Library of Congress Cataloging in Publication Data
Language and Literacy in Bilingual Children/Edited by D. Kimbrough Oller and Rebecca E. Eilers
Child Language and Child Development: 2
Includes bibliographical references and index.
1. Bilingualism in children. 2. Children–Language. I. Oller, D. Kimbrough. II. Eilers, Rebecca E. III. Series
P115.2.L35 2002
404'.2'083–dc21 2001044538

British Library Cataloguing in Publication Data
A catalogue entry for this book is available from the British Library.

ISBN 1-85359-571-3 (hbk)
ISBN 1-85359-570-5 (pbk)

Multilingual Matters Ltd
UK: Frankfurt Lodge, Clevedon Hall, Victoria Road, Clevedon BS21 7HH.
USA: UTP, 2250 Military Road, Tonawanda, NY 14150, USA.
Canada: UTP, 5201 Dufferin Street, North York, Ontario M3H 5T8, Canada.
Australia: Footprint Books, Unit 4/92a Mona Vale Road, Mona Vale, NSW 2103, Australia.

Typeset by Archetype-IT Ltd (http://www.archetype-it.com).
Printed and bound in Great Britain by the Cromwell Press Ltd.

Contents

Chapter 7

Narrative Competence among Monolingual and Bilingual School Children in Miami

BARBARA ZURER PEARSON

The narrative syntax project was designed to evaluate language and literacy development with a single task encompassing both domains. For this purpose, the children were asked to create a story, an extended discourse, which could be evaluated as a whole, but also as the sum of its parts. Thus, the children could demonstrate their strengths and weaknesses at several different levels at once.

Indeed, much of the language development in the ages between 5 and 10 is thought to take place at the level above the individual sentence (Karmiloff-Smith, 1986). That is, children by age 5 have demonstrated the use of the major syntactic structures of their language, but continue to expand the range and complexity of the uses to which those structures are put (Chomsky, 1969; Slobin, 1973). Perhaps the most important linguistic development in the early school period, then, is children's growing ability to create extended texts. They move from the local level of organizing words within a sentence to a global level, where they must organize sentences into a coherent, hierarchical discourse.

This development coincides at the start of school with learning to read. The child, through reading, is increasingly involved with texts longer than the typical conversational turn, and by the end of the elementary school period is expected to be able to write paragraph-length expositions and stories (Hunt, 1977). The sentences in those paragraphs individually contribute information to the whole, but they must also serve to direct the flow of information smoothly and cohesively across sentences. The linguistic devices which accomplish this task for the reader/writer are not usually new forms, but are recruited from the forms already at the child's disposal (Berman & Slobin, 1994) – notably the pronouns, determiners, adverbs, and

conjunctions that can express the temporal, causal, and anaphoric connections between individual phrases.

As attested by the growth in the length of school texts in all subjects as children move up through the grades, the ability to work with longer and longer passages is a key element in academic success. To understand long texts, children must be able to interpret both the 'content' words and the 'function' words. Content words, roughly the nouns and verbs, provide 'information' on topics – 'steel', 'primate', 'evaporates'. Function words convey the relationships between content words, for example the prepositions 'to', 'without', or 'by' – something. Function words show the relations between propositions about the content words. They also signal which words give new information, say with an indefinite article or a full noun phrase, and which words must be interpreted anaphorically, that is with respect to what has come before in the passage. So, knowledge of function words and the grammatical structures they play a role in can be as crucial to understanding and creating texts as the content words that define the subject matter.

Children's ability to interpret these linguistic devices may be tested in reading through standard passage comprehension tasks (cf. Woodcock, 1991). In these tasks, some items are cloze-type questions which probe for a missing function word, thereby sending the child backward or forward in the passage to find the elements that constrain the options for that word. Even items which focus on a missing content word may require the reader to evaluate the relationships established by the conjunctions and prepositions present: ['Therefore, without _____, farmers must rely on wells.'] So typical passage comprehension scores are at least in part measures of children's ability to follow discourse markers embedded in texts.

Testing young children's literacy abilities productively, though, is more difficult. The mechanics of writing are still in the learning phase until middle elementary school, and so tests of children's early writing may underestimate their knowledge of how to structure a discourse. Fortunately, the oral genre of narrative has many features of written discourse (Chafe, 1980, 1982) and employs many of the same distinctive devices that will appear later in children's writing. Indeed, narrative development even at preschool has shown significant prediction of later literacy development (Snow & Dickinson, 1990; Torrance & Olson, 1984). Measurement of oral narrative ability, then, is a promising avenue for understanding children's growth in skills important for literacy.

Children from age 3 can generally be counted on to produce stories in response to a standard prompt. Their stories develop in interesting and measurable ways, especially from ages 5 to 10, in the creation of a unified plot structure, in the motivation of events through reference to internal states of the characters, and in the appreciation of the listener's needs for information which are different from the narrator's. From a discourse perspective, children's stories will provide increasingly more elements of the adult genre: especially, more setting of the scene, more problem-resolution sequences, and more complex and frequent narrator's comments on the action (Kemper, 1984). In terms of the development of discourse markers, children will show increasingly adequate contrast of indefinite and definite reference, less ambiguous pronoun reference, and more frequent and clearer expressions of emphasis (Kemper, 1984: 112). The patterns of that development have been explored in numerous studies, especially since Halliday and Hasan's (1976) groundbreaking work on cohesion. (See Berman & Slobin, 1994; Hedberg & Westby, 1993; Kemper, 1984; Peterson & McCabe, 1983, for syntheses of this literature.) Closely allied to discourse devices are advances with age in the narrative elements the children include in their stories. We see in Berman & Slobin (1994, chap. IIa) that older children's versions of their stories typically include more explicit references to cause and effect, more compound time referencing, and a more complex theory of other minds. So, simple stories like the 'Frog Story' used by Berman and Slobin, give children the opportunity to show evidence of their achievement of these cognitive developments known to be expanding during this age range.

At the same time, from a linguistic point of view, one can expect that as children approach age 10, their stories will become richer lexically and have more embedding syntactically (Karmiloff-Smith, 1986). The stories may contain more complex syntactic structures and more complex combinations of the structures. According to studies reviewed by Scott (1988: 60), post-modified noun-phrases, non-finite verbs, modal auxiliaries, and perfect tenses index increased complexity within clauses; across clauses, one sees an increase with age in the number of low-frequency conjunctions ('although', 'unless', etc.) and a greater density of syntactic units per sentence.

One goal in examining children's narratives, then, will be to evaluate the growth of discourse devices on the one hand, and on the other the specific linguistic structures which the extended narratives afford the opportunity to deploy. By separating the scoring of the stories into independent components and even subcomponents, we can examine the separate contribution of each element to more global measures of the children's growth. This is an especially useful framework for looking at stories from bilinguals, where greater dissociations between component language skills have been hypothesized to exist (Pearson *et al.*, 1996, October). It also provides an opportunity to isolate which elements appear to develop within the context of learning in a specific language, and which are tied to more general growth across languages (Cummins, 1984).

Using the full design at 2nd and 5th grades (Chapter 2), we could see

which combinations of the study's factors – Lingualism, Socio-economic Status (SES), Instructional Method in School (IMS), Language Spoken at Home (LSH), and language of the story – are associated with greater or lesser growth in the two major dimensions outlined above. In addition, with stories in two languages from the same children, we were able to assess the degree to which growth in one language appeared to support or hinder the children's growth in the other language for the two domains of 'discourse' and 'language'.

The hypotheses to be tested were:

(1) Observed differences between bilinguals and monolinguals on a global measure of narrative ability in English will not be equally evidenced in the elements which make up the global measure.
(2) For bilinguals, Two-way instruction will enhance performance on narrative tasks in both languages.
(3) There will be a predictive relationship between narrative abilities children demonstrate in one language on comparable abilities in the other language.

Methods

Participants

The participants for the narrative study were a subset of the full design discussed in Chapter 2 (and below): there were 10 children each in the 8 bilingual groups, and 20 children each in the 2 monolingual groups at both 2nd and 5th grades, 240 children in all. The subject groups are displayed in Table 7.1.

To reiterate the key elements of the selection process (see Chapter 2), all the bilingual children were born in the United States. The OSH (Only Spanish at Home) children lived in homes where primarily Spanish was spoken at least until the child was age 5; the ESH (English and Spanish at Home) group children, in homes where English and Spanish were spoken approximately equally from the time of the child's birth. The bilingual children were in one of two types of schools: in English Immersion schools, instruction was all in English (except for an optional half-hour a day in Spanish); in Two-way schools, both languages were used as the medium of instruction; 60% of each day was taught in English and 40% in Spanish. (All the schools were 'neighborhood' schools; that is, almost no children chose the school because of the language policy.) The English monolinguals were all living in households where only one language was spoken and all but seven children were in schools with a majority of non-Hispanic peers. (The seven monolingual children in Spanish-peer schools were included in the analysis in Chapter 4, which showed no reliable difference between their

Table 7.1 Number of participants by group

	Bilingual								*Monolingual*	
	Eng. Imm. School				*Two-way School*					
	Hi-SES		*Lo-SES*		*Hi*		*Lo*		*Hi*	*Lo*
	OSH	*ESH*	*OSH*	*ESH*	*OSH*	*ESH*	*OSH*	*ESH*		
Grade 2	10	10	10	9	10	10	10	10	20	20
Grade 5	10	10	10	10	10	10	10	11	20	20

Notes: 2nd graders are ages 7and 8; 5th graders are 10 and 11

language achievement and the language achievement of the English-peer children.) Since one goal was to assess the effect of the two educational programs for the bilinguals, all children chosen for the study had been in the same educational program since kindergarten. Further, only children who had experienced the educational programs for at least two years were included. That is, this part of the study did not look at kindergartners.

All children in the narrative probe study (except two) participated in the full design (see Chapter 2). In addition to the narrative task, they were given the full set of eight Woodcock-Johnson oral and written language tests and the Peabody Picture Vocabulary tests, in both Spanish and English in counterbalanced order. All parents filled out an extensive demographic questionnaire. Additionally, many of the children in the narrative probe study also did the phonological translation (Chapter 11) and grammaticality judgment (Chapters 8–10) tasks, so a range of correlational analyses could be done. (The two children noted above as exceptions were selected for the study but were not given the Woodcock Battery because the cells matching their demographics had been filled before the narrative probe study began.)

The narrative probe study subgroups were constituted from the first 10 (or 20) children in each full design subgroup who were tested after the start of the narrative study in the middle of the second semester during which testing was conducted. There were some small adjustments for the balance between schools or occasional technical problems with the taping. In order to estimate how representative the smaller groups were of the groups from which they were drawn, mean scores on the Woodcock tests for the narrative probe study subgroups ($n = 10$ or 20) were compared to those for the subgroups in the larger design ($n = 30$ or $n = 40$, see Results, below).

The task

The children narrated the wordless picture book, *Frog, Where Are You?* by Mayer (1969). This book was chosen because it has been used success-

fully with grade-school children in many countries of the world, including the United States, Spain, and several of the Latin American countries where the Miami children's families came from. The book consists of 24 pictures depicting a little boy's search for a pet frog who escapes from the boy's room at the outset. The boy is aided in the search by his dog, whose adventures and misadventures complicate the story line. In the final pictures, the boy and the dog find a frog family and take one of the baby frogs home. (The book is reprinted in its entirety in Berman & Slobin, 1994, Appendix I.) The culture-specific story frames implicit in the book were explicated by anthropological linguist, D. Wilkins of the Max Planck Institute (quoted in Berman & Slobin, 1994: 21–22). The six or seven frames relevant for the story were judged consistent with the experiences of young children in all those parts of the world from which our subjects' families originated. The pictures are simple line drawings which are ambiguous enough to allow some legitimate differences of opinion about what is happening. The activities of the *two* major male protagonists make adequate pronoun reference and event sequencing quite tricky, even for adults. Finally, there are many more episodes than most people choose to include in their narrations, so although the story is somewhat constrained by the pictures, there is considerable variation in the children's renditions of it.

The stories were audio-taped individually in a quiet room in the children's school on one of the days when several subtests of the Woodcock Battery were also given and recorded. The seven testers were English–Spanish bilinguals recruited to have no non-native accent in either language. They were seven of the same eight research assistants who carried out the full design – standardized tests and other 'probe' studies – and thus were very familiar to the children whom they were recording. The children followed the standard protocol set out by Berman & Slobin (1994: 22), which eliminates any memory demand from the task. Children looked through the book to the end once and then again as they told the story, turning the pages at their own pace.

The bilinguals told the story in English one day and in Spanish another day. The order of the language was determined quasi-randomly for each child, and storytelling language and standardized testing language were maintained within testing sessions. Rendering of the story in the second language was usually a week or so after the first. Because the narrative study utilized only a subset of subjects from the full design, some imbalance was present in the order of testing. Hence, Language Order was included as a covariate in the analyses. The *F*-statistics are included under Control Variables in Results, below. Similarly, gender was not controlled for in the design. Therefore, gender was also included as a covariate and is reported below, under Control Variables.

As a final control variable, to evaluate the effect on the stories of having the bilinguals tell the story twice, 24 of the monolinguals also told the story a second time. Twelve monolingual English-speaking 2nd graders and twelve 5th graders, (half Low-SES and half High-SES), were asked to retell the story during another test session, one to two weeks after the first. The scores from the 24 Time–1 tellings were compared to the 24 Time–2 tellings with a paired-samples *t*-test and correlation. The *t*-statistic, *r*, and 95% confidence intervals for these analyses are also included under Control Variables in the Results, below.

Transcription procedures

The stories were transcribed by bilingual transcribers following the conventions outlined by Berman & Slobin (1994: 657–9), with one 'verbed clause' per line. (A verbed clause contains just one finite verb; it can be one word (a verb), or a whole sentence.) Standard orthography was used throughout; for unusual pronunciations, the child's production was spelled out phonetically in brackets. Ex: the cliff [clift]. All hesitations and false starts were included in curly brackets, {}. Unintelligible passages were marked as 'xxx'; less clear passages were enclosed in parentheses. The word-processed transcripts (without information as to group membership or grade) were printed with line numbers in a standard manner for the coding (described below). Then standard utterance delimiters were added, and the transcripts were put into CHAT format for analysis with the CHILDES programs (MacWhinney, 1995). (The stories are available through the CHILDES website at http://www.psy.cmu.edu/under frog corpora/miami.)

All the stories were listened to at least twice: 15% of the stories had two independent listenings in accordance with standard reliability procedures. The other 85% were given a non-independent second listening. That is, the second person had the transcription in front of her and listened only for disagreements. Stories averaged six discrepancies per 100 lines. Discrepancies between the transcribers or listeners were resolved by the author, resulting in a third listening when necessary.

Measures

The 400 stories were evaluated with a set of measures devised for the study that combined both analytical and holistic judgments. These measures incorporated the two broad areas outlined above. One primarily holistic set, the Story Score, looked at the child's ability to use a hierarchical story structure, maintain a clear flow of information, and include evaluative and metacognitive statements in recounting the events in the picturebook. The second, the Language Score, was a more analytical measure that examined the more purely linguistic aspects of the children's

performances: counts of selected verb forms, conjunctions, adverbs, and the specialized noun vocabulary of the story.

There is not to our knowledge a recognized rubric – no 'answer key' – for scoring stories, but there is a wealth of descriptive information about this Frog Story in particular, which helps characterize the typical 3, 5, 9-year-old or adult response to it (Berman & Slobin, 1994, chap. IIa and Hoff-Ginsberg, 1997). By comparing elements reported for 'a few 5-year-olds and 90% of 9-year-olds' or 'a few 9's and a majority of the adults', we were able to assemble a developmental sequence for several elements of the domains noted above. For example, with respect to one key plot element, the frog's escape from the jar, according to Berman & Slobin (1994, p. 46), only 50% of the 4-year-olds, but 94% of the 9-year-olds used a 'mental predicate' to remark on the missing frog – 'the boy and the dog *saw* or *discovered* or *were upset* that the frog was gone.' Likewise, 'only a few 5's but most of the 9's' began the story with some type of stereotypical story opener (p. 74). The older children were also more likely to give some background on events that might have led up to the first picture – on how the boy had come to have the frog, and so forth (p. 72). We used the elements described in Berman & Slobin (1994) as the basis for the two-pronged story metric. Our goal was to identify, separately for the two domains of narrative and linguistic development, which stories told by the children resembled a typical second-grader story and which were more or less advanced.

Using a model derived from gymnastics judging and portfolio evaluation (see Pearson, 1996, October, for the full rationale behind the measures), we retrofitted a descriptive framework capable of capturing the differences between the five best and five worst of the monolingual stories as judged holistically by a panel of educators and linguists. Monolingual stories were used in creating the rubric to help the panel focus on the story elements and not the language. (Based on the author's experiences training composition instructors, it was deemed too difficult, even for teachers, to keep surface-level language errors from coloring their overall evaluation of the discourse.)

The gymnastics judging model is especially apt for separating the judgment of a story's narrative elements from the judgment of its language. For each optional gymnastics routine on bars, beam, or floor exercise, the judge first counts the level of difficulty of all the moves and connections in the routine. By adding the values assigned each move (in the standard judging manual, which is always available), one arrives at the highest possible score for the particular routine performed – *if* all the elements were executed perfectly. Then, one subtracts execution deductions according to a general table, also in the manual. One takes off so many points for a fall, so much for bending the arms, in an element or throughout. The execution deductions are subtracted from the 'value' of the routine to arrive at an overall

score. Likewise, the metric that was devised partitioned the narrative or 'Story Score' into five areas – Story Elements, Sequencing, Reference to characters, Reference to internal states, and a quality of Engagement. The Language Score was divided into three areas: Complex Syntax, Lexicon, and Morphosyntactic Accuracy. The first two of these count elements present in the story; the last counts errors to be deducted. A summary of the categories with example passages from three stories is given below. A coding sheet is included in the Appendix to this chapter.

Summary on the Story and Language Scores

STORY SCORE	*Possible points*	*Description of 'midpoint' score (6 or 3) for the 'average' 2nd grade story*
Elements	12	a search story, including losing the frog, setting out in search, and finding a frog.
Sequence	12	sentence-by-sentence, picture-by-picture chain of events (little or no orientation, setting, summary).
Reference	6	use of indefinite article for first mentions; generally adequate pronoun antecedents (or use of 'thematic pronoun strategy' (Karmiloff-Smith, 1986), with some lapses.
Internal states	6	little reference to emotions, reactions, or thoughts of the characters.
Engagement	12	matter-of-fact tone; no 'literary' language
Total	48	
LANGUAGE SCORE		***Midpoint story generally correct, but unelaborated. A string of simple sentences, relatively correct.***
Complex Syntax	24	mostly simple verb phrases; points added for each occurrence, (up to 3) of modals or aspectual markings ('began to,' 'kept on'); (in Spanish, perfect tenses, subjunctive); across clauses, points given for conjunctions other than 'and then' ('y después'); bonus for noun or adjective clauses.
Lexicon	12	uses most of a set of 12 words, specific to the story: ('frog', 'jar', 'bees', 'beehive', etc.)
Morpho-syntactic accuracy	12	(errors deducted from 12) generally well-formed, a few non-prescriptive structures ('a owl,' 'there was bees')
Total	48	
NARRATIVE TOTAL	96	

At the 2nd grade level, all but the very lowest stories introduced the three main characters, related key events (while omitting others), and by the end resolved the search for the frog. The medium stories did so as a straightforward chronicle of events: 'this happened and then this happened and then this happened'. As indicated in the summary table, those stories were given the midpoint score on two of the scales of the Story Score: Frog Story 'Elements' and 'Sequence'. But the weaker stories lost points by getting side-tracked into details that did not advance the story or by lapsing into picture-description. The child could also lose points in the Reference category for failing to introduce one or more of the characters, or for losing track of clear reference when using pronouns (or at least not reserving the pronoun exclusively for the main character or 'thematic subject', as described by Karmiloff-Smith (1981)).

Example Second Grade Stories (Excerpts)

Note: Short pauses are indicated with '-' and longer pauses with '...' Words in curly brackets were judged to be retracings. The division into lines with one verbed clause per line follows the convention from Berman & Slobin (1994: 657–9).

Story Example 1 (Low Average)		
line	1	The dog – looked in – the bottle and looked at the frog. And the boy was sitting on a chair. And his – sock and his shirt was laying on the floor.
	5	And the light was on. And the window was opened . . . When – {the} the boy and the dog were sleeping - the frog – stuck his head out with his head an' his arm – out of the bottle.
Story Example 2 (High)		
line	1	One day a boy and his dog had found a frog. They kept him in the big jar. While the boy was asleep the frog climbed out of the jar
	5	and ran away. When the boy woke up the next morning he was very upset to see his frog missing. He searched everywhere.
	10	In boots . . . And he turned over tables.

Story Example 3 (High, Bilingual)		
line:	1	Once there was a little boy with his little dog. It was already night time. They were looking at the little frog. The little boy – and his dog went to sleep.
	5	The frog – wanted to go out to see {the w} the world. So he came out of the little – can. It was morning already. The puppy and the boy looked to the – can
	10	and saw {that there} that the frog was not there.

Stories that were above average included one or more of the following elements (noted by Berman & Slobin 1994: 82): short summary statements (either prospective or retrospective), comments on the reactions of the characters, clear articulation of the boy's goal (not just a description of his activities in pursuit of that goal), or, toward the end of the story, an explicit mention of the boy's misperception of the deer's antlers as sticks. This last element, for example, was reported by half of the adults in Berman & Slobin's Table 4 (p. 55), but by only a small percentage of school-age children. Then, based on the better stories in our sample, we also added a way for the children to earn credit for remarking on the boy's *lack* of success in the search, which must – like statements about internal states of the characters – be added by the child and not 'read' directly from a picture. We also added the possibility for extra credit to be earned in a category we called 'engagement'. Unlike the more strictly linguistic measures of language structures counted in the separate Language Score (discussed below), this column awarded 'narrative' credit for literary-like language – expressions that made the child's rendition more lively or engaging: using a refrain in the story, or direct speech, or even figures of speech. (The closest we found to that in this sample was stylistic word-order inversion, like 'the frog was nowhere to be found' or 'out popped an owl'.)

From the linguistic perspective, the stories that rated 'high' tended to have more subordination and greater specificity in the nouns and verbs used; the child would say 'deer' or 'antler' instead of 'the large animal with those sticks on his head'. The better stories also encoded more complex temporal and causal relations in their verb and adverb constructions. The Language Score tried to characterize how well the child handled the more advanced grammatical structures. The main thrust of the Language Score was to credit the children's performance for the language elements they demonstrated, not to penalize for mistakes, but it was not always possible

to separate 'positive' and 'negative' scoring. Form errors inevitably had some impact on the Complex Syntax score, for example, because credit was given only when constructions were relatively well-formed. In lexical choice, as well, there were three levels of credit: +1 if the child used a keyword, zero if the concept was not referred to, and –0.5 if the child demonstrated that she did not know the word (saying 'flying things' for the bees, or 'that bee-thing' for the 'beehive').

In addition, a very salient characteristic of the bilinguals' stories (especially in Spanish), was the high number of morphosyntactic mistakes: the use of overregularizations, '*falled*' for 'fell', or the wrong form of an article '*el* ventana' for 'la ventana' [the window] or '*a* owl' instead of 'an owl'. (See also Martínez, 1993, on 'morphosyntactic erosion'.) In order to keep track of the incidence of such errors within the various bilingual subgroups and even among the monolinguals, a count of departures from morphosyntactic accuracy (MS Accuracy) was maintained. The Language Score Total was computed in two alternative ways, one based on only the positive qualities present in the language (Language Total 2, with only Complex Syntax and Lexicon), and another (Language Total 1) which also took into account the morphosyntactic errors, or elements that might draw sanction or correction from a monolingual adult (cf. Ochs, 1985). The tally of MS Accuracy is potentially problematic in both languages, but especially so in Spanish because there is not to our knowledge any well-defined framework for what is considered 'acceptable' in the adult varieties of Spanish spoken in Miami. Consultants disagreed as to whether the child should say 'cayó' or 'se cayó' ['he/she/it fell' or 'fell, reflexive'] when the boy or dog fell, as they did in almost every story. Pilot efforts to characterize the different varieties convinced us that while MS Accuracy was a useful index of 'exposure to literate Spanish', no conclusions could be drawn about a host of structures commonly accepted by Hispanic adults in Miami, whose grammatical status is beyond the scope of this paper.

The two scores, Story and Language, were not orthogonal, but they were distinct. For example, one element of the Complex Syntax score tracked how the child expressed the causes of events – including intentions. The expression of intentions was also counted in the Story Score, but the Language Score credit for this was more specific: it indicated that the child had made an explicit link *between clauses*. If a child got points for an intention structure in the Language Score, she would necessarily have credit in the Story Score as well, but the converse was not true. Likewise, the ability to express simultaneity is a key element in the foregrounding and backgrounding of actions, counted in a global way under 'Sequence' in the Story Store. In the Language Score, though, points were awarded only when the child used specific grammatical devices of the language: e.g. the

conjunctions 'while' or 'mientras', some uses of 'when', or the present participle, as in 'he climbed up a rock, *calling* out to his frog'.

So, the child could get credit in the Story Score for elaborating an episode in several simple sentences, but would not get Language credit unless the sentences were also linked grammatically. The first score was more concerned with capturing the level of the child's conceptualization of the story and how it should be recounted – regardless of the level of language used. The Language Score, by contrast, focused more narrowly on the use of specific later-developing lexical and grammatical constructions.

Coding procedures and reliability

'Hand'-scoring. Using coding sheets for Story and Language Scores, stories were coded independently by two researchers who were 'blind' as to the identity and group membership or age of the child under consideration. Differences in the scoring were resolved by discussion, and a consensus score was reached.

As mentioned above under 'Task', a test-retest reliability scoring using the same procedure was performed for 30% of the monolinguals.

Machine-scoring. The CHILDES programs (MacWhinney, 1995) were used to characterize the quantitative aspects of the stories. A tally was made of length in words, clauses, and the sentence-like 't-units' (following the standard definition [Hunt, 1977], of a minimal terminable-unit, 'a single independent clause and all other clauses that . . . go with it so there will be no [dependent clauses] left over', p. 93). From those measures one can derive 'MLU' (mean length of utterance, here the number of words per t-unit) and a 'subordination index' (the mean number of clauses per t-unit). The number of different word-types in the story, as well as the number of types in the first 100 words, were also counted.

Analytic procedures

In all, there was in each language a summary score, Narrative Total, composed of a Story Score and a Language Score, which was comprised in turn of five and three subcomponents, respectively; finally there were eight variables from the CHILDES analysis. All were linked to the database with the standardized test scores and demographic information for each child (see Chapters 2–6). These 19 measures allowed the Frog Stories to be ranked from a variety of perspectives consistent with the main avenues of evaluation in the narrative literature discussed above. The three hypotheses specified above were evaluated in terms of both between- and within-subject questions, with subhypotheses as follows:

Between-subjects questions

H1. How do monolinguals and bilinguals compare with respect to overall scores in English on the global measure of narrative ability? With respect to the specifically narrative aspects (Story Score and subcomponents)? With respect to linguistic aspects (Language Score and subcomponents)?

H2. How do bilinguals in English Immersion schools and Two-way schools compare to each other on the same sets of measures in English and in Spanish? Is the pattern different at 2nd and at 5th grade?

Within-subjects questions

H3. Within subjects, how strong a prediction of ability in one language is given by ability demonstrated in the other? For the global measures of ability? For the subcomponents of the scores in each language? That is, beyond the demographic factors investigated, was there an influence of skills in a first language on learning in a second language (or of skills in the second on the first)?

Results

Summary of Frog Story Results

The patterns of results between the groups and across languages differed according to the measure being examined. For the Narrative Total scores in English there were main effects of SES, Grade, and Lingualism (monolingual versus bilingual), with a strong interaction of Grade and Lingualism. It is not surprising that there were grade effects, as these are non-standardized scores. However the grade effect was different for monolinguals (MLs) and bilinguals (BLs) with BLs showing relatively more improvement on these measures than MLs in this age range. That is, bilingual children's scores were significantly lower than MLs' at 2nd grade, but closer, and on many variables equal to MLs' at 5th grade. When the Narrative Total score was broken down into its component scores, ML-BL differences were quite small for the narrative and discourse elements, as captured in the Story Scores, but the differences were larger for the Language Scores.

Overall, most of the Hispanic children did better in English than in Spanish, with greater differences between languages on Language Scores than Story Scores. Only one of the 16 bilingual groups had higher mean scores in Spanish than in English: the Low-SES 2nd graders with only Spanish at home (OSH) and Spanish and English in the school (Two-way). The group with these same characteristics at 5th grade, though, appeared dominant in English.

Between bilingual groups, the patterns of effects from the factors nested within Lingualism were different in the two languages. In English, LSH showed a significant effect for Language Score (not Story), with the strongest effects seen in Lexicon and Morphosyntactic Accuracy (MS Accuracy). By contrast, in Spanish there was an LSH effect only in MS Accuracy and no other variable. Unlike in English, there was a significant effect of IMS on most measures in Spanish (but not on MS Accuracy). In Spanish, the Grade effect was strong on all measures (except MS Accuracy). Notably, SES showed no statistically reliable results for any Spanish variable studied.

Finally, between languages for the bilinguals, there was a strong prediction of ability in one language to ability in the other for some elements of the Narrative Total scores, but not for others. In particular, the narrative/discourse elements (Story Score) showed carry-over across languages while specific language elements (Language Score) did not. The degree of elaboration and embedding in complex sentences was similar across languages, but knowledge of vocabulary items and general well-formedness of sentences were not similar.

Descriptives

The descriptive statistics in Table 7.2 help define the information in the database.

One can see that as the monolingual children got older, the stories got a little shorter, and MLU went up. Since one sees that scores went up, one can infer that the older children were able to convey more information more efficiently.

The bilinguals' stories (Table 7.3) were shorter than the monolinguals' stories at 2nd grade, but similar at 5th grade. Whereas the monolingual

Table 7.2 Length of story, descriptive statistics (Monolinguals only) (40 children per grade)

		Mean	*Stdev.*	*Range*
2nd grade	# of words	324	(115)	167–605
	different words	109	(29)	63–174
	# of clauses	53	(17)	28–95
	MLU (mean length of utterance)	7.7	(1.2)	5.6–10.5
5th grade	# of words	259	(78)	68–448
	different words	98	(23)	41–157
	# of clauses	43	(12)	10–71
	MLU	8.4	(1.1)	5.7–10.5

Table 7.3 Length of story, descriptive statistics for English (Bilinguals) (40 children per school type per grade)

		Eng. Imm. Schools		*Two-way Schools*	
2nd grade	# of words	246	(78)	247	(70)
	different words	85	(20)	79	(17)
	# of clauses	43	(14)	41	(10.6)
	MLU	7.2	(1.1)	7.3	(1.1)
5th grade	# of words	269	(76)	309	(112)
	different words	97	(20)	106	(24)
	# of clauses	44	(12)	50	(16)
	MLU	8.3	(1.3)	8.9	(1.5)

Table7.4 Length of story, descriptive statistics for Spanish (Bilinguals) (40 children per school per grade)

		Eng. Imm. Schools		*Two-way Schools*	
2nd grade	# of words	248	(70)	217	(64)
	different words	76	(20)	74	(17)
	# of clauses	39	(13)	38	(11)
	MLU	6.6	(0.9)	6.9	(1.1)
5th grade	# of words	236	(80)	262	(84)
	different words	85	(20)	94	(22)
	# of clauses	42	(14)	46	(14)
	MLU	7.0	(1)	7.5	(1.3)

stories got shorter over this age range, the bilingual stories got longer. Both monolingual and bilingual groups experienced growth in MLU, but bilinguals showed more significant growth (ML: $F = 3.72$, $p = 0.058$; BL: $F = 42.4$, $p < 0.001$).

Comparing the English stories across IMS, we see the school types were similar in 2nd grade; in 5th grade the Two-way children's stories were longer and the sentences were somewhat more complex, according to the MLU measure ($F = 3.44$, $p = 0.065$).

The Spanish stories (Table 7.4) appeared similar across IMS at both grades, with a tendency for more complex sentences (or higher MLU) in the Two-way schools at 5th grade.

Table 7.5 includes the monolingual values at 2nd and 5th grade for the main elements of the Frog Story scoring.

Table 7.5 Selected Frog Story measures: Descriptive and inferential statistics (Monolinguals only)

		Mean	*Std.*	*Range*
2ND GRADE				
Story Total	**(of 48)**	**28.2**	**(5.5)**	**20–43**
Elements	(of 12)	6.7	(2.2)	3–12
Sequence	(of 12)	7.8	(1.3)	5–11
Internal	(of 6)	2.7	(1.2)	0–6
Language Total	**(of 50)**	**34.4**	**(5.9)**	**21–45.5**
Complex Syntax	(of 24)	16.7	(3.7)	11–23
Lexicon	(of 14)	9.1	(2.6)	4.5–14
MS Accuracy	(of 12)	8.6	(2.5)	2–12
Frog Total	**(of 98)**	**62.6**	**(10.7)**	**45.5–85**
(from CHILDES)				
Subordination Index		1.26	(0.17)	1.02–1.72
Type-token Ratio		0.35	(0.06)	0.24–0.47
5TH GRADE				
Story Total		**30.1**	**(4.4)**	**23–44**
Elements		7.6[a]	(1.9)	5–12
Sequence		8.8[b]	(1.1)	7–11
Internal		2.4	(1.3)	1–6
Language Total		**37.3[c]**	**(4.2)**	**27–49.5**
Complex Syntax		17.3	(3.1)	12–24
Lexicon		9.2[d]	(2.2)	3–13.5
MS Accuracy		10.7[e]	(1.2)	8–12
Frog Total		**67.3[f]**	**(7.7)**	**53–93.5**
(from CHILDES)				
Subordination Index		1.37[g]	(0.18)	1.08–1.78
Type-Token Ratio		0.39[h]	(0.06)	0.31–0.6

Notes (Significant effects):
[a] SES, $F = 6.35$, $p = 0.014$
[b] Grade, $F = 13.34$, $p < 0.001$
[c] SES, $F = 11.21$, $p = 0.001$; Grade, 6.68, $p = 0.012$
[d] SES, $F = 5.04$, $p = 0.028$
[e] SES, $F = 14.72$, $p < 0.001$; Grade, $F = 24.72$, $p < 0.001$; SES by Grade, $F = 5.81$, $p = 0.018$
[f] SES, $F = 8.18$, $p = 0.005$, Grade, $F = 5.48$, $p = 0.022$
[g] Grade, $F = 7.94$, $p = 0.006$
[h] Grade, $F = 9.03$, $p = 0.004$

Specific results from multivariate analysis of variance

To evaluate differences among group means, the data were analyzed via 6-way multivariate analysis of variance using Type III sums of squares (General Linear Model, SPSS 7.1 for Windows). Three between-subjects factors were fully crossed with one another: SES (high vs low), Grade (2nd or 5th, and Lingualism (bilingual vs monolingual). In addition, three between-subjects factors were nested within Lingualism – IMS (English Immersion or Two-way), LSH (OSH vs ESH), and Language Order (English 1st vs Spanish 1st).

The design yielded a total of 35 effects. To protect against inflation of Type I error rate, a Bonferroni correction was applied to yield a groupwise alpha of 0.05. As this corresponded to a very strict comparisonwise alpha of 0.0014 for the English scores (and 0.0033 for the bilingual-only comparisons with 15 effects [$2^4 - 1$]), some small effects due to Type II error might have been neglected. Nonetheless, the procedure resulted in many effects which remained statistically significant. (It should be added that for tests of simple effects, or *post hoc* comparisons involving only two groups – that is, in cases with only a single comparison – the Bonferroni groupwise correction was not used. Therefore, there may be some anomaly when *F*-statistics with Bonferroni correction were compared to those without the correction. In such cases the alpha value is explicitly stated so that appropriate caution can be used in comparing the simple effect to the original main effect.)

Tests of control variables: Language order, gender, and tester

As noted above in Methods, for the bilingual children, who told the story twice, the language order (English or Spanish first) was slightly unbalanced in the groups. There were 69 who told the story in English first and 91 in Spanish first. Therefore Language Order was included as a factor nested within Lingualism. For 15 of 16 dependent variables, Language Order yielded no significant main effect. For example, with Narrative Total, $F = 2.39$, $p = 0.123$; for Story Scores, $F = < 1$, $p = 0.4$. The single exception was for the English Lexicon scores, with no Bonferroni correction, $F = 7.46$, $p = 0.007$. In addition, Language Order entered into an interaction with SES and grade for Morphosyntactic Accuracy, again in English, such that the Low-SES children in the 5th grade spoke English more accurately in their stories (an average of 2 errors rather than 4) when they told the story in English first than when they told it in Spanish first. This difference was not evident for their High-SES classmates, nor was there an SES difference at 2nd grade. Language Order showed no significant effects for any of the Spanish measures. Still, it was included as a factor in all the analyses to insure that its contribution to the error term was correctly partitioned.

Similarly, gender was not controlled in the assignment of subjects, and

in fact, there were 54% females in each of the three main groups, English Immersion bilinguals, Two-way bilinguals, and monolinguals. When gender was entered in the analysis as a covariate, it was not a significant factor for any dependent variable and thus was not included in any further analyses. Nor did Tester produce any significant effects.

A test for the effect of telling the story twice

To ascertain whether there was a practice effect from telling the same story twice one to two weeks apart, 24 monolinguals also repeated the story after a similar interval. The mean scores for the Time–1 and Time–2 tellings in English are found in Table 7.6.

The 24 Time–1 tellings were compared to the 24 Time–2 tellings with a paired-samples *t*-test, $t < 1$ (n.s.) for both Story and Language scores. Mean differences on the measures between tellings (for the whole group) were small, always less than 0.7 (compared to standard deviations around 6 points). Correlations between times of telling were 0.77 for the Story Scores, 0.76 for the Language Scores, and 0.8 for the two summed together (Narrative Total).

An additional check was made for stability with respect to three highly valued items in the Story Scores – the mention of a mental verb for the discovery of the missing frog (+2), mention of the boy's misperception of the deer's antlers as sticks, or similar misperception (+4), and the inclusion of more than one comment on the 'internal' state of the characters. Between

Table 7.6 Test-Retest Reliability: Story and Language Scores by time of telling and grade means (and standard deviations)

		N	*(Time 1)*	*(Time 2)*	*t*	*95% Conf. Interval*
Story Score						
All		24	29.17 (6.0)	29.17 (6.3)	0.0	±1.76
Grade	2	12	27.9 (6.6)	26.8 (5.6)		
	5	12	30.5 (5.4)	31.5 (6.2)		
Language Score						
All			35.6 (6.8)	36.3 (6.0)	–0.7	±1.88
Grade	2		34.1 (7.8)	34.0 (5.2)		
	5		37.2 (5.5)	38.6 (6.0)		
Narrative Total (Story + Language)						
All			64.8 (12.1)	65.4 (11.8)	–0.4	± 3.22

83% and 92% of the children were consistent across tellings. Of the three, four, and two children for each element, respectively, who were not consistent, there was an equal tendency to embellish the story more or elaborate it less in the second telling. There does not, then, appear to be significant systematic bias in the group scores as a result of telling the story twice. In addition, the lack of reliable differences on the Language Order variable helps confirm that there was no systematic bias according to which language the story was told in first.

Principal analyses

Table 7.7 shows the means and standard deviations for the 10 subgroups at each grade for Story, Language, and Narrative Total scores, in English and in Spanish.

Bilinguals' values on the Frog Story summary variables are shown in Figures 7.1 (English) and 7.2 (Spanish), for both 2nd and 5th grades. The scores are given separately for the eight bilingual subgroups at each grade, with the monolinguals' included for comparison in Figure 7.1.

In general, Story Scores in English were more consistent than Language Scores across the subgroups listed on the x-axis in Figure 7.1. The ML groups' Language Scores were higher than their Story Scores at both 2nd and 5th grade, but the same was true of BL groups only at 5th grade. ML Language Scores were dramatically better than BL at 2nd grade, but not at 5th grade.

In Spanish (Figure 7.2), Story Scores were better at 5th grade than 2nd grade, especially in the Two-way schools, and Language Scores were lower than Story Scores at both grades. Improvement in Language Scores

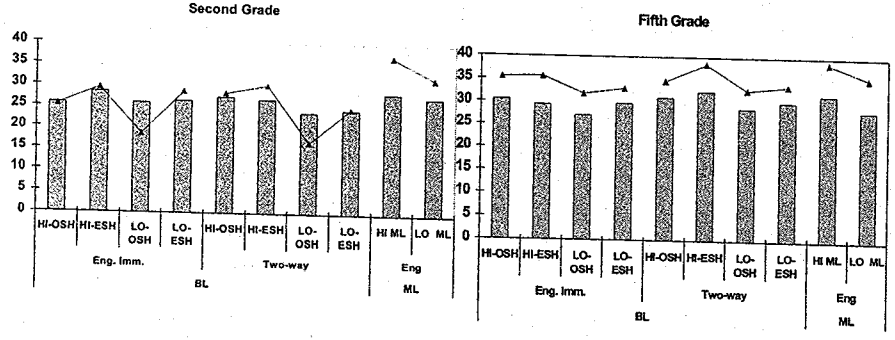

Figure 7.1 Story and Language Scores (English) (bars represent Story Scores; lines represent Language Scores)

Table 7.7 Descriptive statistics for Story, Language, and Narrative Total scores in English and Spanish by Grade, IMS, LSH, and SES

				Scores											
				English Story		English Language		English Total		Spanish Story		Spanish Language		Spanish Total	
Grade	IMS	LSH	SES	n	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
2nd	Eng Imm	OSH	hi	10	26	(3)	26	(4)	51	(6)	26	(5)	22	(6)	47	(10)
			lo	10	26	(3)	19	(6)	44	(7)	25	(3)	17	(7)	42	(8)
		ESH	hi	10	29	(5)	30	(5)	58	(8)	26	(5)	13	(9)	40	(13)
			lo	9	26	(4)	29	(5)	52	(7)	25	(4)	20	(11)	45	(14)
	Two-way	OSH	hi	10	27	(5)	28	(8)	56	(12)	28	(5)	20	(10)	48	(13)
			lo	10	24	(4)	17	(6)	40	(8)	24	(5)	22	(7)	46	(10)
		ESH	hi	10	27	(4)	30	(6)	57	(9)	27	(7)	19	(7)	46	(13)
			lo	10	24	(4)	25	(5)	49	(7)	26	(3)	16	(9)	42	(11)
	Mono-ling	Eng	hi	20	29	(5)	37	(5)	66	(9)						
			lo	20	28	(6)	32	(6)	60	(11)						
Grade	IMS	LSH	SES	n	M	SD	M	SD	M	SD	M	SD	M	SD	M	SD
5th	Eng Imm	OSH	hi	10	31	(4)	36	(4)	66	(7)	28	(4)	22	(7)	50	(10)
			lo	10	27	(4)	32	(6)	59	(9)	27	(6)	23	(10)	49	(11)
		ESH	hi	10	29	(5)	36	(6)	65	(9)	25	(5)	20	(8)	45	(10)
			lo	10	30	(3)	33	(9)	63	(11)	28	(4)	14	(11)	42	(14)
	Two-way	OSH	hi	10	31	(4)	35	(7)	66	(9)	30	(4)	22	(9)	51	(13)
			lo	10	29	(4)	33	(3)	62	(6)	29	(4)	27	(4)	56	(7)
		ESH	hi	10	33	(2)	39	(6)	71	(7)	33	(4)	25	(8)	58	(8)
			lo	11	30	(2)	34	(8)	64	(9)	29	(4)	25	(7)	55	(10)
	Mono-ling	Eng	hi	20	32	(5)	39	(4)	71	(8)						
			lo	20	29	(4)	36	(4)	64	(7)						

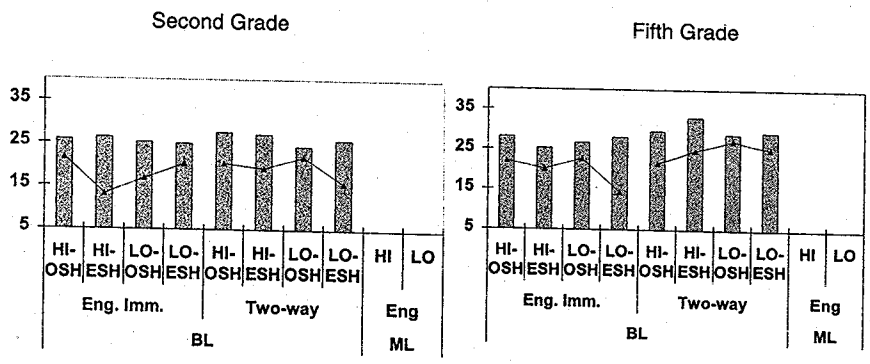

Figure 7.2 Story and Language Scores (Spanish) (bars represent Story Scores; lines represent Language Scores

across grades was found even in some English Immersion groups. The only subgroup whose Spanish Language Score was lower for the 5th graders than for the 2nd graders was the group with economic disadvantage and with the least Spanish input: Low-SES children in English Immersion Schools who had English and Spanish (hence less Spanish) in the home.

As described above, the six-way multivariate analysis of variance using Type III sums of squares was run first for the summary variables and then for the component scores. There were three fully crossed between-subjects factors: SES (high vs low), Grade (2nd vs 5th), and Lingualism (bilingual vs monolingual). In addition, three between-subjects factors were nested within Lingualism – IMS (English Immersion vs Two-way), LSH (OSH vs ESH), and Language Order (English 1st or Spanish 1st). A Bonferroni correction was applied to yield a groupwise alpha of 0.05 (comparison alpha = 0.0014).

Narrative Total score

Table 7.8 gives the Narrative Total values for the Monolinguals and the two groups of Bilinguals according to IMS. There were significant main effects of SES, Grade, and Lingualism, and an interaction of Grade by Lingualism (Table 7.9).

When the interactions were followed up with tests of simple effects (Table 7.10), the Grade effect was significant at both levels of Lingualism (ML: $F = 5.75$, $p = 0.022$; BL: $F = 93.94$, $p < 0.001$). However, the Lingualism effect was seen to diminish greatly from 2nd grade to 5th grade (Figure 7.3),

Table 7.8 Narrative Total Scores by Grade and IMS mean and (standard deviation)

Grade	*English Immersion*	*Two-way*	*Monolingual English*
2nd (7–8 years old)	51.5 (8.5)	50.5 (11.2)	62.5 (10.7)
5th (10–11 years old)	63.2 (9.1)	65.6 (8.6)	67.3 (7.7)

Table 7.9 Significant effects for Narrative Total score (English)

	F	p
SES	30.2	< 0.0001***
Grade	75.4	< 0.0001***
Lingualism	34.3	< 0.0001***
Grade by Ling	11.13	< 0.0010**

** $p < 0.01$, *** $p < 0.001$

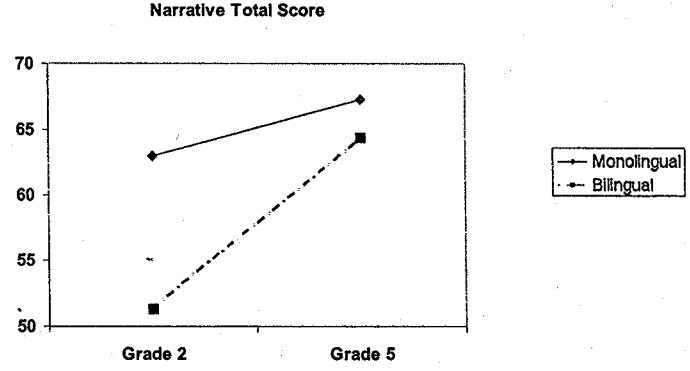

Figure 7.3 Narrative Total (English), Grade by Lingualism

and it was non-significant when the comparison was made between the High-SES/ESH bilinguals and the comparable groups of monolinguals (see Table 7.10).

The Hi-SES 2nd graders, unlike the 5th graders, showed a Lingualism effect, favoring MLs: $F = 12.93$, $p = 0.001$, and for children who were both Hi-SES/ESH at 2nd grade, the values were also still significant with no Bonferroni correction: $F = 7.11$, $p = 0.012$.

Table 7.10 Simple Effects of Lingualism on Narrative Total Score

	ML		*BL*			
Tested at	n	M (SD)	n	M (SD)	F	p
2nd grade	40	63.0 (10.5)	79	51.3 (9.9)	35.6	0.01*
5th grade	40	67.3 (7.7)	81	64.4 (8.9)	3.29	0.07
5th grade, Hi-SES only	20	70.6 (7.7)	40	67.0 (8.3)	2.59	0.11
5th grade, Hi-SES, & ESH(English & Spanish at home)	20	70.6 (7.7)	20	68.0 (8.4)	0.98	0.33

Contribution of component scores

The two component scores which comprise Narrative Total responded differently to the independent variables of the study (Table 7.11). Both showed high correlations to the Narrative Total score, and thus both contributed to the global measure. Nonetheless, the overall correlation of Language Score to Total Score, $r = 0.95$ was higher than for Story Score, $r = 0.83$, and it was consistently higher across the different Lingualism, IMS, and SES levels.

It is important, therefore, to examine both scores in analyzing the different groups' performance on the Frog Stories.

Story Score. For Story Score (see Table 7.12), with Bonferroni correction, there was a main effect for Grade, but not for SES, Lingualism, nor for any interactions. These same factors had different effects for the subcomponents of the score. There was a strong trend for the SES effect on all the scores as well as a Grade by Lingualism interaction, $F = 9.145$, $p = 0.003$.

The Grade effect seen in the global Story Score was most noticeable in the Sequence and Frog Story Elements subscores. The trend toward an SES effect for the global Story Score was seen as a significant effect in the Sequence subscore, and as a trend as well in the Elements and Internal States subscores.

Table 7.11 Correlation of Narrative Total score with Component scores by Lingualism, SES, and IMS

	Story Score	*Language Score*
Monolinguals ($n = 80$)	0.92	0.92
SES-High ($n = 120$)	0.86	0.93
SES-Low ($n = 120$)	0.80	0.95
BL, Eng. Imm. ($n = 79$)	0.75	0.93
BL, Two-way School ($n = 81$)	0.87	0.97

Table 7.12 Main effects for Story Score and Components

	STORY		*Elements*		*Sequence*		*Internal States*	
	F	*p*	F	*p*	F	*p*	F	*p*
Grade	30.3	< 0.001	21.78	< 0.001	54.4	< 0.001	5.19	< 0.024
SES	11.7	0.0017	8.04	0.005	13.42	0.001	8.13	0.005
Lingualism	4.0	0.047	1.50	0.223	2.93	0.088	4.24	0.041

*Significant with Bonferroni correction (comparison alpha = 0.0014)

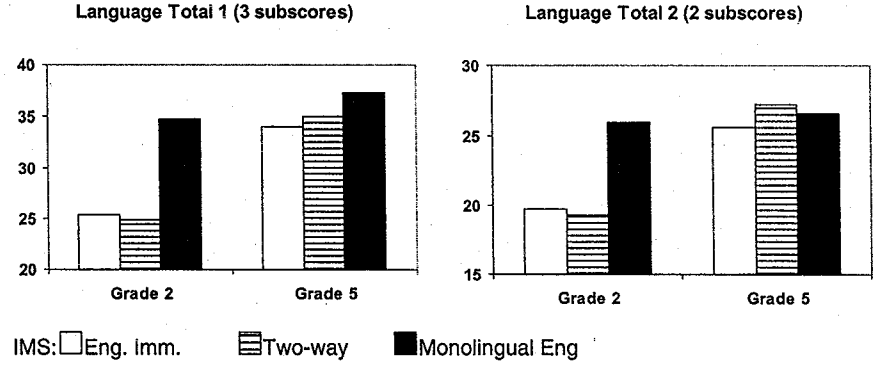

Figure 7.4 Language Totals by School Type and Grade

Language Score. The main effects in English of SES, Grade, and Lingualism observed for the Language Scores were significant for both Language Total 1 (with 3 elements), $F = 34.4$, 84.8, and 56.7, respectively, and Language Total 2 (without MS Accuracy), $F = 22.5$, 62.7, and 25.8, all p *values* < 0.001 (see Figure 7.4).

As seen above for Narrative Total, there was a strong Grade by Lingualism interaction for both Language Totals (see Figure 7.5).

The two versions of the Language Score were highly correlated, $r = 0.90$ and the significance of results was rarely different when using one score instead of the other. The relative contribution of MS Accuracy to evaluations of general language performance is highlighted by the correlation of that measure with the Language Total 2, the one that did not include it. For monolinguals, the correlation was negligible, $r = -0.05$, compared to $r = 0.37$ for the bilinguals. A similar disparity was observed in the Language Total

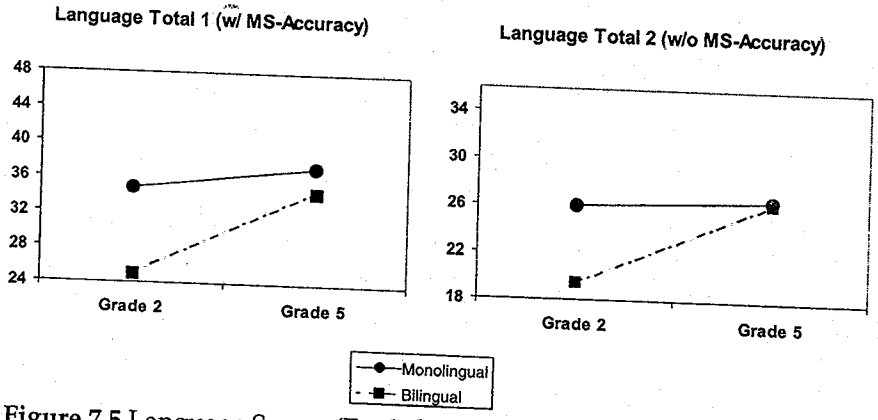

Figure 7.5 Language Scores (English), Grade by Lingualism

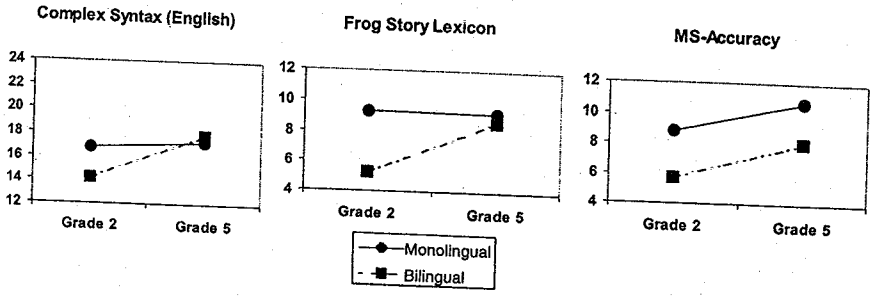

Figure 7.6 Language Score Subcomponents (English), Grade by Lingualism

1, where the correlation was 0.37 for monolinguals versus 0.73 for bilinguals. When used as part of the metric, surface well-formedness of the children's output, then, figured more heavily in the assessment of bilinguals' stories than of monolinguals'.

When one examines the subcomponents comprising the Language Total, one sees patterns of Grade by Lingualism (Figure 7.6) and by LSH (Figure 7.7).

In examining the Grade effect for the three subcomponents, one can see three different patterns of relation between the monolinguals and

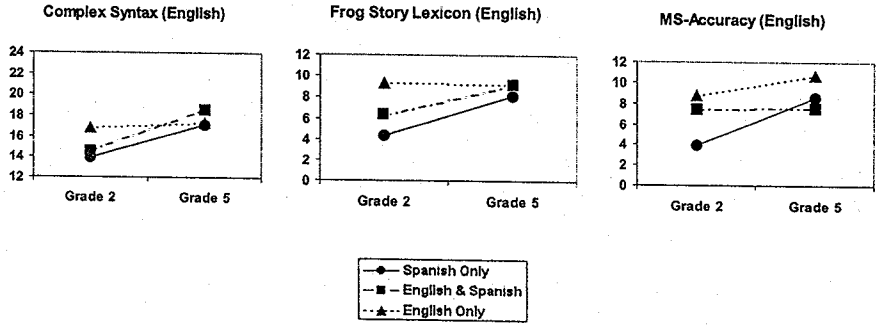

Figure 7.7 Language Score Subcomponents (English), Grade by LSH

bilinguals across grades. In Complex Syntax the 2nd grade bilinguals were well below the monolinguals, but the 5th grade bilinguals were equivalent to the monolinguals; in Lexicon, the gap narrowed but did not close, whereas in MS Accuracy, the 5th grade bilinguals were as far behind the monolinguals as the 2nd graders were. Not surprisingly, both Complex Syntax and Lexicon showed a Grade by Lingualism interaction, but MS Accuracy did not, $F = 12.53$, 18.41, and 0.11, $p < 0.001$, 0.001, and 0.748 respectively.

Curiously, the ESH bilinguals, unlike both the monolinguals and the OSH bilinguals, showed no change in MS Accuracy over this time period. Since these data were cross-sectional, the lack of 'improvement' may have been a cohort effect. Alternately, it may indicate that these children's language patterns in English for MS Accuracy were set by 2nd grade and then were less susceptible to being influenced by factors outside the home, or perhaps that the children were concentrating on gains with the greatest functional significance (which might be presumed to be Lexicon and Complex Syntax).

MS Accuracy alone showed a main effect of Lingualism and no interaction of Grade by Lingualism. Unlike Lexicon and Complex Syntax, grammatical well-formedness, then, appeared to be robustly different between MLs and BLs at both grades. The information in Figure 7.7 collapses across SES, but it should be noted that there were disproportionately large gains from 2nd to 5th for MLs among the Low-SES subgroup. MS Accuracy did enter into an interaction of Grade by LSH. The contrasting Grade by LSH patterns for the Language Score subcomponents are shown in Figure 7.8.

These results were paralleled by effects for the CHILDES measures (see Table 7.13 and values reported in Tables 7.2, 7.3, 7.4 and 7.5). For the Subor-

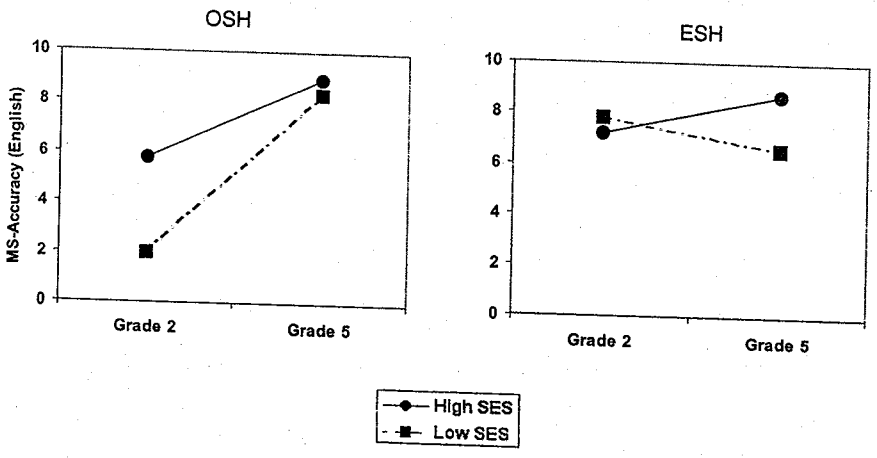

Figure 7.8 MS Accuracy (English), Grade by LSH by SES

dination Index (number of clauses per sentence) and MLU (number of words per utterance), there was also a Grade effect, but no Lingualism, LSH, or SES effect, although for SES there was a trend. For the lexical measure, Type-token ratio, there were both Grade and SES effects; for Types, a Grade effect.

Here the Type-token ratio followed the pattern for the Lexicon measure with significant effects of both Grade and SES. Subordination Index was most like the Complex Syntax measure and showed a Grade effect for Lingualism, but only a trend for SES. In general, despite statistical significance, the differences on these measures even across three grades were relatively small, and so the measures did not appear sensitive enough to characterize perceptible differences between 2nd and 5th grade stories. 'Clauses' and 'words', two basic measures of length, were particularly

Table 7.13 Main effects for 'CHILDES' measures

	Grade		*SES*	
	F	p	F	p
Subordination Index	35.58	< 0.001	6.51	< 0.011
MLU	38.41	< 0.001	6.88	< 0.009
Type-token ratio	14.25	< 0.001	22.5	< 0.001
Types	12.78	< 0.001	8.17	< 0.005

* Significant at the groupwise level, p < 0.0014

problematic: for the monolinguals, the older children's stories were superior to the younger children's, but they were 75 words and 10 clauses shorter on average. Among the bilinguals, the stories appeared to get better as they got longer, but even among bilinguals, the correlation between story quality ('Narrative Total') and length was relatively low, $r = 0.41$. While the correlation was significant, it explained less than 20% of the variance in story quality scores, so no further analyses were done with the CHILDES variables.

Differences between bilingual groups

English Scores. The six-way MANOVA showed no main effects of IMS in English for the bilinguals. Instead, LSH differences appeared more significant (Table 7.14). There were significant effects favoring the children with ESH in the Language Scores (both of them) and in the Narrative Total, but not in the Story Score. Of the subcomponents of the Language Score, there was a strong LSH Effect in Lexicon, but with the Bonferroni correction, neither the MS Accuracy nor the Complex Syntax score reached significance.

The main effect of LSH for Lexicon can be gleaned from Figure 7.7, which shows the LSH differences by Grade with ML English data for comparison. For MS Accuracy, there was an LSH by Grade interaction ($F = 23.49$, $p < 0.001$) and also a three-way interaction ($F = 10.21$, $p = 0.002$*) with LSH by Grade by SES which is graphed in Figure 7.8.

In MS Accuracy, both Hi and Low-SES groups of OSH children had higher scores at 5th grade than the 2nd grade groups. For the Low-SES children with ESH, the scores of 5th graders were lower than the scores of 2nd graders with respect to MS Accuracy. It is interesting to note that MS Accuracy showed a simple SES effect among the English monolinguals, also a decrement from 2nd to 5th grade. This difference between the LSH groups may indicate that the OSH children were still, at this stage, similar to second language learners, still improving their MS Accuracy, at both levels

Table 7.14 Home Language (LSH) effects in English Frog Story measures among Bilinguals

	F	p		F	p
Story Score	2.72	0.133	Complex Syntax	6.45	0.071
Language 1	15.36	0.000*	Lexicon	15.08	< 0.001
Language 2	10.44	0.001*	MS Accuracy	6.05	0.015
Frog Total	10.76	0.001*			

*Significant groupwise alpha (0.0033)

Table 7.15 Significant effects for Spanish Story measures

	Grade		*IMS*		*LSH*	
	F	*p*	F	*p*	F	*p*
Summary Scores						
Frog Total	15.64	< 0.001	9.84	0.002*		
Story Score	13.90	< 0.001	6.99	0.009		
Language Score	9.88	0.002*	7.14	0.008		
Component Scores						
Complex Syntax	22.55	< 0.001	2.20	0.141		
Lexicon	8.15	0.005	11.25	0.001*		
MS Accuracy	0.68	0.411	1.67	0.199	6.94	0.009
CHILDES Measures						
Subordination Index	11.86	0.001*	4.11	0.044		
MLU	9.85	0.002*	0	n.s.		
TTR	1.96	0.164	0.11	0.741		
Word Types**	21.86	< 0.001	2.09	0.151		

* Significant at groupwise alpha level (0.0033)
** Also in an LSH by Grade interaction, $F = 9.30$, $p = 0.003$.

of SES, whereas the ESH children, whose English was perhaps stronger to begin with, were more like the first language learners of English, showing an SES decrement.

Spanish Scores. When testing just the bilingual children with the General Linear Model, the SES effect was not significant for any of the Frog Story measures in Spanish, by contrast to the English analyses where strong SES effects pervaded the results. LSH, like SES, which showed an effect for four of the seven variables in English reported in Table 7.9, was significant (without Bonferroni correction) in Spanish only for MS Accuracy, F = 6.94, p = 0.009 (Table 7.15). Grade was a significant factor for several of the other measures, as was IMS. Thus, when other factors were held constant, Two-way schooling had a greater effect on Spanish proficiency than did the presence or relative lack of English in the home.

Figures 7.9 and 7.10 show the Spanish Narrative Total and Language Score subcomponents by Grade, IMS, and LSH, respectively.

It appears that the bilingual children in both learning situations improved in Spanish as they got older, but they did better in general in the Two-way schools. They did not appear to improve in MS Accuracy, even in

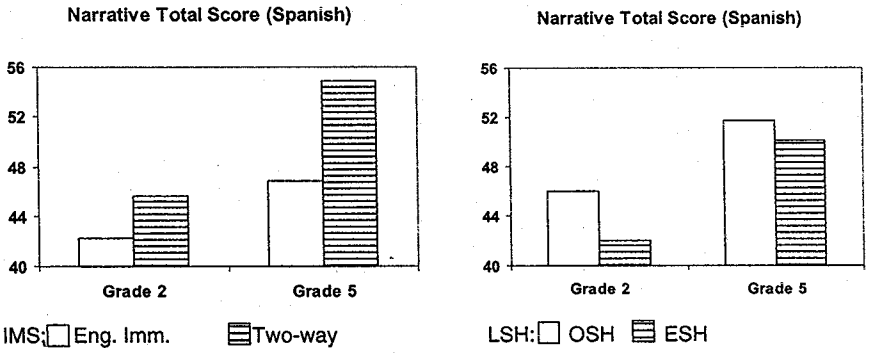

Figure 7.9 Narrative Total (Spanish), Grade by IMS and Grade by LSH

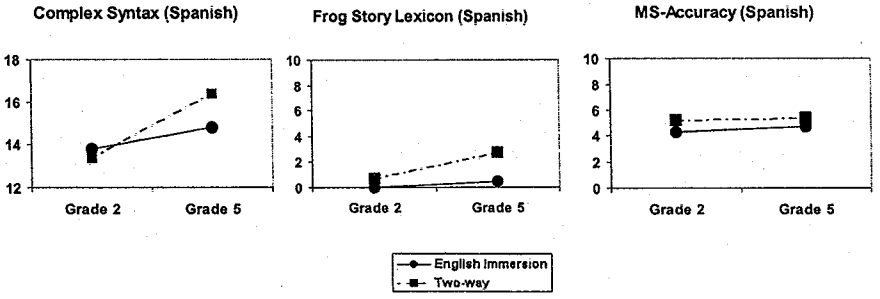

Figure 7.10 Subcomponents of Language Total Scores (Spanish), Grade by IMS

Two-way schools. Rather, IMS seemed to contribute especially to growth in Lexicon and Complex Syntax over this period. The instructional approach of the Two-way schools seemed to produce best results in the area of Lexicon and aspects of language that are tied to general intellectual development: lexical diversity and the expression of causes and effects, intentions, and complex temporal relations; in addition, positive effects were seen in the subordination index and MLU, reflecting the ability to link ideas in multiple clauses. Figure 7.9 illustrates the relative similarity of the two LSH groups at 5th grade in contrast to the relative difference between the IMS groups at 5th grade. Figure 7.9 also highlights the higher level of performance in narrative of the Two-way children relative to the English Immersion children especially at 5th grade. Figure 7.10 illustrates that it was primarily the Two-way children who showed growth across the

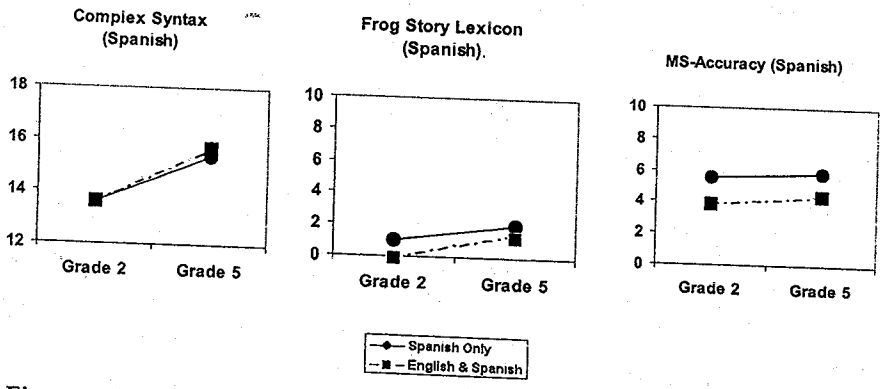

Figure 7.11 Language Score Subcomponents (Spanish), Grade by LSH

grades in Spanish for Complex Syntax and Lexicon. LSH, on the other hand, as illustrated in Figure 7.11, showed no strong effects on Complex Syntax or Lexicon. The OSH children, however, showed a notable advantage over ESH children on MS Accuracy at both grades, while MS Accuracy seemed little affected by IMS (Figure 7.10).

Following up on the one main effect of LSH, the one on MS Accuracy, it is notable that even though the differences were statistically reliable, there was considerable overlap in the two distributions, which both showed large standard deviations relative to the means for the two groups, $M = 6.27$, $SD = 4.5$ for the OSH group; $M = 7.55$, $SD = 3.3$ for the ESH group.

Within subjects results from correlation analysis

From the separation of the Frog Story measures into the two main aspects of the performance, narrative and linguistic, we can see the relation of the specific scores in one language to the same scores in the other. The correlation for the Narrative Total scores in English and Spanish was moderate, $r = 0.36$. However, some subcomponent scores showed much higher correlations than others. Correlations for selected subcomponent scores are presented in Figure 7.12, along with several additional correlations for reference and contrast, some of them based on data from the standardized tests (Chapters 4–6).

The data in Figure 7.12 suggest that Story Score and Complex Syntax correlate more highly across languages than Language Score or its subcomponents, Lexicon and MS Accuracy. On the other hand, the English Complex Syntax Score showed a significant degree of association to the Spanish Story Score (although the similar comparison within language is

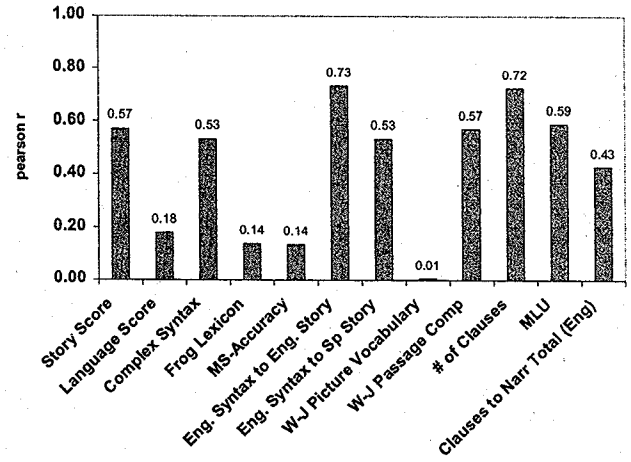

Figure 7.12 Correlations between Frog Story measures, standardized test scores, and CHILDES data (Spanish and English)

still higher, $r = 0.7$). This pattern is reminiscent of that found in Chapter 6, where literacy scores on standardized tests were highly correlated across languages while oral language scores were not so highly correlated. Literacy scores may have reflected more language-general factors because they involve general intellectual and conceptual capabilities while oral language scores may involve the learning of many individual items that are language-specific. In the case of the Frog Story data, it may be that Story Score and Complex Syntax represent relatively general capabilities that span languages, while it appears that Lexicon and MS Accuracy represent factors that need to be learned item by item in each language. While the cross-language correlations of Lexicon and MS Accuracy were very low, they were not negative, which would be one logical possibility if the two languages were in competition with each other, say for lexical storage area. Again the lack of negative correlations is consistent with the results in Chapter 6 which showed no significant negative correlations across languages.

Other objective measures, number of clauses and mean length of utterance (MLU), showed moderate correlation across languages. The correlation of number of clauses to Narrative Total was moderate, $r = 0.56$ for English and $r = 0.32$ for Spanish (not pictured). While they were both significant at the 0.01 level, they accounted for only 25% and 10% of the variance respectively in the Narrative Total scores. The correlations thus underline the need for more nuanced assessments of story quality than those quantitative measures can provide.

Table 7.16 Correlations between Frog Story measures and standardized test scores by IMS

		Eng. Imm. School	*Two-way School*	*Mono-lingual*
Story Score to Passage Comprehension		0.11	0.23	0.24
	to Dictation	0.06	0.03	0.24
	to Picture Vocabulary (production)	0.48	0.57	0.25
	to PPVT (receptive vocabulary)	0.39	0.48	0.21
Language Score to Passage Comp		0.25	0.27	0.31
	to Dictation	0.16	0.23	0.27
	to Picture Vocabulary	0.51	0.76	0.35
	to PPVT (receptive vocabulary)	0.64	0.65	0.37

Correlation analyses also addressed the question of how strongly the Frog Story measures were associated with other measures taken on the same children. Table 7.16 shows that the correlation was moderate at best (as it was for monolinguals). This indicates that the Frog Story measures gave different information from that given by the standardized scores from the Woodcock-Johnson Battery, and thus they are a useful adjunct to standardized tests in evaluating performance levels of monolingual and bilingual children. The patterns of relation were somewhat different for the bilinguals, with greater correlation for vocabulary and lower for Dictation (which may stress MS Accuracy). Correlations computed by splitting 2nd and 5th grade were very similar to these, split by IMS.

Discussion

Overall, these measures underline the multi-faceted nature of language development. With respect to the hypotheses stated in the Introduction, H1 and H3 were supported by these data; H2 was supported for Spanish, but not for English.

H1 posited that BL and ML differences would not be uniform across language behaviors. Indeed, the data showed that the Lingualism differences in favor of the monolinguals were much smaller for the narrative and discourse measures (Story Score) than for the Language Scores.

H2 posited that schooling in Spanish and English would enhance performance on complex tasks in both languages. The data indicated that bilingual children in Two-way schools showed enhanced performance in Spanish on both simple and complex language tasks. Their performance in English, however, was comparable in almost all areas to bilinguals' schooled in English Immersion, and superior in only a few domains.

H3 posited that narrative and linguistic abilities in one language would predict children's narrative and linguistic abilities in their other language. The results indicated a significant correlation between narrative skills (as indicated by the Story Score) and Complex Syntax scores across the two languages. By contrast, other language scores, MS Accuracy and Lexicon, were not reliably correlated across languages.

The bilingual children's stories exhibited age-appropriate skill in difficult tasks like creating a unified plot, motivating events through reference to internal states, and providing narrator's comments on the unfolding story. The BL children were accurate in using compound time referencing and embedded structures which distinguished their own thoughts from those of the characters. In these ways, their responses to the complex demands of the story genre were comparable to those of their monolingual peers'.

The BLs showed the greatest weakness relative to MLs in Lexicon and MS Accuracy. Except for the youngest children with the least exposure to English, who at 2nd grade were still very much like second-language learners of English, there was no implication that skill in English was negatively correlated with skill in Spanish. For the most part, Lexicon especially was low in both languages, suggesting that more time and input were needed in each language for the children to approach ML levels in vocabulary knowledge.

This is exactly what was observed. Although these are not longitudinal data, the older children among the bilinguals, the ones who, based on demographic information (Chapter 2) had more cumulative time and input in English, showed fuller lexicons than the younger children. The younger bilinguals appeared to have had, on average, less time and input in English environments than monolingual peers and the older bilingual cohorts. Measures that were low relative to those of monolinguals in 2nd grade tended to be comparable to those of monolinguals at 5th. For the most part, this Grade by Lingualism interaction on a number of measures can be traced to the lower starting point for bilinguals at the outset, and thus greater growth to reach a similar end point at 5th grade. Some of this 'catching up' might be attributed to relatively higher scores on these measures for the monolingual 2nd graders, and thus, less room for growth. But that does not appear to be the whole story. In fact (as seen in Table 7.5), simple effects when tested for Monolinguals-only showed significant grade effects, and thus growth among monolinguals, in each domain, in Story Score, in Language Score, and in the CHILDES measures. There appeared, then, to be growth in these

measures over this age range for all children, but faster growth for bilinguals, as seen by the numerous Grade by Lingualism interactions.

With respect to non-significance of several simple effects, as seen, for example, in the relative equivalence of performance of High-SES children across the Lingualism groups when tested only at 5th grade, one cannot discount the possibility that the failure to find a reliable difference between Lingualism groups may have been the result of making a comparison with relatively small groups. Keeping this limitation in mind, it is worth noting that MLs showed a large advantage on the Narrative Total score at 2nd grade, but not at 5th grade. Let us reconsider the hypothesis expressed in Chapter 1, that bilingualism of the sort studied here might show its greatest advantages for English learning among children of High-SES who also had ESH. For the children in these specific subgroups, the results indicated no reliable difference on the Narrative Total between MLs and BLs at 5th grade, and this was true even if the Bonferroni correction was waived. The scores across ML and BL children were very similar among High-SES/ESH children.

Conclusions about growth drawn from these data, though, must remain tentative as they were based on a cross-sectional design. Every effort was made to equate the educational, social, and linguistic background of the participants and to sample a large enough group that individual differences in ability and volubility could be adequately randomized. Still, without longitudinal information, one cannot rule out the possibility of a cohort effect.

Further, one might question whether the findings based on the Frog Story groups are generalizable even to the larger design, much less to bilingual children in general. The difficulty of data transcription and coding of these semi-naturalistic data restricted the size of the groups to 80 bilinguals at each age, plus 4 monolingual control cells equaling another 80 children. To test the representativeness of the story subgroups with respect to the groups of 30 and 40 in the full design, the mean scores for the story groups on the Woodcock-Johnson/Muñoz and Peabody standardized measures were compared to the mean scores for the larger groups from which they were drawn. In all cells the subgroup mean was within one-third of a standard deviation of the mean for the larger group (and most were within a standard error), with no pattern as to whether one was higher or lower. Therefore, it appears that these groups adequately sampled the groups of the full design.

It is not surprising then, that the overall results are similar in their broad outline to the results from the Woodcock and Peabody measures (Chapters 4 and 5). These include, in English, main effects of SES, Grade, and Lingualism, and on some variables, a Grade by Lingualism interaction. Likewise, for the factors nested within Lingualism, there were few Frog Story differences in

English between the levels of IMS; the strongest effects were seen on some subcomponents of the Language scores from the LSH factor, favoring children with ESH. In the Spanish stories, there were no differences by IMS at 2nd grade; but by 5th grade there were differences that favored the Two-way schools, especially in Spanish Lexicon and Complex Syntax (although Spanish scores in general were considerably lower than they were in English for all groups, even those weakest in English). Notably, there were no SES effects in Spanish and surprisingly few LSH effects.

As in the standardized tests, the bilinguals were less disadvantaged with respect to the more complex narrative scores and more disadvantaged with respect to the monolinguals on the more language-specific measures (especially MS Accuracy). The largest differences between groups in the Woodcock-Johnson Batteries were found in Picture Vocabulary, a productive measure like the Frog Story, a measure that also seems to emphasize language-specific knowledge. Picture Vocabulary was also a task where a Grade by Lingualism interaction obtained (see Chapter 4).

Crucially, the Frog Story measures *add* to the information provided by the other testing of these children. Since the correlations between Frog Story measures and the standardized scores were only moderate (in the neighborhood of 0.3 for the monolinguals, cf. Table 7.16), we can be relatively confident that these data give information about different aspects of the children's performance. Unlike the standardized tests from which we must infer the details about how the different groups talk, the Frog Stories provided samples of their productive language for direct comparison.

One surprise in listening to the children's stories was how restricted their Spanish was overall, even among the children for whom Spanish was the major language of the household. Only a handful of children were so weak in Spanish as to speak it with an English accent, but many of the children with ESH found it difficult to speak Spanish for the space of the whole story, without lapsing into English for lexical items and even whole clauses. Those with OSH were more fluent in their Spanish stories, but also resorted to paraphrase for lexical gaps, avoided the subjunctive in common expressions, and made salient L2 morphosyntactic mistakes. This weakness in Spanish may have been a result of our selection criteria which stipulated that all participants had to have been born in the United States, thus perhaps restricting the range of Spanish ability in our sample. This narrowed basis for selection was intended to permit a stronger assessment of the effects of bilingualism independent of immigration. However, it eliminated from consideration the estimated 25% of the student population at each grade level who were entering the country and the Dade County school system during the time of the study and, according to information on the DCPS website, continue to enter each year in even larger numbers

(http://www.dade.k12.fl.us/bfls/index.htm). This may be a group for whom native language instruction is more essential and more facilitating than for our subjects, and thus merits further research. Another possible explanation for the low Spanish scores observed in this study concerns the comparison of the contact variety of Spanish spoken in Miami by the largely bilingual Hispanic population with the dialects of Spanish spoken in the countries of origin. Even if specific elements of the language are the same, they may differ in their distribution in the speech of bilingual Spanish speakers as compared to monolingual Spanish speakers in monolingual contexts (Toribio, 2000). The effects of such differences on children's learning of the language are largely unknown.

For our subjects, the strongest educational implication of these findings is that the time spent learning in Spanish does not appear to harm the students' progress in English, but provides significant support for them in Spanish. The time spent learning Spanish in school seems to translate into learning of skills – reading, dictation, strategies for narratives – which appear available to the child in either language. But it does not seem to translate into exposure to more double-language vocabulary, for example. Vocabulary seems to be learned item by item, and has to be done in each language separately. With respect to language-specific learning of morphosyntax (such as irregular forms, or verb complement patterns), here IMS appeared to be a less potent factor than LSH. The children may have spent more time at home than at school. The effect of schooling on language-specific learning may have been weaker than that of home language because the same language-specific structures were being learned in both places, but more of the learning seems to have occurred in the home.

Importantly, the LSH effect was different for the children's two languages. English language skills were highest in children from High-SES families and ESH. IMS was a less crucial factor for English than for Spanish language development as seen in the data from this chapter, as well as Chapter 4. Two-way schooling was helpful for Spanish language development in all demographic groups, High and Low-SES, ESH and OSH.

The separate contributions of LSH, IMS, and SES background to children's language and literacy performance created a complex pattern of effects and non-effects. By using the factorial design of the larger study, which balanced the effect of each factor, we have enhanced our ability to generalize findings from the children's stories. Likewise, by expanding the performance demand on the children through the story task, we have provided an 'auditory snapshot' of each individual, to add to the perspective provided by the standardized test scores. This snapshot enriches our ability to understand what the test scores are saying and to have greater confidence in the messages they convey.

Appendix

FROG STORY CODING SHEET

Coded Subject No. _____ Lang/Order _____ Coder _____ No. of Clauses _____

Part 1. Story Score (Narrative/Discourse Elements)

Each category is worth 12 points. In each column, choose the 'highest' behavior that describes the story being coded, deducting for behaviors NOT shown, according to the scale in that column.

Frog Story Elements	*Sequence*	*Perspective/Affect*	*Engagement*
Mentions discovery of missing frog +2 Uses mental verb+1 'looks' only	0 Picture description	–2 Uses 'here' 'there' 'now' to refer to her own reference frame (not the story's)	
+2 Initiates search	3 Gives isolated events	Poor first mention (uses pronoun right away) –1 Main characters –1 Other characters	–2 Vague or confused (in parts)
+1 Finds frog	Sequential events (some, not all)	Lapses in reference (reader must ask 'who?') –1 1 defective reference –2 2 defective references –4 5 or more	Disfluencies –2 (grave, interfere with listener's ability to follow story)
+1 Takes frog (home)	–1 for picture description –1 for irrelevant details		
(6) AVERAGE	(6) FACTUAL STORY	(6) NEUTRAL OBSERVER	(6) MATTER OF FACT TONE
+1 Articulates goal (each, up to 2)	7 Elaborated episodes	7 Ascribes intention	Attempts to be lively or engaging +1 use of 'refrain' +1 appropriate exclamations
+1 Articulates lack of success		8 Gives internal state info (affective statements, 1 or 2)	
	9 Hierarchical structure (beginning/middle/end) (highlighting of an event)	9 3 or more affective statements	+1 extensive direct speech
+4 Notes character's misperception (branch/antler or other)	12 Retrospective or prospective summary +2 for summary statement +3 for 2 or more	12 Mentions ironic perspective	Uses figures of speech +2 each (up to 3)
_____/12	_____/12	Persp: __/6 Aff: __/6 _____/12	_____/12
			Total _____/48

Part 2. Language Score

Complex Syntax		*Lexicon*	*Morphosyntactic Accuracy*
Verb Phrase	*Between Clauses*	Circle if present by name	Errors –1 each 'type', not 'token' (note word and line)
	(4) NO CONJUNCTIONS (BEYOND 'AND THEN' 'Y PUES' O 'ENTONCES'	frog jar/bottle bees (wasps)	articles pronouns prepositions
(6) NATURAL FLOW OF EVENTS English: Modals or aspectual auxiliaries +1 each (up to 3) Eng. or Spa.: Complex Ss Relative Clause or Noun Clause +1 for 2, up to 3 points Spa: Perfect Tenses +2 for one +3 for 2 or more Spa: Subjunctive +2 for one +3 for 2 or more	Causal connections/ (or intentions) +1 for 1 or 2 +2 for 3 or more Simultaneous actions ('while'/'when'–simul) +2 for 1 +3 for 2 or more Retrospective reference 'still'/'already' 'todavia' +1 each, up to 3	beehive rodent-type owl trunk (Eng) deer antlers pond/lake log cliff Other 'specific' vocabulary: Other vocabulary mistakes:	verb forms conjunctions word order other
_____/12	_____/12 Complex Syntax: _____/24	Total _____/12 (+1 for each listed word; +½ for other specific words; –½ for other mistakes (0–12 min/max)	Total: Subtract from 12 _____/12
			Total _____/48
			Total _____/96

Chapter 8

Command of the Mass/Count Distinction in Bilingual and Monolingual Children: An English Morphosyntactic Distinction

VIRGINIA C. MUELLER GATHERCOLE

The following three chapters report experiments involving morphosyntactic elements of English and Spanish. The primary question addressed is the extent to which bilingual children follow the same processes of acquisition and the same timing in their acquisition of these elements as their monolingual peers. Central to this question is the extent to which the major variables of the core study – Instructional Method in the School (IMS), Socio-economic Status (SES), and Language Spoken in the Home (LSH) – play roles in affecting patterns or rates of acquisition. A secondary question is whether there is a difference in bilinguals' abilities with the acquisition of relatively superficial elements of the grammar (the mass/count distinction in English, gender in Spanish) and their abilities with an element attributed to Universal Grammar (UG) (*that*-trace phenomena). A comparison of bilingual and monolingual performance across these structures will help illuminate the extent to which bilingual acquisition follows paths that are similar to or different from those followed in monolingual acquisition; the extent to which school experience, SES, and LSH play a role; and the extent to which these effects are present in superficial versus purportedly universal aspects of the grammar.

The structures to be examined here are of interest because with each one, the two languages being learned by these bilinguals are vastly different. English has a mass/count distinction, Spanish does not. Spanish has a system of grammatical gender, English has natural gender. English does not allow *that*-trace structures, Spanish does. Children acquiring these two systems must develop distinct or complementary constructs in the two lan-