UC Riverside
Issues in Applied Linguistics

Title
Cohesion and Coherence in Children's Written English: Immersion and English-only Classes

Permalink
https://escholarship.org/uc/item/3hb2z0s4

Journal
Issues in Applied Linguistics, 12(1)

ISSN
1050-4273

Author
Bae, Jungok

Publication Date
2001-06-30

DOI
10.5070/L4121005043

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Cohesion and Coherence in Children's Written English: Immersion and English-only Classes

Jungok Bae
University of California, Los Angeles

This study investigates the nature of cohesion, coherence, content, and grammar emergent in children's essays, with a greater emphasis given to the understanding of cohesion and coherence. Conceptual definitions of these constructs are summarized based on prior research. The measurement of these constructs is operationalized into a picture-based narrative writing task for elicitation and scoring criteria for quantification. 192 first and second graders from an immersion program and English-only classes participated in the study. The analysis uses percentages, correlations, multiple regression, and qualitative analyses. Main findings include the following: (a) the measurement of cohesion and coherence can be operationalized; (b) referential and lexical cohesion correlate highly with the overall writing quality defined as the sum of the ratings of coherence, content, and grammar; (c) ellipses and substitution show a weak correlation with the overall writing quality; (d) lexical and referential cohesion are significant predictors of coherence while other types of cohesion are not; (e) dominant reference types are pronominal forms and proper nouns, and prominent types of conjunctive relation are temporal and additive; and (f) the most common error in cohesion is inaccurate reference. The substance and method of this study can provide a foundation for investigating subsequent topics with latent variables and different linguistic backgrounds and grade levels.

This study aims to understand certain linguistic and semantic resources for text construction, namely the constructs of cohesion, coherence, content, grammar, and text length in English writing. A greater emphasis in this study is given to the understanding of cohesion and coherence. With this in mind, their relations to other salient writing constructs, such as content, grammar, and text length are investigated.

Generally, the concepts of cohesion and coherence are more technical and relatively uncommon to many people compared to the concepts of other more universally understood language-related components, such as grammar, content, and text length. One of the most significant works that have contributed to our explicit understanding of cohesion is Halliday and Hasan (1976/1993). According to Halliday and Hasan (1976/1993, p. 4), the concept of cohesion is a semantic one, referring to "relations of meaning" that exist within the text, and it "occurs where the interpretation of some element in the discourse is dependent on that of another." Cohesion is expressed "partly through the grammar and partly through the vocabulary" (p. 5). In comparison, coherence, generally defined, refers to the quality of a text when it makes sense or is pleasing because all the parts or steps fit

Issues in Applied Linguistics
ISSN 1050-4273
© 2001, Regents of the University of California
Vol. 12 No. 1, 51-88

together well and logically (Collins Cobuild, 1996). It is the connection that is established partly through cohesion (Halliday & Hasan, 1989) and partly through something outside the text that is usually the knowledge which a listener or reader is assumed to possess (Renkema, 1993, p. 35), such as background knowledge, genre expectations, and reader expectations. Extensive studies are available that offer theoretical discussions of cohesion and coherence (Bamberg, 1984; Brown & Yule, 1983; Connor & Johns, 1990; Cook, 1989; Gundel, Hedberg, & Zacharski, 1993; Halliday, 1985; Halliday & Hasan, 1976/1993, 1989; Koshik, 1999; McCarthy, 1991; Oller & Jonz, 1994; Renkema, 1993; Smith, 1984).

Grounded in current theories of cohesion and coherence, this paper outlines theoretical definitions of cohesion, coherence, content, and grammar. The study then operationalizes the measurement of the postulated constructs of coherence and cohesion, as well as content, grammar, and text length as a first step toward developing assessment measures of the constructs. Specifically, the theoretical constructs are realized into operational definitions of the constructs by means of a narrative writing task and scoring criteria used to elicit and quantify the constructs. This study is essentially about *construct validation* (or to use Nunnally's 1978 alternative term for construct validation suitable to this paper, *construct explication*), which is the process of elaborating and refining the meaning of the constructs on the basis of empirical evidence. In sum, cohesion and coherence, which have been given ample theoretical discussions in prior research, are given an empirical investigation and evidence based on assessment procedures in this study. This measurement-based empirical augmentation is what is different from prior research in the theory of cohesion and coherence.

Context of the Study

This investigation of cohesion, coherence, content, and grammar was undertaken within the context of a two-way immersion program, the Korean/English Two-Way Immersion Program (Campbell et al., 1994; Kim, 1996; Walker, 1992), and English-medium classes. The immersion program consists largely of Korean-language background and English-language background students, and the English-medium classes consist of students who are proficient in terms of English oral skills. Although the present study is not intended to compare writing in these instructional programs (see below), the student populations are of interest in that immersion classes and English-medium classes have rarely been used in the study of cohesion and coherence. The English level of all the students in the English classes and the majority of students from the immersion program was identified as proficient in terms of oral English skills at the time of the study. However, the levels of literacy skills of these children differ widely even among the students who have a fluent command of oral English. Using the additional data from Korean-background students who were initially identified as Limited English Proficient (LEP) but who were becoming proficient bilinguals in the immersion pro-

gram at the time of the study, interesting aspects of the specified constructs on the learning curve could be incorporated in the study.

Purpose of Paper

In the present paper, immersion classes and English-medium classes are grouped together in studying the nature of the specified constructs. The concern in this paper is not which instructional program performs better or whether distinct or same characteristics of text qualities are found between these programs. Rather, by incorporating the two instructional groups, the paper aims to substantiate robust constructs of English writing that reflect the text characteristics across these programs. It can be said that, in the present paper, the robust English writing constructs are established as a basis for evaluating the learning of English writing that is going on and for comparing learning across programs in subsequent studies.

Research Questions

This study addresses the following research questions:

1.What characteristics of cohesion and coherence are observed in the narrative writing of children (including immersion students and students in English-only classes), and how are these two qualities related to each other and to other salient components of language such as content, grammar, and text length?

2.Which types of cohesion are more or less prominent in these children's narrative writing?

3.What are the patterns of errors in cohesion for English-language background children and Korean-language background children?

Significance and Rationale

This section discusses the significance of the approaches used in this study. The significance is addressed in the areas of internal and external aspects of construct explication, rationale for selecting the constructs, and the operationalization of the constructs.

Internal and external aspects of investigation

Extensive prior research that focuses on cohesion specifically is available (Cook, 1989; Cox, Shanahan, & Sulzby, 1990; Freedle, 1991; Halliday & Hasan, 1976/1993; Lindsay, 1984; Norris & Bruning, 1988; Ricento, 1987). Very few studies, if any, have looked into cohesion in relation to constructs external to cohesion especially at the statistical level. While this study gives a greater emphasis to the internal aspect of cohesion, it also introduces external aspects of investigations into cohesion. Specifically, this study examines not only the definition of cohesion and its subdimensions, but also illuminates cohesion in relation to other language-related variables external to cohesion, such as coherence, content, grammar, and text length. Ideally, an unlimited number of linguistic and nonlinguistic variables could be included as external variables in a validation of target con-

structs; each external variable would contribute to our understanding of the focal constructs. In practice, however, it is impossible to include all potential variables in one study. Thus, in this study coherence, content, grammar, and text length have been selected since these variables are all salient components in written texts. Coherence receives special interest because coherence has often been compared to and distinguished from cohesion (e.g., Carrell, 1982; Cox, Shanahan, & Tinzmann, 1991; Enkvist, 1990; Fitzgerald & Spiegel, 1986; Koshik, 1999; Oller & Jonz, 1994; Spiegel & Fitzgerald, 1990). Grammar and content are traditionally considered important objectives in the teaching and testing of writing: Grammar represents a linguistic domain of language and content represents the semantic domain of language. It is of interest to examine the extent to which the two constructs of linguistic and semantic properties are related to coherence and cohesion. In addition, text length is often attended to by readers and writers, and it would be of interest to see if text length has something to do with ability in the other constructs. Since we can easily quantify text length by simply counting words, text length is included in examining its relation to the above constructs. These constructs come into play in the analysis, contributing to the understanding of one another as external variables.

Operationalization of the measurement of cohesion and coherence

The significance of this study also concerns methodology employed for operationalization. In particular, number of subjects, elicitation method, and quantifications used in the process of operationalization are noteworthy.

With regard to *N* (number of subjects), research that analyzes written text qualities typically uses case studies, which usually focus on descriptive, qualitative analysis of one or several study participants. With the *N* of a case study, statistical analysis becomes very limited. For instance, correlations, ANOVA, MANOVA, and regression analyses require a large number of subjects—although precisely how large is "large" is debatable—to apply inferential statistics involving statistical significance tests in generalizing findings. This study uses data collected from several elementary school classes (a total of 192 students). With this *N*, statistical analyses such as correlations, regression, and percentage calculations can be performed. This resource thus contributes to the power of the research by enhancing the reliability and generalizability of findings about the text qualities in a way that has not been possible in prior research on cohesion and coherence that use case study analyses.

Secondly, methods for eliciting and quantifying the writing constructs introduced in this study, consisting of a writing task and scoring criteria, are notable. Most studies that use children's data have used spoken data. Consequently, there is a dearth of written data collected from children. This is regrettable because children's writing can provide a promising area of research in language acquisition and assessment. For instance, it is easier to observe language use in writing than in speaking and listening. Children as learners of literacy skills also reveal

interesting developmental features in their writing. The lack of research using children's writing is also due to the fact that young children develop oral skills before they are exposed to written materials. Due to this trend, writing assessment for children has only been in its preliminary stage, leading to a paucity of methods designed to elicit and assess writing samples from them, which, in turn, results in the lack of research on children's written data. In addittion, during recent decades, educators have been aware of the limitation of multiple-choice testing and assiduously called for performance-based assessment to reform education (Arter & Spandel, 1992; Baker, O'Neil & Linn, 1993; Herman, Aschbacher & Winters, 1992; Linn, Baker & Dunbar, 1991; Mehrens, 1992). To date, measures of productive language skills such as writing have not been available as part of standardized tests due to the demands of large-scale testing, which favors multiple-choice test items. This study employs a performance-based task based on an original picture series via group testing in which students are required to produce written texts in the form of stories. This procedure makes the task not only appropriate for children but also educationally beneficial. The task with the writing prompt can also provide a potential resource for future research in children's writing.

In regards to quantification, extensive scoring rubrics are available for measuring grammar and content. Organization, which may be considered similar to or part of coherence, is often scored based on scoring rubrics in assessment projects. I know of no studies to date that have offered a quantification procedure for coherence. This study takes a proactive stance to formulate rating criteria for measuring coherence based on our theoretical understanding of coherence. It also introduces a simple scoring method of counting markers of cohesion to assess the extent to which cohesion is expressed in children's compositions; this quantification method is designed to be sensitive to the unique characteristics of cohesion, that is, distinct subdimensions and overt markers of cohesion, as we shall see below.

Construct Definitions

The constructs (or ability components) being investigated in this study are cohesion and coherence emerging in children's narrative storywriting. These two components will receive special consideration and elaboration in providing definitions. In addition, definitions of content, grammar, and text length are also provided.

Cohesion

Cohesion refers to the range of grammatical and lexical possibilities that exist for linking an element of language with what has gone before or what follows in a text: This linking is achieved through relations in meaning that exist within and across sentences (Halliday & Hasan, 1976/1993, p. 10, 33). Cohesion is confined to the specific, micro-local level of organization between and within individual clauses, thus creating connections between parts. Beginning with Halliday and Hasan, many researchers have identified several types of cohesion. The types

included in this study are reference, lexical ties, conjunction, ellipsis, and substitution, which have been traditional topics in theories of cohesion (Brown & Yule, 1983; Cook, 1989; Halliday, 1985; Halliday & Hasan, 1976, 1989; Renkema, 1993). Definitions and examples of these types of cohesion are presented in Table 1 below.

Coherence

Table 1: Types of Cohesion*

REFERENCE: Items that refer to something else in the text for their interpretation.

(1) pronominal: e.g., *he, her, they, theirs*
(2) proper nouns: e.g., *Brandon, Ms. Sharon*
(3) demonstratives: e.g., *this/these, that/those, here/there*
(4) comparatives: identity/similarity/difference/ordinals/comparatives/superatives, e.g., *same, similar, such, bigger*

CONJUNCTION: Connectors between two independent sentences.

(1) Additive: e.g., *and, or, by the way*
(2) Adversative: e.g., *but, yet, however, rather*
(3) Causal: e.g., *so, therefore, thus*
(4) Temporal: e.g., *and then, then, after that, soon, finally*

ELLIPSIS: Elements left unsaid or unwritten but are understood by the reader/speaker.

(1) Noun ellipsis: delete nouns, e.g., *He liked the blue hat; I myself liked the white.*
(2) Verbal ellipsis: delete verbs, e.g., *Tom drew a small boat and April a big boat.*
(3) Clausal ellipsis: delete clauses, e.g., *A: Will you go? B: Yes.;*
A: Would you like something to drink? B: Sure.

SUBSTITUTION: The replacement of word or structure by a "dummy" word.

(1) Noun substitution: e.g., *Tom drew a big boat and April drew a small one.*
(2) Verb substitution: e.g., *He wanted to draw pictures there, and they really did.*

LEXICAL TIES:

(1) Collocation: e.g., *go home, have fun, rain/rainy/wet/umbrella/soaked*
(2) Repetition: e.g., *drew/draw/drawing, rain/raining/rainy*
(3) Synonym: e.g., *sad/unhappy*
(4) Antonym: e.g., *boy/girl, big/small*
(5) Hyponymy (general-specific relations): e.g., *fruit/banana, apple*
(6) Meronymy (part-whole relations): e.g., *house/door, room, wall, kitchen*

* Summarized from Cook, 1989; Halliday & Hasan, 1976, 1989; McCarthy, 1991; Renkema, 1993.

Coherence is a plot-motivated overall structure (in narrative) or plan on the macro level (Berman & Slobin, 1994, p. 67). It is an overall discourse-level property that makes a text hold together (Fitzgerald & Spiegel, 1990, p. 263).

Coherence, according Halliday and Hasan (1989), can be created by cohesive markers that are appropriately used. Halliday and Hasan (p. 95) comment that early discourse of students in a new field is relatively less coherent than their later discourse because the semantic relations between the key concepts, that is,

cohesion, are not yet clear (see also Halliday & Hasan, 1976/1993, p. 4, 8). Cohesive markers alone, however, do not necessarily make the text coherent and comprehensible. A text full of cohesive markers that are locally correct could be incoherent and incomprehensible as a whole (Oller & Jonz, 1994) as in the following example from Enkvist (1990):

> My car is black. Black English was a controversial subject in the seventies. At seventy most people have retired. To re-tire means "to put new tires on a vehicle." Some vehicles such as hovercraft have no wheels. Wheels go round (Enkvist, 1990, p. 12).

The text in this example has plenty of lexical cohesion (lexical repetition), but it is difficult to imagine any consistent plausible text world (Enkvist, 1990; Oller & Jonz, 1994).

By the same token, a text with missing or misused cohesive devices may still be seen as coherent and comprehensible through means other than cohesion (Koshik, 1999). The following example was composed by a nonnative speaker of English:

> Someone come my house. Says give me money. Husband take gun shoot. Go outside die. Call police. Emergency 911. Policeman come. Take black man go hospital die. (Koshik, 1999, p. 11).

Koshik comments that this story has no grammatical cohesive devices and only instances of lexical cohesion (e.g., *someone/black man*, which are coreferential, and *police/policeman*, which contribute to topic continuity), but that the story is still seen "as a coherent whole" by readers who are native speakers of English (p. 12).

Means other than cohesive devices for establishing a sense of coherence include reader expectations of finding coherence and the frames provided by genre expectations, such as predictability inferred on setting and time sequence in narratives and the structure of ordinary conversation, such as the adjacency pair, question-answer and request-compliance (Koshik, 1999; Sacks, Schegloff, & Jefferson, 1974; Schegloff, 1990). For instance, in a conversation or a written story, even when the speaker or writer misuses a reference marker or a conjunction (e.g., *he* versus *she*, *one* versus *it*, *and* versus *but*; see text example in Appendix C, line 2 and example in Koshik, 1999, p. 14-15), the listener or reader may often have no problem with understanding the referent and a correct supra-sentential connection because the conversational structure and the setting and characters introduced in a narrative provide "a strong predictability even before the wrong form is used" (Koshik, 1999, p. 14).

Coherence is also established by the mutual interaction of the writer and reader to make sense of the text based on their shared background knowledge outside the text (Bamberg, 1984; Koshik, 1999; Renkema, 1993; Smith, 1984). Let us look at the following example used in Enkvist (1990):

> The net bulged with the lightning shot. The referee blew his whistle and signaled. Smith had been offside. The two captains both muttered something. The goalkeeper signed for relief (Enkvist, 1990, p. 12).

This text has coherence, although it lacks overt grammatically describable cohesion markers such as repetition and reference markers (Enkvist, 1990; Oller & Jonz, 1994). The text becomes coherent when certain knowledge of the world, that is, knowledge of a soccer game in this case, is applied (Renkema, 1993, p. 35). Hence, a coherent text conforms to a consistent world picture for the reader, and therefore the meaning in such a text is summarizable, comprehensible, and interpretable (Oller & Jonz).

In summary, cohesion is a more grammatical, formal, and explicit property, whereas coherence is a matter of relevance, more pragmatic in nature, and a more global property (M. Celce-Murcia, personal communication, 1996). While cohesion is easily divisible into distinct subdimensions, coherence is not susceptible to subdimensions characterized by overt markers.

Content

Content is the semantic domain of language. In this study content is defined as the relevance of a written text to a given task, as well as thoroughness, persuasiveness, and creativity consistent with task expectations. The quality of content is thus viewed as the degree to which the writing impresses the reader in terms of these criteria.

Like coherence, content is not divisible on the basis of overt grammatical markers. The quality of content can be evaluated within a phrase or a sentence, but it can also be evaluated in a more global, holistic context such as many pages taken as a whole. An incoherent text with disjointed connections cannot communicate content effectively. For these reasons, content and coherence are thought to be closely related.

Grammar

Grammar refers to morphology and syntax and best represents the linguistic domain of language. Grammar is evaluated by the range of grammatical features and the extent of grammatical errors in the text. In evaluating the extent of grammatical errors, it is useful to classify grammar errors into critical errors and minor errors depending on the seriousness of the effect on reader communication. Critical errors are defined as errors that seriously impede communication, for example, a sentence-level structure and a syntactic chunk missing. These would have to do with grammar at a global level. Minor errors, on the other hand, are defined as errors at local levels such as incorrect or omitted morphemes, for example, third person agreement and tense agreement at the local level, which do not cause difficulties in comprehension.

Viewing grammar as a global text quality makes sense because learners' lack of grammar leads to a limitation on producing a text that has a quality and

length reasonable enough to get ideas across, while serious errors in grammar are likely to cause communication obstacles. Grammar is related to coherence and content in that grammar is also a global quality and because content and coherence do not exist in a vacuum but are represented in forms realized by grammatical structures. Grammar is also connected to cohesion in that grammar can also be a local text quality as indicated in the examples of minor errors (above) and in that cohesion is partly expressed through grammar (Halliday & Hasan, 1976/1993) and is a more formal and explicit property.

Text length

Text length defined in this study is the total number of words in a writing sample.

METHOD

In this section, the method is described in terms of study participants, study variables, the writing task, and administrative and scoring procedures.

Study Participants

A total of 192 students (97 first graders and 95 second graders) participated in the study. They were students enrolled in the Korean/English Two-Way Immersion Program and students in typical English-only classes. While this study does not address a cross-group comparison, it still may be informative to outline group characteristics to illustrate the diversity of learning contexts included in this study. Group characteristics are thus outlined in Table 2.

Table 2: Group Characteristics

Group		# of Subjects	Level of English Oral Proficiency	Curriculum Instruction in English
Immersion Program	Korean-Americans	66	100% LEP upon entering Kindergarten	30% (Kindergarten) to 50% (Grade 2)
	Non-Korean-Americans (EP**)	45	95% EP	
English-only classes	EP	81	100% EP	100 % (all grades)

* LEP: Limited English Proficient
** EP: English-Proficient

Each of the three groups shown in Table 2 had first graders and second graders. The curriculum difference between the immersion classes and English-only classes was the percentage of instruction conducted in English. Specifically, in the

immersion program, 30% (in Kindergarten) to 50% (in Grade 2) of instruction was conducted in English. In contrast, in regular English-only classes, 100% of school instruction was in English for all grades. The following paragraphs will give further information about the subject characteristics of the groups.

Immersion groups

Three elementary schools in the Los Angeles Unified School District (LAUSD) participate in the immersion program, and each school includes one immersion class per grade level. The immersion groups of this study consisted of the three first-grade classes and the three second-grade classes that come from these schools.

The Korean-American students in the immersion program: This group consisted of Korean-American students whose home language is Korean. Upon entering kindergarten, they were identified as *Limited English Proficient* (LEP).[1] This identification was based on the district-administered test of oral proficiency in English, called the pre-Language Assessment System (pre-LAS), which was designed to identify an initial oral proficiency level at the time of entering schools.[2] Although the Korean-American students were identified initially as LEPs, it is noted in Bae (1997) that at the end of the second grade, the English writing skills were on par with their peers in typical English-only classes.

English-dominant students in the immersion program: Non-Korean-American students are from Euro-American, Hispanic, Tagalog, Chinese, or Japanese background. 58% of the students in this group used English-only at home (EO students), 42% of this group used both English and another language (Spanish, Tagalog, Japanese, or Korean) at home with the exception of two students who had only Chinese or Spanish as a home language. All of these non-Korean-American students (except the latter two students, who were LEPs) were classified as English-Proficient (EP) upon entering Kindergarten, according to the LAUSD's language classification criteria (see note 2).

English-only classes

The students from English-only classes came from two schools where the immersion program is operating concurrently. The two schools had middle to upper level status with regard to general academic achievement based on the results of a national test, the Comprehensive Tests of Basic Skills, or CTBS (LAUSD Information Technology Division, 1996).[3] The English-only classes from these schools consisted of two first-grade classes and two second-grade classes. These classes are typical English-only classes in that they receive regular curriculum instruction common to English-medium classes at LAUSD, and, like other English-only classes at LAUSD, these students are English-Proficient students in terms of an English oral proficiency level. In essence, all of the students from English-only classes were English-Proficient students in terms of their English

oral proficiency based on the district's classification criteria (A. Shoji, M. Hicks, & R. Rudnick, personal communication, July, October, 1996).

Study Variables

The theoretical definitions were articulated in the previous section "Construct Definitions." The measurement of these constructs was operationalized into observed variables by implementing a writing task and the scoring criteria to be described below. The observed variables produced in this process are coherence, content, grammar, text length, and the five subconstructs of cohesion, which includes reference, lexical ties, conjunction, ellipsis, and substitution.

Writing Task

Writing prompt

All subjects were given a series of pictures and were instructed to make up and write a story based on the picture series (see the instructions in Appendix A). The picture series is given in Figure 1.

Figure 1*



* Story and illustrations by Jungok Bae and Hyesug Lee.

Genre of the writing test

The writing prompt was intended to elicit narrative, which was considered useful in this assessment for the following reasons. Narrative is a socially and academically valued skill, and children are often called upon to read and tell stories at home and in school to improve reading and writing skill development (Peterson & Dodsworth, 1991). Narration is thus a common experience to children, and the ability to narrate develops early in childhood. In particular, a picture-based narrative task with connected scenes was considered relevant since the scene connection provides a useful means to examine language skills beyond the

sentence level (Ripich & Griffith, 1990). Through the pictures, both the writer and the reader share maximum background knowledge consistent with what is being written, thus minimizing possible comprehension obstacles between the writer and the reader. All in all, the visually connected set of pictures provides a common contextual ground for comparing narrative production across the subjects (Berman & Slobin, 1994, pp. 41-42).

Administrative Procedure

To promote consistent test administration, written and spoken guidelines for delivering instructions for test administration were given to the teachers on the day of the testing at each separate class-by-class test administration. The teacher-delivered instructions were intended to give students their own teacher's language and delivery in a style familiar to them, thus creating a comfortable testing environment. To assure consistent test administration across the classes, the same test coordinator (the author) was present at all class test administrations. The test was given toward the end of the school year, in late May through July, 1996. The time allotted to the actual writing was up to 30 minutes, with most students finishing the story-writing in 20 to 25 minutes, while some students required the maximum time.

Scoring

Coherence, content, and grammar

Dimensions such as coherence and content do not have overt linguistic markers that are countable. Grammar shows overt linguistic markers, but the range of grammatical features is countless. Therefore, a holistic judgment based on the rating scale was made in scoring these three dimensions of language. In other words, the scoring uses a componential or analytic scoring method with a holistic judgment made separately for coherence, content, and grammar within each component. Each dimension was scored independently by two graduate students from the Department of Applied Linguistics and TESL at the University of California, Los Angeles who were native speakers of English. Their scores were averaged to be the score for each individual. For ratings that showed discrepancies with more than one scale point, a third rater assigned a rating: The closest two of the three ratings were averaged to be the score for each individual.[4] All samples were shuffled together before rating, and students were identified only by their identification numbers to prevent raters' possible bias concerning any ethnic group and school grade.

Rating scale

The following generic scale (adapted from Bachman, 1989) formed the basis of the rating scale for scoring coherence, content, and grammar:

0	1	2	3	4
Zero	Limited	Moderate	Extensive	Complete

The characteristics for the five scale points were specified for each component of language. The two ends of the above scale were defined to provide an absolute scale (Bachman, 1989, pp. 251-258; 1990 pp. 340-348) as follows: The one end point (0) represented zero or very little ability; the other end point (4) represented a complete level of the written English language ability for second graders. The second graders were the highest grade used in this study. Thus, the characteristics for determining the scale point of 4 for each of the language components, including benchmark samples, were the ideal level of language use observed in the best second graders' written samples. In addition, ratings with a 0.5 decimal point for each scale point (that is, 0.5, 1.5, 2.5, and 3.5) were incorporated (see Bae, 2000, p. 82, for advantages in allowing decimal points). Expectations relevant to these student levels were also considered in forming the scale descriptions (for example, colloquial expressions, childlike expressions, and errors in spelling were considered acceptable at this level). These characteristics thus formed the criteria. The second graders under consideration consisted of the groups from the immersion program and the English-only classes, which included native and nonnative speakers of English at grades 1 and 2. (See the scoring criteria and selected best sample in the Appendix.) The absolute scale was intended as a common metric scale for all subjects in this study. This common scale is used to make the results of the assessment apply not only to the immersion students but also to the students from English-only classes.

Cohesion

In contrast with coherence, content, and grammar, cohesion has explicit linguistic markers that are countable; thus, counting the number of markers was considered a method that would give a more accurate account of the dimensions of cohesion demonstrated in the writing samples. Thus, appropriately used cohesive markers were counted in each of the following areas of cohesion: reference, lexical ties, conjunction, ellipsis, and substitution (See Table 1 for types of cohesion with examples). Two raters counted the number of these cohesive markers independently, excluding errors, for the randomly selected one-third of the entire samples. Since counting the frequency of something was an objective procedure, compared to judgment against the rating scales, after ensuring the high interrater agreement (over .950, see Table 4), one rater counted the number of cohesive markers for the rest of the samples.

Analysis

Descriptive statistics, interrater reliability estimates, correlations, and multiple regression were calculated on SPSS release 9.0 and SAS release 6.11. Qualitative, descriptive analyses were conducted independently of ratings.

RESULTS

Descriptive Statistics

Table 3 reports descriptive statistics for the scores for each variable. The entire data as a single group (N = 192) was used for this purpose.

Table 3: Descriptive Statistics

	Measurement unit	N	Mean	S.D.	Skewness	Kurtosis
Grammar		192	2.61	1.00	-0.88	0.24
Content	0 to 4 Rating Scale	192	2.48	1.04	-0.71	0.02
Coherence		192	2.20	1.11	-0.28	-0.69
Cohesion						
Reference		192	14.29	9.65	1.68	6.11
Lexical ties	Frequency of Correct Usage	192	24.72	13.48	0.98	2.13
Conjunction		192	5.18	3.88	1.22	2.54
Ellipsis		192	0.31	0.75	3.85	20.35
Substitution		192	0.10	0.34	3.40	11.79
Text length	Number of Words	192	67.51	35.53	1.20	3.81

As shown in Table 3, all variables, except for ellipsis and substitution, showed skewness and kurtosis within or slightly greater than +/- 2, thus showing a normal or approximately normal distribution of scores for these variables. (Skewness greater than +/- 2 indicates that the tail of the distribution is to the right or left compared to a bell curve or a normal distribution; kurtosis indicates that a distribution is either peaked or flat compared to a normal distribution.) The examination of a bar graph for each variable (not provided due to space limitation) confirmed the approximately normal distributions for these variables.

Looking at the means taking all students together, coherence, content, and grammar showed a mean score of around 2.5 on the 0-to-4 scale. The average number of words used in the essays was about 67; the average of number of occurrences of reference markers was 14, and the average number of occurrences of conjunction markers was around 5.

In contrast, ellipsis and substitution were found to have far fewer occurrences in the compositions. The average occurrence in the essays was less than 1 (with a mean of 0.3 for ellipsis and 0.1 for substitution). Ellipsis and substitution showed a deviation from a normal curve as indicated by a very small degree of dispersion of frequencies of occurrence (near zero SD's: 0.75 and 0.34). These

two variables showed a somewhat positively skewed curve (with a skewness of 3.85 and 3.40, respectively), indicating that the frequencies of occurrence clustered around the fewer occurrences. These variables also showed high kurtosis (20.35 and 11.79 each), indicating a sharp, peaked curve around the low frequencies of occurrence. These results indicate that, compared to other components that appeared extensively in the essays, in the vast majority of essays ellipsis and substitution had extremely few occurrences.

Rater Correlation and Reliability

Since two raters independently assigned a rating for each of the components of coherence, content, and grammar, rater agreement and the consistency of ratings are of concern for these components. Rater agreement is represented by the interrater correlation, and consistency (reliability) of two ratings by rater reliability. Thus interrater correlation coefficients and rater reliability coefficients were estimated for the entire data for each variable (See Table 4).

Table 4: Rater Correlations and Rater Reliability

		Interrater correlation		Alpha interrater reliability
		Pearson	Spearman	
Coherence		.907	.904	.951
Grammar		.881	.830	.937
Content		.920	.899	.958
Cohesion:	Reference	.999		.980
	Lexical	.998		.993
	Conjunction	.997		.998
	Ellipsis	.994		.997
	Substitution	1.000 *		1.000*

Note

*The perfect rater reliability indices for substitution were due to the easily noticeable nature of substitution markers (e.g., So he *did*.) and very rare frequencies of substitution (only a total of 5 occurrences from the randomly selected 80 samples), which resulted in the easy agreement of two ratings, and in addition, discussions during the rating session.

As an index of rater agreement, Pearson correlation coefficients were calculated for the components coherence, content, and grammar ($N = 192$). The Pearson r indicates the index of the strength of the linear relationship between two variables (in this case, two ratings independently assigned by two raters). This coeffi-

cient (*r*) is appropriate for interval or continuous data and assumes normality of data. The variables for these three components become continuous data by allowing a decimal point 0.5 on the rating scale (see Bae, 2000, pp. 81-82 for a detailed explanation), and normality is largely met. Therefore, the Pearson *r* can be used as a proper measure of rater agreement for these components. Pearson *r* ranged from .881 to .920. As a supplement, the Spearman's rank *r* is reported. This index provides another measure of the linear relationship between two variables (two ratings) for ordinal or interval data that do not satisfy the normality assumption. Spearman's coefficients ranged from .830 to .904.

As an index of rater reliability, alpha interrater reliability coefficients were calculated. The alpha (α) reliability coefficient reported in this table represents the internal consistency of two scores for each variable, and it is based on the average covariance between the two ratings or scores on the variable. Alpha ranged from .937 to .958 for these components. All coefficients thus indicated a highly acceptable degree of rater agreement and reliability or consistency of two ratings for each component.

In addition, Pearson coefficients and alpha were used as a supplement for the cohesion variables for the randomly selected samples (N = 80). As indicated previously (see Scoring section), counting markers of cohesion was a highly objective procedure compared to judgment against rating criteria. As expected, the indices showed near perfect rater agreements and reliabilities (over .980).

Correlations

Understanding a focal construct in relation to another is essential in construct explication. A statistical index to indicate the relationship between two variables is a correlation coefficient (*r*). Since the present study is interested in the extent to which a particular construct is related to another, correlations will be reported and referenced in the discussions of the constructs in the subsequent sections.

Table 5 reports correlations for the variables of coherence, content, and grammar, the five dimensions of cohesion, and text length. Pearson coefficients were used because the variables were based largely on the continuity of scores. The scores of coherence, content, and grammar were summed up, and will be referred to as the "overall" writing quality throughout the paper.

The Constructs

In this section, the results of analyzing the compositions are provided and discussed for each construct. Text length will be discussed first. Cohesion will be discussed with an elaboration made in the sub-dimensions. Subsequently, coherence, content, and grammar will be discussed within a section on overall writing quality.

Table 5: Correlation Matrix

	Grammar	Content	Coherence	Reference	Lexical Ties	Conjunction	Ellipsis	Substitution	Length
Grammar	--								
Content	.797	--							
Coherence	.797	.900	--						
Reference	.650	.750	.766	--					
Lexical	.637	.769	.767	.901	--				
Conjunction	.364	.472	.413	.484	.534	--			
Ellipsis	.303	.343	.372	.567	.414	.108	--		
Substitution	.155	.225	.245	.327	.343	.273	.099	--	
Length	.650	.762	.760	.927	.937	.636	.455	.388	--
Overall *	.914	.955	.957	.768	.771	.442	.361	.223	.770

Note
N = 192 for all variables. All correlation coefficients above were significant ($p < .001$) except .099 for the ellipsis-substitution relationship.
* Overall = Sum of Grammar, Content, and Coherence.

Text length

The length of a writing sample is defined as the total number of words. Length was considered part of fluency, and its relations to other components of abilities were examined. Raters were asked not to look at length in judging coherence, content, and grammar once an essay exceeded a certain threshold-level length (in this study, approximately 30 words out of a range 7 to 247 words), although an essay that was too short was considered limited in coherence, content, and grammar unless the writing was error-free and persuasive in content. From Table 5, we can see that text length was highly correlated with coherence ($r = .760$), content ($r = .762$), and grammar ($r = .650$). This result supports the notion that in general a fluent writer can write a longer essay within acceptable overall qualities. Text length was also highly correlated with reference ($r = .927$) and lexical ties ($r = .937$)—we will discuss this point in the subsequent section on reference and lexical ties. However, text length showed moderate and small degrees of relationships with ellipsis ($r = .455$) and substitution ($r = .388$), indicating that a longer essay does not necessarily exhibit more occurrences of ellipsis and substitution markers.

Dimensions of cohesion

Let us examine the subdomains of cohesion. The frequencies of occurrence of cohesive markers were counted separately for each domain of cohesion, together with the relative percentages, for the entire sample as a whole (see Table 6).

Table 6: Frequencies of Cohesive Devices (for the Entire Data Set)

Dimensions of Cohesion	*Mean Occurrence*	*Percentage of the Total Occurrences*
Lexical Cohesion	24.7	55.6%
Reference	14.3	31.8%
Conjunction	5.2	11.7%
Ellipsis	0.3	0.6%
Substitution	0.1	0.2%
Total		100%

As Table 6 shows, lexical cohesion was the dominant pattern of cohesion observed in the student narratives (55.6% of total occurrences of all cohesive markers), followed by reference (31.8% of total occurrences) and conjunction (11.7%). Instances of ellipsis and substitution occurred relatively less frequently in the story-writing (each type less than 1%).

In the following sections the subdomains of cohesion are discussed in detail.

Reference and lexical ties

Reference: The types of reference and their relative frequencies observed in the writing samples are given in Table 7. The dominant reference type was pronominal forms (58.8% of total occurrences of references), followed by proper nouns (22.8%). The definite article *the*, demonstratives, and comparative reference occurred relatively less frequently (15.1% to 0.5%). The prominent use of pronouns and proper nouns appears to be due to their role as head nouns, primary information for reference, whereas definite articles and demonstratives are modifiers.

Table 7: Types of Reference and Relative Percentage of Occurrences

Types of Reference	*Percentages*
Pronominals: (*he/she/him/their*)	58.8%
Proper nouns: (*Billy, Jack's*)	22.8%
Definite article: *the*	15.1%
Demonstratives: (*this/these, that/those, here, there*)	2.8%
Comparatives (*bigger, the same, both*)	0.5%
Total	100%

Lexical ties: This study echoes the observation that among areas of cohesion lexical cohesion is the most unexamined area since the studies by Halliday and Hasan (1976/1993, 1989). As such, identifying words that contributed to lexical ties was not as clear-cut as analysis of other areas of cohesion; the same is true with the quantification of lexical ties. Thus, there is still a need for future research to refine theoretical definitions and the domain of lexical cohesion. Nonetheless, in the current study, lexical ties were calculated by counting the number of the words in the essays that belonged to any of the domains of lexical cohesion, such as synonyms, antonyms, collocations, part-whole relations, general-specific lexical ties (see Table 1 for examples), and a stream of main ideas represented by main verbs. Phrasal verbs and idiomatic expressions (e.g., *go home*, *put up*, *come in*) were counted as single lexical items.

Reference and lexical cohesion: As seen in Table 5, use of reference and lexical ties were highly correlated ($r = .901$). The use of reference and lexical ties was moderately or highly correlated with other writing qualities such as grammar ($r = .650, .637$), content ($r = .750, .769$), and coherence ($r = .766, .767$). These high-to-moderate correlations suggest that acquisition of reference markers and vocabulary is critical to extending an essay and enhancing one's overall writing quality.

Another notable observation is that reference and lexically-tied words were highly correlated with length (r = .927 with reference; r = .937 with lexical ties). This correlation is not surprising because the longer a writer composes, the more reference markers and lexical ties the writer will naturally use.

A note is also necessary about the high degree of correlation between these two types of cohesion and length. The greater number of cohesive markers could be explained partly by the longer writing samples. The longer length and the greater number of cohesive markers used could be interpreted as an indicator of fluency. However, caution should be taken not to infer that students who demonstrated a smaller number of cohesive markers in short texts would not show competency in utilizing cohesive markers in their longer writing samples.

Conjunction

For the purpose of this study, the definition of *conjunction* as a form of cohesive tie is confined to coordinating conjunctions between two independent main clauses (e.g., *The boy saw a girl, and she was crying*). Subordinating conjunctions connecting a main clause and a dependent clause (e.g., *I saw a girl when I was walking*) are not considered cohesive conjunctions; thus they were not counted. Halliday & Hasan (1976/1993) introduce four types of conjunctive relations: *additive*, *adversative*, *causal*, and *temporal* (see Halliday & Hasan, pp. 238-273 for detailed examples of the words and phrases that express these meanings). These four types of conjunctions and their frequency of use in the children's writing were examined and are reported in Table 8, taking all writing samples together.

Table 8: Types of Conjunctions and Their Relative Percentages of Occurrences (# of total occurrences: 1141)

Additive		Adversative		Causal		Temporal	
and	23.0%	and	0.5%	and	3.9%	and	30.1%
but	1.1%	but	1.8%	so	12.4%	then	15.5%
				other *	0.3%	and then	7.4%
						adverbial	3.9%
Total (100%)	24.1%		2.3%		16.6%		56.9%

* Ordinarily, subordinating conjunctions are not treated as conjunctions in the form of cohesive ties. However, some students in this study used a subordinating conjunction as a reply to a question in a question-answer sequence: e.g., " Jill went to her, Why are you crying? *Because* my umbrella broked." (from a first grader' s sample). These subordinating conjunctions were counted as conjunctions in the form of cohesive ties because they occurred in an independent sentence, creating a connection across sentences, which represented speaker turns in the narrative.

Nearly 57% of all occurrences of conjunction were temporal. The next most frequent occurrences were additive (24.1%), followed by causal (16.6%) and adversative (2.3%). Another noteworthy feature of conjunction use at these grade levels was the extensive use of *and*, which indicates that the coordinating conjunction *and* is an early acquired conjunction, used for multiple functions:

> He saw a girl and her umbrella was broken. (Additive)
> The girl is getting wet and the boy is not wet. (Adversative/Contrastive)
> My umbrella is torn and I'm soaking wet. (Causal)
> They were drawing and Amy's mother came with food. (Temporal)

The children also often used *and* as an "utterance initial filler to indicate more is to come" (Berman & Slobin, 1994, p. 176) as in:

> Paul was walking . . . And it was raining . . . And Paul saw Tiffany. And Paul said, comein.[5] lets walk home. I'll go with you. OK. And they were walking and walking . . . And then finenuly they went to Tiffany house. And Tiffany said, come in . . . And Paul came in and she was happy . . .

It is noted that the use of conjunctions was only moderately related to the overall writing quality, such as grammar ($r = .364$), content ($r = .472$), and coherence ($r = .413$), which means that more frequent use of coordinating conjunctions does not necessarily contribute to the overall writing quality. This was true because many of the conjunctions used in these children's essays, particularly *and* and *then*, were non-essential elements, although not inappropriate. There were 11 students who used few or no coordinating conjunctions but received a high rating (moderate to complete) in coherence and content. On the other hand, one student connected all sentences with a coordinating conjunction, *and*, producing an entire writing sample of 108 words in one sentence, but still exhibiting high qualities of coherence, content, and grammar. In general, second graders' samples demonstrated a more diverse use of coordinating conjunctions beyond *and* to also include frequent use of *but*, *so*, *then*, *and then* and, less frequently, common temporal adverbials such as *first*, *last*, *after that*, *finally*, and *soon*.

Ellipsis and substitution

Typically, ellipsis is known to occur in responses in spontaneous conversations but is seldom used in formal writing. As such, ellipsis had far fewer occurrences than lexical ties and reference (see descriptive statistics presented in Table 3). However, children frequently introduced dialogues and interactional conversational expressions into their story progression, which can be interpreted as young writers' individual choice of rhetorical styles that enriched the narratives. Use of ellipsis was observed in such dialogues as in "You wanto to draw? Sure, Timmy said." and the writing of a few adept students:

(1) "Thomas and April ate some of the fruit but not all"
(2) "Where are you going Cindy?" "To my house!" . . . So they went inside [She] put [her] backpack away in her room. Even Billy too!"

Like ellipsis, substitution is also a speaker/writer choice and not a compulsory feature (McCarthy. 1991, p. 43). Accordingly, substitution seldom occurred in these writings, except that several students demonstrated an elegant use of substitutions as in:

(3) "Dennis went up to her and said, "Let's share my umbrella." So they did."
(4) "Paul's friend is happy and so is he."

Similar to conjunction, ellipsis and substitution showed weak relationships with overall writing quality (r = .361, ellipsis and overall; r = .223, substitution and overall; see Table 5).

Errors in cohesion

Errors in the use of cohesive ties in each sample were marked by two native speakers of English. Table 9 summarizes the patterns of those errors across the entire data set, with examples provided.

As shown in Table 9, the majority of errors involved problems with reference: unclear references (26.5%) and misuse of *a* or *the* and omission of determiners needed for reference (56.4%). The rest of the errors involved conjunctions and minor grammatical and syntactic errors (e.g., in shes house) at a local level. Despite such errors, the general meaning was inferrable in the text and from the shared contextual schemata provided by the picture series.

A note is necessary about the comparison made between the English-language background and the Korean-language background groups in examining errors in cohesion. Unlike coherence and content, errors in cohesion are easily definable, noticeable, and countable; therefore, subjectivity of cross-group comparison is not a problem with cohesion errors. Compared to the vast range of patterns of errors in grammar, patterns of errors in cohesion are easy to categorize, and the range is not unlimited. Although this paper does not intend to make cross-group comparisons, motivated by its relevance to the understanding of the focal construct, dominant error patterns in cohesion were compared between the two language groups. Dominant error patterns differed depending on the language background. For the English-Proficient (EP) students from both immersion and English-only classes, the dominant error pattern was unclear references: 33% of the immersion EP students' total error occurrences and 25% of total error occurrences made by students in English-only classes were unclear reference. In contrast, the dominant pattern of errors for the Korean-American (KA) students involved misuse of either *a* or *the* (about 59% of KA's total errors) and omission of determiners needed for reference (19% of KA's total errors). This was apparently due to the

Table 9: Errors in Cohesive Markers (Total Number of Occurrences: 204)

Patterns of Errors	*Percentages*	*Examples*
Unclear references	26.5%	Sudden switch of references: e.g., from third person pronouns to a first person pronoun; exophoric, unclear use, and wrong use, which are inferable or incomprehensible.
Misuse of 'the'	24.0%	The Dennis was helping the Helen. The boy went to the her home. Once there was the boy.
Misuse of 'a'	16.7%	Mom brout a food and a fruits. Drew a boats. Drew a ship on a paper.
Omission of determiners that make references (a/the/pronominals)	15.7%	__ Girl waved to __ boy. They went to __ boy's house. The picture was __ ship.
Unnatural use of conjunctions	10.8%	A boy named Eddy had a umbrella at school. But when he was going home he saw a girl.
Other	6.4%	Came over shes house. I shared my's with her. She mom gives...
Total	100%	

transfer from the Korean language, which allows null articles in the places where articles are obligatory in English.

Overall writing quality

Let us turn to the overall writing quality, which includes coherence, content, and grammar.

Coherence

Earlier, we discussed potential factors for establishing coherence: (a) shared background knowledge of the world between the writer/reader, (b) cohesive markers, and (c) the frames underlying genre expectations. Results concerning these factors will be discussed below.

First, the pictures were sufficient for providing writers and readers the shared contextual schemata to facilitate comprehension; thus, the shared background knowledge necessary for a coherent text was well pre-established.

Second, with reference to measuring cohesion as a factor for establishing coherence, a question that can be asked is: How much does the use of cohesive devices account for, or contribute to, coherence? To investigate this question, multiple linear regression analysis was used with coherence as the dependent variable and the five types of cohesion as multiple independent variables or predictors. The entire data as a single group (N = 192) was used for this purpose. Multiple regression is a statistical procedure for analyzing the collective and separate effects of multiple independent variables on the dependent variable. With this method, we can determine which of the independent variables best predicts or accounts for the dependent variable (Hatch & Lazaraton, 1991; Pedhazur, 1982). The results are summarized below.

Table 10: Summary of Multiple Regression: Coherence Accounted for by Cohesion Variables

Significant Predictors	*Simple r with Coherence*	*Cumulative R^2*	*Changes in R^2*	*Beta*	*Significance (p)*
Lexical ties	.766	.588	.588	.406	.000
Reference	.767	.618	.030	.400	.000

Referential and lexical ties were found to be substantively significant (p = .000) predictors of coherence. These two types of cohesion, lexical ties and reference, collectively, accounted for 61.8% of the total variance in coherence (R^2 = .618). On the other hand, the other types of cohesion (conjunction, ellipsis, and substitution) were not significant predictors (p = .944, .517, .555, respectively) and did not enter the final regression equation.

The magnitudes of the effects of the two significant predictors were indicated by the standardized regression coefficient, beta. The beta coefficients indicated that lexical cohesion and reference showed nearly the same magnitudes of strength as a predictor of coherence (beta = .406 for lexical; beta = .400 for reference). At the same time, these two variables showed a large proportion of shared variance in predicting coherence, as indicated by their high correlation and the small change in R^2 in the regression.[6]

Thus, for this writing task, it is concluded that the degrees to which the use of cohesive devices account for, or contribute to, coherence vary from highly significant (lexical and reference) to very little (conjunction, ellipsis, and substitution) depending on the subdomains of cohesion analyzed.

Third, it was pointed out that coherence has to do with a plot-motivated overall structure and the frames underlying genre expectations, such as setting, topic continuity, and a coda in narratives. As such, the following attributes characterized a high rating for coherence in this study: a clear presence of elements for

introductory remarks, a coda (conclusion), and elaborated connections filled in at every important local point; elaborated connections between ideas; a comprehensible and consistent idea stream throughout the essay. As indicators of this, the study calculated the percentages of students who wrote an explicit introduction and a coda in the writing samples and compared the grade performance. The results showed that the vast majority of second graders (90.5%) and the majority of first graders (63.6%) used a clear introduction. These results suggested that they possessed a concept of how typical story books begin. Most of those who had an explicit introduction used either classical opening words such as *once upon a time* and *one day* or some form of temporal or locative adverbials to set up the stage and the time (e.g., *one evening*, *on the street*, *yesterday . . . from school*).

The ability to close the writing with an explicit ending was relatively less developed, suggesting that an awareness of conclusions is a cognitively higher-order or later developed skill than that of awareness of introductions. Only 40.8% of second graders and 14% of first graders drew an explicit conclusion from what was given at the final scene. Examples of the presence of a coda are illustrated below (from second graders):

(5) "... and I thought the girl and I are best friends now."
(6) "... And from that very last day the children would play together Nice things together."

Content

The rating criteria in this study specified that a high rating for content was determined by persuasiveness and creativity within task relevance and thoroughness of the content with respect to the picture prompts. Content was analyzed by examining the essays in terms of the following two categories: (a) mere descriptions of what is visually given in the pictures, and (b) something beyond the visually given content, such as interpretations, evaluations, personal feelings, or imagination that enriched the content within task relevance. Such enriched content in these children's writing is illustrated below from different samples. (The spelling errors in the samples below were not judged in any areas of scoring):

(7) "A girl was walking with a broken umbrella and she was sad so the boy share his umbrella with her they where happy."
(8) "They remembered they had a homework. They took out their homework. It was drawing a boat that rowed around the sea."
(9) "I felt sorry for her so I shared my umbrella with her . . . I asked if she wanted to draw pictures . . . My mom thought we should hang the picture's on the kitchen wall."
(10) "Mother said I like you're activities I'll think I'll paste them on the wall."
(11) "So Jill and Sue had a lought of fun And there best friend."
(12) "Amanda and Eric drank all the milk because they were so hungry. And it was a feast."

In the previous section, the presence of an explicit introduction and ending in the stories was compared by grade level, providing a useful means to examine the development of coherence. In examining the development of content, analysis by grade level was also useful. For each grade level, a percentage of students who went beyond mere descriptions of what is visually given in the pictures (that is, category [2] above) was calculated. 34.3% of the first-grade compositions and 72.7% of the second-grade compositions belong to this category. Thus second graders demonstrated a far greater degree of ability to use diverse expressions, make inferences, and make relevant associations about what was not explicitly provided in the pictures than did the first graders. The development of content, like that of coherence, seems to be related to the cognitive maturational development in young children.

Content and coherence: Whether content and coherence are actually qualitatively the same ability traits is a separate question. However, content and coherence showed an extremely high linear relationship with each other ($r = .900$, see Table 5). The two constructs seem to be very closely associated with each other.

Grammar

Relationship with coherence/content: As seen in Table 5, grammar showed a high correlation with coherence (.797) and content (.797). Thus grammar could be treated as a global writing quality factor, together with coherence and content. Viewing grammar as a global writing quality is reasonable because without adequate competency in grammar it is unlikely that learners can produce writing with quality and text length reasonable enough to communicate ideas. Grammar errors were defined as critical errors or minor errors depending on the seriousness of the effect on reader understanding (See "Construct Definitions"). Writing samples with critical errors received low ratings. Minor errors, on the other hand, were considered tolerable enough to allow the writer to get moderate to high ratings.

Relationship with cohesion: Among areas of cohesion, grammar was most highly correlated with reference and lexical ties ($r = .650$ and .637 respectively) and it was moderately or very weakly associated with conjunction ($r = .364$), ellipsis ($r = .303$), and substitution ($r = .155$) (See Table 5).

CONCLUSIONS

In this study, cohesion, coherence, content, and grammar in English writing samples composed by 192 first or second graders were investigated within the context of immersion and English-medium classes. Following are the conclusions to the research questions concerning (a) the characteristics of cohesion and coherence and the interrelations among cohesion, coherence, content, grammar, and text length, (b) prominent types of cohesion, and (c) errors in cohesion.

Characteristics of Cohesion and Coherence and the Interrelations among Cohesion, Coherence, Content, Grammar, and Text Length

As shown in Table 5, the correlations showed that coherence and content were highly related (r = .900). Grammar showed a high correlation with coherence (.797) and content (.797). Among subdomains of cohesion, the use of reference and the use of lexical ties were highly correlated with each other, and they both were highly related with the length of a text. The ratings of coherence, content, and grammar were summed up, and were referred to as the "overall" writing quality. Referential and lexical cohesion showed relatively high correlations with the overall writing quality (r = .768, .771, respectively). However, ellipsis and substitution showed relatively weak correlations with the overall quality (r = .361, .223, respectively). This supports the idea that the subdomains of cohesion, while they play a distinct role within cohesion, are more local-level links than the global quality represented by coherence, content, and grammar.

Multiple regression analysis examined the effects of the cohesion variables on coherence (See Table 10). The results showed that cohesion differed in degree as a contributor to coherence depending on the subdomains of cohesion. Lexical cohesion and referential cohesion were found to be significant (p = .000) predictors of coherence. The two variables showed almost the same magnitudes of the effect on coherence (beta = .406, .400, respectively). They both, collectively, explained approximately 61.8% of coherence, indicating the importance of the acquisition of reference markers and vocabulary as a factor to establish coherence. On the other hand, other types of cohesion (conjunction, ellipsis, substitution) were not significant predictors of coherence.

Awareness of an introduction in narrative writing was well-developed for these grades (grades 1 and 2); however, an awareness of conclusions was less developed at this stage.

Prominent Types of Cohesion

The most prominent types of cohesion observed in the narratives across the children in this study were lexical and referential ties (respectively 56% and 32% of total occurrences of all cohesive markers; see Table 6). Coordinating conjunction, ellipsis, and substitution occurred less frequently in the written narratives. This result, together with the interpretations made above, suggests that reference and lexical ties are more crucial and necessary while the other types of cohesive markers can be present or absent depending on writer/speaker choice. Meanwhile, dominant reference types (see Table 7) were pronominal forms (about 59% of total occurrences of reference) and proper nouns (about 23%). Prominent types of conjunctive relations (see Table 8) were temporal (nearly 57% of all occurrences of coordinating conjunctions) and additive (24%).

Errors in Cohesion

The dominant patterns of developmental errors in cohesion involved inaccurate reference across the grades and the language background groups: unclear

references and misuse or omission of *a/the* and determiners that make referential ties. The most prominent errors for English Proficient (EP) students (that is, English-dominant students) in both immersion and regular English-only classes were unclear references. However, a dominant error pattern for the Korean-American students in the immersion program involved misuse or omission of *a* and *the*, reflecting transfer from their first language, Korean, which typically does not use articles. All of these errors, in general, did not hinder comprehension of meaning in the context of the texts and in the presence of shared schemata provided by the pictures.

This study conducted a theoretically informed empirical investigation of cohesion, coherence, content, grammar, and text length in children's written narratives in English within the contexts of a two-way immersion program and English-medium classes. The data collected from the performance-based story-writing task via group testing and the descriptive and quantitative data analyses shed light on the global interrelationships among these constructs and their dominant characteristics. The substantiated findings provide evidence for the characteristics of these constructs that would not be provided by purely theoretical write-ups and typical case study analyses. At the same time, the constructs established in this study may well provide a basis for subsequent studies (see below). Instruments comprising the writing prompt and the scoring method can be useful in applications to other contexts and students with different linguistic backgrounds.

Limitations of the Study

The schools, classes, and teachers in this study serve as convenient samples; they were not randomly selected. Even when schools and classes are randomly selected, the number of schools and classes is often so small that we cannot say random sampling gives a representative group (Bentler, 1997; Hatch & Lazaraton, 1991). At the same time, the subjects in this study are elementary school children, and the linguistic backgrounds of these students consist mostly of English, Korean, and Spanish languages. The data based on several classes in this study provides a much stronger basis for generalizing findings than data based on case studies that typically involve several students. Considering the above-mentioned limitations, however, caution should still be taken in generalizing the results of the present study.

Implications for Future Research

Several implications and suggestions for future research are made with respect to potential contributions of this study. The first suggestion involves a more sophisticated analysis with latent variables (as opposed to ordinary observed variables). A latent variable modeling approach is used to design and test models of relationships that include latent variables, free of error of measurement. With this approach, more sophisticated relationships can be specified and tested as they may be illustrated as follows.

1. *Constructs used as latent variables*: The variables cohesion, coherence, content, grammar, and text length used in the present study are observed variables, which contain inherent errors in measurement. One of the advantages of a latent variable approach is its capability to decompose observed variables into latent variables free of error of measurement. Inferences from data thus become more refined at the level of latent variables (also called *factors*), controlling for measurement error. Thus, in a future study, cohesion and coherence could be investigated at the level of latent variables.

2. *Directionality*: The correlations used in the present paper give fundamental information about the degrees of relationships between variables, but a correlation does not indicate causal relationship or explanation.[7] With a latent variable modeling approach, the directionality of influence among variables, if not causal relationships, can be specified. To use this approach fruitfully, however, the researcher should have a well-developed a priori theory to test. For instance, an excellent topic where directionality of influence can be specified and tested with this approach would be the contributions of subdomains of cohesion to coherence. Another example would be to use groups (background characteristics) as predictors of the constructs (see below).

3. *Factorial evidence for construct distinctiveness*: A latent variable modeling approach is well suited to perform hypothesis testing, a typical format of construct validation. One important topic with respect to construct validation that is not addressed by the present paper is to test for psychometric evidence for the distinctiveness of the specified constructs. Inquiries that can be addressed in this direction would include questions, such as: Are cohesion and coherence (and the subconstructs of cohesion) statistically found to be distinguishable factors, and are they found to be distinguishable from grammar and content at the psychometric level? A future study could investigate this topic using confirmatory methods that allow hypothesis testing within a latent variable modeling approach. To my knowledge, no studies have provided psychometric evidence of their distinctiveness (see, however, Bae, 2000 for the evidence for distinctiveness of grammar, content, spelling, and text length).

4. *Group comparisons*: Another important topic is the comparison of writing performance on these constructs between immersion and English-medium classes. Group comparisons inform program evaluation in achieving goals with respect to the achievement of students' English writing skills. Group comparison is also valuable because group characteristics used as predictors contribute to our better understanding of the constructs as attainable skills that are influenced (or not influenced) by group characteristics. For this purpose, an ordinary mean analysis or analysis with latent means (which refers to the means at the level of latent variables) can be used.

Second, the present study illuminated the global relation of several selected constructs. A future study could conduct an in-depth analysis of the relationships of these constructs that have only been alluded to here. For instance, an excellent

area for theoretical and empirical research could be (a) a formulation of a more concrete relationship between content and coherence and (b) lexical cohesion, the most uninvestigated area of cohesion.

Third, a future replication may be conducted for students with different language backgrounds and grade levels.

Finally, another contribution of this study is the excellent utility of the writing test instruments: the prompt and the scoring criteria. Researchers can utilize these methods for eliciting and quantifying children's productive language data (written and spoken) to address a variety of topics.

As a final comment, substantively, the theoretically informed empirical investigations carried out in this study make a significant contribution to our understanding of the salient linguistic and semantic resources for text construction, namely cohesion, coherence, content, grammar, and text length. The substance can provide a foundation for conducting subsequent studies such as the ones suggested above. The statistical data yielded by this study and the writing test instruments can make a valuable methodological contribution as resources for undertaking future studies.

ACKNOWLEDGMENTS

This research was conducted partly under the auspices of the UCLA Language Resource Program under the directorship of Dr. Russell N. Campbell. I thank Professors Marianne Celce-Murcia, Lyle F. Bachman, and Russell N. Campbell for their invaluable comments and feedback during the development of this study and on its early drafts. I would like to express my gratitude to colleagues Robert Agajeenian and Kathryn Howard for their valuable participation as raters and to Hyesug Lee for her participation in the picture illustrations. I thank the following people for their assistance with the study: Dr. Jamal Abedi, Adeline Shoji, Maureen Hicks, Ruby-Ann Rudnick, Chin Kim, and the teachers in the Korean/English Two-Way Immersion Program and also in the English-only classes. Finally, sincere thanks go to three anonymous reviewers and the *IAL* Editors for their extremely helpful suggestions and rigorous review of the earlier versions of the paper. Any remaining shortcomings are my own responsibility.

APPENDIX A

GUIDELINES FOR TEACHERS FOR DELIVERING INSTRUCTIONS FOR THE ENGLISH WRITING TEST

Please read the basic instructions below and deliver them using your everyday language style and vocabulary familiar to your students. You can add relevant words to these basic instructions flexibly, but keep the basic content given below. Please allow about 10 minutes for these instructions.

Procedures for the Instructions

1. Warm-up:
"Many people would like to know how well you write a story in English..."

2. Example story-writing (Optional):
(Show a two- or three- picture series that has an example story written below it):
"Let's look at the example story that somebody wrote. (Read the story to the students.) This is only one way of writing . . . You could write differently, e.g., . . ."

3. Actual story-writing task:
"You will NOT write about the example pictures that we just read. You will have NEW pictures to write about. Look at the new pictures." (Please have the students look at the story line depicted in the 7-picture series for a minute or two. To make sure that the story content is not ambiguous to them, and to activate schemata (background information) for the students, please go over the whole story line with the students, by starting, e.g., "Let's see what this story is about . . ." Make sure, however, to say the following after you went over the general story: Please do not copy what I just said about these pictures. You MUST write YOUR OWN story; e.g., you may not use the name I used for the boy . . ."

Also, please instruct them about the following:

* Create your own story/ Write a CREATIVE story, but the story MUST go with the pictures.
* Your story should be in sequence.
* Write as long as you can, and as best as you can.
* Wrong spelling is fine. When you cannot spell a word correctly, you may sound out. Just do your best. Do not pay attention to punctuation.
* I will give you 30 minutes to finish your writing. It is a good idea to think about what you will write before you begin writing your story on the paper.

APPENDIX B
SCORING CRITERIA

Criteria for Rating English Writing Samples for Early Elementary School Graders (K to 2)

These criteria were developed by Jungok Bae, Kathryn Howard, and Robert Agajeenian.

---

* These criteria are primarily based on narrative writing: The scale descriptions may vary and are flexible depending on the specific writing tasks and contexts that are changeable.
* Spelling errors, punctuation, and handwriting will not be judged.
* Length: Length will not crucially affect scores in general. Beyond a certain threshold level of length (e.g., once a sample has more than N words), length alone will not be considered a factor to high or low ratings. That is, other qualities will be considered more important than length. However, length is part of fluency, so an essay with words < N will belong to the category of Zero or Limited ability unless the essay is perfectly fine and sophisticated and with no errors. (N will vary depending on the specific task used.)
* The following absolute scale of ability will be used (Adapted from Bachman, 1989). To provide an absolute scale, the two ends of the scale and the scale points in between are defined as follows for Grades K through 2, Grade 2 being the highest grade in this context:

0_____1_____2_____3_____4

0: Zero or very little ability
1: Limited
2: Moderate
3: Extensive
4: Complete: Ideal level of language ability and use for second graders; Characterized by the writing features observed among best second graders' writing samples available and expectations relevant to these student levels.

Ratings with a 0.5 decimal point for each scale point (that is, 0.5, 1.5, 2.5, and 3.5) will be acceptable.

* For further guidelines for scoring, see section "Scoring" in the main body of the paper.

COHERENCE:
Definitions: See section "Construct Definitions" in this paper.

0 (Zero):	Too short to judge. No evidence of coherence. Totally incomprehensible regardless of the length.
1 (Limited):	Seriously unconnected/isolated series of ideas. Serious lack of relationships between ideas.
2 (Moderate):	Some connections of separate ideas but no global connections of local ideas. There may be some major connection missing in between.
3 (Extensive):	The whole story organized in general. No serious break. All ideas pretty much connected globally. However, sophistication and elaboration for connections not observed.

4 (Complete): Sophisticated and elaborated connection of ideas. Absolutely comprehensible thematically. Ideas absolutely consistent. Clear presence of elements for introductory/opening remarks (e.g., *One day*; *Once upon a time*; *Yesterday*), coda, and elaborated connections filled in at every important local point.

CONTENT:
Considerations will be given to relevance, thoroughness, persuasiveness, and creativity.

0 (Zero): Too short to judge.

1 (Limited): Not thorough at all (Only 15 - 30 % of the content was expressed).
Serious distortion of the picture content.
Large segments of the content missing.

2 (Moderate): Somewhat relevant but not thorough.
Some minor irrelevance/inaccuracy.

3 (Extensive): The story is complete and thorough in general.
Accurate/relevant in general. In general, FINE, but elaboration and sophistication not observed.

4 (Complete): Descriptions of the situations/events just wonderful. Very thorough.
No irrelevance whatsoever. CREATIVE. Persuasive. Convincing.

GRAMMAR:
Grammar refers to morphemes and syntax. Critical errors are defined as errors that seriously impede communication: e.g., a major syntactic chunk missing, incomprehensible word order. Minor errors are defined as errors that do not cause ambiguity in meaning, misunderstanding, and difficulties of communication: e.g., usually errors in morphemes such as third person present suffixes, tense at local level, and plural suffixes.

0 (Zero): Too short to judge.
No evidence of grammatical knowledge/use.
No sentences; only single words.

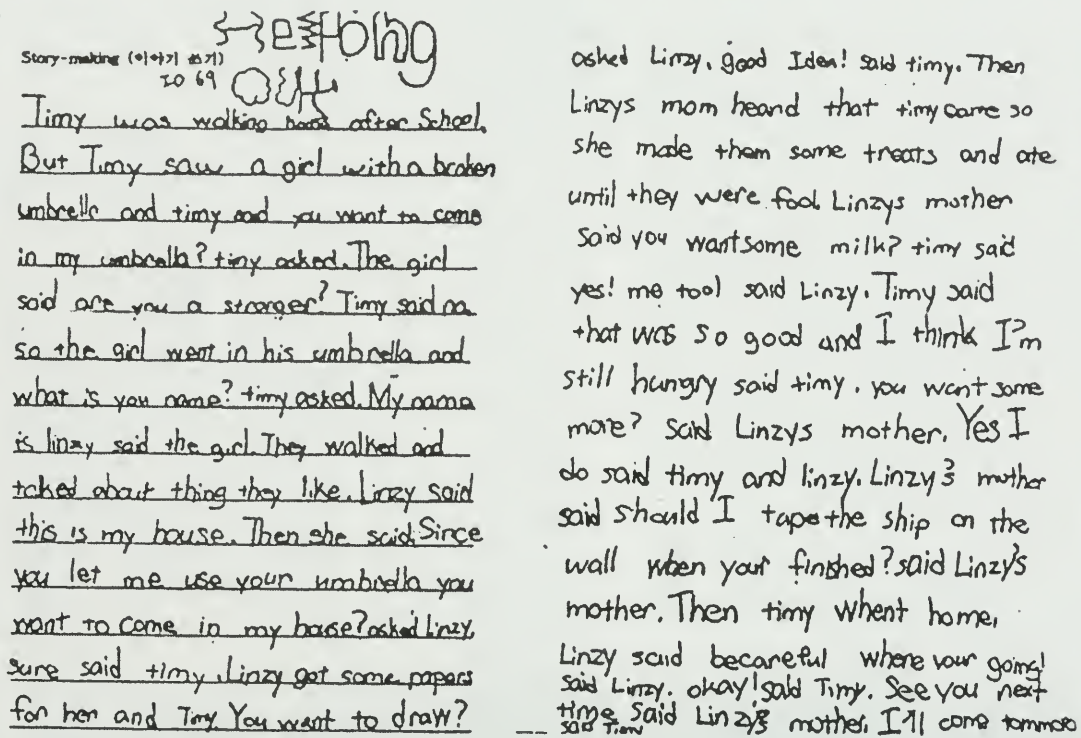
1 (Limited): Frequent critical errors. Extensive minor errors.
Few sentences; only phrases.
A sample with length < N words is considered Limited unless the writing contains complex grammatical features.

2 (Moderate): Some critical errors. Frequent minor errors.

3 (Extensive): Few limitations, no critical errors, occasional minor errors, with no complex sentences.

4 (Complete): Unlimited range. Complex sentences.
A variety of grammatical uses.
Complete control of grammar (Native level). Very few errors.

COHESION:
Definitions and classifications: See section "Construct Definitions" in this paper. Cohesion will be scored based on the number of cohesive markers that are appropriately used.



Story-making (이야기 쓰기)
10 69

Timy was walking hoom after School. But Timy saw a girl with a broken umbrella and timy and you want to come in my umbrella? tiny asked. The girl said are you a stranger? Timy said no. so the girl went in his umbrella and what is you name? timy asked. My name is linzy said the girl. They walked and talked about thing they like. Linzy said this is my house. Then she said Since you let me use your umbrella you want to come in my house? asked Linzy, sure said timy. Linzy get some papers for her and Tiny You want to draw? asked Linzy. Good Idea! said timy. Then Linzys mom heard that timy came so she made them some treats and ate until they were fool. Linzys mother said you want some milk? timy said yes! me too! said Linzy. Timy said that was so good and I think I'm still hungry said timy. you want some more? said Linzys mother. Yes I do said timy and linzy. Linzy's mother said should I tape the ship on the wall when your finished? said Linzy's mother. Then timy whent home. Linzy said becareful where your going! said Linzy. okay! said Timy. See you next time said timy. Said Linzy's mother. I'll come tommoo

NOTES

[1] Six Korean-American students who were initially identified as English-Proficient were classified into the English-Proficient group in this study to be consistent with the District's classification of students.

[2] The English oral proficiency levels based on the pre-LAS test are classified as follows: Non, Limited, Functional, and Proficient. Non and Limited are called *Limited English Proficient (LEP)*, and Functional and Proficient *English Proficient (EP)*. Students whose home language is exclusively English are called *English-Only (EO)*; they are also classified as EPs, exempted from the English proficiency identification test.

[3] In School A, the school's first and second grade students scored at around the middle percentiles on the CTBS/U subsections against the national norm for the most recent 1994-95 school year (the percentile scores in reading, math, and language were 38, 61, and 48, respectively, for first graders, and 54, 66, and 59 for second graders). All first graders and second graders from this group were identified as English-Proficient, and they were all English-Only students. School B scored in the upper percentiles on the CTBS/U subtests for the most recent two academic years 1993-95 (the percentile scores in reading, math, and language were 65, 89, 70 [1994-95], respectively, and 58, 83, 70 [1993-94] for first graders, and 53, 91, 63 [1994-95] and 85, 95, 83 [1993-94] for second graders). All students from this group, as well, were identified as English-Proficient (EP) on the basis of the pre-LAS (pre-Language Assessment System). However, this group included both English-Only students, whose home language is exclusively English, and non-English-Only students who were proficient bilinguals.

According to the school officials, EP students who are non-English-Only students use English and another language at home; they are proficient bilinguals and treated as EOs in terms of English oral proficiency level. On the other hand, the term English-Only (EO) refers to the students whose home language is exclusively English; thus, EOs are proficient in English but not necessarily in another language (A. Shoji, M. Hicks, & R. Rudnick, personal communication, July, October, 1996).

[4] For the cases that had a third rating, rater correlation coefficients and reliability coefficients (Table 4) were obtained using the closest two ratings. Six cases, five cases, and four cases for grammar, coherence, and content, respectively, showed discrepancies with more than one scale point difference, and these cases received a third rating.

[5] Spelling errors, punctuation, and handwriting were not judged in scoring and have been reproduced here as written by the children.

[6] Due to the high correlation between referential ties and lexical cohesion, some multicollinearity (that is, a high correlation between predictors in using regression) was indicated (an index for tolerance was .187 for both lexical cohesion and reference. To be free from multicollinearity, a minimum tolerance .28 is needed based on the criterion $1 - R^2$, with $r = .85$). However, omitting one of these predictors or collapsing them, which is a typical way to handle a multicollinearity problem in regression, was not applied to this case. Instead, these two variables entered as separate predictors in the regression model because they both concurrently appear naturally in the texts and it is clear that they possess their own attributes of cohesion.

[7] If X and Y are related, there are several possible explanations: (a) X may cause Y, (b) Y may cause X, and (c) X and Y may be the result of a common cause (Gronlund & Linn, 1990, p. 499). This type of causal explanation is not provided by a correlation coefficient.

REFERENCES

Arter, J. A., & Spandel, V. (1992). Using portfolios of student work in instruction and assessment. *Educational Measurement: Issues and Practice, Spring*, 36-44.

Bachman, L. F. (1989). The development and use of criterion-referenced tests of language ability in language program evaluation. In K. Johnson (Ed.), *The second language curriculum*, (pp. 242-58). Cambridge: Cambridge University Press.

Bae, J. (1997). Cohesion, coherence, content, and grammar in children's written English: The Korean/English Two-Way Immersion Program and single-language (English-only) classes. Unpublished manuscript, University of California, Los Angeles.

Bae, J. (2000). The construct validation of certain components of English and Korean writing ability in children participating in either a two-way immersion program or monolingual classes: A writing assessment and latent variable approach. Unpublished doctoral dissertation, University of California, Los Angeles.

Baker, E. L., O'Neil, H. F., Linn, R. L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist, 48*, 1210-1218.

Bamberg, B. (1984). Assessing coherence: A reanalysis of essays written for the national assessment of educational progress, 1969-1979. *Research in the Teaching of English, 18*, 305-319.

Bentler, P. M. (1997). Modeling of intervention effects. *Substance Use & Misuse, 32* (12 & 13), 1661-1666.

Berman, L., & Slobin, D. I. (1994). *Relating events in narrative: A crosslinguistic developmental study*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc., Publishers.

Brown, G., & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge University Press.

Campbell, R., Kim, C., Kim, O., Merrill, C., Rolstad, K., & Bae, J. (1994). *Korean/English Bilingual Two-way Immersion Program: Title VII evaluation report* (Federal Grant No.: T003C20062, 1994-1995), Washington, D.C.: U.S. Department of Education.

Carrell, P. L. (1982). Cohesion is not coherence. *TESOL Quarterly, 16*, 479-488.

Collins Cobuild English language dictionary. (1996). New York: HarperCollins.

Comprehensive Tests of Basic Skills. (1990). Monterey: CTB, Macmillan/McGraw-Hill School Publishing Company (Available by written permission to: Curriculum Frameworks Assessment, CTB Macmillan/McGraw-Hill, 2500 Garden Road, Monterey, CA, 93940).

Connor, U., & Johns, A.M. (Eds.) (1990). *Coherence in writing: Research and pedagogical perspectives*. Alexandria, VA: Teachers of English to Speakers of Other Languages.

Cook, G. (1989). *Discourse*. Oxford: Oxford University Press.

Cox, B. E., Shanahan, T., & Sulzby, E. (1990). Good and poor elementary readers' use of cohesion in writing. *Research Quarterly, 15*, 47-65.

Cox, B. E., Shanahan, T., & Tinzmann, M. B. (1991). Children's knowledge of organization, cohesion, and voice in written exposition. *Research in the Teaching of English, 25*, 179-212.

Enkvist, N. E. (1990). Seven problems in the study of coherence and interpretability. In U. Connor & A. M. Johns (Eds.), *Coherence in writing: Research and pedagogical perspectives* (pp. 9-28). Alexandria, VA: Teachers of English to Speakers of Other Languages.

Fitzgerald J., & Spiegel, D. L. (1986). Textual cohesion and coherence in children's writing. *Research in the Teaching of English, 20*, 263-280.

Fitzgerald J., & Spiegel, D. L. (1990). Children's writing. *Research in the Teaching of English, 20*, 263-80.

Freedle, R. O. (Ed.). (1991). *Text and texture: Patterns of cohesion*. Norwood, NJ: Ablex Publishing Co.

Gronlund, N. E., & Linn, R. L. (1990). *Measurement and evaluation in teaching*. (6th ed.). New York: Macmillan.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69*, 274-290.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. London: Edward Arnold.

Halliday, M. A. K., & Hasan, R. (1976/1993). *Cohesion in English*. London and NY: Longman.

Halliday, M. A. K., & Hasan, R. (1989). *Language, context, and text: Aspect of language in a socio-semiotic perspective*. Oxford: Oxford University Press.

Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Boston: Heinle & Heinle.

Herman, J., Aschbacher, P., & Winters, L. (1992). *A practical guide to alternative assessment*. Los Angeles: Association for Supervision and Curriculum Development, UC Regent.

Kim, C. H. (1996, August). T*he Korean/English Two-Way Immersion Program*. paper presented at the Two-Way Bilingual Immersion Summer Conference, Long Beach, CA.

Koshik, I. (1999). A preliminary investigation into the effect of grammatical cohesive devices: Their absence and their misuse on the comprehension of non-native speaker speech and writing. *Issues in Applied Linguistics, 10*, 3-26.

Los Angeles Unified School District Information Technology Division, Testing Unit. (1996). *CTBS test results*. Los Angeles: LAUSD Information Technology Division.

Lindsay, D. B. (1984). *Cohesion in the composition of ESL and English students*. Unpublished Master's thesis. University of California, Los Angeles.

Linn, R., Baker, E., & Dunbar, S. (1991). Complex performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*, 15-21.

McCarthy, M. (1991). *Discourse analysis for language teachers*. Cambridge: Cambridge University Press.

Mehrens, W. A. (1992). Using performance assessment for accountability purposes. *Educational Measurement: Issues and Practice, Spring*, 3-9.

Norris, J. A., & Bruning, R. H. (1988). Cohesion in the narratives of good and poor readers. *Journal of Speech and Hearing Disorders, 53*, 416-424.

Nunnally, J. (1978). *Psychometric theory* (2nd Ed). New York: McGraw-Hill.

Oller, J. W., & Jonz, J. (1994). *Cloze and coherence*. Cranbury, NJ: Associated University Presses.

Pedhazur, E. (1982). *Multiple regression in behavioral research: Explanation and prediction* (2nd Ed.). Orlando: Holt, Rinehart and Winston.

Peterson, C., & Dodsworth, P. (1991). A longitudinal analysis of young children's cohesion and noun specification in narratives. *Journal of Child Language, 18*, 397-415.

Renkema, J. (1993). *Discourse studies*. Amsterdam: John Benjamins.

Ricento, T. (1987). Clausal ellipsis in multi-party conversation in English. *Journal of Pragmatics, 11*, 751-775.

Ripich, D. N., & Griffith, P. L. (1990). Narrative abilities of children with learning disabilities and nondisabled children: Story structure, cohesion, and propositions. *Journal of Learning Disabilities, 21*, 165-73.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking in conversation. *Language, 50*, 696-735.

Schegloff, E. A. (1990). On the organization of sequences as a source of "coherence" in talk-in-interaction. In B. Dorval (Ed.), *Conversational organization and its development*. Norwood, NJ: Ablex.

Smith, R. (1984). Paragraphing for coherence: Writing as implied dialogue. *College English, 46*, 8-21.

Spiegel, D. L., & Fitzgerald, J. (1990). Textual cohesion and coherence in children's writing revisited. *Research in the Teaching of English, 24*, 48-66

Walker, J. F. (1992). Exploring and describing the planning process of a potential Korean/English two-way immersion program with the Los Angeles Unified School District. Unpublished Master's thesis. University of California, Los Angeles.

Jungok Bae recently received her doctorate in applied linguistics at UCLA. Her research interests include language assessment and language acquisition and related applied linguistics topics and research methods.