

# A REVIEW OF THEORIES AND RESEARCH INTO SECOND LANGUAGE WRITING ASSESSMENT CRITERIA

Duong Thu Mai\*

*VNU University of Languages and International Studies,*

*Pham Van Dong, Cau Giay, Hanoi, Vietnam*

Received 19 April 2019

Revised 20 May 2019; Accepted 3 June 2019

**Abstract:** As language assessment in Vietnam is being intensively attended to by the Ministry of Education and Training and is actually critically transformed, criterion-referenced assessment has gradually been a familiar term for language teachers, assessors and administrators. Although the name of the approach has been extensively used, most teachers of English at all levels of language education still face the challenge of identifying “criteria” for writing assessment scales. This paper attempts to provide a reference for teachers and researchers in second language writing concerning on the major development in the field in defining this construct of “writing competence”. The paper focuses more on the existing and published literature globally on English writing teaching approaches, research and practices. These contents are reviewed and summarized into two major strands: the product-oriented considerations and the process-oriented considerations.

*Keywords:* writing assessment, writing teaching approaches, criteria, product-oriented writing assessment, process-oriented writing assessment

## 1. Introduction

For over a hundred years, writing assessment has been considered a significant field, with the increasing participation of researchers and practitioners from many other fields. They contribute voices to sharpen the traditional paradigms and introduce new paradigms of writing assessment. They introduce new theoretical and practical models for writing assessment, both of which hold critical values for the teachers in service. This paper is going to summarize the major findings in this dynamic field to inform the assessment practices of writing teachers in Vietnam in the contexts of substantive assessment reforms in

the orientation of standard-based, competence-based and criterion-based assessment.

## 2. A brief history of writing assessment

Each period in writing assessment history has been dominated by particular assumptions about assessment methods, technical quality and writing competence. Looking through the lens of assessment methods, Yancey (1999) identifies three overlapping paradigms of writing assessment namely *objective testing*, *holistic scoring*, and *portfolio/ performance assessment*. The first era of writing assessment was named objective testing paradigm, in which parametric tests were the reigning educational assessment tool, and the word “writing examination” meant answering selected-response questions

---

\* Tel.: 84-369686968

Email: duongthumai@yahoo.com

in either standardized or locally developed tests (Ruth & Murphy, 1988). Reliability was then supposed to suffice for validity. In short, in this period of writing assessment, testing was separated from classroom activities (Huot, 2002) and had no power (Yancey, 1999). In the second paradigm of writing assessment, direct writing assessment and criterion-referenced test interpretation were the most widely discussed issues. Writing assessment was argued to be more direct than multiple-choice tests, that writing skills could only be assessed with real writing products and that students' mistakes in writing should be investigated to inform followed-up instruction. The development of the holistic method for essay scoring by the educational measurement scholars also emerged, leading to improvements in rater consistency. In the third paradigm, assessing writing means discovering and assessing those processes of self expressing; writing should be assessed through many samples of writing produced at different time and under no pressure, in such forms as projects, portfolios, etc.

It can be summarized that the current context of writing assessment is when the popularity of cognitive learning theory, the attention to learners' and teachers' roles in the classrooms and the development of appropriate assessment methods provide the exact aids for the writing assessment communities to achieve better validity in assessment. On balance, the co-existence of the new and old paradigm in writing assessment can be advantageous, since an application of different methods is bound to bring about the most accurate results in assessment. That there is no single best way to do assessment has become a verity after many ups and downs in assessment history (Brown, 1998). However, the existence of multiple paradigms requires from the assessment instrument

developers the more critical consideration of relevant theories and practices before making hypotheses of their constructs. The following discussions on product-oriented written language production and process approach in writing reflect essential theoretical concerns in defining writing competence as a product and as a process.

### 3. Writing as a product and as a process

#### 3.1. Writing product considerations

In the emergence of the third paradigm in writing assessment, so many different definitions of writing competence have been developed that one author's definition is not general enough for others (Camp, 1993a; White, 1995). One well-structured model of textual construction was proposed by Grabe and Kaplan (1996) based on their review of written language nature, writing studies and popular hypotheses on textual features. Writing ability in this model has seven interacting areas of knowledge:

- syntactic structures;
- semantic senses and mapping;
- cohesion signalling;
- genre and organisational structuring to support coherence interpretations;
- lexical forms and relations;
- stylistic and register dimensions of text structure;
- non-linguistic knowledge bases, including world-knowledge.

Within each of these interacting components are series of other sub-components which also interact with each other. The authors then group these sub-components into four more explicit parts:

*elements of text structure, a theory of coherence, a functional-use dimension of text, and the non-linguistic resources.* Elements of text structure include grammatical features (at sentential level in the forms of semantics and syntax) and some functional features (at both sentential and inter-sentential level in the form of coherence and cohesion). Coherence has a special position in this model, as the authors consider it not only as a textual feature, but also from the reader's perspective, i.e. whether a text is coherent depends not only on the writer's use of cohesive devices but also on the reader's interpretative systems, including their knowledge and their opinions of the relevance of ideas. The balance between textual features and top-down processing is the special point the authors of this model

want to propose in contrast to other authors' claims for the privilege of one of them (such as Halliday and Hasan's claim (1976) for cohesion). *Functional-use dimensions* relate to how the textual features are combined to make a text, such as the logical organization and stylistic features (shown at interpersonal level in the form of stances and postures), which address the appropriateness between texts and writers' goals, and the relation between the writers' attitudes to the readers, the subject, the context, world knowledge, etc. Some examples of stances are personal – interpersonal, distance – solidarity, superior – equal, oblique – confronted, formal – informal, etc.

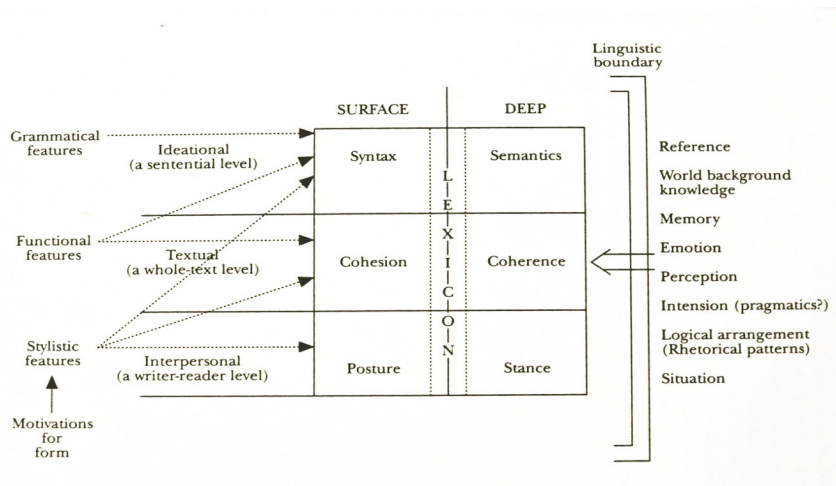


Figure 1: Model of Textual Production (Grabe & Kaplan, 1996)

According to Figure 1, grammatical, functional and stylistic features in a written text are affected by motivation for form and constrained by eight types of non-linguistic knowledge: reference, world knowledge, memory, emotion, perception, intention, situation, logical arrangement. The non-linguistic features may be revealed in the use of lexicon and have strong influence on all the three sets of linguistic features of texts.

Based on a large literature of writing studies and hypotheses, this model aims to clarify the properties of a written text for real use (Grabe & Kaplan, 1996). This purpose seems to have been successfully fulfilled because writing teachers and researchers can easily obtain necessary information on what should be assessed in a writing product, as well as on the linkages between those areas of knowledge. However, the model only provides

the foundation for writing researchers to make hypotheses on writing production knowledge.

### 3.2. *Writing process considerations*

The writing process models by Flower and Hayes (1981), Bereiter and Scardamalia (1987) and Grabe and Kaplan (1996) presented below offer general definitions of the writing process.

Flower and Hayes' model of writing process (1981) (Figure 2) is most frequently discussed in writing-process literature. The authors developed a cognitive process model which assumes that writing includes many distinctive, goal-directed and hierarchical cognitive processes. As seen in Figure 2, the most important element of the model is the rhetorical problem, such as a writing assignment at school, because if student writers cannot understand the problem, they cannot write anything to solve the problem. They need to identify the topic, the audience and their goals in writing. After this initial representation, they deal with constraints such as the amount of text produced, their own knowledge in their long-term memory, and their plans for writing. The process of actual writing starts with planning (the act of building the internal representation: generating ideas, organizing ideas and most importantly, goal setting). Flower and Hayes strongly emphasize goal setting as a continuous phase running through the writing process and as a crucial feature of a creative writer. After planning, the writers move into translating or putting their abstract representations into visible letters. This stage requires them to integrate understanding of all linguistic demands from functional to syntactic. Later, in revising, the writers evaluate what they have written and consider keeping or revising it, which may trigger another cycle of planning, translating

and reviewing. It is important to note that in this model, the three main stages of writing are no longer represented as a linear process. By this, and by stressing that writers differ in their composing strategies, Flower and Hayes have made great contributions to the process-oriented approach in the field of writing. In a later revision of the model (Hayes & Flower, 1987), their argument of the differences in writers' composing processes is even further clarified, when their study found that expert writers composed differently from novice writers in some aspects:

- they take more aspects of the rhetorical problem into consideration;
- they approach these aspects at greater depth;
- they respond to the problem with a fully developed image of what they want to write. They are therefore more creative in solving the problems OR answering the questions;
- they reassess their goals and revise them in the process of writing.

This model, as well as later works of the same authors on the writing process, is well recognized for making a complete representation of the writing process and delivering an influential message on its non-linearity (Weigle, 2002). However, they have paid inadequate attention to the writers' linguistic knowledge and the influence of external facets on students' writing processes (Kaplan & Grabe, 2002; Shaw & Weir, 2006; Weigle, 2002).

The differences between skilled and unskilled writers in terms of composing processes have been more specifically discussed in the work of Bereiter and Scardamalia (1987). The two different models of writing processes proposed by the two authors are appreciated for

coherently covering a wider range of research than previous models (Grabe & Kaplan, 1996). Also, the models confirm the existence of differences between skilled and unskilled writers and bring into clearer focus the problem-solving skills which are required in complicated writing tasks (Grabe & Kaplan, 1996).

The knowledge-telling model (Figure 3) used by less skilled writers is built on the assertion that these writers ignore the more complicated problem-solving strategies skilled writers use. They choose to solve the rhetorical problem through a context-free monologue with their internal knowledge. They only consider the rhetorical problem (topic and genre) in terms of what they know, then write what they know down, examine the text produced and use it to generate new texts. This process works well in writing about simple and familiar topics such as narratives of personal experience because the writers' familiarity with the topics helps them arrange ideas in their mind and hence improves the coherence of their writing.

The knowledge-transforming model (Figure 4) was developed to make up for the disadvantages of the knowledge-telling model in explaining for writers' behaviours in complicated rhetorical problems. These problems always call for higher level thinking skills than memory retrieval and often appear in academic writing assessment. In confronting the task, the writers analyse, set goals for writing and plan the solutions for both content and rhetorical problems. Then there is an interactive stage between content problem solving and rhetorical problem emerging, and vice versa. This stage lasts until both sets of problems seem to be resolved, and the writers then continue with the knowledge-telling model: retrieving solutions from their memory to tell and write. Interestingly, the

knowledge-transforming model also includes the knowledge-telling model because the skilled writers also may use the simpler model in some circumstances. For example, in coping with a task they have met before and the problems they have solved before, they only need to follow the steps in the knowledge-telling model. From this aspect, the complementarity of the two models is unarguable (Shaw & Weir, 2006).

In general, the skilled writers plan longer, produce more detailed pre-writing notes. They consider goals, plans and audience alongside content problems in writing. Their revision covers not only textual elements but also the organisation of the text. They also make use of main ideas as guides for planning and integrating information (Bereiter & Scardamalia, 1987).

Despite their advantages, Bereiter and Scardamalia's models fail to explain the influence of contexts in the writing process, as presented in Hayes and Flower (1987). The authors also did not describe the cognitive development underlying the transformation from knowledge-telling to knowledge-transforming, making it difficult to determine if a writer is in a middle proficiency level between skilled and less skilled (Grabe & Kaplan, 1996)

To suggest a solution for the problems of previous models, and to complement their model of textual production (Figure 1), Grabe and Kaplan (1996) developed a writing process model which considers both external contexts and writers' internal processing. In their model, the situation and performance output are integrated to form the external social context for the writing task. Internally, all the processes of writing happen within the writers' verbal working memory. Based on the contextual features, the writers set goals for writing and generate the first representation of the task which they think fit



well with the goals. This internal goal setting is metaphorically referred to as the “lens” to look at the writers’ products and processes. After goal setting, a circle of metacognitive and verbal processing of linguistic knowledge, world knowledge and online processing assembly (the monitoring of information generated from the other two kinds of knowledge) is triggered and functions in the interaction with the established goals. Only some parts of these components are used in creating the internal processing output, which is compared to the established goals and may be revised as necessary before becoming the textual output in the performance. Even then, this textual output can be compared once again to the goals and another circle of internal processing starts. Writing goals in this model are really important “rulers” for the writer to assess his production at any stage in the process, an idea similar to Hayes and Flower (1987). As regards the differences between writers at different levels of proficiency, besides those in the two previous models, Grabe and Kaplan add that skilled writers:

- review and reassess plans on a regular basis;

- come up with more types of solutions for rhetorical problems;
- plan more perspectives in writing;
- revise according to the goals rather than just language segments;
- own a variety of writing strategies in all stages of writing.

In a condensed comparison, Grabe and Kaplan (1996)’s model is clearer than Hayes and Flower’s (1981) and Bereiter and Scardamalia’s (1987) models in terms of the cognitive and metacognitive processes in writing.

### 3.3. Summary

The models of writing processes and writing products presented in this section have been constructed based on a large reservoir of research results and theories and are still being validated. In the current paradigm, an important point for writing researchers in the validation of models is the need to focus on both product and process writing knowledge. To paraphrase, promoting an enabling process-oriented approach does not imply that the product approach is disabling.

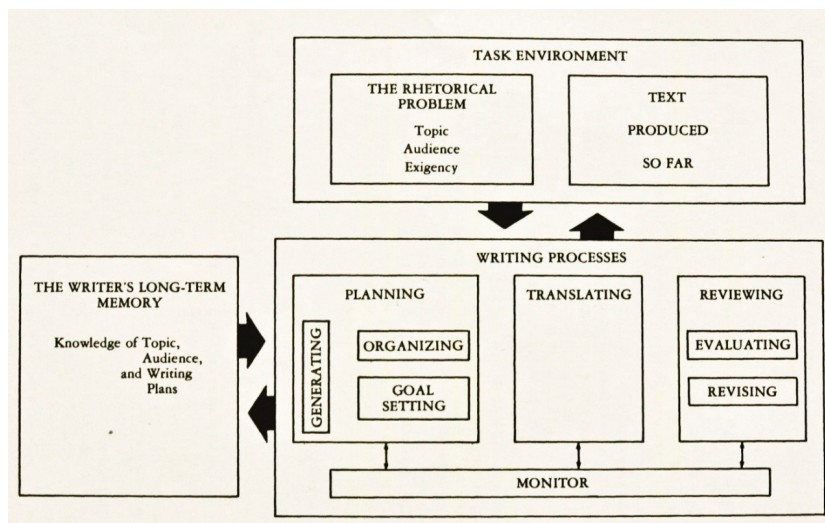


Figure 2: Process of Writing (Flower & Hayes, 1981)

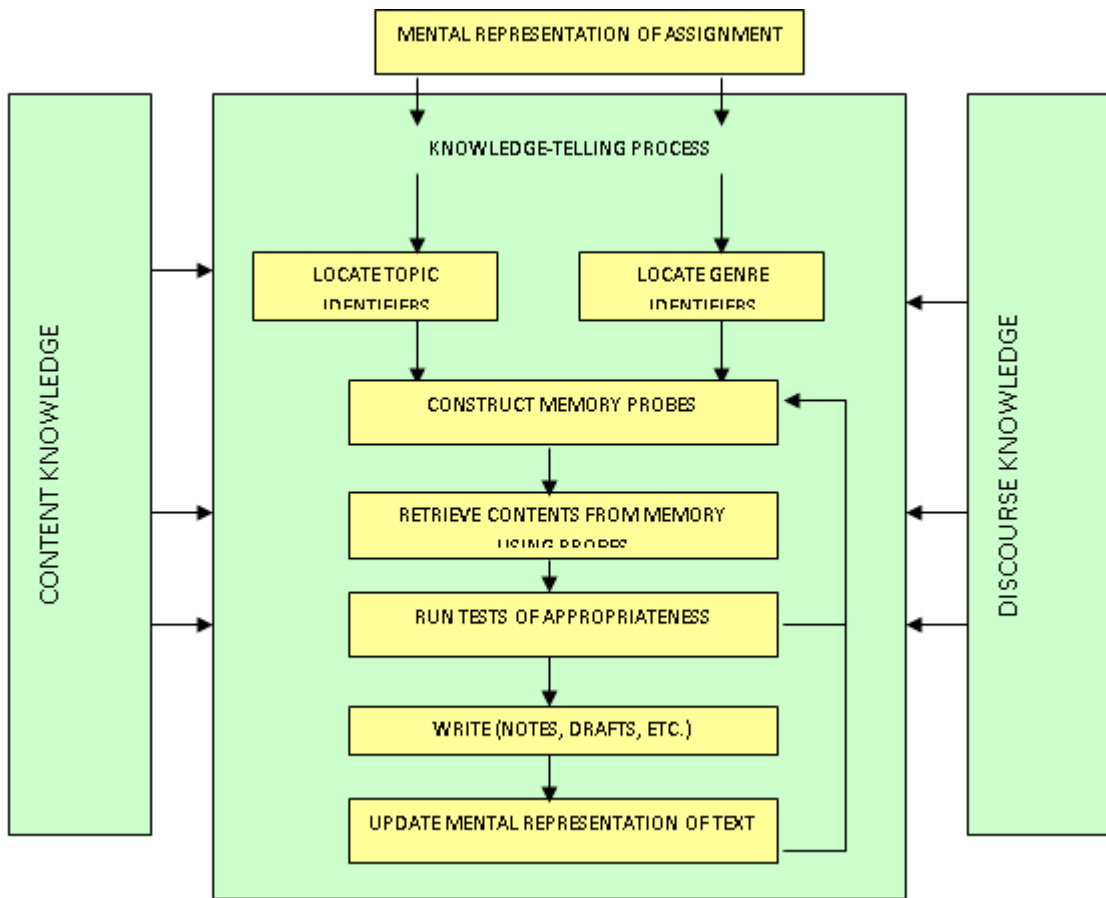


Figure 3: Knowledge-telling Model of the Writing Process (Bereiter & Scardamalia, 1987)

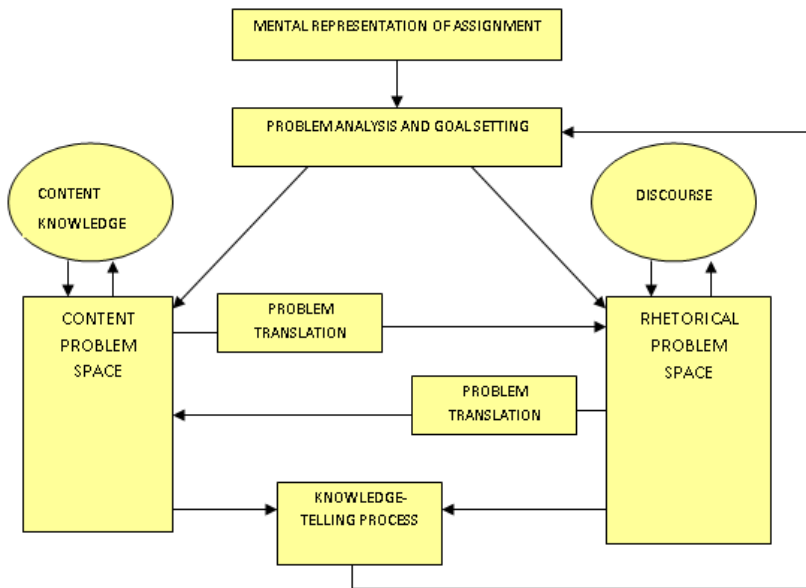


Figure 4. Knowledge-transforming Model of the Writing Process (Bereiter & Scardamalia, 1987)

#### 4. Research on L2 writing products and process

This section provides an overview of the extent to which available research and practices in second language (L2) writing assessment have validated the above mentioned theories of textual productions and writing processes. It is expected that the results of the research can illuminate the short list of criteria which should be employed for measuring students' writing performances.

The section is organized into two areas of writing knowledge which necessarily contribute to a theory of writing (Grabe & Kaplan, 1996): L2 writing production knowledge and L2 writing processes (including L2 writing strategies). Research results on the relation of each type of knowledge to L2 writing proficiency is presented first, followed by a description of currently used L2 writing assessment indicators.

##### 4.1. Research on L2 writing products

Defining what writing ability means is essential to defining the purposes of teaching and assessing writing. In general, the assessment instruments represent what their developers define as the construct they want to measure. This section reviews the research work on factors which affect L2 students' writing ability in three areas of language knowledge: text structure elements, textual knowledge, and sociolinguistic knowledge. These three areas represent current research work in L2 writing production knowledge.

##### 4.1.1 Research on grammatical knowledge/ elements of text structure

Despite the increasing popularity of composing process research, L2 studies on the elements of text structure are still dominant in the literature (Silva & Brice, 2004). Research

by Second Language Acquisition (SLA) researchers in this area often involves the analysis of three important textual features: accuracy, complexity, and fluency.

##### Accuracy

Linguistic units ranging from **T-units** (a grammatical construction with one independent clause (a simple sentence) and/or its related subordinate clauses (a complex sentence), **phrases, clauses, sentences, etc.** are the oldest criteria to judge students' writing accuracy. The majority of papers and studies on linguistic accuracy have been done by composition researchers (Haswell & Wyche-Smith, 1994). The most reliable textual prediction of writing quality was **clause length, T-unit length and the number of clauses per T-unit, the number of subordinate clauses** (Huot, 2002; Veal, 1974). However, studies about the relation between T-unit features and writing competence provided inconsistent results. The calculation of T-units and clause length therefore is not enough (Ruth & Murphy, 1988). The links between **verbal diversity, verb choice, grammatical complexity, freedom of errors and writing quality** were found to be strong (Greenberg, 1981; Grobe, 1981; Witte & Faigley, 1983). Composition researchers were also able to point out other syntactic features which discriminate students' writing quality significantly, including the **increased use of adjectives, nominal complexity, free modifiers, sentence adverbials, relative clauses, finite adverbial clauses, stylistic word-order variation, passives, complex noun phrase subjects, tenses and modes, and unmodified noun phrases** (Grabe & Kaplan, 1996).

The accuracy of the produced language has also been the focus of studies in SLA, L2 writing assessment and L2 writing instruction.



According to Polio (1997), measures can range from holistic scales, error-free units, error-count without classification to error-count with classification. Holistic scales address such indicators as vocabulary, spelling, punctuation, syntax or word forms, which are measured at semantically different levels. Error-free units, including **error-free T-units** and **error-free clauses**, are more objective. Error count without classification involves calculating the ratio of errors and error-free units. Error count with classification seems to be the most advantageous measure for solving the previous problem. Studies on accuracy measures could provide some indications for L2 writing assessment researchers. For example, Brown (2002) investigated the influence of sentence-level errors (sentence structures and grammar/mechanics) on untrained ESL raters' holistic ratings of ESL students by comparing their scores on the original and the corrected essays. The analysis showed a significant difference in the two sets of holistic scores, and a high correlation between the analytic scores for the two types of sentence-level features with the holistic scores. Sentence structures and grammar/mechanics are therefore thought to affect holistic scores. Kennedy and Thorp (2007) found that writers at lower band scores in the IELTS tests made more **lexico-grammatical errors**. Higher-scored IELTS scripts were also found to have fewer **mechanical errors** (Mayor et al., 2007). Vocabulary in L2 academic writing is another instance. The accurate retrieval of **sufficient and diverse vocabulary** is one of the first requirements for success in academic writing and its lack may lead to negative rater judgments. Moreover, **word form accuracy** and **word choice diversity** significantly affect L2 intermediate students' writing scores. Ferris (1994), for example, found that advanced ESL students demonstrated greater

use of some lexical categories (emphatics, hedges) and difficult syntactic construction (stative forms, participial construction, relative clauses, adverbial clauses, etc.). Among textual features, the lower proficiency writers use more lexical repetitions as cohesive devices, while the higher proficiency group chose lexical and referential cohesion devices (synonyms, antonyms, etc.). The more advanced students also use more passives, cleft sentences, and topicalizations. Vocabulary range (including idiomatic language) and grammatical accuracy are also proved to be successful predictors of IELTS band scores by (Banerjee, Franceschina, & Smith, 2007; Kennedy & Thorp, 2007).

#### Complexity

Besides accuracy, SLA and L2 writing researchers also examine the complexity of text structure elements. The importance of lexical and grammatical complexity/sophistication in deciding students' writing quality/scores has been emphasized in numerous studies (Hinkel, 2003; Mayor, Hewings, North, Swann, & Coffin, 2007; Reid, 1993; Vaughan, 1991). Grammatical complexity measures include **t-unit complexity ratio**, the **number of verb phrases per t-unit**, the **number of dependent clauses per t-unit**, etc. Lexical complexity measures include the **ratios between word types over the number of words**, **between sophisticated words and the total number of words**, or between **the number sophisticated words and the number of word types**. In these studies, sophisticated words are often identified by referring to a standard list, such as the Academic Word List (Coxhead, 2000). The best grammatical complexity measures are the **number of clauses per T-units** and the **number of dependent clauses per T-unit** (Wolfe-Quintero, et al., 1998). Other measures such as **passives**, **articles**, **relative clauses**,

**complex nominals** may also be important structures relative to developmental levels.

In investigating the effect of vocabulary on writing proficiency, Zareva, Schwanenflugel and Nikolova (2005) offered sophisticated insight into the role of different aspects of lexical knowledge. Macro-level features of lexical knowledge include three dimensions (quantity, quality, and metacognitive awareness) which can be further categorized into six variables (**vocabulary size, frequency effects, word associations, nativelike associations**, etc.). A vocabulary test was completed by 64 native and ESL students divided into three proficiency groups. The results show that vocabulary quantity and quality can predict native, L2 advanced, and L2 intermediate students' language competence levels. The most important strength of this study is the wide range of examined vocabulary features while the most obvious weakness is the take-home vocabulary test, which may lead to students' reliance on other reference materials.

Complexity measures have been subjected to a number of criticisms, such as its sensitivity to length. Longer texts are often considered more complicated than shorter ones according to these measures, which is not always correct. Moreover, similar to studies on accuracy, studies on grammatical and lexical complexity show inconsistent results (Knoch, 2007; Wolfe-Quintero, Inagaki & Kim, 1998). For example, Banerjee, Franceschina, and Smith (2007) could find no relation between **syntactic complexity** and IELTS band scores. Nevertheless, Laufer (1995) was able to detect a significant relationship over time with an analysis of word type/ sophisticated word ratio. In IELTS studies, Mayor et al. (2007) found that the **complexity of sentence**

**structures** is among the best indicators of writing band scores.

### Fluency

One important aspect of language production besides accuracy and complexity is fluency. Fluency has been defined variously but generally, it is the comfort of language production. English-as-a-second-language (ESL) students often pay attention to accuracy at the expense of fluency and meanings (Knoch, 2007).

Fluency is most frequently measured by the **number of words/structures** the students produce, the ratios of production units (Wolfe-Quintero, Inagaki & Kim, 1998) or the **number of reformulations and self-corrections** (Knoch, 2007). Specifically, measurement may include the number of clauses, sentences, T-units, or ratios of these linguistic units to the text (Chenoweth, 2001). The best indicators of proficiency are **the ratios of T-unit length and error-free T-unit length and clause length**, which linearly increase with proficiency levels across studies (Wolfe-Quintero, Inagaki & Kim, 1998). Among these measures, a serious criticism of the ratio measures of frequency is the failure to take the students' writing process into consideration (Chenoweth & Hayes, 2001). From these ratios, it is hard to imagine how the writers have managed to make their writing fluent. The relations between the number of revisions and writing experience and between the number of revisions and writing proficiency have also been proved positive (Chenoweth & Hayes, 2001). In other words, students' proficiency in writing could be revealed by their frequency of revisions.

**Length of writing** has been considered an important indicator of writing proficiency by Bereiter and Scardamalia (1980). An investigation into IELTS scripts scored from

4 to 8 has revealed that students with score 4 struggled to reach the word limit, while students at score 6 find the word limit feasible, and students with score 8 always exceed it (Kennedy & Thorp, 2007). In the same study, students with higher scores were found to have written longer paragraphs and sentences. Another study on IELTS students which reached similar conclusions on the importance of length is by Mayor et al. (2007), who found longer clauses in higher rated essays. However, opposite results have also been found. Woffle-Quintero (1998)'s analysis of 18 uncontrolled-time writing studies found that only 11 studies found a significant relationship between length and writing development.

#### *Summary of research results on text elements*

The studies show the influence of accuracy, fluency and complexity in discriminating L2 students' writing development and writing quality. These results also present a large resource to validate the L1 and L2 Communicative language competence (CLC) models which have already been established. They must therefore be considered by L2 writing assessment researchers when designing their instruments.

#### *4.1.2 Research into textual knowledge (cohesion, coherence)*

The common thread in composition and applied linguistic research into writing quality, which runs through from the holistic scoring paradigm to the modern assessment paradigm, is the shift from small syntactic units to global syntactic features, or textual knowledge. In this section, the research on these global foci is presented, and various methods of assessing them are discussed.

Firstly, cohesion refers to "the explicit linguistic devices used to convey information,

specifically the discrete lexical cues to signal relations between parts of discourse" (Reid, 1992, p. 81). In other words, cohesive devices include visible linguistic units which connect grammatical units and lexical units. The most popular classification of cohesive devices which L2 writing researchers use in their studies, also the one which most emphasizes the role of cohesion in language production, was given by Halliday and Hasan (1976), who regarded cohesion as the most important feature which defines a text as a text. It "refers to relations of meaning that exist within the text" (p. 4) but is expressed by lexicogrammatical devices, including reference, substitution, ellipsis, conjunction, lexical reiteration, and collocations.

Studies in L1 student essays have found that high-rated essays are generally more cohesive than low-rated ones, especially through the use of reference devices and conjunctions. Lexical cohesion is the most popular type of devices (Witte & Faigley, 1981) but lower-scored essays tend to use more **repetitions** (Witte & Faigley, 1983) and higher-scored essays tend to use more **collocations** and **synonyms** (Crowhurst, 1987). However, other L1 studies could not reach these results; for example, there was no correlation between cohesive density and quality (McCulley, 1983). The relation between cohesion and L1 writing quality is therefore inconclusive.

L2 writing studies show more consistent results. For example, Wenjun (1998) studied how six Chinese ESL students' writing quality was affected by their use of cohesive devices. They found that students of high proficiency do produce more cohesive texts than those with lower proficiency. The ESL students were able to improve their cohesiveness given successful rhetorical transfer and sufficient

exposure to authentic texts. Kennedy and Thorp (2007) reached more detailed results when comparing IELTS writers at different writing band scores: lower-scored students often used obvious transitional signals (such as *however, firstly, secondly*) while higher-scored ones had more indirect ways to link ideas and were able to avoid repetitions. Banerjee, Franceschina and Smith (2007) had similar findings on the use of cohesive devices of students at different band scores in the IELTS test. Thus, cohesion seems to be an essential indicator in the measurement of L2 students' writing ability.

Coherence is one of the most controversial issues in L2 writing assessment. It is considered an elusive, fuzzy concept, with a large number of definitions. Generally, Grabe and Kaplan (1996) stated that coherence occurs not only at the textual level through the creation of a semantic map for the ideas but also depends on the readers. Researchers choose different means to assess coherence according to the characteristics they wish to integrate into their definition (Todd, Thienpermpool, & Keyuravong, 2004). As early as in the 19<sup>th</sup> century, coherence was limited to the use of **sentence connections** and **paragraph structure**. Later measures went beyond sentence limits to texts. Coherence can be defined in terms of **cohesion** and **organization patterns** (Kintscha & Dijk, 1978) and **metadiscourse markers** (Crismore, Markannen, & Steffensen, 1993). As another example, **topical structural analysis**, which involves the analysis of the themes and rhemes of each sentence and how they are connected in the structure of sentences, has been proposed (Lautamatti, 1987) and is still widely used. However, concerns have been voiced for the validity of this approach (Todd, et al., 2004).

Studies on coherence measures show more consistent results than those on textual element measures. In a study on 12 ESL students' essays, for example, Intaraprawat and Steffensen (1995) investigated the difference between good and bad ESL writers' use of **transitional signals** (logical signals, **exemplification signals, reminders, sequencers, topic shift signers**). They found a significant difference between the two groups of learners in using exemplification signals (code glosses) and purpose signals (illocution markers). Kennedy and Thorp (1999) contributed that **organization, argument development and exemplifications** were coherence aspects which could differentiate students' writing scores in IELTS writing test.

#### *Summary of research results on textual features*

In contrast to the inconsistent results of L1 writing studies on the role of cohesion and coherence in distinguishing students, L2 writing researchers are able to draw consistent conclusion. Cohesion and coherence measures should obviously be included in the assessment of L2 writing ability.

#### *4.1.3 Research into sociolinguistic knowledge*

In terms of reader-writer interactions, L2 writing research mostly focuses on the influence of writing styles on writing quality. Knoch (2007, 2008) reviewed this issue thoroughly. Academic writing styles can be more objectively measured through the analysis of **metadiscoursal markers**, including hedges (epistemic certainty markers), certainty markers (emphatics or boosters), attributers, attitude markers, and commentaries (Crismore et al., 1993). EFL writers were found to be more direct by using fewer **hedging devices** than native ones. Similarly, Kennedy and Thorp (2007) found that IELTS writers with high proficiency

used more **hedging, attitude markers, attributors, boosters, and commentaries** than those at lower levels. Similar findings on IELTS writers' stylistic choice were found by Mayor et al. (2007). These writers seem to be confident in their language ability enough to interact with the audience. For another group of ESL writers, Intaraprawat and Steffensen (1995) reached almost similar results. Thus, besides supporting findings by L1 researchers (Crismore et al., 1993), these results added that students at higher proficiency tended to target the audience more while those at the lower end are more prone to use connectives as the main cohesive device. The use of first personal pronoun as the subject in the writing, and second personal pronoun as the object have also been discussed as other indicators of academic styles. Research results (Hyland, 2002; Mayor, et al., 2007; Shaw & Liu, 1998) are inconsistent regarding whether L2 writers are more personal or objective as their writing improves.

In summary, there is still limited research on the sociolinguistic knowledge of writing other than on academic styles. This does not indicate that it is not an important type of knowledge. Regarding its importance, it should be a fertile field for research in the future.

#### *Summary of research results on writing products*

So far, this section has presented an overview of traditional research in L2 writing production. As Cumming (2001b) concludes, the summarized research findings suggest (rather than “confirm”) that student writing proficiency is connected with the complexity of syntax and morphology, the range of vocabulary, the command of rhetorical forms, the control of cohesion and coherence, the use of stylistic devices, amongst others.

#### *4.1.4 L2 writing product assessment criteria in use*

In addition to research results, the criteria used by important L2 testing services in writing assessment practices also serve as important hints for the components of an L2 writing construct. With the emergence of the current writing assessment paradigm, these components have undergone great changes.

The earliest testing organization to reform was the American Council on the Teaching of Foreign Languages (ACTFL). The ACTFL guidelines have been revised (Breiner-Sanders, Swender, & Terry, 2001). The list of writing indicators can be seen in Figure 5.

Changes to the TOEFL test format carried in itself multiple changes of test construct, including the integration of writing skills with other major skills. This can be observed from the changes in a comparison of the writing assessment indicators in the old (Education Testing Service, 2011b) and new instrument (Education Testing Service, 2011a). Figure 2.9 shows that the new instrument is far more complicated and detailed. Content criteria were not mentioned in the old instrument except for the term “addressing the task”, but have been described quite thoroughly in the new one with such aspects as unity, clarity, arguments, and progression. More attention is paid towards syntactic accuracy and complexity than towards range and diversity as before. In terms of textual knowledge, the old instrument only mentioned organization of ideas, while the new one adds the use of examples, explanations, cohesive devices, etc.

Hawkey and Barker (2004) reported the attempt to create a common instrument for the ESOL (Cambridge English for Speakers of Other Languages) and IELTS tests at the



University of Cambridge. First, the criteria used in scoring five ESOL tests and the IELTS test (University of Cambridge Local Examination Syndicate, 2005) were described (as in Figure 5) and regrouped. They are then applied to an extensive corpus of writing scripts of students from different levels of proficiency. The derived four-level draft scale is reported to be applicable for IELTS writing scripts which have received band scores from 3 to 9 in earlier ratings. The criteria on the common instrument can be seen in Figure 5.

The University of Cambridge has also revised their IELTS rating rubrics (S.Shaw & Falvey, 2008). This was done in a highly empirical study involving a thorough literature review, reiterative discussions with experts and sophisticated quantitative validation procedures (Generalizability Theory, Item Response Theory). Apparently, validity and reliability are being seriously reconsidered by the testing system. The results show that in writing task 1, the instrument contains four groups of

criteria: task achievement (task requirement fulfillment, idea development, purpose targeting, format, tone, clarity, illustration, information appropriation), coherence and cohesion (overtness of cohesion devices, paragraphing, sequencing, progression, primary and secondary transition, repetitions, clarity of ideas in each paragraph, references), lexical resources (range, sophistication of control, errors, commonality of vocabulary, precision, formation, communication influence, mechanics, collocation) and grammatical range and accuracy (structure range, errors, appropriation of structures, communication influence). In writing task 2, the three later groups of criteria are the same as in task 1, while the first is changed into task response (response completion, position development, support, focus, generalization, relevance, and format). It is apparent that the revised instruments of the IELTS tests include far more criteria than the old versions (see Figure 5), and the principles underlying the writing

assessment have been specified more thoroughly.

Tests	CPE	CAE	FCE	PET	KET	CELS	IELTS	Draft common ESOL instrument	TWE	iBT T1	ACTFL01
Criteria											
Topic								development		progression	variety, development
Resourcefulness		X									
Coverage		X	X	X							
Content	X	X	relevance, originality,			X	relevance, arguments			argument	ideas
Unity								X		X	
Clarity										X	
Sentence structure							X	length	X, variability	variety, usage	variety
Grammatical structure	Range			range	X	range		X			X
Syntax											variety
Time frames, moods, aspects											X
Vocabulary range	X	X		X		X				Variety	variety
Vocabulary				X	X		X	order, complexity	choice, variability	choice, form	specialization
Idiomtinity								X		X	
Coherence							X			X	transition,
Elaboration				X			support	X		Examples, Explanation	
Organization	X	X	X	X		X	logical	X	X	X	X, prioritization
Links			X			X		X		connections	transitions
Cohesion	X	X	X				X	X			X, devices,

Tests		CPE	CAE	FCE	PET	KET	CELS	IELTS	Draft common ESOL instrument	TWE	iBT T1	ACTFL01
Criteria		Sophistication of language	Naturalness	task realization		Message communication			comprehension	Language use, task addressing, goals	Topic/task addressing	
General impression												
Style									X			X, variation
Register and format	X	X	X	X			X					
Audience target	X	X	X	X				X	X			X, approximation
Presentation				X								
Accuracy				X			X	X				X
Errors	X	X	X		X		X	X	X	X	X	X
Fluency								X				X
Length												
Complexity					X		X		X			X
Mechanics							X					X
Distortion							X				digression	
Format												X
Functions												X
Genres												X

Figure 5. Criteria in Common ESL/EFL Writing Assessment Instruments. X represents that the indicator is used in the test.

Although the criteria vary from one test to another, the table can provide some generalizations of what all the test developers deem to contribute to writing quality. The writing performance indicators vary in different assessment contexts. For example, the ACTFL aims to assess writers' overall ability rather than a one-test writing ability, so they include in their assessment instrument the capability to write with a variety of topics, genres, and functions, which other test instruments do not. Also, unlike any others, the ACTFL highlight students' specialized vocabulary in writing. As another example for the relation between context and assessment criteria, the instruments of both IELTS and TOEFL iBT writing tests include the formation of arguments because both tests require students to write argumentative essays. Interestingly, these variations can serve both as illustrations for the versatile applications of well-established theories in practice and as considerations for designing other studies and forming new theories in L2 writing.

Noticeably, besides the three core areas of text structure knowledge, textual knowledge and sociolinguistic knowledge, one important group of criteria is related to the writing content (topic, content, unity, clarity, resources, coverage, etc). This may relate to the research findings in rater-behaviour studies that raters are strongly concerned with content and organization (Huot, 1990). However, "content" has different meanings in each test. Depending on the contexts, it may include other criteria such as unity, resourcefulness, coverage, and clarity. Furthermore, research on the impact of writing content on writing quality is rather limited in L2 writing assessment (excluding content-based courses) (Knoch, 2007). Two studies involving IELTS scripts by Kennedy

and Thorp (2007) and Mayor et al. (2007) found that students at lower levels tend to write incomplete arguments, fail to elaborate their answers, and produce categorical organization of ideas and paragraphs while higher level students interact with the readers better via a range of rhetorical devices (rhetorical questions, hedging, boosters, downtoners). More studies on content criteria are needed to confirm these results.

Syntactic features are paid equally enormous attention to by all test developers, the primary features being grammatical structure, sentence structure, vocabulary range/diversity and choice. Other aspects which come into focus are word order, word complexity, the use of idioms, the use of time frames and other grammatical aspects. Coherence and cohesion are mentioned in all instruments, but they are often accompanied by more specific indicators such as the use of transitions, links, elaborations (types of elaborations), organization, logic and development. The range of criteria to assess sociolinguistic knowledge in these instruments, which includes writing styles, audience-targeting skill and writing registers, is wider than in corresponding research. Apart from the criteria which apparently relate to specific areas of product-oriented writing knowledge, the instruments also include other more global criteria of accuracy (either directly or indirectly via "errors"), mechanics, fluency (mostly in terms of length), and general impression.

In general, textual structure elements and textual structure are still the main concerns of instrument developers in constructing definitions of L2 writing competence. The criteria for assessing textual features clearly outnumber those for assessing content and sociolinguistic knowledge.

#### 4.2. *Research on L2 composing process*

The increasing popularity of the process approach in L2 writing research is described in many profound historical reflections. The most important reflectioners to whom this study owes considerable interpretative support are Kraspels (1990), Grabe and Kaplan (1996), Larios et al. (2002), Silva and Brice (2004) and Matsuda (2003). The exponential rise of studies in L2 writing may be the most important point in all the reflections. This section presents the key L2 composing research results and discusses how useful they are to studies on writing.

##### 4.2.1 *General composing process*

It has been shown in the early years of L2 writing assessment that L2 writers' composing competence, writing strategies and behaviours can explain their success in writing, rather than their linguistic competence (Raimes, 1985; Zamel, 1984). More recent and more process-oriented studies, such as Sasaki (2000, 2002), found similar results that expert writers plan much longer and with more details and more thorough linguistic preparation, plan their organization better (including several local plans), refine the expressions in their mind before writing, and make fewer pauses than novice ones. Sasaki also stressed that global organization is a difficult skill which may take a less skilled writer a long time to acquire.

Multiple studies reveal that skilled writers can monitor their composing strategies flexibly while unskilled writers' popular pattern is to add new ideas at any stage. However, there are cases when it is hard to clearly define a common profile of composing behaviours for a group of L2 writers (Raimes, 1985).

In line with the theory, audience and purpose orientations are found to be discriminative indicators of writing ability as

seen from the results in a number of different studies (Raimes, 1987; Zamel, 1984).

##### 4.2.2 *Planning*

Planning has apparent effects on fluency and complexity, while having an inconsistent influence on accuracy (Foster & Skehan, 1999). In other words, planning may lead to a trade-off between accuracy and complexity. Larios, Murphy and Marín (2002) reviewed 65 studies on composing processes and concluded that skilled students are more attentive to readers and write for the readers more than the unskilled students. They also have more planning activities, plan according to the goals, and plan in advance as well as when necessary during the writing process. Another comprehensive report of the effects of different types of planning on fluency, accuracy, and complexity has been made by Ellis (2009). Planning activities with different features could lead to a range of subtle effects on writing quality. In other words, the studies on planning are no longer limited to one generally defined planning activity, but have expanded to various sub-strategies in this sub-process. As regards the relation between planning and ESL/EFL writing proficiency, which is not unanimously defined, recent studies showed inconsistent results. Tavokoli and Skehan (2005), for example, reported the positive effects of planning on low proficiency writers' fluency but not on advanced writers, while Mochizuki and Ortega (2008) found no effects of planning for their low proficiency subjects.

##### 4.2.3 *Formulating/ text producing*

According to a report by Larios, Murphy and Marín (2001; 2002), this part of the writing process has received the least research attention. In formulating written texts, skilled writers integrate their knowledge of textual



elements and textual knowledge and control their attention to both. Unskilled writers also pay attention to various areas of knowledge, but they are possibly restricted to structural units of language instead of more global features. Due to this focus on forms, they spend less time on online planning and revising. Inconsistent results were reached regarding the content produced by unskilled writers. In the first place, some studies stated that these students limited their ideas to personal experience, followed a rigid sequence of idea development, and wrote shorter texts than skilled writers (counted as an aspect of formulation). However, other studies rejected these differences. Similar inconsistencies can be found with the studies on pauses during the writing process. So far, it has been agreed that formulating contains many smaller processes of problem solving, on-line planning, on-line revising, adding resources, etc; each varies in time and occurrence in different writers. In their study of formulation, Larios, Marín and Murphy (2001) found that L2 proficiency can determine the dominance and position of formulation in the writing process, with lower proficiency students formulating longer and more frequently at the early stage in the writing process than higher proficiency ones. In other words, higher proficiency students formulate less and at a later stage of their composing process. The strong theoretical arguments and well-defended measures make the results of this study remarkably significant; however, other similar studies are needed so that formulating/text producing can be better understood and can deserve its position in the writing process.

#### 4.2.4. Revising

This is the most well-researched stage in the composing process (Bergh & Rijlarsdam, 2001). The research results are abundant.

Larios, Murphy and Marín (2002) reported that skilled writers distinguish well between revision and writing skills. These students were not as concerned with syntax as with idea, intersentential and paragraph levels. They were conscious of the opportunities they could have in revising drafts and therefore did not stop at revising mechanical mistakes. They can also detect more problems and have more resources to solve the problems than less skilled writers (Kobayashi & Rinnert, 2001). In contrast, unskilled writers do not take revising and editing seriously (Raines, 1985). They are often held back by syntactical revision during the writing process. This may be due to limitations in their linguistic knowledge or the pressure to finish writing in a short time. In general, with quite consistent findings, revising seems to be the sub-process which best discriminates students' writing ability.

#### 4.2.5. Research on writing learning strategies and L2 writing proficiency

Besides the knowledge of the writing process, L2 writing researchers also study processing skills in writing, which have been mentioned as essential in the theory of writing by Grabe and Kaplan (1996). This metaknowledge includes three categories: personal knowledge, task knowledge, and strategic knowledge (Devine, Railey, & Boshoff, 1993). In terms of task and strategic knowledge, studies have yielded consistent results. Specifically, skilled writers are more flexible in their attitude to the tasks of writing. They are more conscious in taking risks with writing complicated structures, as well as in understanding all the skills writing involves. Strategic competence in writing was even found to be the means to acquire better task and personal knowledge (Kasper, 1997). In contrast, unskilled writers are unmotivated and limited their skills to the

writing of grammatically correct sentences. They may not be conscious of their problems and weaknesses in either syntactic or textual knowledge, leading to a lack of confidence in experimenting with new knowledge. In terms of personal knowledge (knowledge held by writers about themselves), studies show various results. Some studies report that skilled writers are more motivated while unskilled writers are not satisfied or happy with writing. Others report an equal level of motivation for both groups. This result has been explained by the limit of time students are allowed to write, which drives all students into the same negative attitude to writing.

In general, considering the research results, it seems that most of what Flower and Hayes (1981), Hayes and Flower (1987), Bereiter and Scardamalia (1987) and Grabe and Kaplan (1996) modelled about the writing process has been supported. The main stages of the composing processes, as well as their complicated occurrence and time allocation, are confirmed. Research results also suggest differences in the ability to plan, revise, edit, search for expressions and attend to ideas between students of different writing levels (Cumming, 2001b). The control of the stages is found to be most indicative of student writing ability. Inconsistent results were only found in certain aspects of the subprocesses. Furthermore, the subprocesses can be ranked along a continuum of difficulty in acquisition (Larios, Murphy & Marín, 2002). This means that some are less difficult for L2 students to acquire than others.

However, as regards methodological issues, results from L2 writing process research should be considered with care. First, the studies are quite limited in their instruments due to the greater time requirement (Polio, 2003). Most studies rely on a small number

of subjects (Cumming, 2001b; Krapels, 1990; Raimes, 1985). In addition, the method of judging, the criteria to judge the proficiency of students, the time allowance for writing, the contexts of assessment (ESL or EFL) and the samples of writing products are not similar across studies (Cumming, 2001). Many researchers also failed to provide reliability estimates, which used to be the most serious criticism of writing assessment (Larios, Murphy & Marín, 2002). Moreover, despite recent geographical growth, there is still a serious lack of research in EFL settings outside the United States (Polio, 2003; Silva & Brice, 2004) and the mismatch between product-focused assessment practices and the process-oriented research is still apparent (Huot, 1996). These features serve as important warnings for L2 writing researchers to be careful with the generalization and the practicality of results. Whether a positive or negative relation is found between composing process(es) and writing quality, it is reckless to suggest the adoption or omission of certain processes for certain groups of L2 students. The range of writing strategies and processes are wide, and students acquire differently from different processes of learning (Polio, 2003).

## 5. Conclusion

The paper begins with an introduction of the developments in writing assessment, including the shift of paradigms. The second part of the paper narrows the discussion to the essential issues in L2 writing assessment, including theoretical models on the construct of writing competence and L2 writing competence, and the current research results which support or dismiss those models. It is apparent from the discussed contents that L2 writing assessment is a complicated, multidisciplinary, fast-growing field and that the teachers and

researchers not only have to read broadly and critically but also to continually update their knowledge to understand the dynamics of the field. Also, L2 writing assessment is obviously developing in the same direction as general writing assessment in the performance-based and process-oriented approach, besides the traditional product-oriented approach. The wide variety of research focus in writing assessment and the differences in the amount of empirical evidence for each research focus should be taken into consideration by teachers and researchers in L2 writing assessment contexts such as Vietnam, so that they can better formulate their writing tests or assignments.

## References

- Banerjee, J., Franceschina, F., & Smith, A. M. (2007). *Documenting features of written language production typical at different IELTS band score levels*. Canberra: IELTS Australia.
- Bereiter, C., & Scardamalia, M. (1987). *The psychology of written composition*. Hillsdale, New Jersey: Lawrence Erlbaum Associates Inc.
- Breiner-Sanders, K. E., Swender, E., & Terry, R. M. (2001). *Preliminary Proficiency Guidelines - Writing revised 2001*. Retrieved January 8th 2011, from <http://www.actfl.org/i4a/pages/index.cfm?pageid=4236>
- Brown, J. D., & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly*, 32(4), 653-675.
- Brown, J. D., & Hudson, T. (2002). *Criterion-referenced language testing*. Cambridge: Cambridge University Press.
- Camp, R. (1993a). Changing the model for the direct assessment of writing. In M. M. Williamson & B. Huot (Eds.), *Validating holistic scoring for writing assessment: Theoretical and empirical foundations*. Cresskill, New Jersey: Hampton Press.
- Camp, R. (1993b). Changing the model for the direct assessment of writing. In B. Huot & P. O'Neill (Eds.), *Assessing writing: A critical sourcebook*. Urbana, Illinois: National Council of Teachers of English.
- Chenoweth, N. A. (2001). Fluency in writing: generating text in L1 and L2. (Statistical Data Included). *Written Communication*, 18(1), 80-98.
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34(2), 213-237.
- Crismore, A., Markannen, R., & Steffensen, M. S. (1993). Metadiscourse in persuasive writing - a study of texts written by American and Finnish University students *Written Communication*, 10(1), 39-71.
- Crowhurst, M. (1987). Cohesion in argument and narration at three grade levels. *Research in the teaching of English*, 21(2), 185-201.
- Cumming, A. (2001b). Learning to write in second language: two decades of research. *International Journal of English Studies*, 1(2), 1-23.
- Cumming, A. (2002). Decision making while rating ESL/EFL writing tasks: A descriptive framework. *The Modern Language Journal*, 86(1), 67-96.
- Devine, J., Railey, K., & Boshoff, P. (1993). The implications of cognitive models in L1 and L2 writing. *Journal of Second Language Writing*, 2(3), 203-225.
- Education Testing Service. (2011a). *TOELF iBT Integrated writing rubrics*. Retrieved May 8th from [http://www.ets.org/Media/Tests/TOEFL/pdf/Integrated\\_Writing\\_Rubrics\\_2008.pdf](http://www.ets.org/Media/Tests/TOEFL/pdf/Integrated_Writing_Rubrics_2008.pdf).
- Education Testing Service. (2011b). *TOELF Paper-based writing course guide*. Retrieved May 8th from [http://www.ets.org/toefl/pbt/scores/writing\\_score\\_guide/](http://www.ets.org/toefl/pbt/scores/writing_score_guide/).
- Ellis, R. (2009). The differential effects of three types of task planning on the fluency, complexity, and accuracy in L2 Oral Production. *Applied Linguistics*, 30(4), 474-509.
- Flower, L., & Hayes, J. R. (1981). A Cognitive process theory of writing. *College Composition and Communication*, 32(4), 365-387.
- Foster, P., & Skehan, P. (1999). The influence of source of planning and focus of planning on task-based performance. *Language Teaching Research*, 3(3), 215-247.
- Grabe, W., & Kaplan, R. B. (1996). *Theory and practice of writing: an applied linguistic perspective*. New York: Longman.
- Greenberg, K. (1981). *The effects of variations in essay questions on the writing performance of CUNY freshmen*. New York: City University.
- Grobe, C. (1981). Syntactic maturity, mechanics, and vocabulary as predictors of writing quality. *Research in the teaching of English*, 15, 75-85.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Haswell, R., & Wyche-Smith, S. (1994). Adventuring into writing assessment. *College Composition and Communication*, 45(2), 220-236.
- Hayes, J. R., & Flower, L. S. (1987). On the structure of the writing process. *Topics in Language Disorders*, 7(4), 19-30.
- Hinkel, E. (2003). Simplicity without elegance: Features of sentences in L1 and L2 academic texts. *TESOL Quarterly*, 37(2), 275-300.

- Huot, B. (2002). *(Re)articulating writing assessment for teaching and learning*. Logan: Utah State University Press.
- Hyland, F. (2002). Authority and invisibility: Authorial identity in academic writing. *Journal of Pragmatics*, 31, 1091-1112.
- Intaraprawat, P., & Steffensen, M. S. (1995). The use of metadiscourse in good and poor ESL essays. *Journal of Second Language Writing*, 4(3), 253-272.
- Kane, M. T. (1992). An argument-based approach to validity. *Psychological Bulletin*, 112(3), 527-535.
- Kaplan, R. B., & Grabe, W. (2002). A modern history of written discourse analysis. *Journal of Second Language Writing*, 11(3), 191-223.
- Kasper, L. F. (1997). Assessing the metacognitive growth of ESL student writers. *TESL-English Journal*, 3(1), 40-53.
- Kennedy, C., & Thorp, D. (2007). A corpus-based investigation of linguistic responses to an IELTS Academic Writing Task. In T. Linda & F. Peter (Eds.), *Studies in language testing: IELTS collected papers - Research into speaking and writing assessment* (Vol. 19, pp. 316-379). Cambridge: Cambridge University Press.
- Kintscha, W., & Dijk, T. A. v. (1978). Toward a model of text comprehension and production *Psychological Review*, 85(5), 363-394.
- Knoch, U. (2007). *Diagnostic writing assessment: The development and validation of a rating scale*. The University of Melbourne, Melbourne.
- Kobayashi, H., & Rinnert, C. (2001). Factors relating to EFL writers' discourse level revision skills. *International Journal of English Studies*, 1(Japanese People), 71-101.
- Larios, J. R. D., Murphy, L., & Marín, J. (2002). A critical examination of L2 writing process research. In S. E. Randell & M.-L. Barbier (Eds.), *New directions for research in L2 writing*. The Netherlands: Kluwer Academic Publishers.
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: Lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322.
- Lautamatti, L. (1987). Observations on the development of the topic in simplified discourse. In R. B. Kaplan & U. Connor (Eds.), *Writing across languages: Analysis of L2 text*. Massachusetts: Addison-Wesley.
- Matsuda, P. K. (2003). Second language writing in the twentieth century: A situated historical perspective. In B. Kroll (Ed.), *Exploring the dynamics of second language writing* (pp. 11 - 15). Cambridge: Cambridge University Press.
- Mayor, B., Hewings, A., North, S., Swann, J., & Coffin, C. (2007). A linguistic analysis of Chinese and Greek L1 scripts for IELTS Academic Writing Task 2. In T. Linda & F. Peter (Eds.), *Studies in Language Testing: IELTS collected papers - Research in speaking and writing assessment* (Vol. 19, pp. 250-314). Cambridge: Cambridge University Press.
- McCulley, G. A. (1983). Writing quality, coherence and cohesion. D. Ed. dissertation, Utah State University, United States - Utah. Retrieved from ProQuest Dissertations and Theses (Publication No. AAT 8313552).
- Mochizuki, N., & Ortega, L. (2008). Balancing communication and grammar in beginning-level foreign language classrooms: A study of guided planning and relativization. *Language Teaching Research*, 12(1), 11-37.
- Polio, C. (1997). Measures of linguistic accuracy in second language writing research. *Language Learning*, 47(1), 101-143.
- Raimes, A. (1985). What Unskilled ESL Students Do as They Write: A Classroom Study of Composing. *TESOL Quarterly*, 19(2), 229-258.
- Reid, J. (1992). A computer text analysis of four cohesion devices in English discourse by native and nonnative writers. *Journal of Second Language Writing*, 1(2), 79-107.
- Reid, J. (1993). *Teaching ESL writing*. Boston, Massachusetts: Prentice Halls.
- Reynolds, C. R., Livingston, R. B., & Willson, V. L. (2006). *Measurement and assessment in education*. Boston: Pearson/Allyn & Bacon.
- Ruth, L., & Murphy, S. (1988). *Designing writing tasks for the assessment of writing*. Norwood, New Jersey: Ablex Publishing Inc.
- Sasaki, M. (2000). Toward an empirical model of EFL writing processes: An exploratory study. *Journal of Second Language Writing*, 9(3), 259-291.
- Sasaki, M. (2002). EFL learners' writing processes. In S. E. Randell & M.-L. Barbier (Eds.), *New directions for research in L2 writing* (pp. 48-80). The Netherlands: Kluwer Academic Publishers.
- Shaw, P., & Liu, E. T.-K. (1998). What develops in the development of second-language writing? *Applied Linguistics*, 19(2), 225-254.
- Shaw, S. & Falvey, P. (2008). *The IELTS Writing Assessment revision project: towards a revised rating scale*. Cambridge: University of Cambridge ESOL Examinations, 1-295.
- Shaw, S., & Weir, C. J. (2006). *Examining writing: Research and practice in assessing second language writing*. Cambridge: Cambridge University Press.
- Silva, T., & Brice, C. (2004). Research in teaching writing. *Annual Review of Applied Linguistics*, 24, 70-106.
- Tavokoli, P., & Skehan, P. (2005). Strategic planning, task structure, and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language*. Amsterdam: John Benjamin Publishing Company.
- Todd, R. W., Thienpermpool, P., & Keyuravong, S. (2004). Measuring the coherence of writing using topic-based analysis. *Assessing Writing*, 9, 85-104.



- University of Cambridge Local Examination Syndicate. (2005). IELTS writing band descriptors: Task 2 (Public version). Retrieved January 8th 2011 from [https://www.teachers.cambridgeesol.org/ts/digitalAssets/113300\\_public\\_writing\\_band\\_descriptors.pdf](https://www.teachers.cambridgeesol.org/ts/digitalAssets/113300_public_writing_band_descriptors.pdf).
- Vaughan, C. (1991). Holistic assessment: What goes on in the rater's mind? In L. Hamp-Lyons (Ed.), *Assessing second language writing in academic contexts*. Norwood, New Jersey: Ablex Publishing Corporation.
- Weigle, S. C. (2002). *Assessing writing*. Cambridge: Cambridge University Press.
- Wenjun, J. (1998). *Cohesion and the academic writing of Chinese ESL students at the graduate level*. Northern Illinois University.
- White, E. (1995). An apologia for the timed impromptu essay test. *College Composition and Communication*, 4(1), 30-45.
- Witte, S. P., & Faigley, L. (1981). Coherence, cohesion and writing quality. *College Composition and Communication*, 32(2), 189-204.
- Witte, S. P., & Faigley, L. (1983). *Evaluating college writing programs*. Carbondale: Southern Illinois University Press.
- Wolfe-Quintero, K., Inagaki, S., & Kim, H.-Y. (1998). *Second language development in writing: Measures of fluency, accuracy and writing quality*. Technical Report Number 17. Honolulu: HI: University of Hawaii.
- Yancey, K. B. (1999). Looking back as we look forward: Historicizing writing assessment. *College Composition and Communication*, 50(3), 483-503.
- Zamel, V. (1984). The composing processes of advanced ESL students - 6 case studies - reply. *TESOL Quarterly*, 18(1), 154-158.
- Zareva, A. (2005). Relationship between lexical competence and language proficiency - Variable sensitivity. *Studies in second language acquisition*, 27, 567-595.

## RÀ SOÁT CÁC CƠ SỞ LÝ THUYẾT VÀ NGHIÊN CỨU VỀ TIÊU CHÍ CHẤM MÔN VIẾT TRONG ĐÁNH GIÁ TIẾNG ANH NHƯ NGÔN NGỮ THỨ HAI

Dương Thu Mai

*Trường Đại học Ngoại ngữ, ĐHQGHN,  
Phạm Văn Đồng, Cầu Giấy, Hà Nội, Việt Nam*

**Tóm tắt:** Đánh giá năng lực ngôn ngữ đã nhận được sự quan tâm ngày càng sâu rộng tại Việt nam trong những năm gần đây và đang được thay đổi tích cực. Trong bối cảnh đó, đánh giá theo tiêu chí đã trở thành một khái niệm quen thuộc với giáo viên, người đánh giá cũng như các nhà quản lý giáo dục. Mặc dù vậy, hầu hết giáo viên tiếng Anh ở các cấp đều vẫn gặp khó khăn khi phải xác định các tiêu chí có thể sử dụng trong thang đánh giá môn viết của họ. Bài viết này nhằm mục đích cung cấp một nguồn tham khảo cho giáo viên và các nhà nghiên cứu có quan tâm tới đánh giá môn viết trong lĩnh vực viết như ngôn ngữ thứ hai trong nhiệm vụ xây dựng thang đánh giá năng lực viết theo tiêu chí. Nội dung của bài viết tập trung vào cơ sở lý thuyết đang tồn tại và đã được xuất bản trên thế giới về các đường hướng dạy viết tiếng Anh, các lý thuyết, các nghiên cứu và thực hành đánh giá viết bằng tiếng Anh. Các nội dung chính này được tổ chức theo hai hướng chính: dựa trên sản phẩm viết, và dựa trên quá trình viết.

**Từ khóa:** Đánh giá môn viết, hướng tiếp cận dạy viết, tiêu chí đánh giá môn viết, đánh giá sản phẩm viết, đánh giá quá trình viết