241=low, 291=high, 467=native writer)) which may not necessarily represent the students’ true writing ability. In other words, what Camp-
bell found may not be the problems of L2, but simply the problems of less-skilled writers in general. In addition, there may be individual differences in the ability to integrate the reading materials. For instance, Block (1986) describes two kinds of ESL readers: (1) integrators, who successfully focus on meaning and integrate new information, and (2) nonintegrators, who rely on personal experience to make meaning of a text and who make few attempts to connect or relate the information they read to their previous knowledge. Block adds that this inability to integrate often occurs in their writing as well. Campbell used only ten subjects in each group. If she used more subjects and if Block is right, then the results may lead in a different direction.

With respect to writing instruction, some researchers (e.g., Young, 1990) suggest using dialogue journals in both reading and writing instruction so as to provide a purposeful activity in which students communicate their thoughts and feelings. Some researchers (e.g., Watkins-Goffman, 1989), proposing that reading and writing share the same underlined process, suggest providing ESL writers with concrete aids to guide them to write in a mode in which they have had little or no experience. One of such concrete aids is to teach ESL students that revising is a process of reading to formulate questions about the text. Watkins-Goffman, hence, illustrate a series of checklists to help students read both readings and their own writings critically. The checklists include those that target surface errors and those concerning the organization and other features of written conventions. They also vary depending on the mode of rhetorical discourses being assigned: narrative, expository, or descriptive. The goal is to teach the students to look at their own drafts from the objective viewpoint of a reader, keeping in mind that the instructor is the final audience.

In general, I agree that as reading is to writing in L1, so is reading to writing in L2 if one realizes how writing is a unique form of communication. But in real life, I would propose that the reading-writing relationship may differ in subtle ways between English natives, ESL students who live in the United States, and EFL students who do not live in the United States. My assumption is that reading would be much more
valued for EFL students to learn writing than for ESL students because, for EFL students who usually do not have the natural environment to acquire English naturally, reading, therefore, is the only source that provides “knowledge of writing and language proficiency in general”; whereas ESL students can more or less acquire their “language” through mass media other than print. English natives, however, unless they are aware that writing needs to be overtly taught (through reading or writing instruction), may not value reading as highly as EFL or ESL students do.

10 MacDonald & MacDonald, 1991, find that a carefully-planned course in ESL based on film can be a highly effective and enjoyable way to introduce students to both language (including writing) and culture.
To write well in English is no easy task not only because there exist many rules required specifically for writing that are different from the familiar linguistic features in speech, but also because formal writing is introduced to children much later than speech. Writing, therefore, is given the second priority as a means of communication so that it naturally obtains less practice both in learning and in use. Without sufficient understanding and practice of the grammatical and rhetorical conventions that writing calls for, one may make errors or deviations when one has to use the written form of communication. This chapter will focus on errors or deviations that the writer makes in writing. Like the other chapters, I will start with error analysis in English (as L1), and then compare the theories and results between L1 and L2. In addition, I agree with Shaughnessy (1977) that there are at least two reasons why the phenomenon of errors should be paid attention to seriously: (1) Error is a barrier that keeps a writer not only from writing “something” in formal English but from having “something” to write about; and (2) without errors, the reader can expect “his investment to pay off in intellectual or emotional enrichment” (p. 12).
4.1 Error analysis concerning L1 writing

Young children before acquiring literacy or in the early grades make some errors\(^{11}\) in speech, but adults usually do not take these errors seriously. Most of these children do not need formal instruction (like attending classes) to correct their language. They only “need more experience in interacting with well-spoken models—parents, teachers, peers—and in listening to a variety of books and other sources of good language so as to acquire good habits” (Nessel, Jones, & Dixon 1989, p. 300). In writing, children in general start with nonstandard symbols, and adults usually do not worry about children’s deviations because they are aware that children do this more for self-satisfaction than for communication (Read, 1981). If children still do not write “well” enough when they are in high school or later, parents and teachers start looking for the reasons because at that age, they are expected to have learned enough to be able to “write” according to conventions. The reasons found are plenty. Some believe that poor writers do not learn how to write because they do not learn how to read (Smith, 1983). Most researchers believe that poor writers do not learn to write well because they are not fully aware of the difference between speech and writing (Shaughnessy, 1977; Danielewics and Chafe, 1985; Horning, 1987). Some also believe

\(^{11}\) Corder (1974) calls attention to the systematic quality of learner’s errors and their difference from mistakes. The distinction is summarized by Sridhar (1980), who maintains that “mistakes are deviations due to performance factors such as memory limitations...spelling pronunciations, fatigue, emotional strain, etc. They are typically random and are readily corrected by the learner when his attention is drawn to them. Errors, on the other hand, are systematic, consistent deviances characteristic of the learner’s linguistic system at a given stage of learning” (p. 105). Horning (1987) agrees that “mistakes are relatively chancy phenomena that are quite simple to deal with and that ought to yield to a few lessons in careful proofreading. Errors, more resistant to instruction,” must be viewed positively, as good signs of systematic attempts to master the redundancy of the system” (p. 43).

Shaughnessy (1977) does not overtly differentiate mistakes from errors, but she does say that some errors are “accidental” (p. 47) so that they are not systematic. She then adds that these accidental errors can usually be corrected by the writers themselves after careful proofreading.
that students make errors because they have different understandings of writing conventions (Shaughnessy, 1977). Any of these reasons may be true to some, but not all, students. Among the reports of error analyses, Shaughnessy provides a most complete list of all possible kinds of errors made by college students. The following discussions will mainly follow the categories defined by Shaughnessy. These categories are: handwriting errors, punctuation errors, syntactic errors, common errors, vocabulary errors, and errors beyond the sentence.

4.1.1 Handwriting errors

Shaughnessy started with the physical limitation that causes basic writers to make errors, that is, handwriting. She found that basic writers are sometimes “hindered not only by the difficulties of articulation that confront most writers but also by their stiffness with the pen (they need more practice until the pen seems a natural extension of the hand, and the hand of the mind itself)” (p. 16). Shaughnessy also found that typing can help students solve the problem from her empirical studies and suggested that high school students should be provided with the chance of learning typing before entering college.

Levin, Riel, Rowe, and Boruta (1985), in introducing writing instruction with the computer, also found that editing with paper and pencil is difficult for elementary school students because it has a small (in some cases negative) value for improving a piece of text, since new errors are produced in the new version at about the same rate as old errors are corrected. On the other hand, editing with the computer has a much more positive value, since “inadvertent new errors are introduced at a much lower rate” (p. 163).

4.1.2 Punctuation errors
Shaughnessy has pointed out that basic writers make punctuation errors mainly because they assume that the uses of spoken language can be transferred to writing without change. Although she found that a basic writer’s punctuation repertoire consists of little more than commas, periods, and capitalization so that he very infrequently uses other marks like semicolons, question marks, etc., she maintained that a basic writer has problem with punctuation “not because he has no competence with sentences at all but because the writing down of sentences introduces new competencies that he has not been taught, including not only a knowledge of the names and functions of the various marks but also an ability to manage the structure that writers depend upon to overcome the redundancy, fragmentation, and loose sequencing that are natural in speech” (p. 27). Danielewicz and Chafe (1985) found the same conclusion in their study on college students’ writing by comparing the ways basic writers use commas and periods with analogous uses of intonation and pauses by educated speakers. They concluded that “the punctuation of writing only indirectly mimics and is often at odds with the spoken model. The deliberateness, slowness, and solitariness of writing have led writers to evolve their own literate structures, and their own rules for punctuating them. The results is that carrying over spoken prosodic habits into the punctuation of writing often leads to nonstandard, and sometimes infelicitous results” (p. 224). To substantiate the nonstandard devices of punctuation, Danielewicz and Chafe provided a very clear list with examples (I will cite only the written examples):

1. The use of rising rather than falling pitch in speaking, and of commas in writing, at the ends of grammatical sentences:

Example\(^{12}\) (a) Language has played a major role in the survival of mankind,

---

\(^{12}\) Each of the expressions led by (a), (b), or other letters in the parentheses represents one “punctuation unit.” These examples are directly quoted from Danielewicz & Chafe (1985).
(b) without such a precise means of communication I doubt that humans would have progressed as quickly as they did (p. 217).

2. The introduction of an intonational boundary in speaking, or of a comma in writing, between the subject and the predicate of a clause:
   Example (a) Changing the meaning of all our words,
   (b) would be like creating a new language (p. 218).

3. The introduction of an intonational boundary in speaking, or of a comma in writing, before a restrictive relative clauses or a complement clause:
   Example (a) One of these other categories,
   (b) that I can be classified in is that of an only child (p. 220).

4. The introduction of an intonational boundary in speaking, or of a comma in writing, after significant pieces of information of other kinds:
   Example (a) It is much more effective to extend the meaning of truth,
   (b) into a sentence,
   (c) rather than to simply define it (p. 221).

5. The failure to set off the beginning of an adverbial phrase with an intonational boundary in speaking, or with a comma in writing:
   Example (a) Language is responsible for describing and communicating ideas,
   (b) and with proper usage,
   (c) they can be the most important link we can have with our society in the past, present, and future (p. 222).

6. The fragmentation of language into a series of brief units bounded with falling pitches in speaking, or with periods in writing:
   Example (a) That night came when we landed at the airport.
   (b) To my amazement there was fog.
   (c) I thought that only San Francisco had a fog.
(d) The airport was full of people and noisy (p. 222).

7. The addition of supplementary, “afterthought-like” material after a falling pitch boundary in speaking, or after a period in writing:
Example (a) For this quality of work she receives much praise from her bosses.
(b) Which is the reason why fall is her favorite season (p. 223).

The evidences provided by Danielewicz and Chafe show that most errors of the basic writers in punctuation are not random errors. If teachers who teach writing are aware of how the impact from speech is likely to be inevitable on writing, teachers can concentrate on pointing out specific ways in which the requirements of writing differ from those of speech.

4.1.3 Syntactic errors

Shaughnessy categorized syntactic errors under the following very general headings:
(1) Accidental errors:
Example Life is really hard today so you can imagine what it will [be] in the coming future and for us generation (p. 48).
Shaughnessy believes this kind of accidental errors can be corrected by the writer after careful proofreading. Therefore, they are not serious or systematic errors.

(2) Blurred patterns, that is, of patterns that “erroneously combine features from several patterns, creating a kind of syntactic dissonance” (p. 49).
Example If they (jobs) decrease in a great number At least I can say is that I will have a college degree.
Blurred forms:
At least I can say is
The least I can say is

(3) Consolidation errors, which have three subcategories:
(a) Coordinate consolidation errors:
Example They believe they can become leaders in their field and a good secure job (p. 55).
The corrected sentence is “They believe they can become leaders in their field and get good secure jobs.” Shaughnessy found that basic writers make errors not only between coordinate clauses, but also between smaller coordinate units like phrases. Among these errors, the most difficult ones are structures of comparison “where elements that are grammatically balanced are semantically unbalanced” (p. 56).
Example When you do graduate from college with a degree your chances of getting that job you want is increased enormously than that of a high school student (p. 56).

(b) Subordinate consolidation errors:
Example Even if a person graduated from high school who is going on to college to obtain a specific position in his career should first know how much in demand his possible future job really is (p. 59).

(c) Juxtaposition consolidation errors:
Example Now mostly every job you go to get worthwhile or making a decent salary something to live off is asking for a college degree (p. 66).

(4) Inversion errors:
Example The things they want you to know you do not know this in high school (p. 68).

Shaughnessy lists three explanations for the reasons why these basic writers fail to manage syntactic complexity. First, “the inexperienced writer is indeed not likely to have command of the language he needs to bring off the consolidations that are called for in writing” (p. 73). Second, “the beginner writer does not know how writers behave” (p. 79). In other words, the beginner writer is not familiar with the composing process. Third, “the student lacks confidence in himself in academic
situations and fears that writing will not only expose but magnify his inadequacies” (p. 85).

Shaughnessy, furthermore, adds that each of these explanations might suggest a pedagogy: (1) “the pedagogy that stresses grammar, whether in the abstract or as a set of forms to be generated through practice with sentences, tends to assume that students do not have command of many of the forms required in written English and must therefore learn them through explicit instruction;” (2) “the pedagogy that stresses process (pre-writing, free writing, composing, re-scanning, proofreading, etc.) tends to minimize the value of grammatical and rhetorical study and assume, rather, that students already ‘know’ the wanted forms but cannot produce them, nor anything resembling their own ‘voices,’ until they are encouraged to behave as writers;” and (3) “the pedagogy that stresses the therapeutic value of writing and seeks the affective response to whatever is read or discussed tends to see confidence as central to the writing act and to dismiss concerns with form or process as incidental to the students’ discovery of themselves as individuals with ideas, points of view, and memories that are worth writing about” (p. 73).

Apparently, none of these three pedagogies can make a student a competent writer alone, for each addresses but one part of the problem. On the other hand, they may be contradictory if implemented at the same time. Instructors should be very careful in identifying students’ errors and be equally cautious in selecting any one of these pedagogies at a time. Whatever pedagogies are used, the business of a writing class is to make writing more than an exercise: it is a place to make the beginner writer aware of how to behave as a writer in all aspects—from forms to thinking.

4.1.4 Common errors

Shaughnessy treated such errors as common errors that are “grammatically less important” (p. 90) and “do not seriously impair meaning” but usually affect ordinary features of written English. She proposes that
these errors are “often rooted in language habits and systems that go back to their [the basic writers’] childhood and continue, despite years of formal instruction, to influence their performance as adult writers” (p. 90). These common errors including the misusage of inflections (e.g., - s, -ed), agreement, tense, and pronouns.

The main sources for common errors are that some inflectional “letters” are not pronounced clearly in speech, or that children find the whole inflectional system alien to them; so, Shaughnessy suggests, children or basic writers make such errors because they fail to grasp the strict requirements of the inflectional system in writing. For example, some children have a hard time distinguishing ing and en by listening to adults’ speaking. In writing, they tend to write:

This is our responsibility giving to you in college (p.100).

Or those who are unaware of the inflectional system would write:

We should be taught how to decided for ourselves (p. 100).

Another reason is that children or basic writers resist accepting the conventional rules in writing. For example, some basic writers or children write

Your would be wasting your time by going to college for four, five, six year when the jobs are in the fields that require a minimum training (p. 107).

Shaughnessy infers that this writer had a sense of “plural nouns” but he did not add s after “year” because he thought it was redundant to do so. The same phenomenon happens in the category of subject-verb agreement. Experienced writers or teachers may think that all these grammatical rules are “mechanical” (Gere and Stevens, 1985) so that children should not have much difficulty acquiring them. Shaughnessy suspects that these rules are not really “easy” in some ways if so many college students have problems applying them correctly. She, therefore, suggests that, theoretically, “throughout instruction, a student should be encouraged not simply to have the right answers but to have grammatical reasons for what he does, for grammar is more a way of thinking, a style of inquiry, than a way of being right” (p. 129).
4.1.5 Vocabulary errors

Vocabulary errors provided by Shaughnessy consist of the following types:

1. A word that does not exist:
   Example: I wish my life to be forfilling with happiness (p. 190).

2. Errors with derivational suffixes:
   Example: He is headed in a destructional way (p. 190).

3. Substitution:
   Example: They used him as an scape goat (scapegoat)(p. 191).

4. Unacceptable words:
   Example: I think psychology has increased my mind (p. 192).

5. Error with prepositions:
   Example: a good opportunity of work (to) (p. 193)

Beyond these explicit errors, there are also stylistic problems that reflected by inappropriate use of vocabulary. The reasons for basic writers’ vocabulary errors are by no means clear. However, Shaughnessy, based on the observed features, make four suggestions for teaching vocabulary properly. First, “it is probably possible to reduce somewhat the number of words students must learn in introductory courses” (p. 224). Second, vocabulary teaching should not be isolated from the “content” of a course and simply serves as a prerequisite part of that course. Third, in teaching specialized vocabulary, it is important “to build criteria for precision--through skillful questioning, through cultivating a sense of audience, through the close examination of precise passages--so that the student begins to push past the limits of his customary vocabulary, searching for better words and phrases” (p. 225). And finally, “the teacher as mediator between the languages students bring to class and the language of the academy must himself serve the students both as translator and model” (p. 225).
4.1.6 Errors beyond the sentence

Shaughnessy also looked at passages rather than individual sentences and found “unmistakable evidences of oral rather than formal written styles of development and of misinformed or partially informed ideas of what the academic model is supposed to be” (p. 273). Among several reasons given, the most important one is that “the teacher often concludes that thinking rather than writing ought to be the focus of his instruction” (p. 236) and assumes that once the student has “learned to think” the task of writing down his thoughts will come easily. Unfortunately this assumption is not often true for most basic writers.

The previous discussions have illustrated that basic writers’ limited linguistic knowledge in academic writing prevents them from writing out their “thoughts” freely in the required form of conventional writing. For this, Shaughnessy suggests three needs—“the need to experience consciously the process whereby a writer arrives at a main idea or point; the need to practice seeing and creating structure in the written language; and the need to recognize specific patterns of thought that lie embedded in sentences and that point to ways of developing larger numbers of sentences into paragraphs and essays” (p. 274). In addition, she also suggests that the instruction of the written knowledge beyond the sentence level should go together with other grammatical rules, spelling, and vocabulary acquisition; it should not be the subject of the “next year” course after the sentence-level writing skills training.

The reason I have given a detailed review on Shaughnessy’s work is that her error analysis on L1 writing has been so far the most representative and most complete one. Most error types discussed in her work are similar to those occurring in the English compositions produced by ESL students, although the sources for these errors may not be the same.
4.2 Applying “interlanguage hypothesis” to L1 writing

Usually, researchers will borrow theories from L1 and apply them to the same field of studies in L2. But in error analysis, some researchers try to apply L2 theories to account for L1 errors. In doing so, these researchers have one assumption in common: they all hypothesize that learning writing is like learning a second language. Bartholomae (1980) in his study of L1 writing errors found that using a second language theory as a base has one benefit: teachers can analyze errors as evidence of the writer’s current system and can teach the writer to alter that system by using reading and editing techniques. Horning (1987) continued developing this notion in his book *Teaching Writing as a Second Language* in which he discussed in detail how Selinker’s (1974) interlanguage hypothesis (for L2 acquisition) can be applied to the learning of L1 writing. Selinker observed what may happen when students attempt to produce sentences in the new language and found the students’ sentences are generally somewhat different from those of native speakers under similar circumstances. “Because of these errors,” summarized Horning (1987), “Selinker hypothesizes that second language learners pass though a stage of language ability in which systematic errors of various kinds occur” (p. 31). This stage is called interlanguage. Two features of the interlanguage concept that attracted Horning to apply them to L1 writing are

(a) the errors in the interlanguage stage are systematic and
(b) students move through the interlanguage stage toward proficiency.

Horning himself also made two other points to support his view that there are similarities between second language learning and learning to write. “First, the goal in both learning situations is to be able to communicate well enough to interact with native speakers” (p. 32). And second, both in second language learning and in learning to write, errors appear from the very beginning, and some even persist after proficiency is nearly achieved.
In analyzing interlanguage errors of L2 learning, Selinker introduces the notion of “fossilization” that Horning thinks can also be applied to learning L1 writing. Selinker says:

Fossilizable linguistic phenomena are linguistic items, rules and subsystems which speakers of a particular native language will tend to keep in their interlanguage relative to a particular target language, no matter what the age of the learner or amount of explanation and instruction he receives in the target language.... It is important to note that fossilizable structures tend to remain as potential performance, re-emerging in the productive performance of an interlanguage even when seemingly eradicated (p. 119).

Horning found the notion of fossilization also “helps account for the fact that composition students can routinely make subjects and verbs agree or place apostrophes correctly in exercises but continue to make errors in writing” (p. 33). Horning, therefore, believes that fossilization in learning to write results from, as Selinker suggests in L2 learning, language transfer, transfer of training, different strategies of learning and communication, and overgeneralization. Persuasive as these theories of interlanguage and fossilization are, both Selinker and Horning admit that it is not clear how fossilization arises or how learners in either L2 or L1 writing move from interlanguage to proficiency.

Unlike Shaughnessy who sometimes sees errors as random ones, Horning’s attitude toward errors made in basic writers’ compositions is to always “look at errors as errors, rather than [random] mistakes, to see if there is a pattern to them” (p. 50). If patterns of errors are impossible to formulate, it is also worthwhile discussing in class each single type of error because “a single type of error may also be a diagnostic key” (p. 50). He also suggests asking students to analyze their own errors and make conscious their unconsciously acquired knowledge.

Although Shaughnessy and Horning look at error from different viewpoints, they have several points in common. Both of them agree (1) that most errors made in basic writers’ compositions are systematic; (2) that many errors in compositions are caused by language transfer (from
speaking to writing); and (3) that basic writers make errors because they are unaware of the rule or because they make them unknowingly.

### 4.3 Error analysis concerning L2 writing

In L2 learning (or acquisition)\(^\text{13}\), it is an undeniable fact that error analyses play a very important role in class no matter what type of pedagogy the instructor employs. Anyone who has ever experienced learning a second language in class must recall how he feels whenever he makes errors. Since errors are the most explicit barriers that cause either misunderstanding or failure of communication, error analysis naturally has become a very important field in L2 research.

Some researchers suggest using a contrastive analysis hypothesis\(^\text{14}\) to predict possible errors that may occur in L2 by comparing the linguistic features of L1 and L2 (Wardhaugh, 1970; Lado, 1957; James, 1971; Cowan, 1977; Weeks, 1983). Some extend a Chomskyan innateness hypothesis regarding first language learning by children to second language learning by adults. Corder (1967), for example, says that “no one has provided evidence that these are fundamentally different processes. And just as errors in first language production provide evidence for the development of that language in the child, so also should such errors in

\(^{13}\) Since the great majority of literature concerning L2 learning is in fact ESL studies, I refer to L2 in this paper as English (as a second language to any non-English speaking subject).

\(^{14}\) Wardhaugh (1970) has a very detailed introduction of the contrastive analysis (CA) hypothesis in which he distinguishes a strong version of CA from a weak version of CA. The strong version proposes that “it is possible to contrast the system of one language--the grammar, phonology, and lexicon--with the system of a second language in order to predict those difficulties which a speaker of the second language will have in learning the first language and to construct teaching materials to help him learn that language” (p. 8). The weak version “requires of the linguist only that he use the best linguistic knowledge available to him in order to account for observed difficulties in second language learning” (p. 10). Wardhaugh comments that the strong version is too idealistic to apply in real world, but the weak version is of more value in L2 learning.
the second language production of the adult” (summarized by Robinett & Schachter, 1983, p. 145). Other researchers (e.g., Selinker, 1974) claim that “the majority of second language learners activate a latent psychological structure which is quite independent of the innate mechanisms activated by children to acquire their first language and which is quite likely to resemble other intellectual (but nonlanguage) psychological structures” (summarized by Robinett and Schachter, 1983, p. 146). Selinker (1974) claims that fossilized errors (see 4.2) result from five central processes that exist in this latent psychological structure: language transfer, transfer of training, strategies of L2 learning, strategies of L2 communication, and overgeneralization.

No matter what approach the researchers take, they all go through one procedure: to collect errors and to categorize them. Conventional classification of L2 errors mainly in terms of error types (e.g., Richards, 1971), frequencies (e.g., Neuman, 1977 for absolute frequency; Agnello, 1977 for relative frequency\(^{15}\)), or sources (e.g., Richards, 1971). Since sources of error are believed to be most useful for L2 instruction designs, it is valuable to discuss them in detail here. Richards in 1971 proposed three categories of the sources of errors: (1) interference errors, (2) intralingual errors, and (3) developmental errors.

(1) The interference errors are those caused by the influence of the learner’s mother tongue on production of the target language in presumably those areas where the languages clearly differ.

(2) The intralingual errors are those originating within the structure of English itself. Complex rule-learning behavior is typically characterized by overgeneralization, incomplete application of rules, and failure to learn conditions for rule application. When the complexity of English structure encourages such learning problems, all learners, regardless of background language, tend to commit similar errors.

\(^{15}\) Relative frequency refers to a fraction obtained by using as numerator the number of times an error was committed and as denominator the number of times that error type could have occurred.
(3) The developmental errors reflect the strategies by which the learner acquires a language. These errors show that the learner—oftentimes completely independent of the native language—is making false hypotheses about the target language based on limited exposure to it. A major justification for labeling an error as developmental comes from noting similarities to errors produced by children who are acquiring the target language as their mother tongue (summarized by Schachter and Celce-Murcia, 1971, p. 274).

These three categories provide rich information that helps the L2 learner to avoid committing errors if researchers or instructors have no difficulty classifying all possible errors in a desirable way.

However, not all researchers interested in error analysis are involved in analyzing error types and how these errors are committed. Some skeptics who have difficulty defining the term “error” start asking the question: “Is this ‘error’ really a deviation from the target language?” If the answer is yes, then comes the following question: “What structure is this error really in?” Hence, the identification and classification of errors may become a powerful factor that decides whether or not an error is an error (Schachter, 1974). In other words, errors are not “committed” by writers alone, they are also “decided” by decoders (researchers or instructors).

In 1982, Ludwig reviewed the literature of second language learning and found that there is a new trend of error analysis focusing on “measuring native speaker reaction by determining which errors interfere with comprehension or are irritating or unacceptable to the receiver” (Vann, Meyer, and Lorenz, 1984; Hwang, 1991). This move takes error judgment from one decoder’s intuition to a group of decoders’ objective responses. Burt and Kiparaky (1974) hypothesized that there is an error hierarchy on the basis of the degree of effect on native English speakers’ comprehensibility. In terms of comprehension, this hierarchy is defined as “A is higher than B in the hierarchy if the correction of A contributes more to comprehensibility than the correction of B” (p. 72). They also hypothesized that “in sentences with more than
one clause, the appropriate overall organization is more important to comprehensibility than the correct formation of each clause with the larger sentence” (p. 73). For this hypothesis, they call errors (they originally use “mistakes”) that confuse the relationship among clauses “global errors” (e.g., word order or connector), and call errors within clauses “local errors” (e.g., articles or subject-verb agreement). Apparently, Burt and Kiparaky (1974) claim that correcting global errors is more important and more effectual in the improvement of comprehensibility than correcting local errors. Tomiyana (1980) increased her sample size (asking 120 graduate students at the University of California at Los Angeles (UCLA) to read two passages) and looked at native speakers’ responses to errors in discourse. Her conclusion supports Burt and Kiparaky’s hypothesis that connector errors (global) are more crucial to communication than article errors (local) although her study is limited to assessing the reaction to only two grammatical items—articles and connectors.

Several similar studies focusing native speaker responses have been conducted with adults learning Indo-European languages other than English (Piazza, 1980 (French), Chastain, 1981 (Spanish)). Though these studies do not use the same hierarchical scheme that Burt & Kiparaky had suggested, the findings all agree that some types of errors impede comprehension to a higher extent than other ones. For example, Chastain (1981) found that diction errors interfere more severely with comprehension than grammatical errors. Piazza (1980) examined errors both in speech and in writing, finding that erroneous sentences occurring in writing were more comprehensible and less irritating to native French speaker than those in speech.

While most error analysis studies emphasize the hierarchy of error gravity for a specific language, some researchers have started shifting the focus onto semantic errors. By definition, semantic errors are the errors (usually at the lexical level) that cause the miscarriage of intended meanings. For example, Politzer (1978) asked 146 German teenagers to respond to 60 pairs of sentences. The results show that vocabulary errors (or semantic errors), are more serious than syntactic errors and
phonological errors. Delisle (1982) replicated the study and provided similar results. Santos (1985) investigated the reactions of 178 professors at UCLA to two 400-word compositions written by two non-English speaking students (one Chinese and the other Korean). The results indicate that among the five most serious and irritating error types, four of them are lexical errors (which imply semantic errors).

Hwang (1991) asked three English instructors at the University of California at Santa Barbara to identify both grammatical and semantic errors in 163 compositions presented by 40 ninth graders, 40 tenth graders, 39 eleventh graders, and 44 college juniors from Korea. She found that as the learning years increase, the students’ writing proficiency increases, and both grammatical and semantic error ratios decrease. In semantic error analysis, the most frequent semantic errors are lexical errors due to wrong choice of words or wrong collocational usages. She, therefore, concluded that the choice of vocabulary both in single and collocational usages is the most problematic area for these Korean students. In this study, Hwang (1991) established a very valuable structure for semantic error analysis. The three-way classification of semantic errors provided by Hwang is as follows:

Category 1: Containing general idea presentation errors which include
   a) incomplete ideas
   b) ambiguous ideas
   c) illogical ideas

Category 2: Containing lexical errors which are usually recognized as pure semantic errors, which consist of
   a) wrong words
   b) collocations
   c) non-English words

Category 3: Containing syntactic errors evoking semantic problems which are generally recognized as the most difficult category distinguishing syntactic errors from semantic ones; this category is found to have three subcategories:
   a) word order
b) word form

c) reference

This category is by no means complete and also can be controversial. For instance, Vann, Meyer, and Lorenz (1984) asked 440 (319 responded) of the Iowa State faculty to rate their tolerance toward 12 types of errors made by ESL students in Iowa State University, finding that the five most intolerable error types belong to, respectively, 1) word order, 2) it-deletion, 3) tense, 4) relative clauses, and 5) word choice. Based on Hwang’s categories of semantic errors, it would be very hard to decide whether errors of tense should be categorized under grammatical errors or semantic errors. Since almost all studies concerning semantic error analysis agree that semantic errors impede comprehension more seriously than grammatical errors, it would be reasonable to consider the question: “Does the misuse of tense interfere with communication?” For example, it is not uncommon to find an ESL student saying or writing the following expression:

I should do that, but I didn’t.

Should we interpret it as “I should have done that but I didn’t,” or “I should do that, but I won’t?” With context in discourse, the meaning may be easier to “guess.” In sentence-level error analysis, this kind of error can be very misleading, because no researchers have ever seen misuse of tense as semantic error. To conclude, Hwang’s semantic error categories, though they successfully serve as a pioneer study, require further studies to make them more distinguishable from grammatical errors.

Vann et al. (1984) also provided another aspect of interest in error analysis. In investigating native speaker response toward errors made by ESL students, the respondents consisted of four age groups (≤ 34, 35-44, 45-54, and ≥ 55) and three academic discipline groups—(1) physical and mathematical sciences and engineering, (2) biological and agricultural science, and (3) social science, education, and humanities. Based on a two-way multivariate analysis of variance, the results indicate that both (age and academic discipline) are significantly different in responding to given errors. Vann et al. found that the oldest group (≥ 55)
showed the most tolerant opinions of the ESL errors (followed by ≤ 34, then 35-44, then 45-54). They also found faculty members from the social sciences, education, and humanities were significantly more tolerant to ESL errors than the faculty members from the other two natural science fields. Vann et al., accordingly, suggested a need to “examine not only the communicative effect of non-native speaker errors, but also their effect on sub-groups of decoders, considered, for example, by age, sex, profession, and linguistic sophistication” (p. 437).

In this chapter, I started with Shaughnessy’s classical error analysis and summarized the basic error types in different grammatical categories with examples produced by college English-speaking basic writers. Then I mentioned Horning who borrows theories of ESL error analysis to support his hypothesis that learning writing is like learning a second language. The ESL error analysis theory Horning applies is Selinker’s theory of interlanguage and fossilization. In the following section I discussed how ESL errors have been treated. I spent relatively fewer words describing the conventional way of analyzing ESL errors, e.g., the contrastive analysis hypothesis, grammatical, or syntactic error analysis. Instead, I gave much space discussing the new trend of error analysis focusing on native speakers’ responses toward errors made by L2 learners because this new trend raises at least three problems that traditional error analysis studies do not deal with. (1) Who decides improper “usage” to be an error, research, an instructor, or a large sample of population? (2) Can grammatical errors be distinguished clearly from semantic errors? Or are these two types of error in effect two ends of one continuum? (3) In addition to the communicative effect, do we also need to look at the background (e.g., education or profession) of those who identify the errors committed by L2 learners?
5.1 From indirect testing\textsuperscript{16} to direct assessment

In the past few decades, teachers, administrators, and researchers have been trying to find a method of assessment that is both reliable and valid. “A reliable method of assessing writing ability would yield a consistent judgment of a student’s abilities if applied again, all else being equal...a valid writing assessment would be sensitive to a writer’s ‘true’ abilities” (Charney, 1984). In real life, it proves to be not easy to assess writing both reliably and validly. In larger-scale testing programs, indirect measures “designed to provide proxies for writing abilities” (Freedman, 1991, p. 2) have been employed for being both reliable and

\textsuperscript{16} In this paper, the term “testing” will, as Freedman (1991) suggests, refer to large-scale standardized evaluation and “assessment” will refer to the evaluative judgments of the classroom teacher. Freedman cited (from Calfee and Drum, 1979, p. 738) that “testing activities are usually group administered, multiple choice, mandated by external authorities, and used by the public and policy makers to decide ‘how the schools are doing’ while assessment activities include the evaluation of individual student performance are based on the teacher’s decisions about curriculum and instruction at the classroom level, and are aimed toward the student’s grasp of concepts and mastery of transferable skills” (p. 1).
economical. However, the validity of using such indirect measures for writing ability is severely criticized for not being able to represent students’ true performance. Indirect measurements are generally multiple-choice tests, typically including questions about grammar, sentence structures, or scrambled paragraphs to be rearranged in a logical order. According to Conlan (1986),

No multiple-choice question can be used to discover how well students can express their own ideas in their own words, how well they can marshal evidence to support their arguments, or how well they can adjust to the need to communicate for a particular purpose and to a particular audience. Nor can multiple-choice questions ever indicate whether what the student writes will be interesting to read (p. 124).

In addition, multiple-choice writing tests, if students are taught or required to take such tests, may have negative effects on instruction since teaching to the test would not include asking students to actually write (Resnick and Resnick, 1990).

Being aware of these presumably serious problems, the College Entrance Examination Board (CEEB) struggled to find practical ways to replace multiple-choice, indirect measures of writing from 1890 on into the 1960s (Freedman, 1991). The goal, of course, was to develop an ideal method to assess actual samples of student writing directly (Hud- dleston, 1954; Meyers, McConville, and Coffman, 1966). To reach this goal, CEEB was faced with several problems. First of all, it cost much more to hire humans to rate or measure actual writing than to score multiple-choice tests. Second, it proved difficult to get the raters to agree with one another on rating criteria, and, consequently, the inter-rater reliability did not come out satisfactorily.

In 1961, Diederich, French, and Carleton at the Educational Testing Service (ETS) also invited “sixty distinguished readers in six occupational fields” (Diederich, 1974, p. 5) to rate 300 papers written by college freshmen and to write, on as many papers as they could, their brief comments about what they liked and disliked. The results showed that, among the 300 papers, “101 received every grade from 1 to 9” (p. 6).
However, the comments written down on the papers helped ETS researchers understand why readers disagreed. They, therefore, started developing ways of training readers to agree independently on “holistic” or general impression scores for student writing. In other words, the readers are trained to evaluate each piece of student writing relative to the other pieces in the set, without considering standards external to the examination itself (Charney, 1984). This seemed to work well and thus solved the reliability problems of direct assessment (Cooper, 1977; Diederich, 1974). In addition, both ETS and CEEB also developed ways to collect writing samples in a controlled setting, on assigned but appropriate topics (cf. Myers, 1980), and under timed conditions.

On the other hand, Freedman and Calfee (1983) conducted an experiment designed as a preliminary study of the importance of three generic categories of factors that experts see as potentially important influences on how evaluators score essays: the essay, the rater, and the context (setting) for the rating. The results indicate that holistic ratings are influenced more by the essay than by the latter two sources, and they concluded that the viability of the holistic rating approach is quite satisfactory. Although the expense of hiring readers and the difficulties of getting them to agree with one another have forced ETS and CEEB to use essay tests as supplementary to multiple-choice testings, with the practical problems solved, the door has been opened to the current, widespread, large-scale, direct assessment of writing (Davis, Scriven, & Thomas, 1987; Diederich, 1974; Faigley et al., 1985; Myers, 1980; White, 1985).

5.2 Holistic scoring

Holistic direct assessment of writing samples, if in a controlled setting, provides teachers, administrators, testing agencies and researchers with the following advantages:

1. Both inter-rater and intra-rater reliabilities are satisfactorily high (from .70 to .80 and over if raters are given special training ses-
(1) Holistic scoring prior to their work (Mellon, 1975; Bauer, 1981; Cooper, 1977; Stiggins, 1982; Moran, 1987).

(2) It provides better face validity (Bridgeman and Carlson, 1983; Criper and Davies, 1988; Keller-Cohen and Wolfe, 1987).

(3) Training time for raters is short (Bauer, 1981; Stiggins, 1982).

(4) First impression holistic scoring requires fast reading without second thoughts, so writing samples are quickly rated and, therefore, more economical (Bauer, 1981; Cooper, 1977; Veal & Hudson, 1983; Isaacson, 1984).

(5) It can be used with large writing samples, so that it satisfies the current need of direct assessment of writing samples (Freedman, 1991).

Having these advantages does not mean holistic scoring solves all assessment problems. Brown (1981) criticizes holistic scoring as “a way of rank-ordering papers” no matter how reliable it is, and “it is inadequate as a measuring tool in itself, because it is relativistic and is not tied to any absolute definition of quality” (p. 36). From this general description of the weakness of holistic scoring, we can specify the following questions to be answered:

(1) What do teachers or researchers as well as writers themselves actually learn from the results of holistic scoring except “the place” where a piece of writing sample fits in the rank-order (Lloyd-Jones, 1977; Mosenthal, Davidson-Mosenthal & Collela, 1987; Wesdorp, Bauer & Purvis, 1982)?

(2) Are students given sufficient time to engage themselves in the elaborate processes that are fundamental to how good writers write and to how writing is ideally taught (Brown, 1986; Lucas, 1988; Simmons, 1990)?

(3) Except for the face validity, what other types of validity can holistic scoring provide us (e.g., construct validity, concurrent validity, or content validity) (Hamp-Lyons, 1990; Lauer and Asher, 1988)?

(4) Will high inter-rater agreement obscure the raters sensitivity to students with a unique usage of language or an imaginative style (Hake, 1986)?
(5) How does an evaluator reach a judgment about a writing sample (Freedman, 1983)?

Since teachers, researchers, and writers themselves no longer believe that indirect multiple-choice testing can serve as a valid method to assess writing ability of any kind, the questions mentioned above become urgent in promoting the strength of the method of holistic scoring.

The following is the basic procedure employed by Zimmer et al. (1990) in rank-ordering English autobiographies produced by 4th and 6th graders. Prior to scoring the written products, three raters were trained on 18 randomly selected writing samples. In order to enhance the consistency across raters from different backgrounds, three training sessions were conducted for two purposes: (1) to narrow the discrepancy on the criterion of “quality” defined by different raters; and (2) to develop a set of rubrics on the basis of the criteria described by these three raters (Zimmer, Ho, Giwuff, Liu, & Tuss, 1990).

The eighteen writing samples were randomly divided into three piles to serve as materials for each training session. In session one, five raters independently rank-ordered the first set of six papers on the basis of “quality,” and to represent the range of quality, according to the notion of holistic scoring assessment or a “General Impression Marking” (Isaacson, 1988; Cooper, 1977). In other words, these writing samples were scored based on “the overall impression of the quality of the writing in relationship to the total population of student writers” (Braungart-Bloom, 1986).

Following the scoring, each rater wrote a description of the characteristics of each of the six rank-ordered articles. These attributes were discussed by the members of the group and a list was made and was served as the initial step in developing the scoring rubric. The synthesized attributes were further developed in the second session.

In session two, the second set of 12 writing samples were distributed to each rater and the raters followed the procedures of the first session. After the raters ranked the second set of writing samples, the chart of results was put on the blackboard and raters discussed reasons for the rankings of the samples in terms of the characteristics and attributes of
the writing samples. During the discussion a consensus for the rubric was reached and a common draft of the rubric was developed. The rubric was described as follows (Zimmer, Ho, Giwuff, Liu, & Tuss, 1990):

(6) A superior response, fluent and coherent. There is development and elaboration of ideas and a clear relationship between sentences and ideas. The writing illustrates movement in time and possibly an understanding of cause and effect. The paper demonstrates a “literary” quality; that is, a sense of the writer as a unique individual telling a personal story.

(5) These papers will have some of the same qualities as a superior paper, but to a lesser degree. While the writing is coherent, the elaboration and development are not as complete as in those in the highest category.

(4) Papers in the mid range will be a mixture of good and bad qualities, for example, elaboration mixed with single unrelated ideas. There are quick changes of topics. The writing is rather mechanical.

(3) More “list like” than a “4” paper, but with some description. Jumps in time with limited connections. Limited sense of the writer as a unique person.

(2) List or list like, a series of isolated incidences. Sentences are relatively short. No sense of the relative importance of the items.

(1) Very limited or no elaboration or descriptions and scattered ideas or none. List-like with no connections or cohesion. No sense of the relative importance of an idea. Short, simple, and shallow sentences.

In session three (one week later), the third and final set of writing samples to be rank-ordered in the training were distributed. The raters used the definitive rubric as criteria for ranking the last set of autobiographies. The results indicated that the previous two sessions of training did help the raters to develop more common criteria for scoring the writing samples.
Following the descriptions in the rubric, eight raters proceeded to score the set of 100 collected written samples. Each paper was evaluated by two different raters. The first score was covered with a double-thickness tag with the rater’s initials on it. The second rater did the same thing but covered his score on the other upper corner of the paper. If two raters’ scores were equal or adjacent in numeric value (for example, two and three or five and six) the papers were considered to be “in agreement” and the score would be the sum of the two scores. Discrepant scores, differing by more than one point (for example, two and four or one and three), were rated by a third person (Paulin, 1980). If the third score fell between the first and second scores, the final score would be twice the middle score (for example, if the first score is 2, the second score is 4, and the third score is 3, the final holistic score will be 6). If the third score was outside the range of the first and second scores, the final holistic score would be scored by adding the two adjacent scores. The agreement between scores for the written products was .704 (N=192).

In this study, Zimmer et al. insisted that the “training” of raters is important because sufficient training definitely narrows the discrepancy of the criterion of “quality” defined by different raters; and different raters may come up with a common set of rubrics to serve as assessing criteria.

5.2.1 How facets of composition relate to holistic scores?

To provide both researchers and writers with more direct information about grammatical and rhetorical facets of student writing, two systems are developed with this goal: the first is the National Assessment of Educational Progress (NAEP) Primary Traits system and the other, the Analytic system. In the mid 1970s NAEP developed “the Primary Traits Scoring method” (Lloyd Jones 1977, p. 33) with a goal to set specific criteria for successful writing on a particular topic ahead of time. Freedman summarized that “the primary trait is determined and defined by the test-maker who decides what will be essential to writing success-
fully on each topic on the test. Traits vary depending on the topics. Tensions arise because the test-makers cannot always anticipate precisely what test-takers will do to produce good writing on a particular topic and what is primary or whether one aspect of writing should be labeled primary is a debatable point” (pp. 5-6). The primary trait system does not prove to be too impressive although Brown (1981) believes it to be promising. In a study comparing holistic rating and primary trait scoring, NAEP finds that primary trait scoring does not correlate particularly well with holistic quality judgments (correlations ranging from .38 to .66 depending on the topic) (Freedman, 1991, data from NEAP, 1986, p. 84).

Not satisfied with what has been done about the written samples analytically, Freedman and Calfee (1983), on the basis of Simon’s (1981) concepts that combining analysis and synthesis in the study of complex systems, examine how various qualities within the essay influence the raters’ decisions. For this, they examined two things that previous studies have never done. First, they examined the relative influence of significant psycholinguistic categories on the holistic score. Based on recent theories of text analysis (de Beaugrande, 1980; Frederiksen, 1975; Kintsch, 1974), the factors they used for this study were propositional (idea development), schematic (organization), syntactic (sentence structure), and conventional (mechanics). Second, they tried to establish a causal relationship between the critical categories of the text and the score, rather than the correlations only. To explain how they perform the analysis, Freedman and Calfee said,

To test the extent to which each of the four categories of the essay...caused raters to vary their holistic scores, we rewrote essays to incorporate strengths or weaknesses within each category and then presented the rewritten essays to raters for holistic scoring. The raters, of course, were not told that the essays had been rewritten (p. 85).

The results indicate that idea development has the greatest effect, followed by organization and mechanics. Sentence structure is not a statistically significant factor. The interactions indicate, according to Freed-
man and Calfee, that “if an essay was well organized, the ratings were clearly higher if the sentence structure and mechanics were also in good shape. However, if an essay had a weak organization, neither the sentence structure nor mechanics factors had much effect on the ratings” (p. 88).

Freedman and Calfee (1983) demonstrated that, if carefully designed, both the Analytic system and the Primary Traits method can provide us more information about how the various facets of compositions affect readers’ rating so as to suggest remedies to students weaknesses.

5.2.2 How much time does a student need to satisfactorily write a sample?

Whether students are given sufficient time to engage in the elaborated processes that are fundamental to how good writers write and to how writing is ideally taught has been discussed by some researchers (Brown, 1986; Lucas, 1988; Simmons, 1990). Any experienced writers are aware that good writing requires “sophisticated higher order thinking” (Freedman, 1991) and such higher order thinking occurs “when there is an increased focus on a writing process which includes encouraging students to take lots of time with their writing, to think deeply and write about issues in which they feel some investment, and to make use of plentiful response from both peers and teachers as they revise (Freedman, 1991). In most test-type writing, students are asked to write their samples in 15 to 30 minutes, depending on the topic. If the topic does not serve as a prompt, as Myers (1980) has suggested, it may take students some time to figure out what they should do about the topic. Before they make clear which stance they should take for the topic, they have probably “wasted” five to ten minutes. NEAP noticed this and in 1988 gave a subsample of the students twice as much as time on one informative, persuasive, and imaginative topic at each grade level. The results summarized by Freedman (1991) show that “with increased time all students scored significantly better on the narrative tasks and fourth-
and twelfth-graders scored significantly better on the persuasive tasks. Only the informative tasks showed no differences. Most disturbing, the extra time proved more helpful to white students than to Blacks or Hispanics, widening the gaps between these groups in the assessment results” (p. 5). How long of a time for producing a representative sample is appropriate for students is no longer an issue of “time” per se, it is closely related to topics selected, background of students, and how “natural” the test setting is. One method to help solving the “time” issues is to collect more samples in different settings with different time spans, to see how consistent these collected samples may be within each time span. The consistency of student writing performance may provide part of the answer. But the cost of this is obviously too high to be used practically in large-scale testings. The other way to solve the issue is, as Freedman (1991) suggests, to use information provided by “writing portfolios” to help understand the real potential of the students’ writing abilities.

5.2.3 Should the validity of holistic scoring be questioned?

Holistic scoring evaluates writing qualities by looking at the writing samples directly; this leaves the face validity unquestioned. The facts prove that the results of this direct assessment have already provided very valuable information for different kinds of purposes (Bridgeman and Carlson, 1983; Criper and Davies, 1988; Keller-Cohen and Wolfe, 1987). But there are also researchers who criticize this, saying that the limited time, unnatural writing conditions and assigned topics may make writing samples fail to reflect students’ true abilities (see 5.2.2).

With respect to other types of validity, Charney (1984) questions the construct validity of holistic scoring because of the variability of topics and types of discourse, the arbitrariness of the criteria, and the failure of raters to adhere to the criteria. Hamp-Lyons (1990) directly points out that “we urgently need serious research into the construct validity of direct tests of writing” (p. 71) because she finds few researchers set up studies that “demonstrate construct validity through empirical
evidence for the hypothesized set of behaviors and relationships, or engage in *a priori* construct validation at such a level of detail and specificity that the constructs described can be recognized in *a posteriori* behavior” (p. 71). Charney (1984) also argues that “the evaluation of writing ability should take the individual’s writing process into account, and the product-based methods of assessment, such as holistic ratings, are therefore invalid *a priori*” (p. 68). It is true that the results of holistic scoring may be useful for administrators and researchers who wish to simply rank order students based on their writing performance or to predict whether a group of students need to go to a remedial English class, but the results do not inform teachers of diagnosis or of the processes a writer experiences during writing. If we believe what Flower and Hayes (1981) claim that the student’s writing process provides the best information about how students write, it is obviously that holistic scoring will not provide any valid information about diagnosis of a student’s writing because the process is not part of the “construct” of the assessment.

Lauer and Asher (1988) suggest using the multitrait-multimethod approach to develop construct validity. They propose that “construct validity, which is typically determined in a multivariate situations, has two components: discriminant and convergent. Convergent validity is the general agreement among ratings where measures should be theoretically related. These ratings are gathered by methods that are maximally independent of one another. Discriminant validity is a low relationship among measures that theoretically should not be related” (p. 144). Indeed, this can be a very useful procedure for researchers to explore methods other than holistic scoring to either supplement the direct assessment of writing or to prove that other methods and holistic methods are in fact the same construct “if the variables (traits) intercorrelate about as high as their reliabilities will allow” (Lauer and Asher, 1988, p. 145).

Content validity will not be discussed because in holistic scoring the assessment does not concern whether the writing sample is representing “the universe” or a certain domain. The large variety of the con-
tents of writing samples make the domain universe almost infinite; thus content validity does not mean much in holistic scoring. Concurrent validity will not be discussed either in that it is the degree of relationship between a known, valid measure of behavior (rating systems, judgments, or tests) and a new measurement system. Since all well known multiple-choice testings (e.g., SAT, ITED (Iowa Test of Educational Development)) have been criticized as not able to reflect true writing ability, there is no well known, valid measure of rating system for holistic scoring to be concurrent with.

5.2.4 Are there negative effects caused by high reliability?

High interrater agreement may obscure the raters sensitivity to students with a unique usage of language or an imaginative style (Hake, 1986). For example, Lloyd-Jones (1977) illustrates one prompt given by NAEP, that is, “Some people believe that a woman’s place is in the home. Others do not. Take ONE side of this issue. Write an essay in which you state your position and defend it” (p. 60). The directions for scoring are as follows: “The writing receives a 0 score if the writer gives no response or a fragmented response; it receives a 1 score if the writer does not take a clear position, takes a position but gives no reason, restates the stem, gives and then abandons a position, presents a confused or undefined position, or gives a position without reasons;...it receives 4 if the writer takes a position and gives two or more elaborated reasons, one elaborated reason plus two or more unelaborated reasons, or four or more unelaborated reasons” (Freedman, 1991, p. 6). If the raters followed the directions closely and consistently, high interrater reliability would be expected. But what may happen to the students who do not take “one” position on a woman’s place but point out the complexity of the issue rather than taking one side? These students may receive 1 score but they also may write a substantially better essay than other students who receive 4 scores by taking one position with 5 unelaborated reasons. The raters are not wrong in tagging along the directions to obtain high interrater reliability, but they may pay the cost of
losing the sensibility to students’ “unexpected, unique, high” quality of writing.

5.2.5 How does an evaluator reach a judgment about a writing sample?

Different methods of measurement are designed for different purposes. For measuring written products, holistic scoring is probably the most efficient way to evaluate writing samples directly for the purpose of rank ordering students writing ability. For this specific purpose, the priority is to reach the highest interrater reliability one can. In other words, if we know how an evaluator reaches a judgment about a writing sample, we may have very substantial directions to train and guide raters so as to reach the goal of high reliability (See 5.2). Freedman and Calfee (1983) claim that the key to understanding how a person performs a complex rating task, following Simon’s (1981) reasoning, is to try to identify the relatively small number of relatively separable processes that underlie the person’s thinking for that task (Calfee, 1981). Three processes are essential in rating a composition: (1) read and comprehend the text to create a text image, (2) evaluate the text image and store impressions, and (3) articulate evaluation. They also claim that high interrater reliability relies on how similar the raters stored images are. They believe that for the research purpose, raters’ rating should reflect their internal decisions without considering writers’ responses. These three processes may be in an ordered sequence, especially in holistic scoring where speed is one of the important matters. These processes may also be recursive if raters have the habit of reading one paragraph, evaluating, and articulating their evaluation before they start reading the second paragraph (See 5.2).

Freedman and Calfee (1983) found empirically that these three processes can be independent of one another. To illustrate how these three processes are influenced by different subprocesses, they placed these subprocesses into two categories: the personal characteristics of the raters and the task environment (See Figure 5.1). Different situations
may have different emphases. For instance, if the result of evaluation is marked by grades or scores instead of verbal expressions, the subprocess of productive ability are not necessary.

Figure 5.1 Variables affecting the rating process (Freedman & Calfee, 1983, p.95)

5.3 Writing portfolios

We have discussed that large-scale assessment, either multiple-choice testing, holistic scoring, or analytical methods of evaluation, is not able to provide teachers and researchers with much information about writing from a developmental perspective (Brown, 1981). Some researchers, therefore, tried to use another approach: they start evaluating portfolios. Portfolios are in fact collections of student writing that is classroom-based and designed to provide information about student growth. Through these procedures, student progress in all aspects is revealed by patterns in behaviors over time (British National Writing Project, 1987; Dixon & Startta, 1986; Genishi & Dyson, 1984; Newkirk & Atwell, 1988). Unlike other types of writing assessments, raters, usually teach-
Evaluating Writing

Evaluators, who evaluate writing portfolios will involve the writers in the evaluation process (Burnham, 1986; Wolf, 1988; Simmons, 1990), discussing with them their ways of writing and their products. Students therefore are helped to formulate concepts of “good” writing, including the variability of “good” writing between situations and audiences (Gere & Stevens, 1985; Knoblauch & Brannon, 1984).

Like any other types of writing samples, writing portfolios also show limitations in some aspects. Freedman (1991) points out that “most portfolio projects lack guidance on several fundamental fronts: what writing is to be collected, under what conditions, for what purposes, and evaluated in what ways” (p. 8). Murphy and Smith (1990) also outline a set of questions that must be answered by anyone who designs a portfolio project: “Who selects what goes into the portfolio?” “What goes into the portfolio?” “How much should be included?” “What might be done with the portfolios?” “Who hears about the results?” “What provisions can be made for revising the portfolio program?” (p. 2).

These questions, however, do not discourage people who believe in the value of writing portfolios. The state of Vermont, for example, started the plan of assessing all students in grades four and eleven (within the United State). The plan includes three parts. (1) Students will write one piece of composition on an assigned topic or prompt (with limited time) which will be holistically scored. (2) With the help of their classroom teacher, students will select and submit a “best piece” from their classroom writing portfolio. This piece will be scored by the same teachers who evaluate the prompted sample. And (3) State evaluation teams will visit all schools “to review a sample of fourth and eleventh grade portfolios” (Vermont Writing Assessment: THE PORTFOLIO, 1989, p. 2). The teams will also look at the range of content, the depth of revision and the student’s willingness to take a risk. The basic ideas are that “scores from the prompted sample and the best piece will indicate each student’s writing abilities; the portfolios will give a picture of the school’s writing program” (cited by Freedman, 1991, p. 12).
Although evaluating portfolios is still a heuristic method, the Vermont plan provides a very good concept that, with the combination of timed and prompted writing samples and classroom-based portfolios, written products need not be restricted to the written products completed under testing and unnatural conditions. In addition, this coordinated plan promises to provide information about the development of individual students, about school programs, and about writing achievement in the state. Freedman (1991) suggests that the new trend of writing assessment should be the combination of holistic scoring and the evaluation of portfolios. But, she also warns that anyone who intends to conduct the portfolio project needs sophisticated designs to solve the problems mentioned by Murphy and Smith (1990).

5.4 Evaluating ESL writing

In general, the issues of evaluating ESL student writing are just like those of English natives’ writing. There are two main purposes of assessing ESL student writing: (1) to decide who should be admitted to enter certain academic institutes, and (2), once admitted, to decide who should be required to go to remedial classes for English writing. For the first purpose, the most well-known, world-wide, large-scale testing is TOEFL test provided by ETS. Because the original goal of the TOEFL test is to measure the general language proficiency of ESL students, the results of this test do not always satisfactorily predict ESL students’ academic achievements as well as their writing ability. Since 1986, the Test of Written English (TWE) has been included in the general TOEFL test but the result of TWE is reported separately from that of the general, multiple-choice proficiency test. For the second purpose, ESL students are often required to write a passage to a prompt under a timed testing condition (e.g., the English placement test for ESL students at the University of California at Santa Barbara). These writing samples are each scored holistically by two raters; then the results will help the ESL program decide who is to go to which linguistic training class. The basic procedures of assessing writing samples are in fact very similar
between what is used for ESL students and those for English natives. In other words, these similar procedures lead to similar issues.

The most commonly used method of directly measuring ESL writing samples is holistic scoring. As I have discussed in the previous sections, one of the most serious problems is to achieve high reliability (Jacobs, et al., 1981; Huot, 1990). To overcome this problem, some researchers suggest homogeneity of readers’ academic and experiential backgrounds (Follman and Anderson, 1967), while others suggest strengthening reader training (McColy, 1970). Most of these suggestions correspond to what researchers have done for assessing English natives’ writing samples (See 5.2).

When Hamp-Lyons (1990) reviewed the assessment issues of ESL writing, she pointed out that, just as the holistic scoring system lacks convincing construct validity in assessing English natives’ writing, it is equally true for the assessment of ESL writing.

As for the analytic studies, the most well known method is measuring the T-unit (Hunt, 1965) and correlating the results with holistic scores (Homburg, 1984). Although the mean length and the number of T-units are found to be highly correlated with scores of holistic assessment (Larsen-Freeman and Strom, 1977; Monroe, 1975), there exists an issue specific to ESL students, that is, the issue of the definition of error. Since the T-units counted should be error free, a number of T-units may fail to reflect the “true writing ability” of the ESL student. One who uses easy syntactic structures without elaborate meanings may be valued much higher than one who tries to use a more difficult syntactic structure, which may be unfortunately imperfect, despite elaborate reasoning. Under this situation rises the question: “How serious does the imperfection affect comprehensibility?”

---

17 T-unit stands for terminable unit, and it is the shortest grammatically complete segment that a passage can be divided into without creating fragments. It may have subordinate clauses, phrases, or modifiers attached to or embedded within it. For instance, a compound sentence is two T-units because it can be divided into two grammatically complete units without leaving fragments (Summarized by Isaacson, 1988).
Nas (1975) classifies errors into three levels according to their effect on comprehensibility. For each level, he outlines spelling, lexical, and grammatical errors that can be summarized as follows:

First-degree errors: Deviations from correct spelling or meaning or grammatical conventions that are minor so that readers have no trouble recognizing them;

Second-degree errors: Deviations from correct spelling or meaning or grammatical conventions that are interpretable with the help of context;

Third-degree errors: Deviations from correct spelling or meaning or grammatical conventions that make it impossible for readers to understand in anyway.

As we have discussed in Chapter four, not everyone tolerates errors to the same degree. To some people, errors are errors, no matter how slight they may be. To others, only the third-degree errors defined by Nas will be treated as errors. The individual difference of error acceptability may cause a large discrepancy in raters’ scores of either holistic scoring and T-unit measuring. Homburg (1984) applied these three classifications of error levels to one of his ESL studies and found that with different levels of errors taken into account, the mean T-unit discriminated the writing groups in different ways.

In addition, a T-unit does not reflect other important components of a written product, such as communicative competence (Hirsch, 1977), organization, cohesion, coherence (Ney, 1966), and other stylistic effectiveness (Hwang, 1991). Few ESL studies have ever investigated these component analytically as Freedman and Calfee (1983) have done for the writing samples produced by English-speaking college students.

Since ESL students are from all over the world, the procedure of assessing writing should always take cultural factors into account. If a topic assigned happens to be a taboo of a certain culture, the students from that culture will not be capable of elaborating on such a topic as well as students from other cultures, where such a topic happens to be a
common daily topic (e.g., criticizing government or political systems). This kind of effect may also influence readers’ judgments if they happen to be intolerant of a certain form of communication. Unfortunately these cultural elements are always too subtle to be measured explicitly. This may be another field that researchers who are interested in the issues of ESL writing assessment need to explore.
In this book, I have synthesized recent studies related to theories of writing developed on the basis of the following five perspectives: the composing process, the speech-writing relation, the reading-writing relationship, error analyses, and writing assessment. My main interest and goal was to probe the recent trend of L2 writing theories. I, however, consistently started each chapter with the discussion of L1 writing models or theories, and then, like most researchers and educators interested in L2 writing research, applied L1 models or theories to L2 writing acquisition. In this chapter, I will briefly summarize the theories discussed from the five perspectives, integrate them, and describe them as theories I myself will follow in my future research on writing.

Anyone learning to write has to follow the grammatical and rhetorical conventions of the target language so as to make the written form of such language fully communicative. To deal with these conventions, a certain set of mental processes must be involved. The general composing model designed by Hayes and Flower (1980) may serve as a universal model for both L1 and L2 writing in English. Besides, it involves a number of recursively operating processes (e.g., planning, translating, reviewing, and monitoring) that may function similarly in different written languages. Possible differences in the operation of these processes
may be ascribed to the genres of written forms, the different discourse types that fit in different writing conventions, or simply the diversity of individual writing habits.

In L2 writing, I assume that a writer also experiences such a general model during composing, but is interfered with by his L1 writing strategies or conventions to various extents, depending on how much he relies on his knowledge in L1 writing during L2 writing. Translating from L1 directly to L2 may or may not exist during L2 writing depending on how well the writer can command the target L2 without interference from L1. Krashen’s (1978) monitoring theory, which enables a basic L2 writer to sense deviations, hence, to be able to correct them during composing, may exist if such a writer is anxious to produce a composition without error. To conclude, any of these theories concerning composing processes contribute some, but not all, solutions to the complicated composing issues. All researchers as well as educators at this moment should not narrow their research or instruction to one specific theory, because each theory I mentioned above has its own potential in helping an L2 writer improve his writing.

Among all possible sources that enrich a writer’s knowledge of writing, the most powerful and efficient one is reading. All studies agree that through a writer’s active reading, his writing becomes better both in quantity and in quality. Smith (1983) suggested that students read as a potential writer, selecting texts that one expects to learn as a model, and acquire the writing conventions in that specific type of writing as one acquires natural, oral language (without anxiety and others’ expectations).

L2 writing instruction has been following this direction and using reading texts as guiding models for different goals of writing. In fact, reading plays a more important role in acquiring L2 proficiency because a person who tries to learn L2 usually does not have the natural intuition of the target language to judge whether or not what he writes makes sense. The most reliable source that helps him establish the capability of judgment is reading texts. Therefore, an L2 writing instructor should
select appropriate reading texts so that they not only serve as writing models, but also attract students to invest their time and energy in them.

I always believe that the reading-writing relationship is interactive in one way or another. Directional influences of reading upon writing may exist in a brief purposeful writing class (e.g., how to write a scientific report), but in the long run, what students have learned from reading, if internalized, will come back to enable such students to predict and construct the meaning of similar reading texts, hence, promote their comprehending ability. And this technique may be applied to writing again when these students learn how to read their own writing critically during revising.

The writing features that differ from those of speech found by linguists during the past decades account for the fact that basic writers tend to fail in distinguishing the different sets of conventions belonging to different types of communication forms—speech and writing. Students’ oral fluency, being acquired earlier than their fluency of writing, makes it common to transfer oral conventions directly to writing conventions, ranging from inflectional letters to idea units, so as to make written products deviate from ideal or expected forms. In addition, speech and writing may represent two different types of cultures which, accordingly, foster two different communities in which varied requirements of living are assumed. To move from speech to writing is more like accepting a brand new culture than simply learning a new set of conventions.

The relationships between speech and writing in L2 are basically in the same situation as in L1. What makes the issues more complicated in L2 writing is the tendency that L2 writers, while at basic or intermediate levels of L2 proficiency, bring in their writing not only the speech conventions from L2 oral cultures, but also their L1 conventions (both from speech and from writing). This tendency will probably stay till an L2 writer writes as (naturally as) a native speaker does, if he can.

L1 and L2 do not differ much in categorizing errors made by basic writers. These errors are also believed (by Horning, 1987) to share the same ascription to Selinker’s (1974) theory of interlanguage and fossilization. However, studies concerning L2 error analyses have developed a
new trend that focuses on measuring native speakers’ responses toward errors made by L2 learners. This new trend has been discussed in detail because it raises at least three problems that traditional error analysis studies do not deal with. (1) Who decides improper “usage” to be erroneous, a research, an instructor, or a large sample of population? (2) Can grammatical errors be distinguished clearly from semantic errors? Or are these two types of errors in effect two ends of one continuum? (3) In addition to communicative effect, do we also need to look at the background (e.g., education, profession) of those who identify the errors committed by L2 learners? Solutions for these problems can be anticipated if this trend continues to attract more studies to join the stream.

Writing assessment in this review focuses on holistic scoring and how this method can be improved. Holistic scoring has already had credits in that it provides high inter- and intra-reliability and face validity; does not need a long time to train raters; and can be used with large writing samples. If researchers or educators want to know only the rank ordering of writing samples, holistic scoring is undoubtedly the best method to assess written products. But if researchers want to measure the real quality of the writing samples, holistic scoring is limited in providing such information. Without the information of specific traits, holistic scoring also fails to serve as a tool for developmental studies. Freedman (1990) suggests using analytic methods to observe each writing sample through different components and see how each component contributes to the total scores each writing obtains. In other words, this analytical method provides better construct validity in interpreting how each component affects writing assessment. Freedman also suggests combining the holistic scoring and portfolio projects to investigate how consistent students can produce their writing samples under different environments. In general, holistic scoring is still considered the most promising method in assessing written products.


Frith, U. (1978) From print to meaning and from print to sound, or how to read without knowing how to spell. *Visible Language, 12*, 43-54.


Huot, Brian (1990) Reliability, Validity, and Holistic Scoring: What We Know and What We Need to Know. College Composition and Communication, 41(2), 201-213.


Leyba, Charles F. *Longitudinal study, Title VII bilingual program, Santa Fe public schools, Santa Fe, New Mexico*. Los Angeles, California: National Dissemination and Assessment Center, California State University, Los Angeles.


Schneider, V. (1971). A study of the effectiveness of emphasizing the teaching of reading skills to improve composition skills in remedial English classes at Kansas City Community Junior College.


Accidental errors, 66
Agnello, 75
Akinnaso, 23, 24
Analytic system, 87, 89
Anderson, 97
Andreouch, 41
Applebee, 1, 6, 46
Arndt, 7
Aronoff, 21
Asher, 84, 91
assessing writing, 81–99
assessment, 81–99
Atwell, 95

Bartholomae, 72
Bauer, 84
Beaman, 22, 24
Beck, 46
beginner writer, 42, 68, 69
Belanger, 41
Bereiter, 6, 9, 10
Berkowitz, 42
Biber, 23, 24, 25, 26, 27, 32
bidirectional hypothesis, 40, 48
Bizzell, 2
Blankenship, 23, 24
Block, 59
Bloomfield, 21
Blurred patterns, 67

Boruta, 63
Brannon, 95
Braungart-Bloom, 85
Bridgeman, 84, 90
British National Writing Project, 95
Britton, 44
Brooks, 9
Brown, 23, 84, 88, 89, 94
Burnham, 95
Burt, 77

CA, 74
Calfée, 81, 83, 88, 89, 93, 94, 98
Calhoun, 38
Campbell, 38, 58, 59
Carleton, 82
Carlson, 84, 90
CEEB, 82, 83
Celce-Murcia, 76
Chafe, 22, 23, 24, 27, 28, 29, 63, 64, 65, 66
Charney, 81, 83, 90
Chastain, 77
Chomsky, 37
Chomskyan innateness hypothesis, 75
Coffman, 82
cognitive, 2, 5, 9, 10, 27, 29, 30, 31, 32, 44, 45, 48, 51
cognitive processes, 2, 5, 45, 51
cohesion, 24, 42, 86, 98
College Entrance Examination Board, 82
Collela, 84
Collins, 43
comma, 65
Common errors, 69
common underlying proficiency (CUP), 57
composing process, 101
composing processes, 1–20
comprehending, 43, 44, 45, 46, 47, 48, 58, 103
Conlan, 82
consciousness, 19, 27, 28, 29
Consolidation errors, 67
contrastive analysis, 74, 80
Cook-Gumperz, 22, 29, 30, 31, 32, 34
Cooper, 83, 84, 85
Corder, 62, 75
Costello, 58
Cowan, 75
Criper, 84, 90
cultural factors, 15, 19, 25, 29, 30, 31, 32, 49, 58, 98, 99
Cumming, 16
Cummins, 57

Danielewics, 63
Danielewicz, 23, 64, 65, 66
Davidson-Mosenthal, 84
Davies, 84, 90
Davis, 83
de Beaugrande, 88
decontextualized, 23, 34
Delisle, 78
derivational suffixes, 70
Diaz, 8
Diederich, 82, 83
differences between spoken and written discourse, 21–35
directional hypothesis, 40
displaced consciousness, 27
Dixon, 44, 51, 52, 54, 56, 62, 95
Donelson, 39
Drum, 81
Dyson, 6, 95
editing, 2, 5, 6, 11, 12, 31, 32, 43, 63, 72
Educational Testing Service, 82
EFL, 14, 32-35, 57, 60
EFL students, i, ii, 32, 33, 60
Eisterhold, 40, 44, 48
Emig, 2
error analysis, 61–81
error gravity, 78
Errors beyond the sentence, 71
ESL, 9, 33-35, 57-59, 60, 72, 74, 79, 80, 96-98
ESL students, 9, 33, 34, 35, 58, 59, 60, 72, 79, 80, 96, 97, 98
Ethnography, 19
ETS, 82, 83, 96
evaluating, 4, 43, 94, 96

Faigley, 83
Farella, 57
Felland, 39
Field, 24
Fillmore, 21
First-degree errors, 98
Fishco, 38
Florio, 6
Flower, 2, 3, 4, 5, 6, 12, 14, 18, 19, 91, 101
Follman, 97
fossilization, 73, 80, 104
Frase, 55
Frederiksen, 88
Freedman, 81, 82, 83, 84, 85, 88, 89, 92, 93, 94, 95, 96, 98, 104
French, 16, 77, 82
Frielderter, 13, 16
Frith, 50

Galvan, 15
Gaskill, 7, 18
General Impression Marking, 85
Generating, 2, 4, 5, 7, 10, 14, 16
Gere, 70, 95
Giwuff, 85, 86
Global errors, 77
Glover, 43
Goodman, 42, 49, 50, 51, 52, 54, 55, 56
Goody, 22
Grabe, 26, 32
Grimm, 21
Grobe, 38
Gumperz, 22, 23, 24, 29, 30, 31, 32, 34
Hafiz, 58
Hake, 85, 92
Hall, 7
Halliday, 22, 24, 42, 54
Hamp-Lyons, 84, 90, 97
Handwriting errors, 63
Hasan, 42
Hayes, 2, 3, 4, 5, 6, 7, 12, 14, 18, 19, 91, 101
Heil, 39
Hiebert, 48
Hilderbrand, 8
Hildyard, 28
Hillocks, 5
Hirsch, 98
Ho, 85, 86
Holistic, 47, 58, 83-98, 104
Homburg, 97, 98
Horning, 22, 32, 62, 63, 72, 73, 74, 80, 103
Horowitz, 58
Huddleston, 82
Hudson, 84
Hunt, 97
Huot, 97
Hwang, 77, 78, 98
Integrators, 59
Interactive model, 50, 58
Interference errors, 75, 76
Interlanguage, 72, 73, 80, 103
Intralingual errors, 75, 76
Inversion errors, 67
Iowa Test of Educational Development, 92
Isaacson, 84, 85, 97
It-deletion, 79
ITED, 92
Jacobs, 8, 97
Johnson, 16, 39
Jones, 8, 9, 10, 11, 17, 44, 51, 52, 54, 56, 62, 92
Keller-Cohen, 84, 90
Kintsch, 88
Kiparaky, 77
Knoblauch, 95
Krapels, 12, 18, 19
Krashen, 16, 35, 102
Kucer, 48
Kulhavy, 43
L1, 7-19, 56-62, 72-75, 101-103
L1 Composing process theories, 7
L2, 7-19, 56-61, 72-76, 80, 101-103
L2 Composing processes, 7, 12, 16, 18, 19
Lado, 75
Lakoff, 32
Langer, 6
Language transfer, 73-75
Lanier, 45
Larsen-Freeman, 97
Lauer, 84, 91
Lay, 15, 16, 18
Levin, 63
Leyba, 57
Linear, 2, 6, 12, 32, 47
Liu, 14, 85, 86
Lloyd Jones, 87
Lloyd-Jones, 84, 92
Loban, 38
Local errors, 77
Lomax, 50
Lorenz, 77, 79
Index

LTM, 3
Lucas, 84, 89
Ludwig, 76

Maloney, 38, 39
Martín-Betancourt, 13
Matsushashi, 6
Matt, 40
McColly, 97
McConville, 82
Mellon, 16, 84
mental processes, 5, 29, 48, 101
Meyer, 77, 79
Meyers, 82
MF/MD, 26, 27
Miller, 51, 53
Mills, 41
Monitor, 4
monitoring theory, 102
Monroe, 97
Moran, 84
Mosenthal, 84
multi-feature/multi-dimension, 26
multi-functional continuum, 24
multitrait-multimethod approach, 91
Murphy, 95, 96
Myers, 83, 89

NAEP, 87, 92
Nas, 98
National Assessment of Educational Progress, 87
Neilson, 22, 32
Nessel, 44, 51, 52, 54, 56, 62
Neuman, 75
Newkirk, 95
Nisbett, 5
Nold, 6
nondirectional hypothesis, 40, 44
nonintegrators, 59

O’Donnell, 22, 23
Ochs, 23
Olson, 23, 28
organization, 16, 59, 77, 88, 89, 98
organizing, 4

Paulin, 87
period, 66
Pfingstag, 18
Philips, 39
Pianko, 6, 11
Piazza, 77
Planning, 4, 5
Politzer, 78
Poole, 24
portfolio, 90, 95, 96, 104
portfolios, 94–96
pre-writing, 2, 68
Primary Traits, 87, 89
Punctuation errors, 64
Purvis, 84

Read, 37, 50, 51, 62, 94
reading and writing, 37–60
reading-writing relationship, 37–60
recursive, 5, 47, 93
relative clauses, 79
Reviewing, 4, 5
Richards, 75
Riel, 63
Robinett, 75
Rohman, 2
Rorschach, 17
Rose, i
Rosier, 57
Rowe, 63
Rubin, 23
rubric, 85, 86, 87

Santos, 78
Sapir, 21
SAT, 59, 92
Scardamalia, 6, 10
Schachter, 75
Schneider, 40
Scriven, 83
Second-degree errors, 98
Selinker, 72, 73, 75
Sentence structure, 89
setting up goals, 4
Shanahan, 48, 49, 50
Shaughnessy, 61, 62, 63, 64, 66, 67, 68, 69, 70, 71, 73, 74
Simmons, 84, 89, 95
Simon, 88, 93
Smith, 42, 43, 45, 51, 52, 53, 54, 55, 56, 62, 95, 96, 102
social factors, 30
Sommers, 6, 11
Spack, 58
speech, 21–35
Spoken language, 21–35
Squire, 44, 45, 46, 47
Sridhar, 62
Startta, 95
Stein, 46
Stevens, 70, 95
Stiggins, 84
Stotsky, 37, 40-42
Strom, 97
Stubbs, 23
Substitution, 70
Sulzby, 31
Sweet, 21
syntactic complexity, 37, 39, 68
Syntactic errors, 66

Tannen, 23, 24
task environment, 3, 94
Taylor, 38, 41, 42
tense, 79
Test of Written English, 96
Tetroe, 9
thinking-aloud protocols, 5, 7, 10
Third-degree errors, 98
Thomas, 39, 83
TOEFL, 96
Tomiyana, 77
Translating, 4, 5, 14, 15, 102
translation, 5, 12-16
Tudor, 58
T-unit, 39, 97, 98
Tuss, 85, 86
TWE, 96

Urzua, 11
Vann, 77, 79, 80
Veal, 84
Vermont Writing Assessment, 95
Vocabulary errors, 70
Wardhaugh, 74, 75
Watkins-Goffman, 59
Watt, 22
Weeks, 75
Wesdorp, 84
White, 83
Wilson, 5
Witte, 2, 6, 19
Wlecke, 2
Wolf, 95
Wolfe, 84, 90
Woodward, 39
word choice, 79
word order, 79
writing processes, 4, 6, 7, 10, 12, 13, 17, 19, 43, 54
writing relationship in L2 acquisition, 57
written language, 21–35
Young, 59, 62
Yule, 23
Zamel, 7, 11, 13, 18
Zeman, 39
Zimmer, 85-87