茨城大学教育実践研究 **33**(2014), 107-121

The Role of Pictures in Children's Comprehension of Repeated Read-Alouds of Picture Books in an EFL Setting

Hidetoshi SAITO* , Yumi KASAKURA**, Miyuki SATO, and Mayumi YAGIOKA***

(Accepted Sept. 16, 2014)

Keywords: children, read-alouds, picture books, learning English

Abstract

Visual information may play a key role in language learning, especially at the early stages of children's EFL learning. We hypothesized that novice learners turn mainly to visual cues to comprehend stories. To test this, seventeen novice Japanese children went through two English read-aloud conditions, normal read-aloud and pictures only. One book was read aloud to the children over five sessions. The other book was read aloud in Sessions 1 and 5, but only pictures from the book were shown in Sessions 2 through 4. Two raters assessed the children's story retelling using two rating forms. The results show that although comprehension in both conditions increased after the five sessions, there was no statistically significant difference in the story comprehension between the two conditions, although result on one measure approached statistical significance. Children's interest in the two books remained stable across the five sessions, regardless of conditions.

Introduction

When reading aloud a picture book to EFL novice children, we wondered to what extent they actually understand the story. Or if they understand it at all? We have always assumed that pictures help them understand the story, even if they could not understand the language they heard. This pilot study began from this assumption and sought to answer a question: How much do pictures help the children to understand a story?

Parents and teachers have recognized read-aloud as a fundamental language activity that helps children learn a first language (L1). Research also suggests that read-aloud has a positive influence on children's L1 literacy, vocabulary (e.g., Whitehurst et al. 1988), reading comprehension (e.g., Wells 1986), and oral skills (e.g., Blok 1999) development. In reading aloud picture books, pictures seem to play more than an assistive role in both perception and comprehension. Beyond the positive influence on lower level cognitive processes such as recovery of

*Ibaraki University **Boyo Elementary School ***Nakaminato Junior High School

degraded sound (Riseborough, 1981), visual information including gestures and facial cues seem to improve second language (L2) listening comprehension of the story (Sueyoshi and Hardison, 2003; Wagner, 2010).

However, there is little basis for generalizing these results to novice EFL learners' comprehension of picture books read-aloud. This is because existing studies often assume that L2 learners can still comprehend the text through auditory input alone. In the very early stages of L2 learning, this assumption is not totally warranted. True novices may face considerable difficulty in extracting even a small piece of information in the auditory input. It can be expected then that the novices in the early stages of L2 learning turn heavily to visual cues, if available, to make sense of mostly incomprehensible auditory input.

The fundamental tenet of the picture books read-aloud is that pictures make the input more comprehensible for L2 learners because of visual cues. Despite all of the potential benefits of read-alouds including receptive and productive vocabulary, grammar, comprehension, and repetition through oral imitation (Amer 1997; Dhaif 1990; Elley 1989; Romney, Romney, and Braun 1988; Silverman 2007; Vivas 1996), previous studies have not clearly delineated the role of pictures in comprehension in the very early stages of L2 learning.

In this study, we propose the Visual Dominance Hypothesis, suggesting that learners in the early stages of L2 learning rely on visual cues (including pictures in picture books) in order to comprehend the auditory text. That is, we suggest that visual cues mediate the relationship between the meaning and form through which auditory input is made comprehensible. We hypothesize that role of visual cues in comprehension is much greater during the very early stages of L2 learning because the learner in these stages possesses limited preexisting linguistic knowledge to help map out the form-meaning relationship. Although no direct evidence exists to support this hypothesis, Tomasello's theory on the role of comprehension and attention in child L1 acquisition strongly supports this proposition. Based on empirical evidence, Tomasello (2003) theorized that children's comprehension of language, hence L1 acquisition, takes place in the joint attentional frame in which the adult, the child, and the event or object—such as pictures—share the referential event of immediate relevance, and the child attempts to read the adults' communicative intentions. In this case, the child associates the linguistic form with intended meaning by attending to the event or object at hand, and visual input seems to be the sole provider of the meaning of the linguistic form.

In the present study, we tested the following working hypothesis based on the Visual Dominance Hypothesis, that is, we suggest that children's comprehension of a story will increase both when the picture book is read aloud normally and when only the pictures are shown. We expected that if children in the early stages of L2 learning were incapable of making full use of auditory information because of their limited proficiency, comprehension based on seeing only pictures would be no different from comprehension based on normal read-alouds. In other words, if auditory input contributed very little to story comprehension, the listener would use pictures as the default main source of comprehension, regardless of the availability of auditory input. If so, the

power of pictures or visual input—the roles of which much underrated in language learning—will be supported.

Method

Participants

A total of seventeen children (6 second- and 11 third-graders, ages 8-10), all native speakers of Japanese studying English as a foreign language in public elementary schools, participated in the study (Table 1). The children's experience in learning English, both in and outside elementary school, was limited to an average of 3.45 years (*SD* = 1.35). The children in Groups 1 (n = 6) and 2 (n = 5) had one 45-minute lesson and one 20-minute lesson each week at their school while the children in Group 3 (n =6) had only one 45-minute lesson per week at their school. Two children reported that they took an extracurricular English course, which met once a week for less than one hour. None of the participants had spent extensive time overseas and none reported that they could read English words. In an informal test given to five of the children, none could recognize simple words such as "apple." Three senior English major university students served as readers of the books.

Table 1. *Profile of the Children.*

	Group 1	Group 2	Group 3	Average
Grade	3	3	2	
Gender	3 (M) 3 (F)	2 (M) 3 (F)	4 (M) 2 (F)	
Mean Length of Study (*SD*)	3.58 (1.20)	4.26 (1.60)	2.66 (0.98)	3.45 (1.35)
Meeting Site	public school	public school	private institute	

Note. The mean length of study is the average cumulative years of studying English.

Procedure

Data collection took place during after-school child-care programs offered in their own school to Groups 1 and 2 on separate occasions and in a private institute to Group 3 after regular school classes had ended for the day. Grouping was arbitrary, and all groups participated in both task conditions. The data collection procedures were identical for the three data sets, with the exception of slight differences in the read-aloud session schedules.

All children in the study were exposed to enlarged copies of two picture books during five sessions conducted over a period of approximately four to five weeks. Each group read-aloud session (including individual recalls) lasted approximately thirty to forty minutes. For scheduling convenience, the intersession intervals were six days except for two one-day intervals for Group 3. In three cases, participants missed one of the sessions, and individual make-up sessions for these

children were held prior to the following scheduled group session.

Before the first session, two books—not those used in the actual experiment—were read aloud to Groups 1 and 3, and the children individually retold the text. Group 2 did not have such rehearsal due to scheduling conflicts. This might have lowered their scores in Session 1. Non-parametric Kruskal-Wallis tests were run to check group differences in Session 1. Group 2 had statistically significantly worse scores than Group 1 on two out of four measures. On the other hand, there were no statistically significant difference between Groups 2 and 3, while the scores on all four measures were statistically significantly different for Groups 1 and 3. Because of these differences among groups in Session 1, we decided to use difference (or gain) scores for analysis.

For Groups 1 and 3, the dino book (*Happy Birthday, Danny and the Dinosaur* by Hoff, 1995) was read aloud while the horse book (*The Horse in Harry's Room* by Hoff, 2001) was chosen for the silent-read. The books were reversed for Group 2 so that the horse book was used for the read-aloud and the dino book for the silent-read. They seemed appropriate for the study because they were short, easy, and reflective of everyday experience while still containing fantasy elements. Both books belong to the lowest level of the reader series as graded by the publisher. Readability indices, such as Flesch grade, indicate that the two books are approximately equal in terms of the surface textual features. The book used in the read-aloud condition will henceforth be referred to as Book A and that of the silent-read condition as Book B.

In the read-aloud condition, Book A was read aloud, and the pictures were shown to children in all five sessions. In the silent-read condition, the reader read aloud Book B with pictures shown only in Sessions 1 and 5. In Sessions 2 through 4, the reader did not read the text, but only showed the pictures in Book B. To keep the pace of showing the pictures consistent with the normal read-aloud, the reader in the latter condition subvocalized the text, i.e., read in their mind. To control for potential factors that might influence comprehension, the reader avoided making any gestures or asking any questions in either condition.

In Sessions 1, 3, and 5, each child individually retold the stories in Japanese to three assistants, who began by prompting the child to "Tell the story as much as you can remember" and ended with "Anything else can you remember?" Recalling was done outside of the classroom, and the assistants made sure that children would not overhear one another. Each story retelling took approximately one to three minutes. Since there were only three assistants at each site, two or three children had to wait a few minutes for their turn to retell the story. To correct for this waiting time, the order of retelling was adjusted so that each child had to wait for at least one of the two retellings in each session.

If the child would not say anything or remembered very few story elements, the assistant prompted by asking, "Who was in the story?" and "What were they doing?" Other than these questions, the assistants refrained from assisting children with retelling. All retellings were audio-recorded and transcribed for analysis. A major reason for choosing oral story-retelling over written recall is that children at this age have limited writing skills, even in L1, hence it may take

them longer, and it may be more difficult for them to fully express their story comprehension in writing.

Scoring

Two scoring instruments, the story structure analysis (SSA) form and the key concept analysis (KCA) form, were developed to analyze the children's comprehension of the stories in the books. Both forms were expected to detect different aspects of story retelling, SSA for structural organization of the story and KCA for the storylines in detail. The SSA was adapted from a form developed by Morrow (1986). The story was first parsed into four story structure elements: setting, plot episodes, resolution, and sequence. The elements were determined before the data analyses based on the discussions among the researchers. Scoring the SSA form involves identifying the presence or absence of these story structure elements in the retelling (Appendix A). For example, if the child successfully mentioned the main character's name and began the retelling with an introduction, the child earned 2 points for setting. There were seven items in total, with each rated either dichotomously (items a, b, f) or on a four-point (items c, e, g) or five-point scale (item d).

The second analysis, the KCA form, was based on Bailey's (1998) approach to scoring the dictogloss. Each item on the KCA form refers to a key concept that combines and summarizes several sentences in the book (Appendix B). The analysis involves identifying the presence or absence of these pieces of information in the retelling. The KCA form was constructed in two phases. In the first phase, two research assistants independently grouped all the sentences in each book into thirteen concept groups. Then they independently chose four major concept groups, which were assigned a weight of 2 in the subsequent analysis. Any disagreements were discussed and reconciled. In the second phase, two of the authors independently summarized and shortened each of the thirteen concept groups and reached a consensus regarding the final form. All items were rated on a three-point scale ranging from perfect, to relevant, to not relevant. Using both the SSA and KCA forms, two raters independently rated all of the transcripts in random order, with the children's names replaced by numbers.

Although irrelevant to the main research question, a concern arose after launching the data collection that children might get bored from listening repeatedly to the same story. Thus, the level of interest in the books was monitored for six children in Group 3, as this decision was made before collecting their data. While waiting for their turn to retell, they completed a four-item questionnaire assessing how well the book maintained their interest during repeated listening.

Analyses

All data from the three groups were combined for the statistical analyses. The data from Group 1 collected during Session 3 was inadvertently lost; therefore, all Session 3 data were eliminated in the subsequent analyses, although they were used to generate comprehension measures in the Rasch analyses. Difference (or gain) scores between Sessions 1 and 5 were used as dependent

variables because of group differences in Session 1 and the ease of interpretation. Because the homogeneity of variance assumption was violated, three non-parametric analyses were done with a Bonferroni corrected p-value set at (.016). First, the length of studying English in months (Length), the sole between-subject independent variable based on a median split, was tested with a non-parametric MANOVA on SSA and KCA comprehension measures. This potential moderating variable that could reflect proficiency turned out to be statistically non-significant, F(1)=.24 (p = .89). Thus, the main results reported below came from Friedman's ANOVA of the SSA and KCA measures with one within-subject independent variable, Condition (read-aloud and silent-read).

Multi-faceted Rasch analyses were used to generate the comprehension measures for both the SSA and KCA. Rasch analysis is capable of modeling different measurement facets, such as rater severity (differences in the severity of raters), item facility (differences in the ease of rating items), and book difficulty (differences in the difficulty of books), into a logit linear function. Use of Rasch analysis allows the researcher to generate measures that take into account differences in these measurement facets. This is advantageous for analysts because, as these measurement facets are statistically already controlled for, the subsequent ANOVAs can safely disregard them, which could otherwise confound the results. In short, both the SSA and KCA measures were derived from the raw scores by taking into account rater severity, item facility, and book difficulty.

Rasch analysis has several strong assumptions, one of which is unidimentionality. The Rasch model assumes that the items in the analysis measure a single latent variable and that the fit statistics flags any aberrant item (or person). Although Rasch analysis requires a large sample size to achieve reliable generalizability of the test results, descriptive use is permissible for a small sample size study such as the present case.

Results

The results of the Rasch analyses show (1) strong exact agreement between the raters (93.8% for the SSA and 92.3% for the KCA); (2) high person and item reliabilities for the SSA (.92, .95 respectively) and the KCA (.85, .95, respectively); and (3) no difference in the difficulty of the two books, $t = 2.53$, $p > .05$, for the SSA, and $t = 2.60$, $p > .05$, for the KCA. All participants and items remained in the subsequent analyses, although one child and two items clearly misfit.

Table 2 presents the descriptive statistics for the raw scores, Rasch measures, and difference scores for the SSA and KCA. As seen, both SSA and KCA increased from Sessions 1 to 5, regardless of conditions. The results of Friedman's ANOVA suggest that Condition (read-aloud versus silent-read) is not statistically significant, $\chi^2(1) = 2.88$ ($r = 0.69$), p > .01 for SSA and $\chi^2(1) = .05$ ($r = 0.01$), p > .01 for KCA, respectively. Although the results support the hypothesis, the probability for the result of SSA is approaching statistical significance (p = .09) with a moderate effect size *r*.

Table 2. Descriptive Statistics for the Story Structure Analysis (SSA) and Key Concept Analysis (KCA).

Table 2. *Descriptive Statistics of Rasch and Raw Scores.*

		Story Structure Analysis						Key Concept Analysis					
		Session 1		Session 5		Difference		Session 1		Session 5		Difference	
		M	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*	*M*	*SD*
Read-aloud	Rasch	-2.55	2.98	-.00	3.62	2.55	1.67	-2.43	1.86	-1.79	1.97	.63	1.49
	Raw	4.85	3.27	7.50	3.55	2.64	1.75	4.76	3.88	6.52	4.29	1.76	3.47
Silent-read	Rasch	-1.87	3.04	-.46	3.21	1.41	1.38	-2.63	2.10	-2.12	2.29	.51	.82
	Raw	5.94	3.76	7.76	4.23	1.82	1.91	5.29	4.17	7.00	5.28	1.70	2.59

Note. $n = 17$. The maximum raw scores are 16 for the SSA and 34 for the KCA. Rasch is logit measures generated by the Rasch analysis. Raw is raw scores. Difference is difference scores between Sessions 1 and 5.

Follow-up Analyses

In both the SSA and KCA measures, the comprehension of read-aloud and silent-read books increased over the five sessions. This suggests the possibility that the children recalled the most information from the pictures, and thus could correctly retell the stories without understanding a great deal of the auditory information. It is possible that the scoring items are biased in favor of visual information. Estimating the extent to which the scoring items reflect visual information thus sheds light on this bias. For this goal, four raters employed a modified Angoff procedure—a standard-setting method in licensure exams—by which the raters estimated the probability with which children answered each item in the two scoring forms successfully based on inferences from the pictures alone. The probability of "inferability" of the item ranged from 0 (no picture based inference possible) to 100% (perfect picture based inference possible).

The interrater reliabilities (intraclass correlations) of item inferability were high among four raters (.88 to .95). The average SSA item inferability was found to be approximately the same for both books (.47), but the average KCA item inferability was less for the horse book (.37) than for the dino book (.50), indicating that it was more difficult to infer the horse book story from pictures alone. The correlations between the SSA item facility measures and inferability were moderate (.52 and .64 for the Dino and Horse books of SSA and .38 and .66 for the Dino and Horse books of KCA, respectively) (Table 3). Although statistical significance was not so meaningful here due to the small number of items in each analysis, the KCA items for the horse book were statistically significant ($p < .05$). These medium-sized correlations indicate that the item inferability could account for approximately 15 to 44% of the variance in item facility of the two analysis forms and that the item facility reflects, to a moderate degree, how easy the children found it to respond to the items in the two forms using pictures alone.

Table 3. *Interrater Reliability (intra-class correlation), Means and Standard Deviations of Item Inferability, and Correlations between Item Inferability and Rasch Item Facility.*

Form	SSA		KCA	
Book	Dino	Horse	Dino	Horse
Interrater	.92	.95	.88	.89
Mean (SD)	.47 (.33)	.47 (.37)	.50 (.33)	.37 (.32)
Correlation	.52	.64	.38	.66*

Notes. SSA =Story Structure Analysis Scoring Form, KCA = Key Concept Analysis Scoring Form. *p < .05

Other Analyses

As seen in Figure 1, the interest of Group 3 in the books remained stable across the five sessions—with an average interest of over three on the four-point scale. Although their interest in the silent-read condition slightly decreased towards the end, the effect of Condition on the children's interest in the books seems to have been trivial, at least for Group 3.

Two additional observations reveal the advantages of combining the visual and auditory cues. First, in the later sessions, Groups 1 and 2 remembered and spontaneously uttered aloud, along with the reader, the salient phrases in the read-aloud condition. Group 1 counted aloud "1, 2, 3, 4, 5, 6," in response to a scene in the story where the children are counting the candles on the birthday cake. Group 2 shouted "Giddyap" in response to a scene where Harry, the protagonist, shouts this word while riding his invisible horse. This suggests that words were easier to remember in read-aloud conditions than were the phrases in the silent-read condition, although talking-aloud with the reader was not equivalent to comprehension.

Second, one child in Group 3 made two noteworthy retellings. His retellings did not accurately reflect the story; instead, he reported what he understood the meaning of certain words to be. In the first session of the read-aloud condition, he said, "At first, I didn't understand what *dinosaur* meant, but …when *kyoryu* [the Japanese word for *dinosaur*] appeared, I could hear that English word, so I thought that the word means *kyoryu*" [authors' translation]. In the fifth session of the read-aloud condition, the same child said, "*Children* is where [sic] many people are. So I thought the word means *everybody*…" [authors' translation]. These statements reveal the moment in which the child noticed the meaning of the word by matching the picture and sound in the picture book. Together with the observation of children simultaneously talking aloud with the reader, these examples emphasize the unsurprising possibility that more opportunities for matching the text and pictures may lead to better comprehension, and possibly acquisition, of certain aspects of the text. However, this apparent advantage of the read-aloud condition induced a statistically significant difference only for SSA, not for KCA.

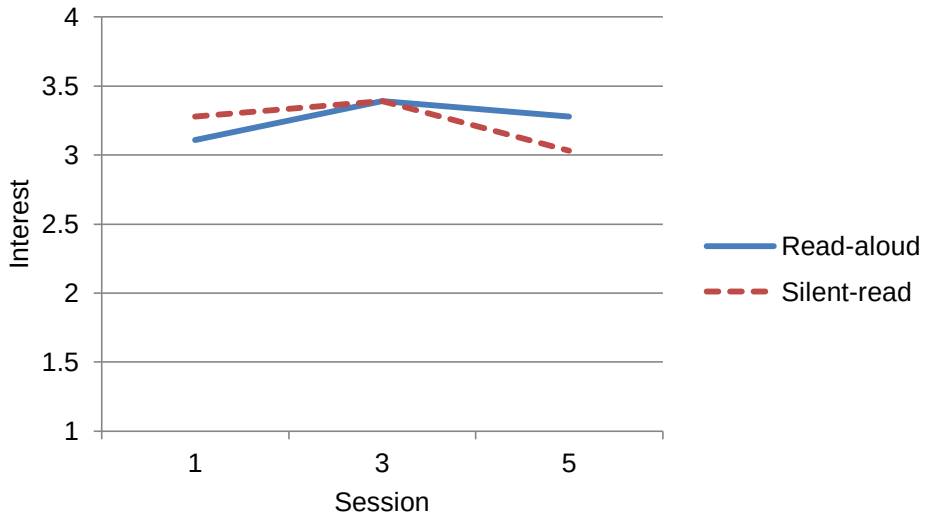

Figure 1. Mean scores of group 3's (n = 6) interest in books at sessions 1, 3, and 5.

Discussion and Conclusions

With respect to the hypothesis testing, the results of this study revealed that although children with limited L2 proficiency in an EFL setting demonstrated increased comprehension of picture books following both repeated read-alouds and silent-read, their comprehension following read-aloud condition did not differ from that following the silent-read condition in Session 5. Although the hypothesis was supported, the results of the SSA analysis approached statistical significance, thus implying the advantage of read-aloud condition over silent-read. A follow-up analysis, demonstrating the moderate inferability of the two books and its moderate shared variance with item facility, suggests that the pictures substantially increased the children's understanding of the stories, leading to increased scores on both the KCA and SSA. Additional evidence, that is, children talking aloud with the reader and one child reporting on the match between pictures and sounds, is in favor of normal read-alouds.

These results conform to the findings of previous research that showed steady increase in comprehension through repeated listening to read-alouds with L1 children (Dennis and Walter, 1995) as well as through repeated L2 listening in L2 (Cervantes and Gainer, 1992). However, in the present study, pictures, not just texts, seemed to be the major source of comprehension as indicated by increased comprehension in both conditions. This is consistent with Tomasello's (2003) theory of child L1 acquisition, which postulates that children interpret the intended meaning from the immediate object or event shared with the adult. Although the results of the present study support the Visual Dominance Hypothesis, there are several issues to consider.

First, SSA and KCA detect different aspects of story retelling performance. The fact that the results of SSA approach statistical significance suggests that pictures might help understand a basic storyline but understanding structural organization of the story requires auditory input. If this

interpretation is the case, it has tremendous implications for theory building. Alternatively, however, KCA may simply be less sensitive, for reasons unknown, to changes in comprehension than SSA. This issue in sensitivity of measures is a perennial problem in educational research, and it is the very reason for recommending the use of plural measures. These two should be compared to other measures of comprehension to test their sensitivity.

Another issue is the number of repetition of read-aloud. Although there was no difference in comprehension between the two conditions after three repetitions of silent-read in the present study, it is possible that as the repeated read-alouds continue, the combination of auditory and visual input could provide more clues for story comprehension than pictures. Research with longer periods of data collection is definitely needed to confirm this, although there is a concern that children might become bored if silent-read lasts for an extended period.

Moreover, the Visual Dominance Hypothesis itself needs critical appraisals. For example, the hypothesis predicts that the dominant use of visual information over auditory information happens in the early stages of L2 learning but the "earliness" is not clearly specified. In the present study, the estimated average number of hours of learning is 138.35 (SD = 54.27), thus, at least the initial exposure of up to around 200 hours could be covered in this study. Another daunting task is to define the term "dominance." The extent to which each of the visual and auditory inputs contributes to comprehension is far from clear-cut.

The present study has several limitations in addition to the small sample size. First, the lack of a true control condition—such as no read-aloud—precludes a straightforward interpretation of the results. The silent-read in the present study is only a partial silent-read.

A second limitation of the study concerns the possibility that methodological variations—varied intersession intervals, the lack of rehearsal in one group, and three different readers— across the three groups might have added some noise to the data. In addition, although the three measurements facets—book difficulty, item facility, and rater severity—were statistically controlled for by using the Rasch model to generate the two comprehension measures, the inferability of the scoring items was not.

Another limitation is the possibility that information leaked among the children if a child told the story to others outside of the sessions. Although it was impossible to control for this possibility in a naturalistic setting such as this, we strongly doubt that this leakage occurred because the children did not receive any incentives for improved retelling and did not know the research objectives. In addition, none of the children showed inflated improvement in their retelling scores. With these limitations, the results of the present exploratory study should be interpreted with caution. Further research with methodological improvement should confirm or disconfirm the preliminary results of the present study.

Finally, we hope that the results of the present study will encourage EFL teachers to explore the possibility of repeated read-alouds from the same book without fear of inducing boredom or an unfruitful outcome. Repeated read-alouds of the same book do not just allow them to relive the

same experience, but also give them a new experience each time the book is read. This happens probably because they do not have to allocate their limited attention and memory to the already-known information in the book, thereby freeing up these cognitive resources for new discoveries and deeper understanding. Six children in Group 3 who remained interested in the stories over the five sessions, regardless of condition, support this conjecture. Even after five sessions, their comprehension had possibly not yet saturated—leaving more gaps to fill in their understanding of the stories.

Authors' note

The order of the second, third, and fourth authors is alphabetical.

References

Amer, A. A. (1997). The effect of the teacher's reading aloud on the reading comprehension of efl students. *ELT Journal, 51*(1), 43-47.

Bailey, K. M. (1998). *Learning about language assessment: Dilemmas, decisions, and directions*. Cambridge, MA: Heinle & Heinle.

Baltova, I. (1994). The impact of video on the comprehension skills of core french students. *The Canadian Modern Language Review, 50*(3), 507-531.

Blok, H. (1999). Reading to young children in educational settings: A meta-analysis of recent research. *Language Learning, 49*(2), 343-371.

Brett, P. (1997). A comparative study of the effects of the use of multimedia on listening comprehension. *System, 25*(1), 39-53.

Cervantes, R., & Gainer, G. (1992). The effects of syntactic simplification and repetition on listening comprehension. *TESOL Quarterly, 26*(4), 767-770.

Coniam, D. (2001). The use of audio or video comprehension as an assessment instrument in the certification of English language teachers: A case study. *System, 29*, 1-14.

Dennis, G., & Walter, E. (1995). The effects of repeated read-alouds on story comprehension as assessed through story retellings. *Reading Improvement, 32*(3), 140-153.

Dhaif, H. (1990). Reading aloud for comprehension: A neglected teaching aid. *Reading in a Foreign Language, 7*(1), 457-464.

Elley, W. B. (1991). Acquiring literacy in a second language: The effect of book-based programs. *Language Learning, 41*(3), 375-411.

Ginther, A. (2002). Context and content visuals and performance on listening comprehension stimuli. *Language Testing, 19*(2), 133-167.

Gruba, P. (1993). A comparison study of audio and video in language testing. *JALT Journal, 16*(1), 85-88.

Hoff, S. (1995). *Happy birthday, Danny and the dinosaur*. New York: Harper Collins.

Hoff, S. (2001). *The horse in Harry's room*. New York: Harper Collins.

Morrow, L. M. (1986). Effects of structural guidance in story retelling on children's dictation of original stories. *Journal of Reading Behavior, 18*(2), 135-152.

Morrow, L. M. (1988). Young children's responses to one-to-one story readings in school settings. *Reading Research Quarterly, 23*(1), 89-105.

Mueller, G. A. (1980). Visual contextual cues and listening comprehension: An experiment. *The Modern Language Journal, 64*, 335-340.

Riseborough, M. G. (1981). Physiographic gestures as decoding facilitators: Three experiments exploring a neglected facet of communication. *Journal of Nonverbal Behavior, 5*(3), 172-183.

Romney, J. C., Romney, D. M., & Braun, C. (1988). The effects of reading aloud in French to immersion children on second language acquisition. *The Canadian Modern Language Review, 45*(3), 530-538.

Rost, M. (2002). *Teaching and researching listening*. Harlow, UK: Pearson Education.

Silverman, R. (2007). A comparison of three methods of vocabulary instruction during read-alouds in kindergarten. *The Elementary School Journal, 108*(2), 97-113.

Sperber, D., & Wilson, D. (1995). *Relevance: Communication & cognition* (2nd ed.). Oxford, UK: Blackwell.

Sueyoshi, A., & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning, 55*(4), 661-699.

Tomasello, M. (2003). *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.

VanPatten, B. (2007). Input processing in adult second language acquisition. In B. VanPatten & J. Williams (Eds.), *Theories in second language acquisition: An introduction* (pp. 115-135). Mahwah,

NJ: Lawrence Erlbaum.

Vivas, E. (1996). Effects of story reading on language. *Language Learning, 46*, 189-216.

Wells, G. (1986). *The meaning makers: Children learning language and using language to learn*. Portsmouth, NH: Heinemann.

Whitehurst, G. J., Falco, F. L., Lonigan, C. J., Fischel, J. E., DeBaryshe, B. D., Valdez-Menchaca, M. C., et al. (1988). Accelerating language development through picture book reading. *Developmental Psychology, 24*(4), 552-559.

Yaden, D. B. J., Smolkin, L. B., & Conlon, A. (1989). Preschoolers' questions about pictures, print conventions, and story text during reading aloud at home. *Reading Research Quarterly, 24*(2), 188-214.

Appendix A. Scoring Form for the Story Structure Analysis of Hoff (1995).

Directions: Place a 1 next to each element if the child includes it in his/her retelling (Credit 1 a gist as an obvious recall).

Setting

a. Begins story with an introduction
(e.g. Danny meets his friend the dinosaur) _____

b. Names main character (Danny) _____

c. Number of other characters named (The dinosaur, Danny's friends, Danny's father, Danny's mother, a girl, and a boy) 1 (1-2) 2 (3-4) 3 (5-6) _____

d. Number of statements about time and place (maximum raw score 4)
(Danny and the Dinosaur's birthday, the museum, the birthday party, and Danny's home) _____

Plot episodes

e. Number of episodes recalled (maximum raw score 3)
(1: Danny invites the dinosaur to his birthday party. They pick up Danny's friends on the way to the party. 2: The dinosaur does a lot of funny things at the party, such as singing a song poorly, playing the game in a funny way, putting his feet out the window, and eating a lot of ice cream. 3: As soon as the birthday cake comes, the dinosaur tries to eat it, but Danny stops him.) _____

Resolution

f. How the story ends (maximum raw score 1)
(They make a wish that they can be together next year and feel that it's the best birthday party they have ever had.) _____

Sequence

g. Retells story in structural order: Setting, plot episodes, resolution (Score 3 for complete, 2 or 1 for partial, 0 for no sequence evident). _____

Highest score possible 16

Child's score _____

(Adapted from Morrow, 1986)

Appendix B. Scoring Form for the Key Concept Analysis of Hoff (1995).

	2 Perfect	1 Relevant	0 Not relevant	Total
1. Danny goes to see his friend the dinosaur.				
2. Danny invites the dinosaur to his sixth birthday party, and the dinosaur says "I'd be delighted."				
3. Danny leaves the museum, riding on the back of the dinosaur. On the way, they pick up Danny's friends.				
4. The dinosaur says its his hundred millionth birthday, and Danny says "It's your birthday too."				
5. The children and the dinosaur help Danny's father hang balloons.				
6. The dinosaur puts on a party hat from Danny's mother and asks "How do I look?"				
7. A boy and girl and the dinosaur sing. Everybody covers their ears.				
8. The dinosaur pins the tail on his tail during the game.				
9. While the children are taking a rest, Danny says "Don't put your feet on the furniture." The dinosaur puts his feet out the window.				
10. Danny's parents give ice cream to all the children. They have to give a lot to the dinosaur.				
11. The birthday cake arrives, and they count the candles "1, 2, 3." The dinosaur begins eating it.				
12. Danny stops and makes a wish with the dinosaur, saying "I wish we can all be together again next year."				
13. They blow out the candles and sing Happy Birthday. Danny and the dinosaur say that it's the best birthday party ever.				

Note. Items 2, 4, 12, and 13 were assigned a weight of 2. The highest possible score was 34.

Appendix C. Questionnaire on interest in books.

Circle the one that best matches your feeling.

1. Did you enjoy hearing this story?

I enjoyed it a lot.	I enjoyed it some.	I didn't enjoy it much.	I didn't enjoy it at all.

3. Do you want to hear this story again?

I really want to.	I want to.	I don't want to.	I never want to.

5. Are you tired of hearing this story?

Not at all.	Not really.	A bit tired.	Really tired

6. Were you able to concentrate on listening to the story?

Very much so	Yes, most of the time	Not much	Not at all

Note. The original form was written in Japanese and included two additional questions about comprehension of the story.