

Introducing Writing Prompts: A Resource for Teaching, Evaluating, and Researching Children's Writing

Introduction

Children's writing provides a promising area of research because writing (as opposed to reading, listening, and speaking) yields texts through which observations of language and thoughts are easily assessable and enduring, so that texts composed by children can reveal interesting developmental features of writing acquisition. A myriad of research topics can be addressed utilizing these advantages.

Scarcity of Data from Children's Writing

The wealth of children's writing, however, has been underutilized for a long time. Research in writing and second language acquisition in general has mainly been focused on adult learners and students in late adolescence (Braine, 2001; Campbell, 1990; Dorothy, 2000; Fahy & Morgan, 1999; Friedlander, 1990; Hamp-Lyons, 1991; Hamp-Lyons & Condon, 1993; Kazemek, 1999; Kroll, 1990; Peyton, 2000). Consequently, there is a dearth of studies that use writing data collected from children. In addition, research that examines children's language has used mostly spoken data (Corsetti, 1996; Moran, 1993; Nicholas, 2000; Stow & Pert, 1998; Verhoeven, 1989; Yaruss & Conture, 1995). Hence, the value of children's written data for providing valuable resources for research has largely been hidden, and research in children's writing is only in its preliminary stages.

Methods for Writing Assessment

Crucial to carrying out research with children's writing has been the issue of where people can find good methods for eliciting and evaluating writing samples from children. We have noted the need to disseminate prompts to elicit language samples from children. For instance, the electronic Forum for Discussion of Research on Bilingualism and Bilingual Education (BILING) has had many postings from people who have had difficulty finding prompts to collect language data from bilingual speakers. Many ask how to get a copy of Mayer's (1969) "frog story."¹

At the same time, good methods for assessing children's writing are in great demand. A few decades ago, educational measurement began to see the drawbacks of multiple-choice testing. Aware of the influence of testing on

curriculum, educators, evaluators, and policy makers have turned to performance-based assessment as a tool for educational reform (Arter & Spandel, 1992; Aschbacher, 1991; Baker, O'Neil, & Linn, 1993; Linn, Baker, & Dunbar, 1991). One prominent area that has called for such performance-based assessment is writing assessment (Cumming, 1998; Hughes, 1989; Mehrens, 1992). In performance-based writing tasks, students are required to demonstrate that they can write extended stretches of discourse in an open-ended setting, while drawing on prior knowledge and experience, recent learning, and relevant skills to accomplish the task (Baker, Freeman, & Clayton, 1991; Cumming, 1997; Herman, Aschbacher, & Winters, 1992; Kane, Crooks, & Cohen, 1999). Implementing these types of tasks is instrumental in improving the construct validity of writing assessment and the quality of writing skills instruction.

Rationale for Picture-Based Narrative Writing Tasks for Children

Within the framework of performance assessment, this article presents a method of writing assessment that utilizes a picture-based narrative task, which is considered useful and appropriate for assessing children's writing skills for several reasons. Generally, the ability to construct a narrative is a socially and academically valued skill, and children are often asked to tell and read stories to improve literacy skill development (Peterson & Dods-worth, 1991). Narration is thus a common practice to children, and the capacity to narrate develops early in childhood.

To facilitate a narrative task, a set of pictures with connected scenes is considered especially useful because the scene connection provides a convenient means to examine language skills beyond the sentence (Ripich & Griffith, 1990). In addition, graphic illustrations provide the participants with a common schema. Because they share the same contextual medium, the writer and the reader bring into play similar background knowledge required for performing the task, thus reducing possible comprehension barriers between the writer and the reader. Furthermore, since the basic content is provided visually, students can concentrate on the presentation of language (Celce-Murcia, 1996) without having to create content from scratch. This advantage is especially useful for children, who have not yet developed enough cognitive skill to engage in writing based on abstract, nonvisual, verbal prompts or topics. It is clear from these points that picture-based narrative writing tasks provide the best means for observing and assessing children's writing.

Purpose of the Paper

As discussed previously, performance-based narrative writing tasks using a series of pictures can provide an excellent means for eliciting and evaluating writing by children. The purpose of this article is to share several pictorial prompts developed for performance-based narrative tasks that require elementary school students to produce narrative discourse in writing. Guidelines are provided to help apply the method. This provision will facilitate productive research using children's compositions as a rich database. We

wish to make clear that our purpose with this paper is not, however, to analyze written data produced by using the instrument.

Pictorial Prompts

The pictorial writing prompts discussed in this article were developed and used in the context of the Korean/English Two-Way Immersion Program (Campbell, C. Kim, O. Kim, Merrill, Rolstad, & Bae, 1994), in which students learn curricular content in both languages. The prompts were given to students in Grades 1 and 4 to elicit and measure their writing in English and Korean. However, these prompts can be used flexibly for different grade levels and languages. Generally, these prompts can be appropriate for the following settings: (a) Grades: K-5, (b) Classes: Bilingual or single-language classes, (c) Testing: Groups or individual students.

The prompts presented below can easily be used to measure writing skills in one language in monolingual-class settings. These prompts can also be used in bilingual classes to measure writing skills in two languages (see example: Bae, 2000). The instrument is appropriate not only to study individual students' texts but also to conduct group testing based on a consistent procedure. For a comparison of writing performance across individuals and groups or over time, it is important that the same test and the same test administration and scoring procedures apply across all subjects, groups, and times to ensure equal measurement conditions across groups and over time.²

Writing Prompts

Below are six series of pictorial prompts that we developed to elicit narrative.

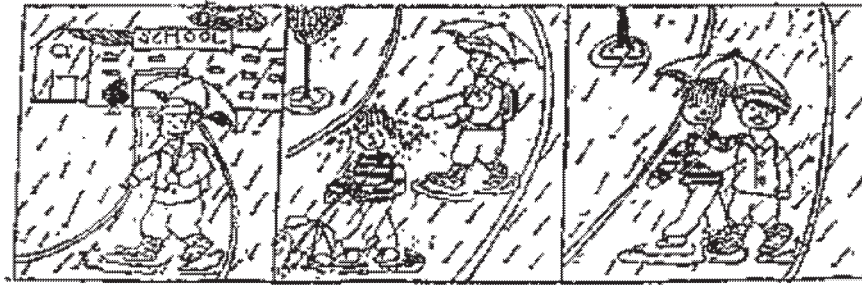


Figure 1. Shorter prompt: *A Rainy Day*. Story and illustrations by Hyesug Lee.

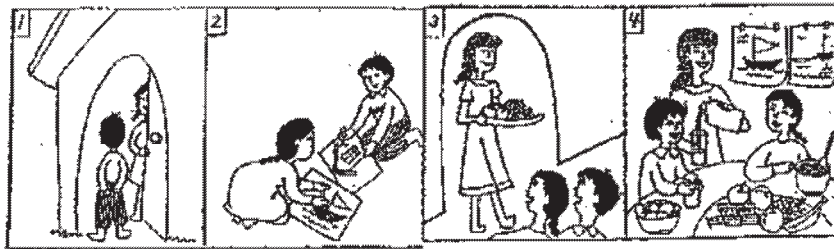


Figure 2. Shorter prompt: *Friends at Home*. Story and illustrations by Jungok Bae.

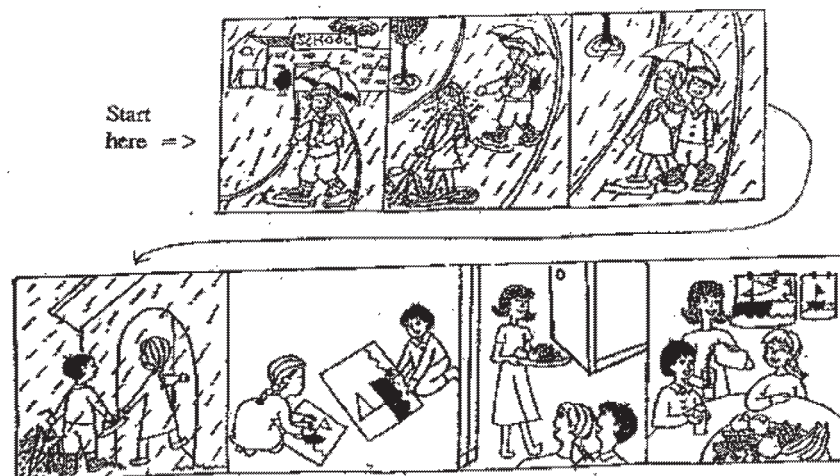


Figure 3. Longer prompt (Two shorter prompts combined). Story and illustrations by Jungok Bae and Hyesug Lee.

Figure 3 reprinted with permission from *Issues in Applied Linguistics*. Source article: Bae, J. (2001). Cohesion and coherence in children's written English: Immersion and English-only classes. *Issues in Applied Linguistics*, 12, 51-88.

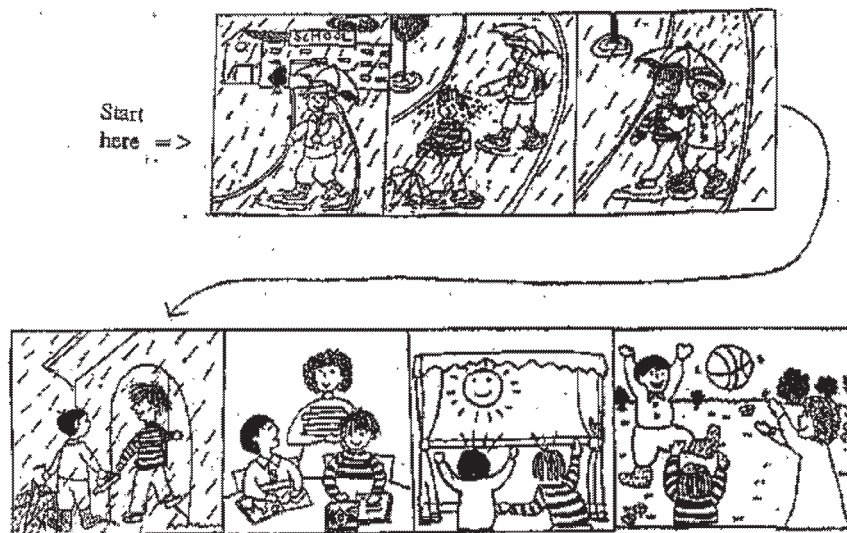


Figure 4. Longer prompt (parallel form of Figure 3). Prompt jointly produced by Hyesug Lee, Minah Lee, and Jungok Bae.

Form 2



Figure 5. Longer prompt. Story and illustrations by Jungok Bae.

Form 1



Figure 6. Longer prompt. Story and illustrations by Hyesug Lee.

Figures 5 and 6 reprinted with permission from 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.

Figures 1 and 2 are picture prompts for short-story writing; each prompt produces a separate story. Figure 3 combines the two prompts, serving as another separate, but longer, prompt. Figure 4 can serve as a parallel or alternate form of Figure 3. Figures 5 and 6 provide another set of parallel forms of prompts for longer stories. Longer prompts are useful for producing a more extended discourse. Depending on the needs and purposes, users can adopt one or more of these picture prompts.

Background of Picture Development

A brief episodic background behind the development of these picture series is given below.

Figures 1 and 2 (shorter series). The shorter prompts (Figures 1 and 2) were first developed in 1995 to assess the Korean writing skills of children enrolled in the aforementioned immersion program. These prompts were used to measure discourse qualities of grammar, content, and cohesion emergent in story writing in Korean.

Figures 3 and 4 (longer series). In the following year, a need arose to test these students' English writing skills to conduct a project for a graduate course, TESL/AL250: Seminar in English Cohesion, taken at UCLA by one of the authors. For this project, a longer prompt was sought to obtain extended writing samples. It was determined from informal discussions with classmates that combining the two shorter prompts would result in a story that would be long enough. Thus, the two prompts were combined into the longer prompts (Figure 3), which was used for the students in this program. In the next year of implementation, it was noticed that students remembered the picture strip when it was reused. Thus, to prevent the influence of familiarity with the pictures and the practice effect on students' compositions in the subsequent years, a prompt that would keep the basic format and method but would contain different content with the same degree of complexity was used. A teacher came up with a new idea for the second half of the story content as illustrated in Figure 4. Thus, Figure 4 was used as an alternate form of the other longer prompt (Figure 3).

Figures 5 and 6 (longer series). Figures 5 and 6 provide another pair of longer prompts. This set was used in Bae's (2000) dissertation and is reprinted in this paper. Compared to the other pairs of prompts presented in this article, these two forms were developed with an explicit a priori design and the purpose of making them parallel from the very beginning of the picture development. The content in the two forms is parallel in complexity and type of activities. One of the purposes of Bae's dissertation was to compare the degrees of Korean and English writing performance as developed by the bilingual immersion students. To achieve this purpose, the two parallel picture forms were used in combination with a counterbalanced design with students in the immersion classes, who were required to produce a story in one language followed by the other. This design preempted problems with students' possible translation tendency and the effect of picture practice when the students composed in one language followed by the other. The evidence that the two prompts function as parallel forms is provided in Bae (2000). It is not the

part of the scope of this article to provide details of these administrative design and statistical procedures.

Instructions

This section presents instructions for using the prompts. The instructions below can apply across the different prompts (Figures 1 to 6).

1. Cue the students to “look at the pictures” and follow the general story line.
2. To ensure that the story content is not ambiguous and to activate schemata, go over the whole story line with the students.
3. Begin by stating, for example, “Let’s see what this story is about.”³
4. Instruct students to “Please not repeat what I just said about these pictures. You must create your own story. For example, you may not use the name I used for the girl.”
5. In addition, incorporate the following statements in the instructions. These statements may be modified according to the needs of a given assignment:
 - “Feel free to make up your own story as long as your story matches the pictures. Be creative.”
 - “The events in your story should be told in order.”
 - “Write as long as you can for the time given.”
 - “I will give you *N* minutes to finish your writing. It’s a good idea to think about what you will write before you begin writing on the paper.”
 - “Incorrect spelling is fine. When you cannot spell a word correctly, you may sound it out. Do not worry about punctuation.” (Use this statement when accuracy of spelling and mechanics is not the object of measurement.)⁴

Instructions should be given in the student’s first language. When there are two target language groups in the same class, the instructions should be given in both languages; for example, in English then in Korean.

Oral instructions should be given for students in lower elementary school grades, such as kindergarten and first grade. Orally delivered instructions help make children attentive to the instructions and provide effective guidelines with sufficient details. Interactions occur naturally with spoken instructions; thus, it becomes immediately clear to the teacher/tester whether the students understand the instructions. Most importantly, oral instructions are essential because students at these grade levels usually lack the ability to comprehend written instructions. For students in the second grade or higher, concise written instructions may be given with supplementary oral instructions.

Time Allotment and Text Samples

The following table gives recommended time allotments for the actual writing, excluding the time for instructions for these prompts. We note that children often have difficulty sitting for a long time. A longer time allotment

is unworkable because they easily lose concentration after twenty or thirty minutes.

Table 1
Time Allotment and Expected Text Length

<i>Picture prompts</i>	<i>Writing time</i>	<i>Text length</i>
1 and 2 (shorter prompts)	Up to 5 – 10 minutes	½ to 1 page
3 and 4 (longer prompts)	Up to 20 – 30 minutes	1 to 2 pages
5 and 6 (longer prompts)	Up to 20 – 30 minutes	1 to 2 pages

Note. Flexibility may be used.

Appendix A contains a sample composition selected from among actual essays. The sample gives an idea of the kinds of compositions the prompts can generate and how the stories can be assessed.

Scoring Criteria

Scoring criteria are important to a writing assessment method. Criteria for scoring based on a picture-based narrative writing task are available in Bae, Agajeenian, Han, Howard, & Lee (1999) and Bae (2001). The salient writing components discussed in Bae et al. (1999) consist of content, grammar, spelling, and text length while the writing components in Bae (2001) are coherence, content, and grammar. These criteria contain 1) refined definitions of the writing components being measured, 2) descriptors for each scale point, and 3) guidelines for scoring. We reproduce the criteria appearing in Bae et al. in this paper (see Appendix B). These criteria were designed to apply across English and Korean writing and to be applicable, with appropriate adaptations, to different task types and to both bilingual and monolingual students in different contexts and with different target languages.

Conclusions and Suggestions

This paper has introduced the writing test instruments. The written data that these prompts yield can provide excellent resources for both quantitative and qualitative analysis of children's essays. Statistical evidence that would confirm the validity of these instruments is very supportive, provided that they are used with appropriate students and reliable test administration and if adequate rater selection and training are implemented. One interesting topic would be to further test whether the two short prompts (Figures 1 and 2) and the two longer prompts (Figures 3 and 4) are empirically found to be parallel. Positive statistical evidence and supportive content analysis would confirm their equivalence. (See Bae, 2000, for evidence that Figures 5 and 6 are parallel).

We hope the writing test instruments introduced in this paper are helpful to those interested in writing pedagogy, assessment, and research using children's compositions. Users can apply these assessment tools with appropriate adaptations to bilingual and monolingual students to elicit and assess writing in any language.

Authors

Jungok Bae recently received her doctorate in applied linguistics from UCLA with a specialty in language assessment. Her research interests include language testing, bilingual immersion education, computer programming, and research methods.

Hyesug Lee is currently working as a high school English teacher in Seoul, Korea. She is interested in developing teaching and testing materials for the Korean high school context. She has written test items for elementary school students learning two languages in the US and for Korean high school English reading references.

Acknowledgments

The picture prompts and the scoring criteria appearing in this paper were developed with support from the UCLA Language Resource Program during the mid-1990s. Thanks are due to Dr. Russell N. Campbell, Dr. Lyle F. Bachman, Ms. Minah Lee, Ms. Chin Kim, and Ms. Shelley Kwock for their support and contribution during the development, and to Catherine G. Gunn, Debra Friedman, and three anonymous reviewers for their helpful comments on early versions of the paper.

Endnotes

¹ Mayer's frog story (1969) is a well-known children's picture story that is often used by researchers collecting data on narrative skills.

² This condition is referred to as *measurement equivalence*. For the topic of measurement equivalence, see, for example, Cole & Maxwell (1985), Byrne (1989), Bae & Bachman (1998), and Bae (2000).

³ Further story content is not provided to allow flexibility.

⁴ Spelling accuracy may not be the object of measurement for lower level elementary school students such as kindergarteners and first graders. At this grade level, creative spelling and acquisition of the correspondence between written form and meaning of new words are encouraged rather than accuracy of spelling. That is, errors in spelling are considered acceptable at this level. However, for students in higher school grades such as second, third, and fourth graders, spelling with accuracy is clearly consid-

ered an important objective in teaching and testing. (See Appendix B for criteria for scoring spelling.)

References

- Aschbacher, P.R. (1991). Performance assessment: State of activity, interest, and concerns. *Applied Measurement in Education*, 4, 275-288.
- Arter, J.A., & Spandel, V. (1992). Using portfolios of student work in instruction and assessment. *Educational Measurement: Issues and Practice*, Spring 1992, 36-44.
- 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.
- Bae, J. (2001). Cohesion and coherence in children's written English: Immersion and English-only classes. *Issues in Applied Linguistics*, 12, 51-88.
- Bae, J., Agajeenian, R., Han, N., Howard, K., & Lee, N. (1999). The criteria for scoring some salient writing components for elementary school students (K to Grade 4). *Language Testing Update*, 26, 72-82.
- Bae, J., & Bachman, L.F. (1998). A latent variable approach to listening and reading: testing factorial invariance across two groups of children in the Korean/English two-way immersion program. *Language Testing*, 15, 380-414.
- Baker, E.L., Freeman, M., & Clayton, S. (1991). Cognitive assessment of history for large-scale testing. In M.C. Wittrock & E.L. Baker (Eds.), *Testing and cognition*. Englewood, NJ: Prentice Hall.
- Baker, E.L., O'Neil, H.F., & Linn, R.L. (1993). Policy and validity prospects for performance-based assessment. *American Psychologist*, 48, 1210-1218.
- BILING. Forum for Discussion of Research on Bilingualism and Bilingual Education. Internet discussion group. Archives of BILING@ASU.EDU available online: <http://lists.asu.edu/archives/biling.html>.
- Braine, G. (2001). When an exit test fails. *System*, 29, 221-234.
- Byrne, B.M. (1989). *A primer of LISREL: Basic applications and programming for confirmatory factor analysis*. New York: Springer-Verlag.
- Campbell, C. (1990). Writing with others' words: Using background reading text in academic compositions. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 211-230). Cambridge, England: Cambridge University Press.
- Campbell, R.N., 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). Washington DC: US Department of Education.
- Celce-Murcia, M. (1996). Notes from course lecture. TESL/AL 250: Seminar in Cohesion Analysis, Spring 1996, University of California, Los Angeles.

- Cole, D.A., & Maxwell, S.E. (1985). Multitrait-multimethod comparisons across populations: A confirmatory factor analysis approach. *Multivariate Behavioral Research*, 20, 389-417.
- Corsetti, R. (1996). A mother tongue spoken mainly by fathers. *Language Problems and Language Planning*, 20, 263-273.
- Cumming, A. (1997). The testing of writing in a second language. In C.M. Clapham & D. Corson (Eds.), *Language testing and assessment. Encyclopedia of language and education* (pp. 51-63). Dordrecht, Holland: Kluwer Academic Publishers.
- Cumming, A. (1998). Theoretical perspectives on writing. *Annual Review of Applied Linguistics*, 18, 61-78. Cambridge, England: Cambridge University Press.
- Dorothy, L. (2000). Mixed metaphors for reading and writing the qualitative thesis in adult education. *Studies in the Education of Adults*, 32, 148-65.
- Fahy, P.J., & Morgan, D. (1999). Current writing attitudes and practices in selected adult literacy programs in western Canada. *Alberta Journal of Educational Research*, 45, 209-211.
- Friedlander, A. (1990). Composing in English: effects of a first language on writing in English as a second language. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 109-125). Cambridge, England: Cambridge University Press.
- Hamp-Lyons, L. (Ed.), (1991). *Assessing second language writing in academic contexts*. Norwood, NJ: Ablex.
- Hamp-Lyons, L., & Condon, W. (1993). Questioning assumptions about portfolio-based assessment. *College Composition and Communication*, 44, 176-190.
- Herman, J., Aschbacher, P., & Winters, L. (1992). *A practical guide to alternative assessment*. Los Angeles: Association for Supervision and Curriculum Development, UC Regents.
- Hughes, A. (1989). *Testing for language teachers*. Cambridge, England: Cambridge University Press.
- Kane, M., Crooks, T., & Cohen, A. (1999). Validating measures of performance. *Educational Measurement: Issues and Practice*, 18, 5-17.
- Kazemek, F.E. (1999). A gathering of individuals: A longitudinal study of a writing workshop for older adults. *Adult Basic Education*, 9, 3-20.
- Kroll, B. (1990). What does time buy? ESL student performance on home versus class compositions. In B. Kroll (Ed.), *Second language writing: Research insights for the classroom* (pp. 140-154). Cambridge, England: Cambridge University Press.
- Linn, R., Baker, E., & Dunbar, S. (1991). Complex performance-based assessment: Expectations and validation criteria. *Educational Researcher*, 20, 15-21.
- Mayer, M. (1969). *Frog, where are you?* New York: Dial Books.
- Mehrens, W.A. (1992). Using performance assessment for accountability purposes. *Educational Measurement: Issues and Practice*, Spring 1992, 3-9.

- Moran, M.J. (1993). Final consonant deletion in African American children speaking black English: A closer look. *Language, Speech, and Hearing Services in Schools, 24*, 161-166.
- Nicholas, J.G. (2000). Age differences in the use of informative/heuristic communicative functions in young children with and without hearing loss who are learning spoken language. *Journal of Speech, Language, and Hearing Research, 43*, 397-415.
- Peyton, J. (2000). *Dialogue journals: Interactive writing to develop language and literacy*. Washington, DC: Office of Vocational and Adult Education. (ERIC Document Reproduction Service No. ED450614)
- 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.
- 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.
- Stow, C., & Pert, S. (1998). The development of a bilingual phonology assessment. *International Journal of Language & Communication Disorders, 33*(Suppl.), 338-342.
- Verhoeven, L.T. (1989). Language monitoring in children's second language speech. *Second Language Research, 5*, 141-155.
- Yaruss, J.S., & Conture, E.G. (1995). Mother and child speaking rates and utterance lengths in adjacent fluent utterances: Preliminary observations. *Journal of Fluency Disorders, 20*, 257-278.

Appendix A
Sample Composition
 (based on the prompt shown in Figure 3)

Story-writing (1991) 227
 16 124

One Monday morning a boy
 name Thomas was going home
 from school. He had a umbrella
 While he was walking he
 saw a girl that broke her
 umbrella. The girl name was
 April. April asked the
 Thomas if she could have
 Thomas umbrella. Thomas
 said yes. So Thomas walk
 April home. When they
 got there April asked
 Thomas if he could come
 in. Thomas said yes. April

said to Thomas close up your
 umbrella before you go in the
 house. said April, Thomas
 said I'll keep my umbrella
 outside by the door. So
 Thomas put he umbrella
 by the door and came in
 April's house. When Thomas
 came in the house April
 and Thomas drawed pictures.
 April drew a small sailboat
 and Thomas drew a big sail-
 boat. When They where
 finish Mrs. brown brought

April and Thomas a big tray with
 fruit in it. Thomas and April ate
 some of the fruit but not all. When
 they were done they drank milk. Mrs.
 brown drank milk to.

Appendix B

The Criteria for Scoring Four Salient Writing Components for Elementary School Students (K to Grade 4)

Introduction

The criteria defined below were developed to assess four salient components of writing observed in stories and letters composed by elementary school students. The assessment context for the criteria was the Korean/English Dual Language Program established at the Los Angeles Unified School District in 1992. The criteria as presented below are the results of a rigorous, iterative rating process by several raters who participated in the writing assessments for the elementary school students of this program.

The following scale was used to assess the writing quality of the elementary students. This scale is a common scale, so that it applies to both bilingual and monolingual students and to all specified grades irrespective of regions. The grade levels used in this particular assessment includes Kindergarten through Grade 4, Grade 4 being the highest grade in this assessment. The scale is an absolute scale in the sense that the ability levels (especially the ends of the ability scale) are defined and anchored to specified, not arbitrary, levels of ability: the ends are zero (no ability) and the perfect level that reflects the writing features observed in the best students' writing. The ends of the scale and the scale points in-between are defined as follows:

0 ____ . ____ 1 ____ . ____ 2 ____ . ____ 3 ____ . ____ 4

- 0: Zero or no ability
- 1: Limited
- 2: Somewhat insufficient
- 3: Good but not excellent
- 4: Perfect (for 4th grade level)

A score of 4 represents the ideal level of language ability and use for 4th graders; it is characterized by the writing features observed among the best 4th graders' writing samples available and evaluators'/teachers' expectations relevant to the students at the 4th grade level. The term "perfect" used for scale point 4 may sound like a point hard to reach for most students compared to alternative terms such as excellent. However, this choice is to ensure that the one end of the scale provides an absolute scale of ability for the subjects who participated in the study (see Introduction). In addition, defining the one end to a perfect state with everything else in between increases the discriminating power of the scale.

In addition, the following characteristics and guidelines apply to the scale introduced above:

- Ratings with a 0.5 decimal point for each scale point (i.e., 0.5, 1.5, 2.5, 3.5) are acceptable.

- The scale is generic, so that it can be applied to writing in any language with appropriate adaptations.
- The scale is applied to the following writing tasks commonly and separately: letter-writing (20 minutes) and story-writing (30 minutes). See chapter 3 in Bae (2000) for tasks and administrative design. The scale can be used for similar types of writing tasks with appropriate adaptations.
- For all components of writing, handwriting and punctuation will NOT be assessed.
- In general the length of writing may not be a crucial factor in scoring other construct. Beyond a certain threshold level of length (e.g., once an essay exceeds N words; N will vary depending on the specific task used), length alone is not considered a factor in high or low ratings. Other qualities are considered more important than length. However, length is part of fluency, so language production with words $<N$ is likely to belong to the category of Zero or Limited ability unless the short writing produced contains no errors.

The general guidelines for scoring introduced above have been adapted from Appendix B appearing in Bae (2000, p. 211). In the following section, the specific criteria for scoring four selected components of writing are presented. For this purpose, part of Bae, Agajeenian, Han, Howard, and Lee (1999, pp. 75- 82) is reprinted with permission from *Language Testing Update*. The reprint below is adapted from the source article for the purpose of condensation.

Criteria for Scoring Four Components of Writing

In the following sections, we provide specific criteria for four selected constructs: content, grammar, spelling, and text length. For each component, the criteria consist of a combination of construct definitions, a rationale for construct selection, general guidelines, and the descriptors for each scale point.

We hope these criteria will be helpful to those interested in assessing the writing ability of children. These criteria are applicable with appropriate adaptations to different task types and bilingual and monolingual students in different settings and different languages.

Content

Content is defined as the relevance of a written text to a given task, as well as thoroughness, persuasiveness, and creativeness consistent with the task. The quality of content is viewed as the degree to which the writing impresses the reader in terms of the above criteria.

Content is a global property, and an evaluation of content is determined on a macro level, holistically. Therefore, in this assessment, several other attributes are considered subsumed under content, e.g., coherence, lexical choices, organization, cohesion.

In relation to the particular tasks used in this study, the content score is based on whether the responses are: (a) relevant to the picture content (if the content goes beyond what is depicted in the pictures, consider the writer's imagination appropriate as long as the imagination is relevant to picture content and task expectations); (b) suited to the assignment/task given: e.g., a letter should be addressed to the sick friend who is absent; see instructions for the letter task (see Bae, 2000); (c) thorough, persuasive, impressive, and creative.

Content descriptors for scale points:

- 0 No ability: No production/content.
- 1 Limited: Not thorough at all (Only 15-30 % of the content was expressed). Serious distortion of the picture content/task, or large segments of the content missing.
- 2 Somewhat Insufficient: Somewhat insufficient in content. Somewhat irrelevant/inaccurate/not thorough as a whole or locally. Or a couple of sentences per major scene with mere (literal) descriptions.
- 3 Good but not Excellent: The story/letter is complete and thorough in general. Accurate/relevant in general. In general, fine, but elaboration and sophistication not observed. Descriptions good (literal) but not impressive. Or, descriptions somewhat insufficient; however, some impressive, relevant elaboration observed locally.
- 4 Perfect for fourth-grade level: Wonderful descriptions of the situations and/or events. Very thorough. No irrelevance whatsoever. Creative. Persuasive. Convincing. Impressive.

Grammar

Grammar refers to morphology and syntax. Grammar best represents the *linguistic* domain. The measurement of grammar is based on the range of grammatical features and the degree/number of grammatical errors that appear in the text. We classify grammatical errors into critical errors and minor errors to operationalize the measurement of grammar: *Critical errors* are defined as errors that seriously impede communication: e.g., a major syntactic chunk is missing or word order is incomprehensible. *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 singular present suffixes in English, tense at local level, plural suffixes, articles, and run-on sentences.

Grammar descriptors for scale points:

- 0 Zero: No production.
- 0.5 No sentences, but only (a list of) single words, phrases, or morphemes.
- 1 Limited: Frequent critical errors. Extensive minor errors. Few sentences. A sample with length $<N$ words is considered Limited unless the writing contains complex grammatical features.

- 1.5 Has aspects of 1, but one occurrence of complex or compound sentence structure.
- 2 Somewhat Insufficient: Some critical errors. Frequent minor errors. +0.5 if the writing sample is long.
- 3 Good but not Excellent: No or few critical errors. Occasional (1-2) minor errors with a few occurrences of complex or compound sentence patterns.
- 3.5 N (7) complex/compound connection observed but some critical errors, or less than N (7) complex/compound sentences but no errors.
- 4 Perfect for 4th grade level: Complete control of grammar (Native speaker level for Grade 4). A variety of grammatical use. A total of N (7) clauses that demonstrate either complex sentence connection (with subordinating conjunctions) or compound connection (with coordinating conjunctions) with or without 1-2 minor errors.

The $N=7$ for an essay (above) is to help determine the score 4; in the cases of 7 or 9, be flexible. Be flexible when determining the number of N , depending on the characteristics of samples arising from the nature of the writing task, e.g., general time allotment and text length of the task.

Guidelines for determining N
(number of occurrences of either complex or compound connections)

A complex sentence which has a main clause and a subordinate clause with a subordinating conjunction (although, if, when, because) is counted as one occurrence: e.g. I went to the store because I needed some milk. A compound sentence which has two independent/main clauses connected with a coordinating conjunction (and, but, or, so, for) is also counted as one occurrence: e.g. I went to the store and bought some milk. Each additional subordinate or coordinate clause within a sentence is counted as an additional occurrence.

Practice. Complex sentences, compound sentences, and simple sentences: How many occurrences of either *complex* or *compounding* (i.e. subordinating or coordinate) connections are observed in the following sentences? For consistency of measurement across all compositions, apply the guidelines above to determine the number (N):

- a. Jim was studying in the classroom, and his friends were playing soccer outside. ($N= 1$ occurrence)
- b. I went to the elephant, gave it peanuts, and touched the trunk. ($N=2$)
- c. I went up to the lamb and gave it some grass, and my friend began to feed the hens. ($N= 2$)
- d. A class was walking to the park, singing merrily. ($N= 1$)
- e. I saw two squirrels sitting* on the tree. ($N = 0$) (Note: S + V + O + O.C*: A participle used as an object complement will not be treated as a subordinating or coordinating connection.)
- f. While they were walking, they saw little squirrels, which were climbing the trees. ($N= 2$)
- g. Before I went to bed, I wrote what I did in the park. ($N= 2$)

Spelling

Writing is also an act of *encoding*. A unique feature that distinguishes writing from other modes of language is that writing involves production in a graphic mode that requires control of orthography, or spelling. Furthermore, one clearly important domain of literacy learning for children is spelling. In lower-grade classes such as Kindergarten, creative (rather than correct) spelling is encouraged; however, in higher grades, correct spelling is encouraged and tested. Therefore, spelling receives special attention and is selected as a construct both unique to writing and appropriate for the measurement of composition by elementary school students.

Spelling is defined as the ability to spell individual letters of a word correctly in terms of form and order (Collins Cobuild English Language Dictionary, 1996; M. Celce-Murcia, 1999, personal communication). A level of spelling ability is determined by the recognizability (or legibility, comprehensibility) of letters in words and the pattern and number of errors in spelling.

Like grammar, in operationalizing the measurement of spelling, it is useful to categorize spelling errors as critical and minor ones. *Critical errors* in spelling are defined as incorrect letters in words that are completely incomprehensible or those comprehensible but only with great effort of the reader to understand the meaning. *Minor errors* are defined as incorrect letters in words that are easily comprehensible (e.g. words spelled as pronounced, homonyms, or homophones: e.g. *our/lare* teacher; *picnic/pig neck*; *we where* going to a zoo).

- Disregard (Scores not to be affected by the following): Handwriting, punctuation, grammar (morphemes and syntax), meaning/content in a global context, dialect, informal words/expressions.
- Estimate the number of spelling errors per approximately one-page composition (= about 100 spacing units for Korean; about *N* words for English) as an index of spelling error. The number of errors specified in the scale descriptions (below) refers to these per-page errors.
- Code-mixed words that appear locally (e.g., more than four English words in Korean writing) should not be counted as errors, but a -0.5 will be subtracted from the score assigned without considering the code-mixed words.
- A repeated error observed in the same word is counted only once in the error assessment. e.g., *recieve...recieve...recieve*; number of errors=1.
- Depending on legibility and comprehensibility of words, the scale is divided into two areas to help determine a score as follows:

Scores:

- *Below 2*: Comprehension impeded owing to frequent minor errors with or without serious errors. Number of spelling errors is often

hard to identify owing to the lack of legibility; therefore, number of errors may not be an important factor in determining a score.

- *2 or above*: Legible and comprehensible with or without minor errors. Number of errors plays an important role in assigning a score.

Spelling descriptors for scale points. Number of errors below represents a degree of errors estimated per page (see above).

- 0 (No production): Nothing written; nothing to judge. Or too little to judge.
- 0.5 (A list or a repetition of) alphabets only, or a list of completely incomprehensible words, regardless of the length of the writing sample.
- 1 (Limited): Over 50% of what was written is incomprehensible.
- 1.5 Words are pretty comprehensible/legible if the reader makes an effort to comprehend the incorrectly written words.
- 2 (Somewhat insufficient): Words legible with frequent minor errors (e.g. over 15 minor errors) with or without occasional critical errors.
- 2.5 Words legible with 8-15 minor errors.
- 3 (Good but not excellent): Words completely legible with 5-7 minor errors.
- 3.5 Words completely legible with 3-4 minor errors.
- 4 (Ideal for fourth-grade level): Words completely legible with no errors or 1-2 minor errors.

Text Length

How much is written often receives attention by the writer and the reader. Text length is thus selected as an attribute of writing. Text length is defined as follows: In English, it is the number of words in the English texts; in Korean, the number of spacing units in the Korean texts.

The total number of words/spacing units is converted to a score on the scale used for scoring the other constructs, as below. Alternatively, a raw score (# of words/spacing units) can be used. Please be flexible when using the range definitions depending on the characteristics of the writing task.

Number of words/spacing units for scale points

- 0 = No words written.
- .5 = 1-25 words/spacing units
- 1 = 26-50
- 1.5 = 51 -75
- 2 = 76-100
- 2.5 = 101-125
- 3 = 126-150
- 3.5 = 151-175
- 4 = over 175 words/spacing units