THE RELATIONSHIP BETWEEN SPEAKING AND

WRITING IN COLLEGE-LEVEL ESL
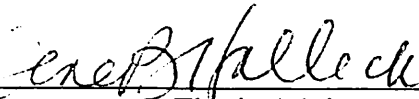
STUDENTS

By

XINHUA ZHU

Master of Arts
Beijing University of Mining and Technology
Beijing, China
2001

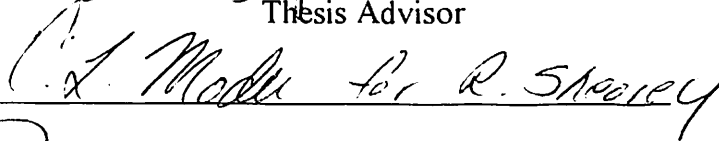Bachelor of Arts
Ningxia University
Ningxia, China
1998

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF ARTS
May, 2004

THE RELATIONSHIP BETWEEN SPEAKING AND

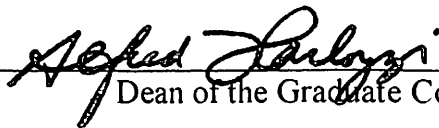# WRITING IN COLLEGE-LEVEL ESL

STUDENTS

Thesis Approved:

_____
Thesis Advisor

_____

_____

_____
Dean of the Graduate College

ACKNOWLEDGEMENTS

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I

INTRODUCTION

Is speaking similar to or different from writing? Does speaking practice improve writing or vice versa? Obviously, these questions are concerned with the relationship between speaking and writing. Further, does the relationship between speaking and writing develop in the same way in first (L1) and second language (L2) acquisition? Answers to these questions would provide insights into educational practice in both L1 and L2 acquisition. This study is, therefore, intended to shed light on the relationship between speaking and writing with particular attention paid to adult L2 acquisition.

A number of studies have been conducted on the relationship between speaking and writing development in L1 acquisition since the 1920s. These studies have provided valuable information for L1 instructors to facilitate their students' language acquisition and help their students to make smooth transition from orality to literacy. Specifically, these studies have investigated the similarities and differences between speaking and writing. Although L1 researchers have different views on whether these two modes of expression are alike or not, the review of literature reveals that most researchers hold that spoken and written languages are obviously different from each other. Traditionally, linguists such as grammarians (e.g. Schafer, 1981) considered writing as the primary mode of communication in comparison to speaking. However, linguistic research (e.g. Emig, 1977) has been recently guided by the assumption that spoken language is primary

compared to written language. Still, some linguists challenged these two views and held that both languages were equally important (e.g. Vachek, 1973).

Although prolific research has been conducted on the relationship between speaking and writing in L1 acquisition, few studies have touched on this relationship in L2 acquisition. Traditionally, L2 instructions separated speaking from writing mainly due to the influence of structural linguists, who believed in the primary importance of spoken language (Mangelsdorf, 1989). Consequently, writing was always considered as a complementary way to reinforce the instruction of speaking. Only in recent years have L2 researchers come to realize the importance of writing and the relationship between speaking and writing. Besides, the lack of the investigation into this relationship would make much of the classroom methodology be based on trial-and-error (Vann, 1979). Further, there is a discrepancy between teacher belief and learner belief in L2 acquisition and learning on the primacy of speaking and writing. Because of the lack of research on this issue in L2 acquisition, L2 instructors have to turn to L1 acquisition research for information on this relationship for their teaching strategies. Nevertheless, Vann (1981) states that in second or foreign language acquisition, many questions need to be answered concerning this relationship. For example, "How much transfer of learning can we expect from one language skill area to another?" "Will fluent speakers make good writers?" "Why do some students have particular difficulty with one mode or another? (p.167). Likewise, Kim (2000) points out the stages in L1 and L2 acquisition in general might not be the same. She asks the following questions: "Are the stages in L1 and L2 acquisition the same?" "Do speech and writing development occur simultaneously?" and "What is the role of monitor in L2 speech and writing?" (p. 73). Thus, research is much needed to

explore this relationship to enrich L2 acquisition theories and provide guidance for L2 instructors when they practice their teaching in classrooms.

The purpose of the present study is to examine the relationship between speaking and writing in college-level ESL students. In order to explore this relationship, the present study has used the measures of syntactic maturity and lexical density to investigate the syntactic and lexical development of speaking and writing.

The first chapter reveals the purpose of this study. The second chapter reviews the literature on the relationship between speaking and writing in both L1 and L2 acquisition, syntactic maturity and lexical density in L1 and L2 acquisition at great length, and the use of the Video Oral Communication Instrument (VOCI) in L2 acquisition. The third chapter describes the methodology used in this study to explore the relationship between speaking and writing in college-level ESL students. The instrument used in this study is the VOCI developed by Halleck and Young (1995). The measures of syntactic maturity and lexical density were adopted to analyze both the spoken and written data of this study. The fourth chapter presents the results and discussion of this study and the final chapter concludes the whole study.

CHAPTER II

REVIEW OF LITERATURE

First Language Acquisition

The relationship between speaking and writing has been successfully and widely studied in L1 acquisition since the 1920s. Most of these studies have been conducted to see whether these two modes of expression are alike or different (e.g. Harrell, 1957; Gleason, 1965). Some researchers place emphasis on the connections between speaking and writing (e.g. Rubin, 1975) while others stress the differences between these two modes of communication (e.g. Vygotsky, 1962). Most research in L1 acquisition, however, has found more differences than similarities between these two modes. The significance of the research on this relationship lies in that it provides important information for classroom discussion and research. In relation to the present study, the review of literature in L1 acquisition may offer some insights into the research in L2 acquisition.

According to the research in L1 acquisition, speaking and writing seem to have a reciprocal influence on each other. Children can build strong written communication skills by drawing on their oral language skills (Tough, 1977). Writing can help children to clarify their oral expression after they have practiced what they intended to express through writing (Lundsteen, 1976). Because of the mutual benefits of speaking and writing for each other, it is necessary to review the literature on their reciprocal influence.

Since the methods of syntactic maturity and lexical density are used to conduct the present study, the studies on syntactic maturity and lexical density have also been

reviewed to investigate their validity as measurements of the relationship between speaking and writing and specific indices within each method.

Therefore, this section will be devoted to the discussion of the following aspects: the relationship between speaking and writing in L1 acquisition, the effects of speaking on writing, the effects of writing on speaking, syntactic maturity in L1 acquisition and lexical density in L1 acquisition.

The Relationship between Speaking and Writing in L1 Acquisition

The relationship between speaking and writing has been a complex issue. More than eight decades ago, Woolbert (1922) drew attention to the relationship between speaking and writing, specifically to whether these two modes of expression were alike or different and commented that the relationship between these two modes had never been adequately stated. Following Woolbert's call, researchers began to conduct research on this particular issue. Some researchers believe that speaking and writing are similar in many ways (e.g. Rubin, 1975). Whereas, others hold that oral and written languages are two different modes of communication, in which written expression is not simply speech written down (e.g. Harpin, 1976). Of these two views, most research on the relation between oral and written language has been focused on the differences between these two modes of communication rather than the similarities (Cambourne, 1981). In the following section, I will discuss the similarities and differences identified by those researchers with a focus on the differences between these two modes so as to provide a better understanding of this relationship in L1 acquisition.

Very few studies have been devoted to the similarities between speaking and writing in L1 acquisition. And these studies have simply stressed the connections

between speaking and writing, but not dealt with the specific similarities between these two modes of expression. Rubin (1975) pointed out that oral language was closely related to written language and written expression was simply speech written down. Likewise, Cramer (1978) emphasized the positive influence of speaking on writing because of the derivation of written language from oral language and claimed that the connections between spoken and written language should be maximized through the teaching of reading and writing. Further, Kroll (1981) proposes a changing relationship between speaking and writing, where similarities and differences between spoken and written languages are emphasized in different phases of development.

Most researchers in L1 acquisition seem to favor the emphasis of differences rather than similarities, because "the written mode is more abstract, secondary form of the language than the oral mode; written language is more difficult to learn and requires special learning conditions" (Cambourne, 1981, p.82). Most of these studies are aimed to help students promote their writing skills. In addition, some studies have examined this relationship with regard to children while others have investigated it insofar as adults are concerned. As noted earlier, the research on this relationship in L1 acquisition could be traced to the 1920s when Woolbert called for attention to this particular issue. Bushnell (1930) studied both the written and spoken samples of 10th graders and found that written themes were richer in thought content and sentence structure and contained fewer grammatical errors.

Although 1930s until mid-1950s did not see much work on this issue, two studies are worth mentioning here. In 1944, Fairbank and Mann conducted two interrelated studies with Fairbank's study focusing on spoken language and Mann's study on written

language. Their results showed that the type-token ratios (the ratio between all the running words [tokens] and all the different words [types]) were higher in writing than in speaking and writing was found to be higher in nouns, adjectives, prepositions, and articles and speaking higher in pronouns, verbs, adverbs and interjections.

Since the mid-1950s, prolific studies have been carried out on this topic. Harrell (1957) compared written and oral narratives of 320 children from nine to fifteen years of age and found the use of more subordinate clauses in these children's writing than in speaking and the use of more adverbial and adjectival clauses in writing than in speech, but the use of more noun clauses in their oral stories. Like Harrell, Blankenship (1962) also compared grammatical characteristics of speech and writing by analyzing the published articles of four persons and their speeches delivered on a university campus and noted little variation in sentence length, some variation in sentence patterns and some differences in the use of word classes in various positions, more transitive verbs in speech and more passive constructions in writing. Newman & Milton (1965) analyzed spoken and written discourse samples produced by college students with the control of topic and time and observed that speaking produced more words and ideas of all kinds and was more repetitious than writing. Speaking seemed to represent more serious commitment to precise language and rational, deliberate thinking.

Similarly, in order to determine more clearly the differences and similarities between oral and written styles, Gibson et al. (1966) compared the spoken and written samples of forty-five beginning speech students by computing three kinds of scores: Flesch Reading Ease (mean sentence length and average number of syllables per 100 words), Flesch Human Interest (the number of the words referring to people and

sentences directly addressed to the reader), and Type-Token Ratios. Their findings indicated that speech was significantly more readable, more interesting with a simpler vocabulary than writing. DeVito (1966, 1967), who was the most productive writer on this subject, studied the differences between oral and written discourse in vocabulary and in level of abstraction by analyzing language samples from ten university faculty members and noted that speech differed from writing in that it contained more words that referred to the speaker, more indefinite quantifying words (e.g. *much, many*), more "allness" terms (e.g. *all, some*) (p.355), more qualification terms (e.g. *if, but*) and more terms that indicate opinion (e.g. *seems, appears*), more finite verbs and fewer nouns of abstraction.

In addition, Britton (1970) noted the differences between the process of speaking and the process of writing and pointed out that the writing process was much more controlled than the speaking process in the sense that people could take more time to plan and organize their thoughts in writing than in speaking. In terms of the functions of these two modes of communication, Halliday (1973) stated that spoken language functioned in an interpersonal way while written language did so in an ideational way. By this, he meant that speaking served the functions of social interaction, which was characterized by conversation, storytelling, verse and song, while writing functioned to package ideas and information, which took the form of statement, argument, and detailed explanation. Poole & Field (1976) compared the spoken and written samples of Australian students from middle-class and working-class backgrounds and noticed that spoken language was more complex in terms of embedding, with more adverbs and personal pronouns, whereas the written language had more adjectives and complex verb structures. Emig

(1977) summarized the differences between speaking and writing in great detail. For example, talking is natural while writing is an artificial process, which is a learned behavior. Talking is much dependent on the environment, whereas writing must provide its own context, to name only a few. Silber (1979) pointed out that accurate diction and sufficient details were required in writing, while the matter of an entire paragraph could be conveyed by a gesture or facial expression in speaking. In addition, Silber also claimed that one had to clearly identify syntactic boundaries and junctions in writing, whereas the surface structure of spoken utterance was often inadequate to produce the deep structure accurately. For example, the spoken language often omits subjects, verbs, or other elements, which results in sentence fragments and comma splice. She further stated that information had to be ordered in terms of chronology, cause and effect, or similar patterns to be readily conveyed in writing. But, one could go back and forth when talking about a topic.

Compared to the research of prior decades that had focused on the similarities and differences between the two modes, research in the 1980s began to look at this relationship from different perspectives rather than the mere comparison of the two modes. To illustrate, the following studies adopted different approaches to study the relationship between speaking and writing rather than the mere comparison of the two modes. Schafer (1981) considered the distinction between dialogue and monologue rather than the difference between speaking and writing and claimed that the opposition between dialogue and monologue were more valid than the opposition between writing and speech. Chafe (1982, 1985) introduced the concept of discourse dimensions, namely, fragmentation/integration and involvement/detachment. His findings showed that written

language generally contained more complex syntax and a richer, more varied, and more formal vocabulary than did spoken language. He attributed this difference to the involvement between interlocutors in speech and the integration of ideas in writing. In particular, written language is featured with greater integration than spoken language because it employs the different devices to compact idea units such as nominalization, complement clauses and so on. On the other hand, in spoken language, the speaker demonstrates the involvement with his or her audience. For instance, the spoken language contains "more references to the speaker, references to the speaker's mental processes, devices for monitoring the flow of information, the use of emphatic particles, fuzziness, and the use of direct quotes" (p. 48). Hidi & Hildyard (1983) reported another different approach, claiming that it was the genres rather than the modality that made the differences in both spoken and written production. They examined three materials produced by children in grades 3 and 5: their spoken and written samples on the same topic and a narrative. Their analysis showed that clear differences existed between the two genres (opinion and narrative) but no differences were found between speaking and writing.

In 1980s, the most extensive quantitative investigation of spoken-written language differences is the research conducted by Biber (1986). His study indicates that there is no single, absolute difference between speaking and writing in English, but there are a number of dimensions, namely, Interactive vs. Edited Text, Abstract vs. Situated Content, and Reported vs. Immediate Style. Specifically, Haynes (1992) provides a detailed explanation of Biber's dimensions (pp. 50-51):

1. Edited versus interactive text: Edited text is concise, possibly indicating more planning than interactive text. Features that indicate interaction are "characterized as verbal, interactional, affective, fragmented, reduced in form, and generalized in content" (Biber, 1988, p. 105).

2. Abstract versus situated content: Abstract discourse focuses on ideas or thoughts and is "semantically complex"; often the "active agentive participant" is lost, and this results in the "promotion of a more abstract concept" (Biber, 1986, p. 395). Situated discourse refers "directly to an external situation" and is more concrete than abstract discourse (Biber, 1986, p. 396).

3. Reported versus immediate style: Reported discourse refers to a removed situation and is narrative in nature. Immediate discourse has little reference to a removed situation.

Biber mentions contradictory findings in other studies on the relationship between two modes, which he attributed to the imbalance in the text types being compared, and to the use of linguistic features belonging to different textual dimensions. To illustrate, one such contradictory finding is Chafe's (1982) study, which used conversation and academic prose respectively to serve as spoken and written samples. However, these two text types are represented in opposite dimensions of Interactive vs. Edited text, and Abstract vs. Situated content. As Biber claimed, "if conversation is taken to represent speech, and academic prose to represent writing, then most linguistic features considered in previous research could be presented as evidence for a spoken/written distinction" (p. 408). However, Biber also points out that individual findings about text types of the previous

studies have been generally valid. Nevertheless, the global conclusions should not be based on them, because "the text types chosen for comparison were too similar or too different and the linguistic features chosen belonged to different textual dimensions" (p. 409). Therefore, he advocates that future studies should attend to the text types and linguistic features chosen to investigate the relationship between speaking and writing. Research of this kind has been very quiescent since the 1990s. Very few studies have been conducted on the relationship between speaking and writing.

With regard to Biber's findings, it seems that it would be meaningless to compare spoken and written languages because of the overlap of some speaking and writing styles. However, Chafe & Tannen (1987) claimed that there is still much to learn about "the cognitive and social processes most typically associated with speaking on the one hand, or writing on the other and about the effects these processes are most likely to have on the language itself" (p.391), although no single feature or dimension can distinguish all of speaking from all of writing. In addition, despite the differences between spoken and written language, Tannen (1982) emphasizes that they are not dichotomous, but rather a continuum, which are "superimposed upon and intertwined with each other" (p.3). It is not in absolute terms for any piece of discourse to be characteristic of spoken or written language, but rather in terms of degree.

Traditionally, linguists have considered writing as the primary mode of communication compared to speaking. For example, as claimed by Emig (1977), writing is a more valuable mode of learning than speaking, since "writing is 'self-rhythmed', represents 'a powerful instance of self-provided feedback,' and 'establishes explicit and systematic conceptual groupings' (as cited in Schafe, 1981, p. 23), to name but a few.

However, linguistic research has been recently guided by the assumption that spoken language is primary compared to written language. According to de Saussure (1916, 1959) and Bloomfield (1933), spoken language took the primary role in the sense that it was older and more widespread. In addition, they noticed that all three of the common systems of writing were based on different units of spoken language, namely, the sound-based alphabetic system, the syllable-based syllabic system and the word-based ideographic system Schafer (1981). Because of the first presence of these units in spoken language, Lyons (1968) concluded that spoken language was the primary while written language was derived from spoken language.

Still, some linguists challenged both views and held that spoken and written languages were equally important. Vachek (1973) contended that neither speaking nor writing was primary, rather there were functionally complementary. By this he meant that speech was employed to meet a society's communicative needs better in some situations than writing while writing was better preferred in other situations.

The Effects of Speaking on Writing

There are two views concerning the effects of speaking on writing. Some researchers, especially composition theorists believe that speaking may interfere with the development of writing, which they term as "oral language interference" while others (e.g. Schafer, 1981) hold that oral language in some ways does contribute to writing. It seems that most researchers concur with the second view that speaking does have a positive effect on writing.

Oral language interference has been a focus of discussion among the researchers in the study of the effects of speaking on writing. Based on Bernstein's (1971) distinction

between context-dependence and context-independence, some composition theorists have tried to explain three common errors in the essays of inexperienced writers: coherence gaps, vague pronoun reference, and premature closure on a point. Since typical oral text depends very much on the situational or cultural contexts compared to typical written text, they attributed these errors to oral language interference, "a transfer of communication strategies that work in most speech situations into writing situations" (Schafer, 1981, p. 20). However, Schafer argues that most errors cannot be attributed to oral language interference, although it does have an effect on some errors inexperienced writers have made. On the other hand, particular oral language interference could help writing. Sometimes, writing may be more effective if students can employ some of the "liveness" (p. 31) of their conversation in their writing. Zoellner (1969) contended that it was the disassociation of the two modes rather than oral language interference that resulted in poor writing. Similarly, Groff (1979) argued that little research had been conducted concerning the influence of oral language on the growth of written language competence. Therefore, oral language interference might not be a valid factor with respect to the effects of speaking on writing.

Tough (1977) commented that children can build strong written communication skills by drawing on their oral language skills, since talking to some extent is the basis for writing. In line with Tough, Kroll (1981) proposes a developmental model for the changing relationship between speaking and writing, which consists of four phases as mentioned earlier. They are 'separation', 'consolidation', 'differentiation' and 'integration'. According to this model, children do benefit from talking in developing their writing skills especially during the first two phases. Specifically, during the first

phase, speaking and writing to children are essentially separate, since children have very minimal writing skills. One factor she proposes to improve children's writing during the first phase of the model is that children may extend their oral communication skills to their writing. The second phase concerns consolidation of oral and written language skills. In this phase, children's writing to a large extent depends heavily on their spoken language repertoire with writing very much like talk written down. As Rubin (1975) stated, children's writing reflects their speaking. During this phase, it is very important for children to learn to draw on their oral language repertoire to produce writing. To promote children's writing, Kroll proposes three important elements in this stage, each of which is concerned with the use of children's oral language skills. The first element is that children need to continue developing their oral language skills, since writing is derived from speaking (Smith, Goodman & Meredith, 1976). Lundsteen (1976) shared the same view and suggested that it would be beneficial for the teacher to help children with poor writing abilities to work on their oral language. A second element Kroll proposes is that teachers need to make the forms and functions of speech and writing as similar as possible in language activities. In a specific manner, children can be involved in producing fewer conversation-like oral utterances such as oral monologue, since writing in a sense is a monologue because the writer shoulders the sole responsibility of getting the message across. In addition, children can be encouraged to produce the kind of writing that is closer to speaking such as personal writing because this kind of writing contains expressive language, which can help children to make the easy transition from speaking to writing (Fisher & Terry, 1977). The third element is that speaking can be employed to prepare children for writing assignments such as oral discussion prior to

writing, which is generally agreed upon by language arts specialists (e.g. Cramer, 1978; Petty, Petty, & Becking, 1976; Burton, Donelson, Fillion, & Haley, 1975 and Marcus, 1977). Harpin (1976) commented that beginning writers "are likely to be aided by the opportunity to rehearse in speech their own ideas and to overhear the thoughts of others" (p.135). In sum, according to Kroll's model, children can draw on their oral language resources to gradually promote their writing abilities.

From the above research, it seems that speaking has a positive effect on writing. However, Harpin cautions that oral preparation before writing is only a general principle; it should not be taken as a universal or habitual practice, because he notes that with regards to creative writing, the piece of writing with verbal preparation is less mature than the writing without such preparation. In a similar vein, Purcell-Gates (1989) points out that the overemphasis on improving children's speaking skills in order to promote their writing abilities might be potentially confusing to children, especially those with limited experience with written language, because written language uses different vocabulary and syntax and is rather decontextualized in nature. Accordingly, she advocates the important role of reading in bridging the gap between speaking and writing, since reading can help those children to learn the differences between speaking and writing. Specifically, she suggests that teachers should encourage students to produce relatively formal speech such as "pretend reading" (p.292) so as to enhance their writing. In particular, the children may be encouraged to pretend to read a self-chosen book. In this way, the teacher can assess whether the children have learned the vocabulary and syntax of the written language.

Because of the fundamental differences between speaking and writing, some researchers have attempted to find ways to help students to make a smooth transition from speaking to writing without feeling frustrated as in traditional writing classes. DeVries (1970) introduced The Oral-Aural-Visual Stimuli Approach to Teaching Written Composition in elementary school, which used oral-aural-visual procedures to teach written composition to students. This approach is based on the theoretical idea proposed by modern linguists that talking should serve as a prelude or the basis to and for the writing as a natural mode of expression. Snipes (1973) proposed an oral approach to the composing process: a talk-retalk-write-rewrite method and claimed that the method has many advantages over the standard prewriting-writing-rewriting process. Specifically, this method is student-centered and audience-oriented, which offers the student an opportunity to study his or her use of the language and play with the language to experience self-discovery. As indicated by its name, this method consists of four stages, the talking, retalking, writing, rewriting stages. With a tape-recorder, students first talk to themselves on the topic or converse with classmates in the talking stage. During the retalking stage, students reorganize their ideas and thoughts by selecting the ideas they wish to present, choosing a particular pattern or order through which they wish to present these ideas and recording these ideas through this patterns. After that, students write a first draft by taking down exactly what they said in the last recording. Finally, students revise their draft by incorporating more carefully selected details, clearer examples, and illustrations and so on.

The Effects of Writing on Speaking

Although a significant amount of research has been carried out on the positive effects of speaking on writing, very few studies have shed light on the effects of writing on speaking. In fact, some researchers believe that a reciprocal relationship exists between speaking and writing. In other words, writing and speaking, as two modes of expression, mutually benefit each other. Just as appropriate speaking practice can enhance and promote students' writing skills, appropriate writing practice can enhance and promote students' speaking skills.

Greenfield (1972) attributed the context-dependent speech produced by some people to the lack of exposure to written texts and stated that people practiced using linguistic contexts as independent of immediate reference when writing. So, Greenfield claimed that learning to write would improve a student's capacity to produce context-independent speech and enhance his or her ability to think abstractly and logically. Lundsteen (1976) agreed that writing could help children to clarify their oral expression after they have practiced what they have intended to express through writing. Cramer (1978) also shared the same view and claimed that writing could promote the growth of oral language. Donovan (1986) proposes that speaking and writing should be taught together on the assumption that they are both modes of communication and equally important. Specifically, as modes of communication, they both have to be adapted to an audience to meet some sort of audience expectation.

Overall, research in L1 acquisition shows that speaking and writing as two modes of expression are of mutual benefits to each other. As Kantor and Rubin (1981) stated, "these two modes of communications develop concurrently and exert mutual influences"

(p.56). Although a significant amount of research has been conducted on the effects of speaking on writing, much research is needed to investigate the effects of writing on speaking.

Syntactic Maturity in L1 Acquisition

Since Hunt (1965 a) developed T-unit analysis as a measurement of syntactic maturity, it has been successfully used in L1 research to measure the overall syntactic complexity of speech and writing samples. Hunt defined syntactic maturity as the ability to manipulate the syntax of the language and claimed that students' syntactic complexity increased with their age. According to Hunt, syntactic maturity can be measured by objective measures such as mean T-unit length (MTUL), mean error-free T-unit length (MEFTUL), percentage of error-free T-units (%EFTU), subordination ratio (SR), mean clause length (MCL), mean sentence length and coordination ratio. According to Hunt, T-unit is "minimal terminal syntactic unit" which means any independent clause with all its dependent clauses (as cited in Cooper, 1976, p.177). Therefore, syntactic maturity can be examined with T-unit analysis with above-mentioned measures like MTUL. Basically, the longer the T-unit, the more syntactically mature it is.

As a method of measurement and as a theoretical hypothesis, T-unit analysis has been regarded as an objective, reliable, and valid measure of overall syntactic complexity (O'Donnell, et al. 1967). Many L1 researchers favor the use of syntactic maturity to conduct research, because syntactic maturity is easy to count and requires no elaborate training in grammar or statistics. In addition, it provides more objective and quantifiable measures of significant structural characteristics. Therefore, it serves as an explicit model for measuring written performance and "thus imposes on writing research a welcome

added discipline, control, and degree of accountability" (Kerek, 1979, p. 8). O'Donnell's (1974) study demonstrates that T-unit analysis is an effective method of analysis compared to other methods like using the sentence as a syntactic unit. With other methods, it is obviously difficult to objectively identify speech. Furthermore, despite the objective identification of writing when based on periods, spaces, and capitals, it may result in "arbitrary and trivial distinctions" (p.109), because punctuation alone may not be a good indicator of syntactic maturity. As a result, it is impossible to guarantee the comparability of speech and writing and the equivalence of the findings of one investigator to that of another. Unlike the sentence, the T-unit can be objectively identified in both speech and writing and its length reflects "generalized (sentence-combining) transformations applied to its underlying structure"(O'Donnell, 1974, p. 104). In regard to traditional grammar, the number of words in the unit in a sense is the reflection of the number of dependent clauses, nonfinite verbals, and modifiers of nouns. The indices of syntactic maturity seem to indicate language ability and are not subject to cultural influence (Price & Graves, 1980). However, T-unit analysis has not been without its criticism. Ney (1966) argued that T-unit analysis did not take excessive coordination within a sentence into consideration. Moffett (1968) showed concern regarding the validity of the T-unit in that sentence maturity was solely based on sentence complexity with the T-unit analysis. Kerek (1979) also pointed out that syntactic maturity could not provide valid estimates to account for the inconsistencies in dealing with the extraneous, leftover material in sentences like "garbles, fragments, ellipses, speaker tags and the like" (p. 6). Despite the criticism of syntactic maturity, syntactic maturity has been successfully and widely used in L1 acquisition.

Insofar as the indices of syntactic maturity are concerned, L1 researchers have agreed that mean T-unit length is the most significant index of syntactic maturity (Hunt, 1965 b; Mellon, 1970; O'Hare, 1973; Stewart, 1978; Tomlinson & Straehley, 1978). For instance, Hunt (1965b) found that mean T-unit length is the most significant index of syntactic maturity for school groups, followed by mean clause length and subordination ratio. But for skilled adults, mean clause length seems as significant as mean T-unit length. Mendelsohn (1983) looked at the previous research (e.g. O'Donnell, et al.1967; Harrell, 1957) and noted that mean T-unit length and subordination ratio are two indices of syntactic maturity that have been popularly used to measure syntactic maturity. And mean T-unit length, in particular, has been found to be a sensitive index of syntactic maturity in analyzing both written and spoken data. In regard to subordination ratio, researchers seem to have different views on its sensitivity as an indicator of syntactic maturity. Witte & Sodowsky (1978) examined the first in-class essays and the final exam essays of 24 freshman writing students to determine the syntactic maturity these students gained during an eight-month period and found that the subordination ratio seemed to be a more precise measure of syntactic maturity gains than other indices. However, Charles (1978) analyzed the papers written by 57 high school students to determine patterns of student writing maturity in comparison with research by Hunt and discovered that there was no direct correlation between a paper's quality grade and the subordination ratio.

Syntactic maturity has been used in L1 acquisition to examine the development of both writing and speaking. Several studies found some age-related trends using the method of syntactic maturity, which is that as children become more mature, their syntax would get more complex. Hunt (1965a) analyzed the writing samples of school children

in grades four, eight and twelve and found that students' age and syntactic maturity have a positive correlation with each other. Simply put, as children get older, they tend to use longer T-units in their writing. Like Hunt, O'Donnell et al. (1967) adopted the approach of syntactic maturity to study the development of children's oral and written language and found that the mean T-units length was significantly greater in the speech than in the writing of third graders, but not the fifth and seventh graders. It seems that when children get older, their written units become longer and more complex than spoken ones and there is more frequent appearance of modifiers of nouns, coordinate predicates, and adverbial clauses. Their findings suggested that children learn to be more careful in writing than in speaking as they make progress in school. Another important study using the measures of syntactic maturity on oral and written language data was conducted by Loban (1976), who studied the language development of a group of kindergarten children through 12th grade. He agreed with Hunt and O'Donnell et al. that subordination was an important index to demonstrate syntactic development and stated that skilled speakers and writers used phrasal structures or other strategies instead of dependent clauses to compact ideas. Stewart (1978) investigated the written syntactic maturity of students from high school to university and found that syntactic complexity did increase with the age of students. Specifically, those in upper levels of the various grades tended to produce significantly longer clauses, T-units containing a greater number of clauses, and longer T-units.

Some studies have examined gender differences in syntax and usage in both modes of communication. One of these studies was carried out by Price and Graves (1980), who examined sex differences in syntax and usage in oral and written language of

students in the middle school years with indices of syntactic maturity. Their findings indicated no apparent innate differences in language abilities between males and females.

In sum, syntactic maturity has been successfully used in L1 acquisition. Among the measures of syntactic maturity, mean T-unit length has been proved to be a sensitive indicator of L1 learners' syntactic development. L1 research using the measures of syntactic maturity has revealed that there is an age-related trend in children's syntactic development, namely, as children get older, their syntax will become more mature.

Lexical Density in L1 Acquisition

Since Ure (1971) first coined the term, lexical density (LD), it has been recognized as a reliable indicator for distinguishing between oral and written language. Research has indicated that texts can be ranked in relation to each other with LD (i.e. ratio of lexical to grammatical items within a text) (e.g. Halliday, 1985). In other words, the ratio of lexical to grammatical items shows how lexically dense one text is as compared to another.

There are two approaches researchers have used to calculate LD on spoken and written discourse: one is manual and the other is automatic. For the first approach, the analyst categorizes the lexical versus the grammatical words in a text and then works out the percentages. Regarding the second approach, the calculation is mainly done with computer programs such as the one devised by Stubbs (1986) to run on the London-Lund corpus of spoken English. Despite its low efficiency, the manual approach is more accurate since each problem is treated by an analyst in its real context. To illustrate, model verbs may function as nouns in certain contexts, which can be identified with the manual approach, but not with the automatic one (Zora & Johns-Lewis, 1989).

According to Halliday (1985), there are two ways to define LD. One is that LD is "the number of lexical items as a proportion of the number of running words" (p. 64). However, as he claims, this definition is a measure of words against words and cannot serve as a more valid measure of LD. On the other hand, he holds that the package of lexical words into a grammatical structure can indeed "determine the informational density of a passage of a text" (p.66). Therefore, he offers the other way of defining LD, which is the number of lexical words per clause. Nevertheless, his definition of "clause" is rather complicated, which is based on the locus of transitivity, mood, and theme. To simplify it, Harrison (1998) takes 'clause' to include both finite and non-finite clauses. However, Sardina (1996) reviews these two definitions and prefers the first one. For one thing, there is no need to mark up clause boundaries with the first definition, i.e, the number of lexical items per the total number of words. For another thing, the variation of density levels can be flexibly investigated at short intervals.

Determining LD depends on distinguishing lexical and grammatical items in a text. Halliday's (1985) study contributes to this distinction, in which he defines lexical words as content words and grammatical words as function words. Specifically, lexical words consist of nouns, main verbs, adjectives, and adverbs. On the other hand, grammatical items are made up of determiners, pronouns, most prepositions, conjunctions and some classes of adverbs. However, Halliday states that there is a continuum from lexis to grammar. To illustrate, some words may be classified either as lexical items or grammatical items such as the modal adverbs like *always* and *perhaps* (p. 63). As Halliday proposes, it does not matter how to classify the words into either category. The main requirement is consistency. Halliday further refines this distinction

by differentiating between high and low frequency lexical items. High frequency lexical items refer to those which have a common occurrence in the language in general such as "people", "thing", "way" and so on. High frequency words are on the borderline with grammar, which are weighted the half value of the low frequency items. O'Loughlin (2001) argues for this method of calculation stating that it "provides a truer, more fine-grained estimate of the overall lexical density" (p.102).

L1 researchers using LD have reached the agreement that lexical density can sensitively and reliably distinguish between spoken and written data. Specifically, spoken language is lower in lexical density than written language. DeVito (1965) found written language was much higher in lexical density than spoken language. Ure (1971) calculated LD in 34 spoken texts and 30 written texts with the manual approach and concluded that spoken texts tended to have lesser density (with a LD of less than 40% [of a total of 100 words, 40 words are lexical words]) than written text (with a LD of greater than 40%). So, she concluded that lexical density was a sensitive measure for distinguishing between spoken and written data. Halliday's (1979) study indicated that spoken language contained more complex structures with low lexical density (more clauses, but fewer high content words per clause), while written language was characterized by simple sentence structures with high lexical density (more high content words per clause, but fewer clauses). Halliday's findings are somewhat surprising concerning the difference between spoken and written language. Chafe & Tannen (1987) attributed Halliday's findings to the lack of a systematic quantitative study of spoken and written samples and the use of isolated examples.

However, Beaman (1984) investigated Halliday's assertions with a detailed study of coordination and subordination in spoken and written accounts of a film and found that spoken language was in some ways more complex than written, which supported Halliday's assertions. She stated that the long-held opinion on the differences in syntactic complexity between two modes were actually resulted from differences in the formality, or register of the discourse rather than true differences between spoken and written language. Halliday (1985) studied the difference between spoken and written discourse with lexical density and reached the same finding as Ure did that written discourse typically has a higher lexical density than spoken discourse. In other words, a higher level of information is contained in written text than in spoken text for a given number of words. So, he stated that the complexity of written language is lexical while that of spoken language is grammatical. Eggins's (1994) study echoes with previous research in that lexical density is a good measure to distinguish spoken and written texts. His study shows that the spoken text is 33% lexical while the written version is 42% lexical. In general, L1 research with lexical density has revealed that spoken language is lower in lexical density while written language is higher in lexical density.

More recent studies on lexical density have paid special attention to sources of variation in spoken discourse (Stubbs, 1986; Hasan, 1988; Zora and Johns-Lewis, 1989). Stubbs (1986) used a computer program to analyze six spoken sub-texts of the London-Lund corpus produced by highly educated informants. In his study, telephone conversation has the lowest lexical density and radio commentaries have the highest LD. His findings showed a significantly higher LD ranging from 44% to 56% (the percentage of lexical words to the total number of words) than that reported by Ure. He states that

the calculation method and the nature of corpora and his subjects (all are educated native speakers of English) might have an effect on LD.

Zora & Johns-Lewis (1989) examined lexical density in interviews and conversation with both undergraduate and graduate students and found that lexical density is not a sensitive measure to differentiate between two discourse modes in a global way. Specifically, with regard to graduate students, the two modes were differentiated with the higher density in interview than in conversation. But for undergraduate students, lexical density remained the same. A possible explanation might be that graduate students are more mature and have more exposure to higher education, so they realize the difference between the formal and informal discourse modes and adjust their lexical density accordingly. Zora & Johns-Lewis in their study discussed the sources of variation for different researchers to reach different percentages of lexical density despite the identical speaking tasks. Eight sources of variation are pointed out in their study including basis for calculating LD, expected interruption and length of speaking turn, function of component units of text, self-consciousness/self-monitoring, personal attribute, group attributes, planning time and topic. Additionally, their study demonstrated that the presence or the absence of 'feedback' might be a more powerful factor in determining lexical density than discourse modes. Likewise, Ure's study also provided the evidence that plannedness might be a factor in determining lexical density. In her study, all prepared spoken texts have a lexical density of 37% or higher.

Sardina (1996) argues that lexical density has been done in previous research only on monologues with respect to spoken data. Therefore, he conducts a study on the analysis of dialogues with lexical density, specifically, the dialogue of two professional

encounters. In addition, unlike previous research, he computes lexical density at intervals of the dialogue as well as the lexical density of the dialogue when taken as a whole text. His study indicates that the dialogue has very high-density portions, as opposed to what the ratios for the whole text would suggest and it is redundancy, or the use of many repeated words near each other that raises density. However, Halliday in his study claims that repetition reduces the effect of density. As Sardina explains, when a lexical word is repeated in short succession, it is conveying a different kind of information. In addition, although lexical density is supposedly an indicator of the amount of information packaged into a text, Sardina holds that lexical density is not equal to information density. Specifically, that the interactants are using more lexical words does not mean that more information is conveyed. The difference between his finding and that of previous research may be because the spoken data studied in previous research is mostly monologues but his data are concerned with dialogues. In his study, he calls for more research on the analysis of dialogue with lexical density so as to have a better understanding of lexical density.

To conclude, lexical density has been viewed as a sensitive and reliable indicator for differentiating between oral and written language in L1 research. It has been agreed that the written language is more lexically dense than the spoken language. More recent studies on lexical density have attended to sources of variation in spoken language. Two approaches are used to calculate lexical density, namely, the manual and automatic approaches. In addition, there are two ways to define lexical density: one is the lexical items per the total number of words and the lexical items per clause according to Halliday (1985).

Second Language Acquisition

As noted earlier, a number of studies have been conducted on the relationship between speaking and writing development in L1 acquisition. Composition theory and research seem to have been almost exclusively studied to investigate the products and processes of native-speaker writers rather than L2 students (Raimes, 1983). These studies have provided valuable information for L1 instructors to facilitate their students' language acquisition. However, fewer studies have been written about this issue on L2 acquisition.

This section is devoted to reviewing the research studies conducted on the relevant aspects of this issue in L2 acquisition. Specifically, the relationship between speaking and writing in L2 acquisition will be discussed, as well as the studies on syntactic maturity and lexical density, and the use of the VOCI in L2 acquisition.

The Relationship between Speaking and Writing in L2 Acquisition

Due to the influence of structural linguistics, which placed too much emphasis on spoken rather than written language, speaking and writing had been traditionally separated in L2 curricula (Diller, 1978). Another influence has been due to child language acquisition such as the natural approach, which holds that L2 acquisition should be like L1 acquisition, since children learn to speak before they write (Krashen & Terrell, 1983). In addition, speech act theory also contributes to this separation, since it considers the teaching of common speech act as an important goal in courses based on functional syllabi (McKay, 1979). As a result of these influences, writing is always regarded as a complementary way to reinforce the instruction of speaking (Kim, 1998). Hudelson

(1984) claimed that writing was strictly controlled and reduced to a limited range of exercises and activities because it had been believed that L2 acquisition should simulate patterns of L1 acquisition by adhering to children's learning sequence: listening, speaking, reading and writing. However, L2 researchers have recently come to realize the importance of writing and the differences between L1 and L2 language acquisition. Now, more and more research has been carried out to study L2 writing (Valdes, Haro, & Echevarriarza, 1992). Still, very few studies have been conducted on the relationship between speaking and writing in L2 acquisition.

One of the earliest studies of this kind was conducted by Gipps & Ewen (1974), who studied the development of the Mean T-unit Length (MTUL) of Asian children. These children had been in Britain for different lengths of time. They found that MTUL correlated positively with these children's length of stay in Britain. That is to say, the longer they have stayed in Britain, the longer T-unit they use.

Two years later, Vann (1979) analyzed this relationship with measures of syntactic maturity. One of the purposes of her study was to inform teachers and researchers in English education about the relationship between oral and written languages, particularly among adult foreign language learners. Accordingly, in her study, she examined paired oral and written discourse of a group of 28 adult native speakers of Arabic studying English in the United States with the methodology based on previous research using native speakers. She first asked her subjects to watch a short film and then interviewed half of these subjects, but assigned the other half to write a composition on the film. She then analyzed her subjects' oral and written discourse with five indices of language proficiency, namely, mean T-units length, percentage of dependent clauses to

T-units, percentage of "mazes (p.6)" (i.e. linguistic tangles, false starts and redundancies) to total discourse, mean error-free T-unit length, and percentage of error-free T-units to total T-units. Her findings are as follows:

1. "Oral compositions were almost twice as long as written ones, although oral segments represented only one –fourth the time allotment of written ones (5 minutes of oral taping versus 20 minutes of writing).

2. Mazes were plentiful in speech and rare in writing.

3. Both mean T-unit length and mean error-free T-unit length were longer in written than in oral discourse.

4. The ratio of adverbial and adjectival clauses to T-units was higher in written than in oral language, especially in the case of adjectival constructions, the percentage of which more than doubled in written expression" (p.9).

Mangelsdorf (1989) described the parallels between speaking and writing and claimed that the integration of these two modes of language can strengthen L2 acquisition. In addition, she also proposed some ways to integrate speaking and writing in L2 classrooms and curricula. According to her, speaking and writing are similar in the way that they both function as ways of communication, which needs to be achieved through the construction and negotiation of meaning. Secondly, they both have the nature of interaction, because communication is interactive. As students go through several drafts, they may clarify their ideas from the perspective of a reader. Further, speaking and writing are both processes of testing hypotheses. For example, students test and revise language according to audience feedback in the context of interaction. Finally, speaking and writing can both be treated as dialogues, because "all speakers and writers have

listeners and readers" (Berthoff, 1981, p. 119). Because of the similarities between speaking and writing, she suggested the integration of speaking and writing by peer review, oral presentation prior to writing, dialogue journals and so on.

Seda & Abramson's (1990) study highlighted the transactional nature of oral and written language, which L2 learners can benefit from regardless of their speaking proficiency in the target language. Their conclusion in fact is in line with other work in the emerging literacy in L2 (Edelsky, 1986; Hudelson, 1984). Haynes (1992) investigated target language variability between speaking and writing in the L2 acquisition of non-native English speakers. She compared spoken and written narratives from three groups of non-native English speakers and spoken and written narratives of native English speakers based on the multi-feature/multidimensional approach developed by Biber (1986). The results of her study indicate that as L2 learners progress in English proficiency, they develop more abstract content and more reported style in both speech and writing. On the other hand, both speech and writing become more interactive as L2 learners develop in English proficiency. That is to say, both speech and writing of L2 learners become more verbal, interactional and affective as their English proficiency improves. Further, results indicating variability between spoken and written narratives demonstrate that non-native speakers develop systematically toward native English variability between speaking and writing.

Kim (1996, 1998 & 2000) conducted a study to compare speaking and writing development in a L2 college student over a two-year period. In this study, she analyzed errors and syntactic maturity in the subject's writing development and his formal language institute oral interviews to determine whether the L2 college student followed

the same pattern of writing and speaking development reported for L1 acquisition. Her findings indicate that errors decrease and syntactic maturity increases in the same way that it occurs in L1 development. Additionally, her study indicates that writing is of greater syntactic maturity than speaking in an L2, which agrees with the research in an L1 (e.g. Harrell, 1957).

Vann (1981) proposed a model for the relationship between speaking and writing among second/foreign language learners. Based on Weaver's (1979) study on the writing development of children, she described second/foreign language learners at three levels. For level one writers, like children at this stage, they cannot differentiate writing from speaking, so their writing reflects the influence of their speaking. Concerning level two writers, they place much emphasis on producing correct surface forms. Like children, some adult ESL/EFL learners may skip this level and directly go to level three. On the third level, writers are able to synthesize spoken and written forms. For the writers at different stages, Vann provides classroom teachers with some practical suggestions. When students are at the early stage of development, the teachers must not place too much emphasis on the correct use of forms. Rather, they should help students to produce meaningful and organized discourse and notice the difference between speaking and writing. In regard to level two students, the teachers should expose students to the form and function of the target language and help them to make connections between form and function. At the third stage, students should be provided with opportunities to see and hear different registers of English. Vann's model in fact is similar to Kroll's (1981) in moving through stages of differentiation to synthesis, which provides some insights into the relationship between speaking and writing in adult L2 learning.

On the basis of the relationship between speaking and writing in L2 acquisition, some researchers have tried to create ways to improve L2 learners' speaking and writing. One of these ways is the use of dialogue journal writing to improve L2 student's writing skills. Some researchers have proposed the application of dialogue journal writing to the writing class, because this form of writing is close to the spoken register (Shuy, 1988) and can function as a bridge from expressive to academic writing (Kreeft, et al. 1984; Blanton, 1995). Weissberg (1997) advocates the importance of dialogue journal writing for adult ESL learners to acquire English syntax, since it bridges the gap between speaking and writing and provides the learners with a non-threatening, supportive social environment. His study also indicates that dialogue journal writing is the most-favored modality by his subjects for syntactic innovation and accuracy compared to speech and academic writing. Shuy (1987) claims that the communication between the teacher and student via dialogue journal writing "scaffolds intellectual problems for their students that promote higher-order thinking, such as constructing logical arguments and providing extended elaboration on a topic" (as cited in Weissberg, 1997, p. 3).

Syntactic Maturity in L2 Acquisition

As in L1 research, the measures of syntactic maturity have been successfully used in L2 research, since L2 researchers assume that the growth of L2 learners' syntactic repertoire and their appropriate use of the repertoire in a variety of situations reflects their language development among other processes (Ortega, 2003). The use of syntactic maturity measures in L2 research has been favored by many L2 researchers "to characterize the syntactic nature of linguistic input and to assess the syntactic maturity of learners' written production" (Gaies, 1980, p. 2). The main strength of T-unit analysis

lies in "its very ease and alleged objectivity of T-unit segmentation" (Gaies, 1980, p. 4). In addition, Gaies (1980) argues that syntactic maturity, as a measure of linguistic development, can be applied to any particular set of data and it allows for meaningful statistical comparison between L1 and L2 data. Further, syntactic maturity has been successfully applied in several languages to discriminate among L2 learners at different levels of proficiency such as German (Cooper, 1976), French (Monroe, 1975), Spanish (Thornhill, 1969), Arabic (Vann, 1979), and ESL (Gaies, 1980; Halleck, 1995).

In regard to the indices of syntactic maturity, similar to what has been found in L1 research, mean T-unit length has been found to be a useful measure of L2 development (Thornhill, 1969; Gaies, 1980; Larsen-Freeman & Storm, 1977), though Witte and Davis (1983) questioned its reliability. Kameen's (1983) findings indicate that mean T-unit length and mean clause length are two more reliable indices of rated quality than the index of mean sentence length. In particular, Kameen states that mean T-unit length, especially longer T-units (21 words or more per T-unit) is considered as a powerful index for differentiating between the writing of 'good' and 'poor' college-level ESL writers. According to Kameen, 'good' writers wrote an average of 18.40 words/T-unit while 'poor' writers used 14.30 words/T-unit. Concerning mean error-free T-unit length, L2 researchers have paid more and more attention to its use as a valid measure of L2 development because of the particularities of L2 development. Unlike L1 data, errors are frequently found in adult L2 data, so the incidence of developmental errors should be reflected by the index of syntactic maturity (Scott and Tucker, 1974; Gaies, 1976; Larsen-Freeman and Strom, 1977; Larsen-Freeman, 1978; Vann, 1979). Therefore, mean error-free T-unit length has now been viewed as a more valid measure of growth in L2

research. Larsen-Freeman (1978) found mean error-free T-unit to be a more powerful index to discriminate between different levels of proficiency, while mean T-unit length was not a good index to discriminate at the top level of proficiency. Similarly, Kyle (1980) evaluated advanced ESL students' compositions with ten objective measures of writing proficiency and claimed that only those objective measures which take the absence of errors into consideration can discriminate among holistic evaluations of the compositions by experienced ESL teachers. In addition, Sharma's (1979) study also indicated that percentage of error-free T-units and mean error-free T-unit length are the most productive measures in separating ESL students' proficiency levels. Ortega's (2003) study shows that mean sentence length, mean T-unit length, mean clause length, mean number of T-units per sentence, subordination ratio and mean number of dependent clauses per clauses are the six most frequently used syntactic maturity measures across the twenty-one studies reviewed in his study.

However, T-unit analysis is not without its limitations. Most L2 research using the T-unit analysis is aimed to investigate the extent to which syntactic maturity measures derived from L2 writing can indicate the students' overall proficiency in the target language. Gaies (1980) pointed out two limitations of T-unit analysis. One is that it might not be appropriate for T-unit analysis to be applied to the data from subjects with relatively low proficiency, because these subjects make so frequent grammatical and lexical errors that these errors seem to interfere not only with the reader's or listener's understanding, but also with the researcher's ability to tabulate T-units. The other limitation is concerned with the discriminating power of the T-unit, because mean error-free T-unit length and related measures cannot function as a sensitive indicator of L2

development though it can discriminate between learners with low proficiency from those with high proficiency. Specifically, these measures cannot discriminate between adjacent groups. In a similar vein, Wolfe-Quintero et al. (1998) examined cumulatively the strength of the relation between a number of syntactic maturity measures and proficiency levels across thirty-nine L2 writing studies. They conclude that mean T-unit length, mean clause length, subordination ratio, and dependent clauses per clause could discriminate between proficiency levels, because these measures demonstrated a linear and consistent relationship with program, school, and holistic rating levels across the thirty-nine primary study reports. However, they caution that these measures nevertheless are not good indices to discriminate between adjacent levels of proficiency and that statistically significant relationships only appear in the studies involving holistic ratings. Based on Wolfe-Quintero et al.'s study, Ortega (2003) investigates the same issue and claims that this relationship differs systematically across studies according to whether the study is concerned with a second or foreign learning context and whether proficiency is defined by program level or by holistic rating.

The review of literature on syntactic maturity indicates that most L2 research using the T-unit analysis has focused on writing rather than speaking. For instance, the thirty-nine studies examined in Wolfe-Quintero et al. are all about the analysis of L2 writing. Likewise, twenty-one studies reviewed by Ortega (2003) are also concerned with writing. Very few studies have paid attention to speaking in L2, which should be an important part in L2 development. The earliest study which examined the spoken data with syntactic maturity measures was conducted by Thornhill (1970), who discovered that remarkable parallels existed between the trends in the language behavior in both L1

acquisition and L2 learning. Cooper (1976) called for further research on the use of the method of syntactic maturity to measure speech to know how development in the two modes of communication compares. Halleck (1995) compared holistic ratings of ACTFL Oral Proficiency Interviews (OPIs) and objective measures of syntactic maturity to investigate the syntactic basis of ACTFL's holistic ratings and examine the relationship between oral proficiency and syntactic maturity. Her findings demonstrate that levels of syntactic maturity vary according to interview tasks and objective measures of syntactic maturity overall correlate with holistic evaluation. In other words, her study provides some empirical evidence that measures of syntactic maturity can be used to demonstrate second and foreign language learners' oral proficiency. Therefore, further studies should attend to the use of syntactic maturity measures in L2 speaking in order to get an overall picture of L2 development.

Lexical Density in L2 Acquisition

As noted earlier, lexical density has been widely used in L1 acquisition to study the differences between spoken and written data. However, few studies have been conducted in L2 acquisition.

In terms of the use of lexical density in L2 acquisition, several studies have discussed the lexical development of Swedish learners by comparing Swedish L2 learners and Swedish native-speakers. These researchers seem to agree that lexical density is not a good indicator of L2 proficiency and L2 learners should at least reach the norm of native speakers. Linnarud (1986) measured the difference between the lexis in English composition by seventeen-year-old Swedish L2 learners and those by native speakers by counting the lexical density of the texts. The results of the study indicated a small, non-

significant difference between the native speakers and Swedish L2 learners in lexical density and no correlation between a high density and a high grade. She then concluded that L2 learners should at least produce the same lexical density as native speakers do and "a higher lexical density does not always accompany increasing success, once this threshold is passed" (p.82).

Likewise, Hyltenstam (1988) studied bilingual Finnish-Swedish and Spanish-Swedish as against monolingual Swedish secondary-school pupils' oral and written production and agreed with Ure (1971) that spoken data contained a lower lexical density than written data with the lexical density of oral texts ranging between 35.9 for the monolingual Swedish control group and 38.3 for the bilingual Finnish-Swedish group of pupils. And the lexical density of the written language for these groups was about 45%. Lauren (2002) uses lexical density to study the immersion pupils' progress in the development of their Swedish as a L2, compared to corresponding Swedish-speaking groups of the same age. She finds that the average value of lexical density in spoken data is around 35% both for the immersion groups and for the control groups and lexical density is higher in written data than in spoken data for both groups, which is in agreement with Ure's (1971) study. However, she notes that the lexical density of the immersion pupils in both spoken and written data demonstrate a higher percentage than that of the Swedish-speaking control group. This is an interesting finding. As she states, a possible reason might be interpreted as a lower degree of nativeness in the immersion pupils. That is to say, the native group can vary their use of words like pronouns, adverbs and prepositions instead of meaningful lexical words to pad out their narration. Various non-meaningful form words in fact provide cohesive textual links for the meaningful

lexical words. In agreement with Linnarud, she also suggests that a higher lexical density in the immersion group may indicate their deviation from the pattern if the Swedish-speaking control group is viewed as a norm. This finding is important in L2 acquisition in the sense that lexical density may be used as a measure to test the language development of L2 learners. To illustrate, in Lauren's study, if the Swedish-speaking control groups are taken as a norm, it is clear that the immersion groups do not follow this pattern and thus need to further improve their language skills to arrive at the norm.

Like LD studies in L1, a few studies in L2 also shed light on the sources of variation in spoken discourse. Hassan (1988) analyzed the spoken data produced by native and non-native speakers in five types of formal and informal types of spoken discourse with the mean LD per T-unit and reported that his formal native-speaker interviews had a relatively higher (47.02%) LD than informal conversation (42.48%). Conversely, non-native speakers showed a substantially lower LD in formal interviews (33.67%) than in formal and informal classroom interactions, informal classroom discussions and informal conversations with native speakers outside the classroom. Generally speaking, this study indicated that the context or setting, the type of interaction and the role of the speakers all had an effect on determining the degree of lexical density. In a study of the validity of direct versus semi-direct oral tests, Shohamy (1994) compared ten of twenty oral language samples by counting the lexical versus grammatical items. Her results indicated that the ratio of lexical items to grammatical items in the Oral Proficiency Interview (OPI), a direct oral proficiency test, was approximately 4:6, while in the Simulated Oral Proficiency Interview (SOPI), a semi-direct oral proficiency test, the ratio is reversed. Similarly, O'Loughlin (2001) pays

special attention to the equivalence of direct and semi-direct oral interviews and demonstrates that a significantly lower lexical density is found in the live version than in the taped version in all tasks. In conformity with Ure, he argues that the presence of interaction may have an effect on this difference. In other words, the more interaction there is, the lower the lexical density will be. In addition, he explains that these differences might be attributed to the language abilities these two versions of test try to elicit: interactive versus monologic speaking ability. However, in agreement with Linnarud (1986) and Lauren (2002), he cautions that LD does not distinguish between language proficiency levels.

In sum, the LD research in L2 acquisition has indicated that it can distinguish between spoken and written data as agreed in research in L1 acquisition, but it can not sensitively indicate L2 learners' proficiency levels. In addition, LD may be used to measure L2 learners' language development by the comparison of LD produced by L2 learners with that by native speakers, which serves as the norm of language development.

The Use of the Video Oral Communication Instrument (VOCI) in L2 Acquisition

The VOCI is used as an instrument in this study, which is a semi-direct and tape-mediated speaking test. It was originally created as an alternative for the OPI (Oral Proficiency Interview) to determine level of oral proficiency. Unlike other semi-direct speaking tests such as Simulated Oral Proficiency Interview (SOPI), the VOCI has only been developed for a relatively short time. Therefore, empirical evidence is needed to prove its validity as an assessment of oral proficiency. However, the VOCI has its advantages as a semi-direct speaking test. Firstly, it controls the variables found in the OPI such as the role relationship, personality and gender of testers and respondents, the

purpose of interaction, the topic and the setting. These variables have been shown to affect the oral language score (Shohamy, 1983). Secondly, it allows the test-takers to respond to perform all the speaking tasks, which provides an opportunity for the test-takers to demonstrate their abilities in responding to more challenging questions (McCrackin, 1997). In addition, it can be administered to a group of test-takers in a language laboratory, which is clearly more cost efficient than direct tests (O'Loughlin, 2001). Another benefit of the VOCI is that all test takers respond to the same questions and so perform similar language tasks, which improves the comparability of test-takers' oral performance and therefore "standardize the assessment of speaking while retaining the communicative basis of the OPI" (Shohamy, 1994, p.101).

The use of the VOCI in this study increases the comparability of the spoken and written data in this study. With the VOCI, the specific type of oral language, namely, extemporaneous monologue can be elicited, since the subjects respond to each question in the test without interaction with other people. Because monologue bears some similarities to writing, the spoken sample is comparable to the written sample in this study. In a specific manner, as previously mentioned, writing is attributed to be "self-rhythmed", which represents "a powerful instance of self-provided feedback" and "establishes explicit and systematic conceptual groupings" (Emig, 1977, p. 128 as cited in Schafer, 1981, p. 23). Schafer (1981) claims that an oral monologue contains all these attributes the written discourse has except that the monologue in a conversation may be interspersed with "intermittent audience feedback in the form of 'continuers'- 'mmhmm', or 'yeah,' or 'right'-and an occasional request for clarification", since "the speaker has to come up with some logic to hold his text together" (p.23). Likewise, Moffet (1968)

contends that a speaker in a monologue can no longer depend on social interaction to

organize his discourse and must begin to manage his utterances logically. Schafer has

criticized other studies that have compared and contrasted speaking and writing by using

other types of oral language text such as dyadic conversation and prepared speech. He

believes that "one particular type of speech of a particular sample of the population

differed in certain ways from the written sample to which it was compared" (p. 12). In

line with Schafer, Biber (1986) points out that it is the imbalance of text types that results

in the contradictory findings obtained from the studies that investigated the relationship

between speaking and writing. On the other hand, Kroll and Lempers (1981) attended to

this problem by eliciting both spoken and written data in a more similar context than they

typically were. As far as the VOCI test is concerned, the monologue produced by the

subjects is a bit unlike the monologues mentioned by Schafer, because there is not any

audience present to give any feedback or request any clarification. So, the monologue

produced in the VOCI is more ideal than the monologue mentioned by Schafer to be used

as the spoken text to be compared with the written text. In addition, the VOCI was

originally designed as an alternative for the OPI, which takes the form of interview. So,

the spoken text elicited by the VOCI test is in a sense more of an interview type, which is

relatively formal, compared to conversation. Since the written data in this study are

academic prose, which is a formal type of written text, the spoken data elicited by the

VOCI test in this sense is of a comparable nature with the written data in this study.

Another justification for the use of the VOCI is that taped monologue elicited through the

VOCI enables the students to plan and monitor their speech (Weissberg, 1997). In this

way, it reduces the differences between speaking and writing given writing is a planned

mode while speaking is an unplanned mode, which makes the spoken and written sample more comparable.

In short, the use of the VOCI has its advantages as an oral language test. Additionally, it increases the comparability between the spoken and written data in the present study.

Summary

The above review of literature shows that a significant amount of research has been carried out in L1 acquisition on the relationship between speaking and writing. Most of the research has focused on the differences between these two modes of expression. L1 research of this kind indicates that speaking and writing have a reciprocal influence on each other. However, few studies have been conducted on this relationship in L2 acquisition. Since the studies of this kind provide valuable information for L2 instructors to facilitate their students' language acquisition, much research is called for on L2 acquisition to investigate this relationship. In addition, the literature review demonstrates that the measures of syntactic maturity have been successfully used in both L1 and L2 acquisition to examine the syntactic development of both speaking and writing. Like the measures of syntactic maturity, lexical density has also been adopted in both L1 and L2 research to investigate the lexical development of the two modes. The studies in both L1 and L2 show that lexical density can sensitively differentiate between speaking and writing. However, L2 research, in particular, indicates that lexical density cannot differentiate between L2 learners' proficiency levels. Finally, it is also shown in the literature review that the use of the VOCI increases the comparability of the spoken and written data in the present study.

CHAPTER III

METHODOLOGY

Overview

As noted earlier, this study explores the relationship between speaking and writing in adult L2 learners. Accordingly, Chapter III presents the research methodology used in this study to investigate this relationship. This chapter will begin with the explanation of the instrument used in this study and the measures adopted to analyze the data elicited by this instrument. Then, it will describe subjects, procedures, and research hypotheses and data analysis, respectively.

Methods

The instrument, the Video Oral Communication Instrument (VOCI) for ESL/EFL was used to assess the oral proficiency of the subjects. The measures of syntactic maturity and lexical density were employed to examine the syntactic and lexical development of both the subjects' oral and written language. The methods of this study will be explained in detail in the following paragraphs.

As mentioned in Chapter II, the VOCI is a semi-direct and tape-mediated speaking test, which is used as an alternative for the OPI (Oral Proficiency Interview) to determine level of oral proficiency. This study particularly used the English version, the VOCI for ESL/EFL, which was developed at San Diego State University's Language Acquisition Resource Center (LARC) by Halleck and Young (1995). Like other semi-direct speaking tests such as Test of Spoken English, the VOCI uses video stimuli to elicit samples of oral performance from the subjects. In the VOCI test, the participants create a context or situation and then ask the test-taker a question related to the situation

or the context. The test-taker then has to respond to the question accordingly and the response is recorded with a manually operated audio recorder. Specifically, the test-taker watches the audio-visual stimulus and pauses the VCR with a remote control to respond to the question asked by the participant. After the test-taker finishes answering the question, he or she restarts the video and watches the next video. The VOCI takes two forms: one with time constraints (a timed version) and the other without time constraints (an untimed version). The present study used the untimed version.

The VOCI for ESL/EFL used in this study consists of a total of 23 questions (Please see Appendix A for a complete transcript of the VOCI questions), which assess the four proficiency levels: novice, intermediate, advanced, and superior levels defined by the ACTFL (American Council on the Teaching of Foreign Languages) guidelines (1986). The speech tasks elicited by the VOCI vary from describing, comparing and contrasting to supporting an opinion and hypothesizing. Table 1 presents examples of such questions and the level of difficulty represented by each question. Table 2 shows the generic descriptions of the ACTFL Proficiency Guidelines of all four levels.

The measure of syntactic maturity was used to investigate the relationship between speaking and writing in college-level ESL. Seven measures of syntactic maturity were used in this study to examine the syntactic development of both the subjects' speaking and writing. Specifically, these measures are total number of words (TNW), mean T-Unit length (MTUL), mean error-free T-Unit length (MEFTUL), and percentage of error-free T-unit (%EFTU), subordination ratio (SR), mean clause length (MCL) and dependent clauses per clause (DC/C). Of these measures, TNW, MTUL and MCL tap

TABLE 1

Examples of Questions on the VOCI

Question Level	Example
Novice	▪ What's your name? (Q1) ▪ Where are you from? (Q2)
Intermediate	▪ Tell us about your hometown. (Q3) ▪ Describe one of your friends. (Q5)
Advanced	▪ Compare your hometown with a city you visited or know well. (Q7) ▪ Describe an unforgettable experience. (Q9)
Superior	▪ If you were a teacher and you discovered one of your students had cheated on a test by copying from another student's paper, what would you do? (Q 17) ▪ What do you think about the portrayal of violence and crime on TV? (Q21)

(Adapted from McCrackin, 1998)

length of production at the clausal and phrasal level. MEFTL and %EFTU measure the

accuracy of both spoken and written performance. SR and DC/C gauge the amount of

subordination in spoken and written data. As far as the calculation of these indices were

concerned, the total number of words included in each language sample, and the total

number of T-units were counted, as well as the total number of dependent clauses and the

total number of error-free T-units. On this basis, six indices were then determined

through the following formulas:

MTUL=number of words/number of T-units

MCL = number of words/number of T-units+ number of dependent clauses

MEFTL=number of words/number of error-free T-units

%EFTU=number of error-free T-units/number of T-units

SR=number of T-units+number of dependent clauses/number of T-units

DC/C=number of dependent clauses/ number of T-units + number of dependent

clauses

TABLE 2

Generic Descriptions of ACTFL Proficiency Guidelines-Speaking

Novice	Speakers at the novice level are able to:
■	Speak mostly in isolated words and phrases
■	Deal with topics of immediate daily concern
■	Be understood with difficulty by a person accustomed to non-native speakers.
Intermediate	Speakers at the intermediate level are able to:
■	Speak primarily in sentences and strings of sentences
■	Create with the language using learned elements
■	Ask and respond to questions
■	Deal with survival situations and topics primarily related to self
■	Be understood best by a person accustomed to non-native speakers
Advanced	Speakers at the advanced level are able to:
■	Speak in paragraph length discourse
■	Describe and narrate in past, present, and future time/aspects
■	Discuss topics of personal or public interest (i.e. school, work, current events)
■	Compare and contrast or deal with situations with a complication
■	Be easily understood by a native speaker
Superior	Speakers at the superior level are able to:
■	Speak in extensive discourse
■	Support opinions and hypothesize
■	Participate in both formal and informal conversations
■	Deal with topics of general interest and some special fields of expertise
■	Discuss abstract and unfamiliar topics
■	Speak with a high level of accuracy
■	Be easily understood by a native speaker

(ACTFL Proficiency Guidelines, 1986)

When computing these indices, the researcher followed the guideline for counting

T-unit, clauses, word and errors provided by Polio (1997), which will be explained later.

Please see Appendix B for this guideline. To illustrate the calculation of each index,

please see the following passage taken from one subject's written data.

There was a small country in Asia. */The country was too weak to protect

themselves, *so it was invaded a lot by other countries.* / People tried so hard to

overcome these situations /and finally, they got freedom and development of their

country. /The country is Korea; */ In the history, Korean made their own cultural

characteristics, like patience, diligence and cooperation. /

This passage consists of 60 words, 6 T-units (between the virgules), 2 error-free

T-units (with an asterisk), 1 dependent clause (*italicized*), and 7 clauses, both dependent

and main. From these figures one can determine the above-mentioned seven indices:

1. Total Number of Words (TNW) (60 words)

2. Mean T-unit length (MTUL) (60 words/6T-unis=10 words per T-unit)

3. Mean clause length (MCL) (60 words/7 clauses= 8.6 words per clause)

4. Mean error-free T-unit length (MEFTUL) (60 words/2 error-free T-units=30 words per error-free T-unit)

5. Percent of error-free T-unit (%EFTU) (2 error-free T-units/6 T-units=33% of error-free T-unit)

6. Subordination ratio (SR) (6T-units +1 dependent clause/6 T-units= 1.2 dependent clauses for every main clause or clauses per T-unit)

7. Dependent clauses per clause (DC/C) (1dependent clause/1dependent clause+6 T-units=0.14 dependent clause per clause)

Lexical density was the third measure used in this study to investigate the lexical development of the subjects' speaking and writing. As noted previously, there are two approaches to counting lexical density: one is manual and the other is automatic. In comparison with the automatic approach, the manual approach is more accurate since each problem is treated by an analyst in its real context despite its low efficiency. Therefore, the present study used the manual approach to calculate lexical density in both spoken and written data.

As mentioned earlier, Halliday (1985) proposes two ways to define lexical density: one is that lexical density is "the number of lexical items as a proportion of the number of running words" (p. 64) and the other is the number of lexical words per clause. These two definitions both have their advantages and disadvantages. In order to get a complete picture of the relationship between the subjects' speaking and writing and investigate the validity of these two definitions, the present study used both definitions to analyze the spoken and written samples. In addition, unlike other studies, Hasan's study (1988) uses the mean lexical density per T-unit to analyze the spoken data. Since this definition in a sense combines the measures of both lexical density and syntactic maturity, the present study also calculated the lexical density with this definition. In sum, the present study used three indices of lexical density to analyze the subjects' spoken and written data: the number of lexical words per the total number of words (LD/TNW), the number of lexical words per clause (LD/C) and the number of lexical words per T-unit (LD/TU). As far as the calculation of these indices was concerned, the total number of words included in each language sample, and the total number of T-units were counted, as well as the total

number of lexical words, and the total number of dependent clauses. On this basis, three

indices were then determined through the following formulas:

LD/TNW=number of lexical words/number of words

LD/C=number of lexical words/ number of T-units+ number of dependent clauses

LD/TU=number of lexical words/number of T-units

When counting lexical density, it is important to distinguish lexical and

grammatical items in a text. As noted earlier, Halliday (1985) defines lexical items as

content words and grammatical items as function words. Specifically, lexical items

consist of nouns, main verbs, adjectives, and adverbs and so on. Grammatical items are

made up of determiners, pronouns, prepositions, conjunctions and some classes of

adverbs. O'Loughlin (2001) provides a detailed definition of lexical items and

grammatical items. In his study, he also makes a distinction between "linguistic items"

and "linguistic words" stating that they do not correspond to each other. For example, one

item may consist of two words such as phrasal verbs like *look up*. On the other hand, one

word may be made up of two items such as contractions like *they're*. He claims that the

linguistic items are more appropriate units of measurement than linguistic words.

However, it might be confusing to differentiate linguistic items from linguistic words

when counting lexical items in a text. Although Halliday (1985) also distinguishes

between "item" and "word", he states that it does not matter how to classify the words

into either category. What matters is consistency. Besides, other studies (Ure, 1971;

Hasan, 1988; Zora & Johns-Lewis, 1989) adopted the word rather than the item as the

basic unit of lexical density. Therefore, the present study counted lexical words instead of

items.

Another factor worth noting when determining lexical density is the further classification of lexical items, namely, high and low frequency lexical items as refined by Halliday. Since high frequency lexical items occur more often in the language in general, they are on the borderline with grammar. Therefore, they are given half the value of the low frequency items. O'Loughlin states this method of calculation provides a more refined estimation of the overall lexical density. However, O'Loughlin used both methods to calculate lexical density: the methods with and without the distinction

TABLE 3

Classification of Lexical and Grammatical Words

Grammatical Words	
	• Verbs: "to be"; "to have"; modals and auxiliaries • Determiners: articles, deomonstratives adj (e.g. this, that, these, those) and possessive adj (e.g. my, your...) • Quantifiers (e.g. some, any) • Numerals: cardinal and ordinal • Pro-forms: pronouns (e.g. she, they, someone, something); proverbs (e.g.A: Are you coming with us? B: Yes, I am); pro-clauses (this, that when used to replace whole clauses) • Interrogative adverbs (e.g. what, when and how) • Negative adverbs (e.g. not, never) • Contractions (e.g. they're) • All prepositions and conjunctions • All discourse markers: conjunctions (e.g. and, but, so); sequencers (next, finally); lexicalized clauses (e.g.y'know, I mean); meta-talk (e.g. what I mean, the point is); temporal deictics (e.g.now, then); spatial deictics (e.g. here, there); quantifier phrases (e.g. anywhere, anyhow, whatever) • All lexical filled clauses (e.g. well, I mean, so) • All interjections (e.g. gosh, really, oh) • All reactive tokens (e.g. yes, no, ok, right)
Lexical Words	Adj, nouns, main verbs, adverbs of time, manner and place.

(O'Loughlin, 2001, p.107)

between high and low frequency items and suggests that the results are similar but not identical for the two methods. Therefore, in the present study, for the convenience of statistical analysis of lexical and grammatical words, the distinction between high and low frequency words was ignored in favor of a crude classification.

According to the above reasons, the classification of items given by O'Loughlin for analysis of lexical density was modified for the present study as shown in Table 3. The passage below taken from a subject's written sample illustrates the analysis of the data with the measure of lexical density.

> There was a *small country* in *Asia*. The *country* was too *weak* to *protect themselves*, so it was *invaded* a lot by other *countries*. *People tried* so *hard* to *overcome* these *situations* and finally, they got *freedom* and *development* of their *country*. The *country* is *Korea*; In the *history*, *Korean made* their *own cultural characteristics*, like *patience, diligence* and *cooperation*.

As noted above, this passage consists of 60 words, 6 T-units and 7 clauses. The italicized words are counted as the lexical words and the remaining words are the grammatical words. Of the total number of words, there are 28 lexical words and 32 grammatical words. From these figures one can determine the above-mentioned three indices of lexical density:

1. Number of lexical words/number of words (LD/TNW) (28 lexical words/60 words=0.47 lexical words per number of words)

2. Number of lexical words/ clauses (LD/C) (28 lexical words/7 clauses=4 lexical word per clause)

3. Number of lexical words/number of T-units (LD/TU) (28 lexical words/6 T-units=4.7 lexical words per T-unit)

In the present study, the same method of calculation was adopted in measuring syntactic maturity and lexical density in both spoken and written data with an attempt to increase the comparability of the subjects' spoken and written language. Besides, a guideline adapted from Polio (1997) for counting T-unit, clauses, word and errors was followed in order to obtain a consistent calculation of syntactic maturity and lexical density in both spoken and written data.

In addition, in the written data, corrections were made to the words wrongly written as two words or as one word since they directly affect the number of words. For example, if the word "notebook" is written as the two words, "note" and "book", it will be combined and counted as one word. On the other hand, the words, "middle class" wrongly written as "middleclass" will be separated and counted as two words. In regard to the spoken data, non-lexical filled pauses (e.g. er, um) were not counted for the sake of analysis because their inclusion as grammatical items may "obscure the relationship between lexical and grammatical items in the sample" (O'Loughlin, 2001, p. 105). Further, the words indicating hesitation (hmm…) and repetitions were left out when counting the number of words in the text (Laura, 2002). Finally, when the subjects made self-repair, only the final version of a word or utterance was figured in the analysis. As an illustration, in the sentence, "*And when I know, when I knew I pass the TOEFL, I was, I*

was so happy", the repeated words, "*I was*" were not counted and only the final version of the utterance, "when I knew" were counted.

Subjects

Originally, 20 international students out of a total of 40 participated in this study. The students were enrolled in a college-level ESL composition course, *International Freshman Composition I* at Oklahoma State University. Later, the number of subjects was narrowed to 20 according to the rating of the second writing assignment for the course: 10 high-rated students and 10 low-rated students. In order to further increase the comparability of high- and low-rated students, these 20 students' first writing assignments were then graded. Finally, 10 students were chosen to participate in the study: 5 highest-rated and 5 lowest-rated students accordingly.

These subjects came from different countries such as Korea, Japan, Uzbekistan, Taiwan, Nigeria, Indonesia, Myanmar, Nepal and Bangladesh. They were of diverse cultural backgrounds. There were nine native languages reported by the subjects: Korean, Russian, Uzbek, Chinese, English, Indonesian, Burmese, Nepali and Bengali. The length of stay in the United States ranged from 5 days to 1.5 years with a mean of 4.3 months. The average number of years spent studying English was 8.1 years with a range of 2 to 15 years. Their majors varied from public administration, international business, communication sciences and disorders, electrical engineering, architecture, computer science and accounting.

Procedure

Since the purpose of this study is to explore the relationship between adult L2 learners' speaking and writing, both the written and spoken samples were collected from

the subjects. Permission to conduct the investigation was obtained from their instructor before the study was conducted. Each subject signed a consent form before his or her participation, which was completely voluntary. Please see Appendix C for Approved IRB Review Form.

The written samples of this study were the regular writing assignment of the subjects, which were taken from the diagnostic essay the subjects wrote for the above-mentioned composition course. Compared to other essays written for this course, the diagnostic essay is written at the beginning of the course as the first written assignment for the instructor to examine the strengths and weaknesses of the subjects' writing skills. For this course, every essay requires two drafts, which receive two different evaluations. For the diagnostic essay, the first draft is not graded, but receives comments from the instructor. According to the comments, the second draft is written and graded. The instructor evaluates students' essays with regard to content, organization, grammar, word usage and mechanics. Students basically spend two weeks, about six class periods writing the diagnostic essay in class. The essay examined in the present study was the first draft of the diagnostic essay the subjects wrote during the fall semester of 2002. The essay prompt required them to describe three underlying rules shaping student behavior in American classrooms, which they had discovered since they came to Oklahoma State University. Since the diagnostic essay is written in class, it has less planning time than other essays that are written after class by the students. Therefore, the use of the diagnostic essays increases the comparability between the spoken and written data.

As noted earlier, 40 students from two classes taught by the same instructor participated in this study, so 40 papers were originally collected from the second written

assignment of the class. Then, according to the rating of the instructor, 20 students took part in the study: 10 high-rated and 10 low-rated students. After the written samples were collected, three raters graded the papers holistically on a scale of 100. A higher interrater reliability was found between Raters 1 and 2 ($r=0.91$; $p=0.0002$) than between Rater 3 and the other two raters, respectively. As Davies, et al. (1999) suggests, a correlation coefficient of more than 0.8 indicates a good interrater reliability. So, the rating between Raters 1 and 2 was used as the evaluation of the written samples. Specifically, the ratings of these two raters were averaged and the averaged score was taken as the final score. The averaged score of Raters 1 and 2 were then transferred to the rank-order of the written samples from 1 to 10. Of these 20 students, 5 high-rated and 5 low-rated students were finally chosen to participate in this study. The average of 5 high-rated students' written score was 90.1 and that of the 5 low-rated students' written score was 77.1. So, there were 13 points apart between the written samples of 5 high-rated and 5 low-rated students.

After the written data of these 10 students were collected, the spoken data were gathered through the students' participation in the VOCI. The subjects took the VOCI in a testing room alone in order for them to feel less nervous. As noted earlier, the VOCI is a semi-direct, tape-mediated oral proficiency test, which lasts approximately 30 minutes and consists of a total of 23 questions (Please see Appendix A for the VOCI questions). Of these 23 questions, the subjects' answers to 17 questions were used for analysis mainly according to the familiarity of the topics of these questions. These 17 questions are questions 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 15, 16, 17, 18, 21, 22 and 23. In a specific manner, questions 1 and 2 function as the practice questions in order for the students to

familiarize themselves with the use of the machines. Questions 13, 14, 19 and 20 were not used for data analysis because the researcher felt that the subjects might not be familiar with the topics such as "lasting peace" (Q 13), "abstract painting" (Q 14), "free trade" (Q 19), "televising trials" (Q 20). After the collection of the spoken samples, they were rated and rank-ordered from 1 to 10 by a certified OPI tester according to the ACTFL Guidelines (1986).

After the spoken samples were gathered, they were transcribed. In regard to the spoken sample, each subject's answers to all the above-mentioned 17 VOCI questions were taken as a whole to be analyzed in comparison with the subjects' written samples.

Research Hypotheses

The hypotheses of this study consist of three parts: one kind is the hypotheses on the relationship between speaking and writing regarding high-and low-rated subjects. Another kind is the hypotheses on the correlations between the spoken and written samples of all subjects. And the third kind is the hypotheses on the differences between speaking and writing.

Hypotheses on the Relationship between Speaking and Writing regarding High-and Low-rated Subjects

These hypotheses examine the relationship between speaking and writing with respect to high-and low-rated subjects. The categorization of the subjects into high-and low-rated subjects depends on the rank-order of the subjects' spoken and written data, which will be revealed in hypothesis 1. Specifically, they study whether students with good writing skills have good speaking skills and vice versa. Hypothesis 1 concerns the rank-order of the subjects' spoken and written data.

Hypothesis 1: *Students who were ranked highly in the spoken samples will be those who were ranked highly in the written samples and vice versa.*

If hypothesis 1 is sustained, there will be 5 high-rated subjects and 5 low-rated ones in regard to both samples. Specifically speaking, 5 high-rated students will be ranked highly in both samples and 5 low-rated students will be placed in the low rank area in both samples. Hypotheses 2-5 are concerned with the measures of syntactic maturity and lexical density.

Hypothesis 2: *High-rated students will produce more words and longer T-units and clauses in a) the spoken and b) written samples than low-rated ones.*

Hypothesis 3: *High-rated students will produce longer error-free T-units and more error-free T-units than low-rated ones in a) the spoken and b) the written samples.*

Hypothesis 4: *High-rated students will produce more dependent clauses in a) the spoken and b) the written samples than low-rated ones.*

Hypothesis 5: *High-rated students will produce more lexical words than low-rated ones in a) the spoken and b) the written samples.*

Hypotheses on the Correlation between the Spoken and Written Samples of All Subjects

Hypotheses 6 and 7 are concerned with the correlation between each subject's spoken and written sample in relation to the indices of both syntactic maturity and lexical density. These two hypotheses also explores whether students who are found to speak well are judged to write well too and vice versa. But they take all subjects' written and

spoken performance into account instead of considering high- versus low-rated subjects as demonstrated in hypotheses 1-5.

Hypothesis 6: *The written samples will positively correlate with the spoken samples with regard to the measures of syntactic maturity for all subjects.*

Hypothesis 7: *The written samples will positively correlate with the spoken samples in relation to the measures of lexical density for all subjects.*

Hypotheses on the Differences between Speaking and Writing

With respect to these hypotheses, they are set to examine the differences between speaking and writing, which is part of the relationship between speaking and writing as noted in Chapter I. Here are these hypotheses:

Hypothesis 8: *The written samples will contain longer T-units and longer clauses than spoken samples.*

Hypothesis 9: *The written samples will have longer error-free T-units and more error-free T-units than the spoken samples.*

Hypothesis 10: *The written samples will contain more dependent clauses than the spoken samples.*

Hypothesis 11: *The written samples will contain more lexical words than the spoken samples.*

Data Analysis

Three procedures were carried out to analyze statistical data in regard to the purpose of this study. First, the objective measures of syntactic maturity and lexical density were marked and tallied in both the spoken and written samples. With respect to syntactic maturity and lexical density, means of each measure were computed separately

concerning high- and low-rated samples and also were calculated separately regarding the spoken and written samples. Second, the data were analyzed using version 3.03 of the statistical software, GraphPad Prism. Pearson product-moment correlation was carried out to demonstrate how each measure in the spoken and written data correlates with each other so as to investigate the relationship between speaking and writing. Finally, the level of significance was computed with version 8 of SAS software to examine whether the results obtained were statistically significant or not. Results were considered significant at the $p < .05$ level. Chapter IV presents and discusses the results of the above analyses.

CHAPTER IV

RESULTS AND DISCUSSION

This chapter presents and discusses the results from the assessment measures

introduced in Chapter III. The present study examines the relationship between speaking

and writing, which consists of three kinds as noted earlier. One is about the relationship

between speaking and writing with regard to high-and low-rated subjects according to the

rank-order of their spoken and written samples. Another is concerned with the correlation

between the spoken and written samples for all the subjects. And the last one deals with

the differences between these two modes. Therefore, this chapter will report and discuss

the findings of these three kinds of relationships between speaking and writing.

The Relationship between Speaking and Writing regarding High-and

Low-rated Subjects

This relationship is mainly concerned with whether students with good writing

skills have good speaking skills too and vice versa. This section first demonstrates the

rank-order of the subjects in terms of their speaking and writing performance. And then,

results of the measures of syntactic maturity and lexical density are analyzed to further

explore this relationship.

The Rank-Order of the Subjects

Table 4 provides an overview of how the subjects' speaking and writing related to

each other in terms of the rank-order of the subjects in accordance to their speaking and

writing performance. Both the spoken and written samples were ranked from 1 to 10. In

order to protect the confidentiality of the subjects, pseudonyms were used to identify each

subject. With regard to the spoken samples, Melody, Andrea and Polina were rated as No

1, since these subjects had the highest level of oral proficiency. Similarly, John and Kala were both ranked as No 4 and Anna and Tom were rated No 8. As revealed in the table, Melody, Andrea, Polina, John and Kala were rated highly in both spoken and written samples as compared to Anna, Emily, Tom, Jack and Mike, who were ranked in both

TABLE 4

The Rank Order of Subjects (Identified by pseudonyms) in both Spoken and Written Samples (N=10)

Ranking	Spoken Sample	Written Sample
High-rated Ones		
1	Melody, Andrea, Polina	Melody
2		Andrea
3		Polina
4	John, Kala	Kala
5		John
Low-rated Ones		
6	Emily	Jack
7	Jack	Tom
8	Anna, Tom	Emily
9		Anna
10	Mike	Mike

samples as the low-level group. In other words, Melody, Andrea, Polina, John and Kala belonged to the high-rated group while Anna, Emily, Tom, Jack and Mike were in the

low-rated group in both samples. In general, this finding indicates that subjects who write well also speak well and vice versa.

Concerning the spoken samples, the difference is related to the holistic nature of the oral proficiency rating scales. With a close look at the table, one could find that Melody, Andrea and Polina were rated No1 in the spoken sample as compared to No 1, 2 and 3, respectively in the written sample. So, they were almost ranked the same in both the spoken and written samples. The same observation could also apply to John and Kala, who were ranked No 4 in the spoken sample and No 4 and 5 in the written sample. As far as the low-rated subjects were concerned, Mike was ranked No 10 in both samples. Thus, the rank-order of the subjects in light of their speaking and writing performance indicated that there seemed to exist a positive relationship between the subjects' speaking and writing. Stated differently, if students are found to write well, they are judged to speak well too and vice versa.

The Relationship between Speaking and Writing regarding High-and Low-rated Subjects with Measures of Syntactic Maturity

As introduced in Chapter III, the measures of syntactic maturity were used to investigate the syntactic development of the subjects' speaking and writing. Specifically, seven objective measures were employed, which were total number of words (TNW), mean T-unit length (MTUL), mean clause length (MCL), mean error-free T-unit length (MEFTUL), percentage of error-free T-units (%EFTU), subordination ratio (SR) and dependent clauses per clause (DC/C). As noted above, the same subjects were ranked as high- and low-rated groups in both spoken and written samples. So, the analysis of this relationship with measures of syntactic maturity was particularly concerned with the

comparison of high- and low-rated groups in terms of the above-mentioned seven measures.

Tables 5 and 6 show the means of these measures of both high- and low-rated groups in the spoken and written samples as well as the level of significance for each measure. As noted in Tables 5 and 6, the high-rated group demonstrated a higher index in all the measures than the low-rated group in both samples. Specifically, the mean value of TNW was 1620 words for the high-rated group as opposed to 1607 words for the low-rated group in the spoken samples, and the written sample demonstrated a mean value of 550 words versus 363 words for the high-and low-rated groups, respectively. Regarding MTUL, the high-rated group showed a mean value of 12.82 and 17.28 words per T-unit while the low-rated group exhibited a mean value of 9.40 and 12.46 words per T-unit in the spoken and written samples, respectively. Insofar as MCL was concerned, the means of the high-rated group were 7.28 and 9.07 words per clause as opposed to 6.94 and 7.13 words per clause for the low-rated group in the spoken and written samples, respectively. In relation to MEFTUL, the means of 14.49 versus 13.96 words per error-free T-unit for the high-and low-rated students respectively were found in the spoken data and the means of 25.30 versus 22.49 words per error-free T-unit for the two groups were noted in the written data. %EFTU displayed a mean value of 88% and 71% of error-free T-units for the high-rated group versus 68% and 60% of error-free T-units for the low-rated group in the spoken and written samples, respectively. For the last two measures of syntactic maturity, SR showed a mean value of 1.74 versus 1.35 dependent clauses per T-unit in the spoken data and a mean value of 1.90 versus 1.73 dependent clauses per T-unit in the written data for the two groups, respectively. Concerning DC/C, the high-rated group had

a mean value of 0.41 and 0.42 dependent clauses per clause as opposed to 0.26 and 0.34

dependent clauses per clause for the low-rated group in the spoken and written samples,

respectively.

TABLE 5

Average of Objective Measures of Syntactic Maturity of both High- and Low-rated

Groups (Spoken Sample) *(* Significant at p <. 05)*

Measures	High-rated Group	Low-rated Group
TNW	1620	1607
MTUL	12.82	9.40
MCL	7.28	6.94
MEFTUL	14.49	13.96
***%EFTU**	0.88	0.68
***SR**	1.74	1.35
***DC/C**	0.41	0.26

According to this result, it might be said that the high-rated subjects do produce

more words, longer T-units, longer clauses, longer error-free T-units, and more error-free

T-units and incorporate more dependent clauses than the low-rated ones. However,

insofar as the level of significance is concerned, three measures, %EFTL, SR and DC/C

showed a significant difference in regard to the spoken sample. Concerning the written

sample, the mean differences between the high and low-rated groups were significant in

relation to MCL and TNW. As mentioned earlier, SR and DC/C measure the amount of

subordination and TNW and MCL tap the quantity of both spoken and written samples.

Thus, it can be concluded that high-rated students do speak with more dependent clauses than the low-rated group, which, however, may not be the case when they write. Besides, when writing, high-rated students write more words than low-rated ones. But, this might not hold true when they speak. In light of %EFTU, high-rated students speak significantly with more error-free T-units than low-rated ones. Nevertheless, they may not write with more error-free T-units than low-rated ones.

TABLE 6

Average of Objective Measures of Syntactic Maturity of both High- and Low-rated Groups (Written Sample) *(* Significant at p<.05)*

Measures	High-rated Group	Low-rated Group
*TNW	550	363
MTUL	17.28	12.46
*MCL	9.07	7.13
MEFTUL	25.30	22.49
%EFTU	0.71	0.60
SR	1.90	1.73
DC/C	0.42	0.34

As mentioned in the above paragraph, MCL and TNW were found statistically different regarding the written sample. So, the present study indicated that TNW and MCL were better indices to differentiate between good and poor students concerning writing. It has been noted from the other studies that MTUL is a useful measure of L2 development (e.g. Gaies, 1980; Larsen-Freeman & Storm, 1977). Since MTUL did not

indicate significance in the present study, the present study did not confirm the findings

of other studies. Besides MTUL, Kameen (1983) also considered MCL as a reliable index

to distinguish between L2 learners' written proficiency levels, which did hold true for the

present study. In the present study, a significant difference was found between the high-

and low-rated groups in terms of MCL (p <.05). Kameen claims that good writers write

an average of 18.40 words/T-unit while poor writers did 14.30 words/T-unit. According

to this criterion, the high-rated students in the present study can be grouped into good

writers, since their average MTUL of them is 17.28 words/T-unit with a range from 10.50

to 20.12 words/T-unit. On the other hand, the low-rated students have the average MTUL

of 12.46 words/T-unit with a range from 8.78 to 15.76 words/T-unit. Although the

average MTUL of high- and low-rated students in the present study is slightly lower than

Kameen's figures, they are approximately close to these figures. Thus, it may be said that

the present study agrees with Kameen on the criterion to distinguish between good and

poor writers among college-level ESL students.

An important finding of L2 research is that MEFTUL and %EFTU are believed to

be more reliable and powerful to discriminate levels of proficiency regarding writing,

since errors frequently appear in adult L2 data (e.g. Scott & Tucker, 1974; Gaies, 1976;

Vann, 1978; Sharma, 1979). However, the present study indicated that MEFTUL and

%EFTU could not serve as a good index to distinguish between high-and low-rated

students, because they did not reach the significant level. Therefore, the finding of the

present study does not support those of L2 studies. As far as speaking is concerned, there

are not many L2 studies using the measures of syntactic maturity. One study was carried

out by Halleck (1995), who found that objective measures of syntactic maturity could be

used to demonstrate L2 learners' oral proficiency. In regard to the present study, %EFTU

indicated a significant difference between the high-and low-rated groups. Hence, this

finding of the present study has partially supported Halleck's study.

The Relationship between Speaking and Writing regarding High-and Low-rated

Subjects with Measures of Lexical Density

As introduced in Chapter 3, three measures of lexical density were used to explore

the relationship between speaking and writing in the present study, which were number of

lexical words/number of words (LD/TNW), number of lexical words/number of clauses

(LD/C) and of lexical words/number of T-units (LD/TU). Like the above section, the

results in this section will be reported and discussed in regard to the average of each

index in terms of high- and low-rated groups and the correlation between the spoken and

written data concerning each index for all the subjects.

Tables 7 and 8 indicate the comparison between high and low-rated groups in

relation to the mean of each measure of lexical density in the spoken and written samples,

respectively. As observed from the two tables, LD/TNW showed a lower index for the

high-rated group (LD/TNW=0.34) than for the low-rated group (LD/TNW=0.37) in the

spoken samples. The written samples also demonstrated the same finding with LD/TNW

for the high-rated group being lower (LD/TNW = 0.42) than that for the low-rated group

(LD/TNW=0.43). On the other hand, in regard to LD/TU, the finding is reversed. In both

samples, LD/TU exhibited a higher index for the high-rated group (LD/TU=4.40 in the

spoken sample; LD/TU=7.17 in the written sample) than for the low rated group

(LD/TU=3.45 in the spoken sample; LD/TU=5.36 in the written sample). Concerning the

significant level, only one index, LD/C in the written sample was found statistically

significant at p<.05.

TABLE 7

Average of the Measures of Lexical Density of both High- and Low-rated

Groups (Spoken Sample) (*significant at p<.05)

Measures	High-rated Group	Low-rated Group
LD/TNW	0.34	0.37
LD/C	2.50	2.56
LD/TU	4.40	3.45

TABLE 8

Average of the Measures of Lexical Density of both High- and Low-rated

Groups (Written Sample) (*significant at p<.05)

Measures	High-rated Group	Low-rated Group
LD/TNW	0.42	0.43
*LD/C	3.79	3.08
LD/TU	7.17	5.36

Therefore, the results of Tables 7 and 8 indicate that the high-rated students spoke

with a higher density than the low-rated students with regard to LD/TU, but not in terms

of LD/TNW and LD/C insofar as speaking is concerned. In light of the written sample, it

was found that the high-rated group wrote more lexical words than the low-rated group with respect of LD/C and LD/TU, but not in relation to LD/TNW. In addition, of these measures, LD/C indicated a significant difference between the two groups. So, it can be said that high-rated students produce more lexical words than low-rated ones in terms of LD/C. On the whole, concerning the spoken data, none of the measures reached the significance level. With respect to the written data, only LD/C was found statistically different regarding the two groups. In relation to this index, it can be said that students' written samples have a higher density than their spoken samples. Overall, according to Tables 7 and 8, the measures of lexical density cannot distinguish between high-and low-rated groups.

The Correlation between Spoken and Written Samples of All Subjects

In discussing the relationship between the subjects' speaking and writing, it is also useful to investigate whether there is a positive correlation between the spoken and written data in terms of the measures of syntactic maturity and lexical density for all subjects. If this correlation can be found between these two modes, it may be concluded that the one who speaks well writes well too and vice versa.

The Correlation between Spoken and Written Samples with Measures of Syntactic Maturity of All Subjects

Table 9 reflects this correlation with measures of syntactic maturity. In Table 9, a Pearson product-moment correlation coefficient was computed to show the relationship between the spoken and written data. As far as this correlation is concerned, TNW was not considered because the topics of the written and spoken data were different from each other. Besides, there were 17 topics in the spoken data as opposed to 1 topic in the

written data. So, the spoken and written data were not of comparability in terms of this index. Thus, all the other measures except TNW were calculated to exhibit this correlation. Of all the measures, MTUL showed the strongest correlation with a correlation coefficient r being 0.85 followed by DC/C (r=0.75), %EFTU (r=0.72), SR (r=0.71), MCL (r=0.67) and MEFTUL (r=0.67).

As indicated in this table, the spoken and written data were found to positively and significantly correlate with each other in light of all the measures. So, it supports the assumption made above that the one with good speaking skills has good writing skills and vice versa. When examined closely, it can be noted that MTUL demonstrates the strongest correlation between the spoken and written data followed by DC/C, %EFTU,

TABLE 9

The Correlation between the Spoken and Written Data with Regard to the Measures of Syntactic Maturity for All Subjects (*significant at p<.05)

Measures	Correlation Coefficient (r)
MTUL	0.85*
MEFTU	0.67*
%EFTU	0.72*
SR	0.71*
MCL	0.67*
DC/C	0.75*

SR, MEFTUL and MCL, respectively. Therefore, it can be concluded from Table 9 that the students with good speaking skills have good writing skills and vice versa. Besides, MTUL is the most sensitive index to exhibit the relationship between speaking and writing with respect to high-and low-rated students.

The Correlation between Spoken and Written Samples with Measures of Lexical Density of All Subjects

Besides the discussion of the average of the measures of the high- and low-rated students in both samples, the present study also reported the correlation between spoken and written data with regard to measures of lexical density as indicated in Table 10. From the table, the spoken and written data correlated positively with regard to the three indices with the correlation coefficients of LD/TNW being 0.35, that of LD/C being 0.15 and that of LD/TU being 0.75. Of these three measures, only LD/TU demonstrated a relatively strong and significant correlation between the two samples.

TABLE 10

The Correlation between Spoken and Written Data with Regard to the Measures of Lexical Density for All Subjects (*significant at $p<.05$)

Measures	Correlation Coefficient (r)
LD/TNW	0.35
LD/C	0.15
LD/TU	0.75*

From the table, we see that a positive correlation was found between the spoken and written data with respect to one of measures of lexical density. Only LD/TU indicated a relatively strong and significant correlation between two samples. That is to say, students with high proficiency in an L2 produced more lexical words per T-unit when they speak and write than those with low proficiency. Therefore, this finding proves that LD/TU can sensitively demonstrate the relationship between speaking and writing concerning the levels of proficiency. The reason why this index may exhibit the relationship between speaking and writing may be that this index combines the measures of both lexical density and syntactic maturity. As noted earlier in Tables 5 and 6, although the measures of syntactic maturity did not all indicate a significant difference between the two groups of students in both speaking and writing, all the measures of the high-rated group were found to be higher than those of the low-rated group in both samples. Besides, the spoken and written data correlate positively and significantly with each other with respect to all the measures of syntactic maturity. Therefore, LD/TU would probably serve as a sensitive indicator of ESL students' proficiency levels. This result will be supported even further in the later section.

To conclude, significant correlation between the two samples was found in regard to all the measures of syntactic maturity, but not to all the measures of lexical density. Only LD/TU revealed a significance level between these two modes. As noted, LD/TU was a combination of measures of both syntactic maturity and lexical density. Therefore, it can be concluded that measures of syntactic maturity are good indicators of levels of proficiency, which does not hold true for the measures of lexical density. The findings of this section agree with those of L2 studies with the measures of lexical density in that the

measures of lexical density, specifically, LD/TNW and LD/C cannot differentiate between L2 learners' proficiency levels. As noted in the literature review, L2 studies using the measures of lexical density (e.g, Linnarud, 1986; Lauren, 2002) agree that lexical density is not a good indicator of L2 proficiency. To illustrate, Linnarud's (1986) study showed a small, non-significant difference between the Swedish native speakers and Swedish learners in lexical density and a high density did not correlate with a high grade. Lauren (2002) even found that the lexical density of the immersion pupils in both spoken and written samples demonstrated a higher density than that of the Swedish-speaking control group. It is noted that these studies either used LD/TNW or LD/C. In line with these studies, the present study also arrived at the same findings using these two measures: that lexical density cannot distinguish between high and low-proficient ESL students with respect to either their speaking or writing. However, an interesting finding obtained from the present study is that LD/TU may serve as a sensitive indicator of ESL students' proficiency levels. Overall, compared to the measures of syntactic maturity, the measures of lexical density cannot differentiate between proficiency levels and so they cannot serve to measure whether there is a positive relationship between the spoken and written data.

The Differences between Speaking and Writing

The previous section was devoted to the analysis and discussion of the relationship between speaking and writing regarding the correlation between the spoken and written data. Another kind of this relationship investigated in the present study is about the differences between speaking and writing. This relationship will be analyzed

and reported in this section with the measures of syntactic maturity and lexical density, respectively.

The Differences between Speaking and Writing with Measures of Syntactic Maturity

Table 11 presents the means of all measures of syntactic maturity of spoken and written data. From this table, it can be observed that all the measures except %EFTU were higher in the written data than in the spoken data. Specifically, the average of MTUL in the written samples was 14.87 words per T-unit as opposed to 11.11 words per T-unit in the spoken samples. Similarly, the written samples showed a higher index with the mean of MCL being 8.10 words per clause than the spoken samples with that of MCL being 7.11 words per clause. Regarding MEFTUL, the average of it in the written

TABLE 11

Average of All Measures of Syntactic Maturity of Spoken and Written Data

*(*significant at p<. 05)*

Measures	Spoken	Written
*MTUL	11.11	14.87
*MCL	7.11	8.10
MEFTU	14.21	23.89
%EFTU	0.78	0.65
SR	1.54	1.81
DC/C	0.34	0.38

samples (mean=23.89) was also higher than that of it in the spoken samples

(mean=14.21). Likewise, the mean values of SR and DC/C exhibited a higher index in

the written samples (mean for SR =1.81; mean for DC/C=0.38) than in the spoken

samples (mean for SR =1.54; mean for DC/C=0.34). However, given the level of

significance, a significance difference was found only with the measures of MTUL and

MEFTUL. Nevertheless, unlike other indices, %EFTU indicated a lower index in the

written sample (mean=0.65) than in the spoken sample (o.78). Besides, this index

approached the significant difference with p=0.06.

According to the above results, in general, Table 11 demonstrated that the written

samples contained longer T-units, longer clauses and longer error-free T-units and

incorporated more dependent clauses than the spoken samples. However, since only

MTUL and MEFTUL showed a significant difference between the spoken and written

data, it may be said these two indices can function as good indicators of the differences

between these two modes.

However, one thing worth noting is that %EFTU showed that the subjects' spoken

discourse included more error-free T-units than their written discourse. This finding is

very unexpected, because it might be assumed that written discourse should contain more

error-free T-units than spoken discourse since more planning time is available in writing

than in speaking. However, a close look at %EFTU of each subject's spoken and written

data demonstrated that all the subjects produced more error-free T-units in speaking than

writing. Besides, this index approached significance. So, this finding would probably not

be obtained by chance. With a larger sample, it may be found that this index would

indicate a significant difference between these two modes of communication. With a

close examination of the spoken transcripts of the subjects, it was found that the subjects spoke in short sentences, which made it unlikely for them to make errors. However, when they wrote, they tended to write longer sentences with more dependent clauses. In this way, it is very likely for them to make errors in writing. As shown in the table, SR and DC/C exhibited that the subjects used more dependent clauses in writing than in speaking. To illustrate, approximately 100 words from the same subject's spoken and written data were taken to demonstrate this difference between the two modes.

The Written Sample:

In my country, Indonesia, there are many culture characteristics *which are respected. Because Indonesia is an Asian country and we always call Asian countries "East Countries"./* So, there are many "rules" in our country.*/ There are three culture characteristic *which are respected in my country* */and those are always giving something with your right hand, don't ever call a person *who is older than you only with his or her name* and always walk on the left side./ Giving or receiving something with your left hand is unpolite.*/ We are always taught with our parents *since we were a child* to give and receive with right hand./

As observed from the written sample, it consists of 6 T-units (between the virgules), 5 dependent clauses (italicized) and 3 error-free T-units (with an asterisk). According to the guideline for T-units, clauses, word counts and errors provided by Polio (1997), if a subordinate clause is standing alone, attach it to the preceding or following sentence and count as a 1 T-unit with an error. So, the first T-unit (*In my country, Indonesia, there are*

many culture characteristics which are respected. Because Indonesia is an Asian country

and we always call Asian countries "East Countries".) is counted as a T-unit with an

error. As for the third T-unit (*There are three culture characteristic which are respected*

in my country) and the fifth T-unit (*Giving or receiving something with your left hand is*

unpolite.), the spelling mistakes are not taken into account. So, these two T-units are

error-free. In comparison to the written sample, the spoken sample contains 8 T-units, 3

dependent clauses and 6 error-free T-units.

The Spoken Sample

Ten years is a very very long time.*/ I mean, ok honestly, I have nothing in my

mind *what I'm gonna be next ten years, because it's a lot of time.*/ But, my goal

in the next four years absolutely is *I want to graduate from here as soon as*

possible./ I want to graduate from universities.*/ I don't know, I want to

graduate from here or another university./ But I want to graduate as soon as

possible.*/ And I want to get great job in the future./ And I also plan to get my

master's degree.*/

To illustrate and explain the reason why %EFTU is generally higher in the spoken

than in the written samples, Table 12 shows the MTUL, MCL, MEFTUL, %EFTU, SR

and DC/C in both samples. As noted from Table 12, the %EFTU was obviously higher in

the spoken data (%EFTU=0.75) than in the written data (%EFTU=0.50). Concerning

other indices, all except MCL showed a lower index in the spoken sample than in the

written sample. As observed from the table, the subject produced fewer words per T-unit

in the spoken (MTUL =12.50) than the written sample (MTUL=16.70). Regarding

MEFTUL, the spoken sample demonstrated a lower index (MEFTUL=16.67) than the

written sample (MEFTUL=33.33). Besides, the subject produced fewer dependent clause

per T-unit in the spoken (SR=1.38) than in the written sample (SR=1.83). Similarly,

DC/C indicated a lower index in the spoken sample (DC/C=0.27) than in the written

sample (DC/C=0.45).

TABLE 12

Measures of Syntactic Maturity of Selected Spoken and Written Data from the

Same Subject (TNW=100 words)

Measures	Spoken	Written
MTUL	12.50	16.70
MCL	9	9
MEFTU	16.67	33.33
%EFTU	0.75	0.50
SR	1.38	1.83
DC/C	0.27	0.45

In a discussion of the above results, the subject produced fewer words per T-unit

and per error-free T-unit, fewer dependent clauses per T-unit and per clause in the spoken

than in the written samples. So, when the subject speaks, he speaks in shorter sentences

than he writes. Besides, he does not incorporate as many dependent clauses in a T-unit in

speaking as in writing. Therefore, it's less likely for him to make errors when he speaks.

On the other hand, since his writing include many long sentences and dependent clauses,

he is very likely to make errors in writing. That's why his spoken sample shows more

error-free T-units than his written sample.

The Differences between Speaking and Writing with Measures of Lexical Density

This section is particularly devoted to the analysis of the differences between

speaking and writing in relation to the measures of lexical density. Table 13 reflects the

mean differences between these two modes with respect to three measures of lexical

density. As noted in the table, all the measures demonstrated a higher density in the

written discourse than in the speaking discourse. In a specific manner, LD/TNW was 0.36

lexical words/number of words in the spoken data and 0.43 lexical words/number of

words in the written data. With respect to LD/C, the spoken samples showed a lower

index (LD/C=2.53 lexical words/clause) than the written samples (LD/C=3.43 lexical

words/clause). Concerning LD/TU, it was lower in the spoken samples (LD/TU=3.93

words/T-unit) than in the written samples (LD/TU=6.26). Moreover, the differences

between these two modes were found very significant in relation to all the measures of

lexical density.

TABLE 13

Average of All the Measures of Lexical Density of Spoken and Written Data

(*Significant at p<. 05)

Measures	Spoken	Written	Percentage Difference
*LD/TNW	0.36	0.43	19%
*LD/C	2.53	3.43	36%
*LD/TU	3.93	6.26	59%

The results obtained from Table 13 show that written samples overall are more lexically dense than the spoken samples. This finding is in line with L2 studies (e.g. Hyltenstam, 1988) using the measures of lexical density in that written production has a higher density than spoken production. Hyltenstam (1988) reported that spoken data contains a lower lexical density than written data with the lexical density of oral texts being around 38.3% and that of the written language around 45%. It can be noted from this finding that the mean lexical density of the subjects' spoken data was 36% and that of the written data was 43%, which is almost in agreement with Hyltenstam's finding. In short, the present study shows that lexical density does function as a sensitive indicator of the differences between speaking and writing.

As far as these three measures are concerned, it seems that LD/TU is more sensitive than the other two measures, since it indicated the largest difference between the two modes as shown by the percentage difference as indicated in Table 13. The percentage difference is computed by subtracting the means of each measure of the written data from that of the spoken data and then dividing the number by the mean of the spoken data. This figure is used to show how much difference is between the two modes.

As shown in the table, the percentage difference of LD/TU is 59% followed by LD/C being 36% and LD/TNW being 19%. So, it can be said that LD/TU is the most sensitive indicator of the difference between the two modes with LD/C being the second best indicator and LD/TNW the third best indicator. This finding is supported even further by the data shown in Table 14.

Table 14 provides detailed information of LD/TNW, LD/C and LD/TU of both spoken and written data of each student separately. As shown in the table, in general, the written data exhibited a higher density than spoken data as far as each student is concerned. However, one exception was Emily. This subject's LD/TNW was the same in both samples (LD/TNW =0.40) and the LD/C was lower in the written sample

TABLE 14

The Comparison of Spoken and Written data with Respect to the Measures of

Lexical Density for Each Subject Separately

Student Name	LD/TNW		LD/C		LD/TU	
	Spoken	Written	Spoken	Written	Spoken	Written
Anna	0.38	0.44	2.52	3.30	3.22	5.44
Emily	0.40	0.40	2.58	2.12	3.34	3.50
Melody	0.31	0.38	2.53	3.44	5.29	7.44
Tom	0.32	0.40	2.35	3.23	3.57	5.96
Jack	0.37	0.43	2.78	3.53	4.11	6.71
Andrea	0.35	0.43	2.27	3.63	3.57	7.96
Polina	0.38	0.42	3.05	3.84	5.99	8.44
John	0.33	0.46	1.99	3.55	2.87	4.81
Mike	0.38	0.50	2.57	3.20	3.02	5.16
Kala	0.34	0.41	2.67	4.46	4.29	7.20

(LD/C=2.12) than in the spoken sample (LD/C=2.58). Only the LD/TU demonstrated a similar result with other subjects that it was higher in the written sample (LD/TU=3.50) than in the spoken sample (LD/TU=3.34). As far as the literature was reviewed in both L1 and L2 research using lexical density, it has been confirmed that writing contains a higher density than speaking (e.g. Ure, 1971; Halliday, 1979; Hyltenstam, 1988). Besides, all the other subjects in the present study produced more lexical words in writing than speaking. Therefore, LD/TU can most sensitively differentiate between speaking and writing among these three measures of lexical density.

In sum, the spoken samples are shown to have more lexical density than the written samples. And, of all the measures of lexical density, LD/TU seems to be the best indicator of the differences between the spoken and written data.

CHAPTER V

CONCLUSION

This chapter presents the conclusion of the study by discussing outcomes of the hypotheses, implications of the results, the limitations of the study and future research directions. As noted earlier, seven measures of syntactic maturity were used to explore the syntactic development of the subjects' speaking and writing. These measures are Total number of words (TNW), mean T-unit length (MTUL), mean clause length (MCL), mean error-free T-unit length (MEFTUL), percentage of error-free T-units (%EFTU), subordination ratio (SR) and dependent clauses per clause (DC/C). Besides, three measures were adopted to investigate the lexical development of the subjects' speaking and writing. These three measures are lexical density per number of words (LD/TNW), lexical density per clause (LD/C) and lexical density per T-unit (LD/TU).

Hypotheses Outcomes

Hypothesis 1: *Students who were ranked highly in the spoken samples will be those who were ranked highly in the written samples and vice versa.*

This hypothesis deals with the rank-order of the subjects in both the spoken and written samples. The results of the present study show that the same subjects were ranked as high- and low-rated groups in both spoken and written samples. That is to say, the subjects who were ranked highly in the spoken samples were also ranked highly in the written samples and vice versa. Therefore, this hypothesis is proved.

Based on this hypothesis, hypotheses 2-5 are concerned with the relationship between speaking and writing in relation to high-and low-rated subjects.

Hypothesis 2: *High-rated students will produce more words and longer T-units and clauses in a) the spoken and b) written samples than low-rated ones.*

This hypothesis is about three indices of syntactic maturity, namely, TNW, MTUL and MCL, which measure quantity. As far as the spoken data were concerned, the high-rated group demonstrated a higher mean index than the low-rated group regarding all three of the measures. However, all three measures were not found statistically significant between high- and low-rated groups. With respect to the written data, high-rated students were found to write more words, longer T-units and clauses than low-rated ones. Unlike the spoken data, of three measures, TNW and MCL indicated a statistical difference between two groups. Therefore, on the whole, the hypothesis *a* is rejected and hypothesis *b* is sustained with respect to TNW and MCL.

Hypothesis 3: *High-rated students will produce longer error-free T-units and more error-free T-units than low-rated ones in a) the spoken and b) the written samples.*

This hypothesis has to do with EFTUL and %EFTU, which tap accuracy. The findings of the study exhibited that high-rated students did produce longer error-free T-units and more error-free T-units than low-rated ones in both samples. Nevertheless, only %EFTU was found statistically significant in the spoken data and both measures indicated non-significant difference in the written data. Thus, hypothesis *a* is proved with respect to %EFTU and hypothesis *b* is rejected.

Hypothesis 4: *High-rated students will produce more dependent clauses in a) the spoken and b) the written samples than low-rated ones.*

Two measures were used to test this hypothesis, SR and DC/C. The

finding of the study demonstrated that the high-rated students did write and speak more dependent clauses than the low-rated ones in both data. However, in regard to the significance level, a significant difference was found in writing but not in speaking concerning these two measures. So, hypothesis *a* is rejected and hypothesis *b* is sustained.

Hypothesis 5: *High-rated students will produce more lexical words than low-rated ones*

in a) the spoken and b) the written samples.

As for this hypothesis, three measures of lexical density were used: LD/TNW, LD/C and LD/TU. In the present study, none of the measures in the spoken samples was found statistically significant. In regard to the written sample, only LD/C showed that high-rated students did significantly produce more lexical words than low-rated students. Therefore, this hypothesis is rejected.

As mentioned in Chapter III, hypotheses 6 and 7 investigate whether the spoken and written samples for all subjects correlate with each other. They take into account all subjects' written and spoken performance rather than considering the spoken and written performance of high-versus low-rated students as demonstrated in hypotheses 1-5.

Hypothesis 6: *The written samples will positively correlate with the spoken samples with*

regard to the measures of syntactic maturity for all subjects.

Positive and significant correlation was found between spoken and written data with respect to all the measures of syntactic maturity. This meant that students with good speaking skills had good writing skills and vice versa. So, this hypothesis is sustained.

Hypothesis 7: *The written samples will positively correlate with the spoken samples in*

relation to the measures of lexical density for all subjects.

As for the lexical density, the study demonstrated that statistical difference was

only found with regard to LD/TU. Besides, of all the measures, LD/TU exhibited the

strongest correlation between the two samples. Concerning the other two measures of

lexical density, they indicated a low and non-significant correlation. As it was assumed,

the reason why LD/TU showed a significant difference between the two samples may be

that this index combines the measures of lexical density and syntactic maturity. In short,

this hypothesis is proved regarding LD/TU but not LD/TNW and LD/C.

Hypotheses 8-11 are set to examine the differences between speaking and writing

with regard to the measures of both syntactic maturity and lexical density.

Hypothesis 8: *The written samples will contain longer T-units and longer clauses than*

spoken samples.

Concerning this hypothesis, the study showed that the written sample did contain

longer T-units and clauses than spoken samples. In addition, MTUL was found

statistically significant. Therefore, it can be said that this hypothesis is sustained with

respect to MTUL but not to MCL.

Hypothesis 9: *The written samples will have longer error-free T-units and more error-*

free T-units than the spoken samples.

As noted from the findings of the study, more error-free T-units appeared in the

spoken sample than in the written sample regarding %EFTU in both samples. This

finding was unexpected. However, this index was not found statistically significant. With

respect to MEFTUL, the written sample contained more words per an error-free T-unit

than the spoken sample and the difference between the two modes was statistically

significant. So, this hypothesis is proved concerning MEFTUL but not %EFTU.

Hypothesis 10: *The written samples will contain more dependent clauses than the spoken*

samples.

As noted earlier, this hypothesis was concerned with SR and DC/C. The finding

of the study demonstrated that both indices were higher in the written sample than in the

spoken sample. However, neither of these two indices was found to reach the significant

level. Thus, this hypothesis is rejected.

Hypothesis 11: *The written samples will contain more lexical words than the spoken*

samples.

The study showed that all the measures of lexical density were higher in the

written samples than in the spoken samples. Besides, such difference was found

statistically significant in relation to all the measures. Therefore, this hypothesis is

sustained.

Implications

The findings of the present study seemed to indicate that there was a positive

relationship between speaking and writing according to the rank-order of the subjects and

the analysis of syntactic maturity of both the spoken and written samples. However, this

implication cannot be generalized, because other aspects of this relationship have to be

considered to get an overall picture of this relationship.

In addition, with regard to the measures of syntactic maturity and lexical density,

the present study found that the measures of syntactic maturity could distinguish between

proficiency levels but could not indicate the differences between the two modes. On the

other hand, the measures of lexical density are good indicators of the differences between the two modes but cannot differentiate between proficiency levels. Therefore, when we are concerned with the distinction between proficiency levels in L2 research or classroom practice, the measures of syntactic maturity may be used. And when we are concerned with the difference between the two modes, the indices of lexical density can be adopted.

Further, of the three measures of lexical density, LD/TU was found to be the most sensitive indicator of the differences between the two modes. Besides, when investigating the correlation between the spoken and written data, LD/TU was the only index which showed a very significant and relatively higher correlation than the other two indices. As discussed in Chapter VI, the reason for this index to be more sensitive than the other two might be that it combines the measures of lexical density and syntactic maturity. Therefore, this index might be a very good measure to distinguish between the proficiency levels and between the two modes. Since very few studies in L2 have used this index to carry out research, it may be suggested that future studies use this index to further test its reliability and validity.

Finally, MEFTUL was not shown in the present study to be a more valid index than MTUL to discriminate between different levels of proficiency as claimed by other L2 studies (e.g. Scott & Tucker, 1974; Gaies, 1976). Thus, it should be cautioned to view this index as a more valid measure of syntactic growth in L2 research.

Research Limitations and Directions

Research Limitations

Like other studies, the present study has its limitations too. First, the present study has a small sample size considering the large population represented in the study. In

particular, there are only 10 subjects who participated in the present study. However, the population represented in the present study is college-level ESL students. Schafer (1981) pointed out that this was the frustrating aspect of the relevant studies and criticized that a few subjects could not generalize the universal relation between speaking and writing. Therefore, the results of the present study should not be generalized about the relationship between the two modes.

A second limitation concerns the topics used in the present study to elicit both the spoken and written samples. Different topics were used to obtain spoken and written samples in the present study. As mentioned previously, Zora & Johns-Lewis (1989) pointed out eight sources of variation for different researchers to obtain different lexical density despite the same tasks. One of these sources is topic. They claimed that "the same 'genre' with different topic and presumably different textual sub-functions can manifest different LD levels" (p.99). Therefore, it is more desirable to use same topics in different tasks in order to obtain a higher comparability. In the present study, with regard to the spoken data, the subjects' response to 17 questions in the VOCI test were examined, so there were actually a total of 17 topics. And the tasks fulfilled by answering these questions vary from describing, comparing and contrasting to supporting an opinion and hypothesizing. However, with respect to the written data, there was only one topic. Only one task, describing, was performed. Apart from this, no topic of speaking is similar to the topic of writing. Nevertheless, researchers who compared speaking and writing have usually used the same topic to elicit both the spoken and written samples in order to avoid the influence of topic. For instance, Cayer and Sacks (1979) asked their subjects to first

have dyadic conversations on a current social problem (women's liberation and the future of the family) and then write about their reactions to the same topic.

Since the present study has these limitations, the conclusions and implications drawn are tentative and should be interpreted only as such. However, directions for future research could be drawn from these limitations, which can provide some useful insights for L2 research and curriculum.

Research Directions

Given the population represented in the present study, a larger sample should be used. Since the targeted population of the present study is college-level ESL students, 10 participants could not adequately represent this population. If a large sample of students could be involved in the present study, the findings of the study would be of much more reliability and of more confident generalization.

Besides, the same topics should be used to elicit both the spoken and written samples with an attempt to control the influence of topic. It may be suggested that same subjects but different topics with different tasks should be used in this kind of study in both modes. In this way, more comparability can be obtained between these two modes. Besides, the subjects' oral and written proficiency can be better represented.

REFERENCES

Beaman, K. (1984). Coordination and subordination revised: Syntactic complexity in spoken and written narrative discourse. In D.Tannen (Ed.), *Coherence in spoken and written discourse*. Norwood, NJ: Ablex.

Beaugrande, R.de.(1984). *Text production: Toward a science of composition*. Norwood, NJ: Ablex.

Bernstein, B. (1971). *Class, codes and control (Vol. 1)*. London: Routledge and Kegan Paul.

Berthoff, A.E. (1980). *The making of meaning*. Upper Montclair, NJ: Boynton/Cook.

Biber, D. (1986). Spoken and written textual dimensions in English: Resolving the contradictory findings. *Language, 62* (2), 384-414.

Biber, D. (1988). *Variation across speech and writing*. Cambridge, UK: Cambridge University Press.

Blankenship, J. (1962).A linguistic analysis of oral and written style. *Quarterly Journal of Speech, 48*, 419-422.

Blanton, L.(1995). Elephants and paradigms: Conversations about teacher L2 writing. *College English, 5* (10), 1-21.

Bloomfield, L. (1933). *Language*. New York: Holt, Rinehart & Winston.

Britton, J. (1970). Talking and writing. In E.Evertts (Ed.), *Explorations in children's writing* (p.30). Champaign, Ill.: NCTE.

Burton, D.L., Donelson, K. L., Fillion, B., & Harley, B. (1975). *Teaching English today*. Boston: Houghton Mifflin.

Bushnell, P.P. (1930). An analytic contrast of oral with written English. In *Teachers*

   college contributions to education No. 451. New York: Bur.Publ, Columbia

   Teachers College.

Cambourne, B.L. (1977). The processing of text-book prose. Proceedings of 4th

   Australian Reading Conference. Sydney: Ashton-Scholastic.

Cambourne, B.L.(1981). Oral and written relationship: A reading perspective. In B.M.

   Kroll & R.J.Vann (Eds.), *Exploring speaking-writing relationships: Connections*

   and contrasts (pp.82-98). Illinois: National Council of Teachers of English.

Chafe, W. (1982). Integration and involvement in speaking, writing and oral literature. In

   D.Tannen (Ed.), *Spoken and written language: Exploring orality and literacy, Vol.*

   IX, Advances in discourse processes (pp. 35-53).Norwood, NJ: Ablex.

Chafe,W. (1985). Linguistic differences produced by differences between speaking and

   writing. In D. R. Olson, N. Torrance, & A. Hildyard (Eds.). *Literacy, language,*

   and learning: The nature and consequences of reading and writing (pp.105-123).

   Cambridge: Cambridge University Press.

Chafe, W., & Tannen, D. (1987). The relation between written and spoken language.

   Annual Review of Anthropology, 16, 383-407.

Charles, R.C. (1978). A study to determine the relationship between indices of maturity

   and a quality grade in a written examination. Retrieved January, 16, 2004 from:

   ERIC database (ERIC Document Reproduction Service No. ED 178 914)

Cooper, T. (1976). Measuring written syntactic patterns of second language learners of

   German. *Journal of Educational Research, 69*, 176-183.

Cramer, R. L. (1978). *Writing, reading, and language Growth: An introduction to language arts*. Columbus, Ohio: Charles E. Merrill.

Davies, A., Brown, A., Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Studies in language testing 7: Dictionary of language testing*. Cambridge: Cambridge University Press.

De Saussure, F. (1959). *Course in general Linguistics*. (C. Bally & A. Sechehaye, Eds.; W.Baskin, Trans.). New York: McGraw-Hill. (Original work published 1916).

DeVito, J.A. *(1965)*. A linguistic analysis of spoken and written language. *Cent. States Speech J. 18*, 81-85.

Devito, J.A. (1966). Psychogrammatical factors in oral and written discourse by skilled communicators. *Speech monograph, 33*: 73-76.

Devito, J.A. (1967). Levels of abstraction in spoken and written language. *Journal of Communication, 17*, 354-61.

DeVries, T.D. (1970). Writing writing and talking writing. *Elementary English, 47*, 1067-1071.

Diller, K.C. (1978). *The language teaching controversy*. Rowley, MA: Newbury House.

Donovan, O.J. (1986). A delightful couple: writing and speaking together. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 266 512)

Edelsky, C. (1986). *Writing in a bilingual program: Habia una vez*. Norwood, NJ: Ablex.

Eggins, S. (1994). *An introduction to systemic functional linguistics*. London: Printer Publishers.

Emig, J. (1977). Writing as a mode of learning. *College Composition and Communication, 28,* 122-128.

Fisher, C.J.& Terry, C.A. (1977). *Children's language and the language arts.* New York: McGraw-Hill.

Gaies, S.J. (1976). *Sentence-combining: A technique for assessing proficiency in a second language.* Paper read at the conference on perspectives on language, University of Louisville, Louisville, Kentucky, May 6-8. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 130 512).

Gaies, S.J. (1980). T-unit analysis in second language research: Applications, problems and limitations. *TESOL Quarterly, 14,* 53-60.

Gibson, J.W., Gruner, C.R., Kibler, R.J., & Kelly, F.J. (1966). A quantitative examination of differences and similarities in written and spoken messages. *Speech Monograph, 33,* 444-451.

Gipps,G & Ewen, E. (1974). Scoring written work in English as a second language: The use of the T-unit. *Educational Research, 16* (2), 121-25.

Gleason, H.A., Jr. (1965). *Linguistics and English grammar.* New York: Holt.

Gibson, J.W. Gruner., C.R.Kiber.R.J., & Kelly, F. (1966). A quantitative examination of differences and similarities in written and spoken messages. *Speech Monographs, 33,* 444-51.

Greenfield, P. (1972). Oral and written language: The consequences for cognitive development in Africa, the United States and England. *Language and Speech, 15,* 169-178.

Groff, P. (1979).The effects of talking on writing. *English in Education, 13*, 33-37.

Halleck, G.B. (1995). Assessing oral proficiency: A comparison of holistic and objective measures. *The Modern Language Journal, 79* (2), 223-234.

Halleck, G.B. & Young, R.F. (1995). *Video oral communication instrument (VOCI) for ESL/EFL.* Language Acquisition Research Center, San Diego, CA.

Halliday, M.A.K. (1973). *Explorations in the functions of language.* London: Edward Arnold, 1973.

Halliday, M.A.K. (1979). Differences between spoken and written language: Some implications for literacy teaching. In G.Page, J. Elkins, & B.O'Connor (Eds.), *Communication through reading proceedings of the fourth Australian reading conference* (pp. 37-52). Adelaide, SA: Aust. Read. Assoc.

Halliday, M.A.K. (1985). *Spoken and written language.* Deakin University Press.

Harpin, W. (1976). *The Second "R": Writing development in the junior school.* London: George Allen & Unwin.

Harrell, Lester E., Jr. (1957). A comparison of oral and written language in school-age children. *Monographs of the Society for Research in Child Development, 22,* no 3, serial no. 66. Lafayette, Ind.: Child Development Pubs.

Harrisn, S. (1998). Two new readability predictors for the professional writer: Pilot trials. *Journal of Research in Reading, 21*(2), 121-129.

Hasan, A.S. (1988). *Variation in spoken discourse in and beyond the English foreign language classroom.* PhD thesis. Aston University.

Haynes, L.A. (1992). The development of speaking/writing variability in narratives of non-native English speakers. *Issues in Applied Linguistic, 3* (1), 43-67.

Hidi, S., & Hildyard, A. (1983). The comparison of oral and written productions in two discourse types. *Discourse Processes, 6*, 91-105.

Hildyard, A., & Hidi, S. (1985). Oral-written differences in the production and recall of narratives. In D. R. Olson, N. Torrance. , & A. Hildyard.(Eds.), *Literacy, language, and learning: The nature and consequences of reading and writing* (pp. 285-306). London: Cambridge University Press.

Hudelson, S. (1984). Kan Yu Ret an Rayt en Ingles: Children become literate in English as a second language. *TESOL Quarterly, 18* (2), 221-238.

Hunt, K.W. (1965a). *Grammatical structures written at three grades levels.* Research Report No 3. Urbana, Illinois: NCTE.

Hunt, K.W. (1965b). Sentence structures used by superior students in grades four and twelve and superior adults. *Cooperative Research Projects, No.5-0312.* Tallahassee, Florida: Florida State University. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 010 047).

Hunt, K.W. (1970). Syntactic maturity in school children and adults. *Monographs of the society for research in child development,* serial No. 134, 35/1. Chicago, Illinois: University of Chicago.

Hyltenstam, K.(1988). Lexical characteristics of near-native second-language learners of Swedish. *Journal of Multilingual and Multicultural Development 9,1&2,* 67-84.

Kameen, P. (1983). Syntactic skill and ESL writing quality. In A. Freedman, I. Pringle. , & J.Yalden.(Eds.), *Learning to write: First language/second language* (pp. 139-158). London: Longman.

Kantor, K.J. & Rubin, D.L. (1981). Between speaking and writing: Processes of

differentiation. In B.M.Kroll., & R.J.Vann (Eds.), *Exploring speaking-writing*

relationships: Connections and contrasts (pp.55-81). Illinois: National Council of

Teachers of English.

Kerek, A. (1979). Measuring and evaluating syntactic maturity: some questions of

methodology. Retrieved January 16, 2004 from: ERIC database (ERIC Document

Reproduction Service No. ED 173 783).

Kim, A. (1996), A comparison of oral and writing development in a second language

college student. Retrieved March 13, 2003 from ERIC database (ERIC Document

Reproduction Service No. ED169 790.

Kim, A. (1998). A case study of error analysis/syntactic maturity as indicators of second

language writing development. *Research and Teaching in Developmental*

Education, 15 (1), 79-88.

Kim, A. (2000). A comparison of oral & writing development in a second language

college student. *Research and Teaching in Developmental Education, 16* (2), 73-

82.

Krashen, S.D. (1979). *The input hypothesis.* Paper presented at the Georgetown

University Roundtable in languages and linguistics, Washington, D.C.

Krashen, S.D., & Terrell, T. (1983). *The natural approach: Language acquisition in the*

classroom. Oxford: Pergamon Press.

Kreeft, J., Shuy, R., Staton, J., Reed, L., & Morroy, R.(1984). Dialog writing: Analysis of

student-teacher interactive writing in learning of English as second language.

Washington DC: Center for Applied Linguistics. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED252 097)

Kroll, B. (1981). Developmental relationships between speaking and writing. In B.M. Kroll., & R. J. Vann (Eds.), *Exploring speaking-writing relationships: Connections and contrasts* (pp.32-54). Illinois: National Council of Teachers of English.

Kroll, B.M., & Lempers, J. D. (1981). Effect of mode of communication on the informational adequacy of children's explanations. *Journal of Genetic Psychology, 138*, 27-35.

Kyle, P. (1980). Using objective methods of attained writing proficiency to discriminate among holistic evaluations. *TESOL Quarterly, 14* (1), 61-69.

Larsen-Freeman, D., & Strom, V. (1977). The construction of a second language acquisition index of development. *Language Learning, 27*,124-134.

Larsen-Freeman, D. (1978). An ESL index of development. *TESOL Quarterly, 12* (4), 439-448.

Lauren, U. (2002). Some lexical features of immersion pupils' oral and written narration. *Working Papers, 50*, 63-78.

Liggett, S. (1984). The relationship between speaking and writing: An annotated bibliography. *College Composition and Communication, 35*, 334-44.

Linnarud, M. (1986). *Lexis in composition: A performance analysis of Swedish learners' written English*. CWL Gleerup.

Loban,W. (1976). Language development: Kindergarten through grade twelve. *Research report* No. 18. Champaign-Urbana, Ill.: National Council of Teachers of English.

Lundsteen, S.W. (1976). *Children learn to communicate: Language arts through creative problem solving.* Englewood Cliffs, N.J.: Prentice-Hall.

Lyons, J. (1968). *Introduction to theoretical linguistics.* Cambridge: Cambridge University Press.

Mangelsdorf, K. (1989). Parallels between speaking and writing in second language acquisition. In D.Johnson., & D.Roen (Eds.), *Richness in writing: Empowering ESL students* (pp.134-145). New York: Longman.

Mann, M.B. (1944). The quantitative differentiation of samples of written language. *Psychol. Monogr, 56,* 41-74.

Marcus, M. (1977). *Diagnostic teaching of the language arts.* New York: John Wiley.

McCrackin, K. (1997). Challenging questions result in higher proficiency in oral testing. Paper presented at the 16th Annual OKTESOL Conference, Checotah, Oklahoma.

McCrackin, K. (1998). *The effects of learning strategies and motivation on the oral proficiency of ESL students.* Unpublished thesis. Oklahoma State University.

McKay, S. (1979). Towards an integrated syllabus. In K. Croft (Ed.), *Readings on English as a second language* (pp.72-84). Boston: Little, Brown.

Mellon, J. (1969). *Transformational sentence-combining: A method for enhancing the development of syntactic fluency in English composition.* Urbana, IL: National Council of Teachers of English.

Mendelsohn, D.J. (1983). The case of considering syntactic maturity in ESL and EFL. *IRAL, XXI* (4), 299-311.

Moffett, J. (1968). *Teaching the universe of discourse.* Boston: Houghton Mifflin.

Monroe, J. (1975). Measuring and enhancing syntactic fluency in French. *The French Review, 48*, 1023-1031.

Newman, J.G.& Milton, W. H. (1965a). Writing and speaking. *College Composition and Communication, 16*, 160-164.

Ney, J.W. (1966). Review of "Grammatical structures written at three grade levels" (Hunt, K.W.). *Language Learning, 16*, 230-235.

Nystrand, M. (1982). An analysis of errors in written communication. In M.Nystrand (Ed.), *What a writer knows: The language, process and structure of written discourse* (pp. 57-74). New York : Academic Press.

O'Donnell,R.C., Griffin, W.J., & Norris, R.C. (1967). A transformational analysis of oral and written grammatical structures in the language of children in grades three, five, and seven. *Journal of Educational Research, 61*, 36-39.

O'Donnell, R.C. (1974). Syntactic differences between speech and writing. *American Speech, 49*, 102-110.

O'Hare, F. (1973). Sentence combining: Improving student writing about formal grammar instruction. *NCTE Research Report* No.150. Urbana, IL: National Council of Teachers of English.

O'Loughlin, K. (2001). Studies in language testing 13: The equivalence of direct and semi-direct speaking tests. Cambridge: Cambridge University Press.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24* (4), 492-581.

Petty, W.T., Petty, D.C., & Becking, M.F. (1976). *Experiences in language: Tools and techniques of language arts methods* (2nd ed.). Boston: Allyn and Bacon.

Polio, C.G. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning, 47*(1), 101-143.

Poole, M.E., Field, T.W. (1976). A comparison of oral and written code elaboration. *Lang. Speech, 19*, 305-11.

Price, G.B., & Graves, R.L. (1980). Sex differences in syntax and usage in oral and written language. *Research in the Teaching of English, 4* (2), 47-153.

Purcell-Gates,V. (1989). What oral/written language differences can tell us about beginning instruction? *The Reading Teacher, January*: 290-294.

Raimes, A. (1983). Anguish as a second language: Remedies for composition teachers. In A. Freedman, I. Pringle, & J. Yalden, J. (Eds.), *Learning to write: First language/second language.* London: Longman.

Rubin, D.L. (1975). *Teaching elementary language arts*. New York: Holt, Rinehart & Winston.

Sardinha, A.P.B. (1996). A window on lexical density in speech. Paper presented at 8th Euro-International Systemic Functional Workshop, Nottingham, UK, July 24-27. Retrieved January 12, 2004 from http://www.liv.ac.uk/~tony1/homepage.html.

Schafer, J.C. (1981). The linguistic analysis of spoken and written texts. In B.M. Kroll & R.J.Vann (Eds.), *Exploring speaking-writing relationships: Connections and contrasts* (pp.1-31). Illinois: National Council of Teachers of English.

Scott, M., & Tucker, G.R. (1974). Error analysis and English-language strategies of Arab students. *Language Learning, 24*, 69-97.

Seda, I., & Abramson, S. (1990). English writing development of young, linguistically different learners. *Early Childhood Quarterly, 5*, 379-391.

Sharma, A. (1979). Syntactic maturity: Assessing writing proficiency in a second language. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 185 105).

Shaughnessy, M.P. (1977). *Errors and expectations: A guide for the teacher of basic writing.* New York: Oxford University Press, 1977.

Shohamy, E. (1994). The validity of direct versus semi-direct oral tests. *Language Testing, 11* (2), 99-123.

Shuy, R. (1988). The oral language basis for dialogue journals. In J. Station et al. (Eds.) *Dialogue journal communication: Classroom, linguistic, social and cognitive views.* Norwood, NJ: Ablex.

Silber, P. (1979). Teaching written English as Second language. *College Composition and Communication, 30* (3), 296-300.

Smith, E.B., Goodman, K.S., & Meredith, R. (1976). *Language and thinking in school* (2nd ed.). New York: Holt, Rinehart & Winston.

Snipes, W.C. (1973). Oral composing as an approach to writing. *College Composition and Communication, 24* (2), 200-205.

Stewart, N.F. (1978). Syntactic maturity from high school to university: A first look. *Research in the Teaching of English, 12* (1), 37-46.

Stubbs, M. (1986). Lexical density: A computational technique and some findings. In M.Coulthard (Ed.). *Talking about text* (pp.27-48). Birmingham, University of Birmingham: English Language research.

Tannen, D. (1982). *Spoken and written language: Exploring orality and literacy* Norwook, NJ: Ablex.

Taylor, J. (1973). *Reading and writing in the first school*. London: George Allen & Unwin.

Thornhill, D. (1970). *A quantitative analysis of the development of syntactic fluency of four young adult Spanish speakers learning English*. Ph.D dissertation, Florida State University.

Tomlinson, B., & Straehley, M. (1978). Developing and measuring mature syntax. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 158 244)

Tough, J. (1977b). *Focus on meaning: Talking and learning: A guide for fostering communication skills in nursery and infant schools*. London: Ward Lock Educational.

Ure, J. (1971). Lexical density and register differentiation. In G.E.Perren and J.L.M.Trim (Eds.).*Applications of linguistics* (pp.443-452). Cambridge: Cambridge University Press.

Vachek, J. (1973). *Written language: General problems and problems of English*. The Hague: Mouton.

Valdes, G., Haro, P.& Echevarriarza, M. (1992). *The Development of writing abilities in a foreign language: Contributions toward a general theory of L2 writing.* Berkeley: University of California, Center for the Study of writing and literacy.

Vann, R.J. (1979). Oral and written language relationships: The question of universals. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 191 051).

Vann, R.J. (1981). Bridging the gap between oral and written communication in EFL.In B.M. Kroll & R.J.Vann (Eds.), *Exploring speaking-writing relationships: Connections and contrasts* (pp.154-167). Illinois: National Council of Teachers of English.

Vygotsky, L.S. (1962). *Thought and language.* E.Hanfmann & G. Vakar (Eds and trans.). Cambridge, Mass.: M.I.T.Press, 1962. (Originally published, 1934).

Weissberg, B. (1997). On the interface of writing and speech: Acquiring English syntax through dialogue journal writing. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 412 558).

Witte, S. (1983). The reliability of mean T-unit length: some questions for research in written composition. In A. Freedman, I. Pringle, & J.Yalden (Eds.) *Learning to Write: First language/second language* (pp. 172-177). London: Longman.

Witte, S.P. ,& Davis A.S. (1983). The reliability of mean T-unit length: Some questions for research in written composition. In A.Freedman, I.Pringle. , & J.Yalden (Eds.). *Learning to write: First language/second language* (pp.171-177). London: Longman.

Witte, S. & Sodowsky, R.E. (1978). Syntactic maturity in the writing of college freshman. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 163 460).

Wolfe-Quintero, K., S. Inagaki, and H.-Y. Kim. (1998). Second language development in writing: Measures of fluency, accuracy, and complexity. *Technical Report* No. 17. Honolulu, HI: University of Hawaii, Second Language Teaching and Curriculum Center.

Woolbert, C.H. (1922). Speaking and writing-a study of differences. *Quarterly Journal of Speech Education, 8,* 271-85.

Yau, M.S. (1983). Syntactic development in the writing of ESL students. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 234 618).

Zoellner, R. (1969). Talk-write: A behavioral pedagogy for composition. *College English, 30,* 267-320.

Zora, S., & Johns-Lewis, C. (1989). Lexical density in interview and conversation. Retrieved January 16, 2004 from: ERIC database (ERIC Document Reproduction Service No. ED 322 776).

APPENDICES

APPENDIX A

VOCI QUESTIONS

VOCI QUESTIONS

Q1: Hello, my name is Gene and this is Ron. What's your name?

Q2: I am from New York and Ron is from Wisconsin. Where are you from?

Q3: This is a picture of my hometown. Tell us about your hometown?

Q4: Instead of writing letters, you have decided to send a cassette message to a friend back home. Describe where you are living now and what you've been doing recently.

Q5: I'm so happy my best friend just got back from vacation. I really missed him a lot. My best friend moved away and she's impossible to replace because she's so special. Describe one of your friends.

Q6: Because of a last minute problem you missed a dinner engagement with a friend. You called to apologize, but your friend is not yet home, so you need to leave a message on the answering machine apologizing for the date and explaining why you were not there.

Q7: Did you know that I went to New York last month? It sure is an interesting city. What's so special about it? The entire time I was there I tried to compare it with our city. There are lots of differences, but on the other hand, lots of things are similar. Can you compare your hometown with a city that you visited or you know well?

Q8: One thing that I didn't like about New York was that it is so big. I never really feel comfortable in big cities anymore. Really, I love city life. There's nothing more fascinating that a really big city. Not me. There are too many problems I guess. What do you think? What are the advantages or disadvantages of big city life?

Q9: Yes, that's just really unbelievable. It was a really terrific experience. There are some experiences you just can't forget. That's true. Have you ever had such an experience? An experience that you'll never forget. It can be something positive or it can be something negative. Tell us about it.

Q10: So, you finally made up your mind? Yes, and I'm really excited about it. Then you must have pretty concrete plans for the next few years? I have a good idea about what my life might be like. And you, what are your plans? What do you need to reach your goals? How might your life look ten years from now?

Q11: You have a summer job selling great books. I'm a potential customer. Convince me why I should buy the books from you.

Q12: Gene, did you read about the student who took one of these Swiss army knives to school with him in his pocket? No, what happened? Well, when he was using the scissors part of it, his teacher caught him and she took the knife away from him and they expelled

him from school. I don't get it. It looks like an innocent tool to me. Well, their school has a zero tolerance policy and they considered a Swiss army knife as a weapon. If you were the principal of this school, what would you do about this issue?

Q13: Wow, look at the headlines, another war. There have always been wars. It's nothing new. It's just human nature. Not necessarily. How do you feel about this issue? How do you think we could create a lasting peace?

Q14: I really love this painting. I don't understand it at all. Tell us why you think this is or isn't art.

Q15: My computer is broken again. Man, what a disaster. I feel so dependent on this machine. Yeah, modern technology can make life easy, but sometimes it can cause a lot of frustrations too. Discuss the positive benefits and the negative consequences of our dependence on such machines.

Q16: Some undergraduates at American universities think that native speakers of English make the most effective teachers. On the other hand, some people think the advantages of having an international teacher outweigh the disadvantages. What do you think?

Q17: If you were a teacher and you discovered one of your students had cheated on a test by copying from another student's paper, what would you do?

Q18: In many countries, higher education is for an elite group of students. Not everybody can go to the university. That certainly isn't the case in this country. Our universities are open to almost everyone regardless of their background. I can see the pros and cons of both types of educational systems. Discuss the advantages and disadvantages of both types of educational systems.

Q19: You know, I'm reading an article here on free trade in Europe and in America and it says that everybody benefits from having a free trade. No, I don't know. There's still an awful lot of opposition in a few countries to the whole issue of free trade. Take one position and defend your opinion regarding the issue of free trade.

Q20: Did you know that US laws allow trials to be televised? Yes, several high profile trials have been televised recently because of the freedom of information act. I wonder if that's such a good idea? What do you think about televising criminal trials?

Q21: Have you noticed how many shows on TV portray violent crimes? Pretty hard not to notice. Some people feel that this creates violence in our society. Yeah, but other people feel it doesn't have any effect at all on our young people. In fact, they're proud of this country's freedom of expression. What do you think about the portrayal of violence and crime on TV?

Q22: Three must be problems in your country too. What are some of the problems in your country? Suggest some solutions and discuss the implications of these solutions.

Q23: This is the last question. If you've gotten this far, you've probably taken other English tests. If so, how does this test compare to other English tests you have taken.

APPENDIX B

GUIDELINE FOR T-UNITS, CLAUSES, WORD COUNTS, AND ERRORS

GUIDELINE FOR T-UNITS, CLAUSES, WORD COUNTS, AND ERRORS

T-Units

a. A T-unit is defined an independent clause and all its dependent clauses.

b. Count run-on sentences and comma splices as two T-units with an error in the first T-unit.

Ex: My school was in Saudi Arabia, it was the best school there.

T	/	T
1 error		error-free

c. For sentence fragments, if the verb or copula is missing, count the sentence as 1 T-unit with an error. If an NP is standing alone, attach it to the preceding or following T-unit as appropriate and count as an error. If a subordinate clause is standing alone, attach it to the preceding or following S and count as a 1 T-unit with an error.

d. When there is a grammatical subject deletion in a coordinate clause, count the entire sentence as 1 T-unit.

Ex. First we went to our school and then went out with our friends.

e. Count both "so" and "but" as coordinating conjunctions. Count "so that" as a subordinating conjunction unless "so" is obviously meant.

f. Do not count tag-questions as separate T-units.

g. Count S-nodes with a deleted complementizer as a subordinate clause as in : I believe that A and (that) B=1 T-unit.

h. But, direct quotes should be counted as:

   John said, "A and B."

1 T-unit   1 T-unit

i. Assess the following type of structures on a case-by-case basis:

    If A, then B and C.

    As a result, A or B.

j. Count T-units in parentheses as individual T-units.

Clauses

a. A clause equals an overt subject and a finite verb. The following are only one clause each:

    He left the house and drove away.

    He wanted John to leave the house.

b. Only an imperative does not require a subject to be considered a clause.

c. In a sentence that has a subject with only an auxiliary verb, do not count that subject and verb as a separate clause (or as a separate T-unit). (e.g. John likes to ski and Mary does too; John likes to ski, doesn't he?; John is happy and Mary is too.)

Error Guidelines

a. Do not count spelling errors (including word changes like "there/their").

b. Be conservative about counting comma errors; don't count missing commas between clauses or after prepositional phrases. Comma errors related to restrictive/non-restrictive relative clauses should be counted. Extraneous commas should also be considered errors.

c. Base tense/reference errors on preceding discourse; do not look at the sentence in isolation.

d. Don't count British usages as errors, (e.g. "in hospital," "at university," collective nouns as plural)

e. Be lenient about article errors from translations of proper nouns.

f. Don't count errors in capitalization.

g. Count errors that could be made by native speakers (e.g. between you and I).

h. Do not count register errors related to lexical choices (e.g. lots, kids).

i. Disregard an unfinished sentence at the end of the essay.

Word Count

a. Count contractions as one word whether correct or not.

b. Count numbers as one word.

c. Count proper nouns in English and in other languages as they are written.

d. Do not count hyphenated words as single words. (e.g. well-written =2 words)

e. Do not include essay title in word count.

f. Count words as they are written, even if they are incorrect.

From Polio (1997)

APPENDIX C-IRB REVIEW FORM

Oklahoma State University
Institutional Review Board

Protocol Expires: 11/18/2003

Date: Tuesday, November 19, 2002   IRB application No AS0332

Proposal Title: THE RELATIONSHIP BETWEEN ESL STUDENTS' ORAL ENGLISH PROFICIENCY AND WRITING ENGLISH PROFICIENCY

Principal
Investigator(s):

Gene Halleck   Xinhua Zhu
311B Morrill   721 N. Hester
Stillwater, OK 74078   Stillwater, OK 74075

Reviewed and
Proceeded as: Exempt

Approval Status Recommended by Reviewer(s): Approved

Dear PI:

Your IRB application referenced above has been approved for one calendar year. Please make note of the expiration date indicated above. It is the judgment of the reviewers that the rights and welfare of individuals who may be aside to participate in this study will be respected, and that the research will be conducted in a manner consistent with the IRB requirements as outlined in section 45 CFR 46.

As principal investigator, it is your responsibility to do the following:
1. Conduct this study exactly as it has been approved. Any modifications to the research protocol must be submitted with the appropriate signatures for IRB approval.
2. Submit a request for continuation if the study extends beyond the approval period of one calendar year. This continuation must receive IRB review and approval before the research can continue.
3. Report any adverse events to the IFB Chair promptly. Adverse evens are those which are unanticipated and impact the subjects during the course of this research; and
4. Notify the IFB office in writing when your research project is complete.

Please note that approved projects are subject to monitoring by the IRB. If you have questions about the IFB procedures or need any assistance from the Board, please contact Sharon Bacher, the Executive Secretary to the IRB, in 415 Whitehurst (phone: 405-744-5700, sbacher@okstaet.edu).

Sinerely

Carol Olson, Chair
Institutional Review Board

VITA

Xinhua Zhu

Candidate for the Degree of Master of Arts

Thesis: THE RELATIONSHIP BETWEEN SPEAKING AND WRITING IN
              COLLEGE-LEVEL ESL STUDENTS

Major Field: English

Biographical:

Personal Data: Born in Zhongwei, Ninngxia, China, on February 9, 1977.

Education: Received Bachelor of Arts in English Education from Ningxia University,
    Yingchuan, Ningxia, China in July, 1998; Received Masters of Arts in English
    Language and Culture from China University of Mining and Technology,
    Beijing, China in July, 2001. Completed the requirements for the Masters of
    Arts degree with a major in English at Oklahoma State University in May 2004.

Experience: Substitute Teacher, Ningxia University, Ningxia, China, February 1998
    to July 1998; English Instructor, China University of Geosciences, Beijing,
    China, September 2000 to January 2001; English Instructor, China University
    of Mining and Technology, Beijing, China, September 2000 to July 2001; A
    Faculty Member, Beijing University of Aeronautics and Astronautics, Beijing,
    China, September 2001 to December 2001; Tutor, Oklahoma State University,
    Writing Center, January 2002 to present.

Awards: Member of the Honor Society of Phi Kappa Phi, Oklahoma State University