The Efficacy of Fast ForWord Language Intervention in School-Age Children With Language Impairment: A Randomized Controlled Trial

Purpose: A randomized controlled trial was conducted to compare the language and auditory processing outcomes of children assigned to receive the Fast ForWord Language intervention (FFW-L) with the outcomes of children assigned to nonspecific or specific language intervention comparison treatments that did not contain modified speech.

Method: Two hundred sixteen children between the ages of 6 and 9 years with language impairments were randomly assigned to 1 of 4 conditions: (a) Fast ForWord Language (FFW-L), (b) academic enrichment (AE), (c) computer-assisted language intervention (CALI), or (d) individualized language intervention (ILI) provided by a speech-language pathologist. All children received 1 hr and 40 min of treatment, 5 days per week, for 6 weeks. Language and auditory processing measures were administered to the children by blinded examiners before treatment, immediately after treatment, 3 months after treatment, and 6 months after treatment.

Results: The children in all 4 conditions improved significantly on a global language test and a test of backward masking. Children with poor backward masking scores who were randomized to the FFW-L condition did not present greater improvement on the language measures than children with poor backward masking scores who were randomized to the other 3 conditions. Effect sizes, analyses of standard error of measurement, and normalization percentages supported the clinical significance of the improvements on the Comprehensive Assessment of Spoken Language (E. Carrow-Woolfolk, 1999). There was a treatment effect for the Blending Words subtest of the Comprehensive Test of Phonological Processing (R. K. Wagner, J. K. Torgesen, & C. A. Rashotte, 1999). Participants in the FFW-L and CALI conditions earned higher phonological awareness scores than children in the ILI and AE conditions at the 6-month follow-up testing.

Conclusion: Fast ForWord Language, the intervention that provided modified speech to address a hypothesized underlying auditory processing deficit, was not more effective at improving general language skills or temporal processing skills than a nonspecific comparison treatment (AE) or specific language intervention comparison treatments (CALI and ILI) that did not contain modified speech stimuli. These findings call into question the temporal processing hypothesis of language impairment and the hypothesized benefits of using acoustically modified speech to improve language skills. The finding that children in the 3 treatment conditions and the active comparison condition made clinically relevant gains on measures of language and temporal auditory processing informs our understanding of the variety of intervention activities that can facilitate development.

KEY WORDS: Fast ForWord, language intervention, auditory processing, clinical trial
Approximately 7% of all school-age children have unusual difficulty learning and using language despite adequate hearing, nonverbal intelligence, and motor abilities (Tomblin, Records, & Zhang, 1996). This difficulty, which has been referred to by a variety of terms, including language impairment, language-learning disability, specific language impairment, and language-learning impairment, can have serious social, academic, and vocational ramifications (Brinton, Spackman, Fujiki, & Ricks, 2007; Catts, Fey, Tomblin, & Zhang, 2002; Clegg, Hollis, Mawhood, & Rutter, 2005; Conti-Ramsden & Botting, 2004). Unfortunately, a high percentage of school-age children with language impairments do not make gains beyond 1 standard error of measurement on formal tests administered 1, 3, or even as many as 10 years after they were identified (Aram & Hall, 1989; Bernstein & Stark, 1985; Conti-Ramsden & Botting, 1999; Fazio, Naremore, & Connell, 1996; Silva, Williams, & McGee, 1987; Snowling & Hulme, 1989; Tallal et al., 1996; Tomblin, Zhang, Buckwalter, & Rutter, 2005; Conti-Ramsden & Botting, 2004). Such results, although only indirectly a measure of treatment efficacy, are not a strong endorsement of the language intervention services that have been provided to school-age children.

One popular approach to language intervention for school-age children is **Fast ForWord Language** (FFW-L; Scientific Learning Corporation, 1998). FFW-L has received a great deal of attention in the scientific literature and the lay press. The developers of FFW-L assert that the program leads to neural reorganization that causes an increased ability to perceive fast-changing acoustic input and that such improvement leads to subsequent gains of 1 to 1.5 years on standardized tests of language skills after 6 weeks of training (Merzenich et al., 1996; Tallal, 2004; Tallal & Gaab, 2006; Tallal et al., 1996).

Two hypotheses underlying the use of FFW-L in intervention are that language impairments result from difficulty recognizing and sequencing the spectrotemporal structure of speech (Tallal, 1980, 1990, 2004) and that temporal processing abilities improve as a result of computer-assisted instruction in which modifications to the acoustic and temporal properties of the speech signal are gradually reduced as a function of increased performance on sound, syllable, word, and sentence comprehension tasks (Agocs, Burns, DeLey, Miller, & Calhoun, 2006; Merzenich et al., 1996; Tallal et al., 1996). Merzenich, Tallal, and their colleagues have published two influential studies about changes in temporal processing and language comprehension in children with language impairments (Merzenich et al., 1996; Tallal et al., 1996). In the first study, 7 children with language learning disabilities (LLD) between 5 and 9 years of age received language intervention 3 hr per day, 5 days per week, for 4 weeks. The participants were characterized as having normal nonverbal intellectual abilities, delays in receptive and expressive language development, and reading difficulties. The children played prototypes of the FFW-L games called “Circus Sequence” and “Phoneme Identification” and rotated through eight other speech and language exercises that were presented by trained clinicians in individual sessions. In addition, children completed 1 to 2 hr of homework every day that involved listening to stories. After training, the children improved significantly on measures of speech discrimination and language comprehension.

A second study (Merzenich et al., 1996; Tallal et al., 1996) was conducted with 22 children with language impairments who ranged in age from 5;4 (months;years) to 10;0. These children were divided into two matched groups according to nonverbal intelligence and receptive language abilities. Children in both groups attended laboratory sessions for 3.5 hr each day. The children in the second study also completed 1 to 2 hr of listening homework each day. The groups differed according to the auditory stimuli that were presented. Children in one group listened to modified speech as they played revised versions of the Circus Sequence and Phoneme Identification games that were used in Experiment 1 and two additional games (Old MacDonald’s Flying Farm and Phoneme Match) that are part of the current FFW-L package. They also listened to modified speech in the clinician-directed intervention sessions each day and in their daily homework sessions. The children in the second group “received equivalent language training but with natural speech materials,” and they “played video games rather than these adaptive auditory-speech training games” (Merzenich et al., 1996, p. 80). After training, children in both treatment groups improved significantly on all measures. The children whose language intervention included modified speech stimuli evidenced greater improvement on measures of temporal processing, speech discrimination, and grammatical comprehension than children in the natural speech treatment group. The authors concluded that the training remediated the underlying temporal processing deficit that contributed to difficulties with speech perception and language comprehension (Merzenich et al., 1996; Tallal et al., 1996).

Despite decades of providing language intervention to school-age children, few studies have compared outcomes of different intervention procedures. The absence of such data is perhaps one of the reasons for the public and professional excitement about the studies of FFW-L reported by Tallal and colleagues (Merzenich et al., 1996; Tallal et al., 1996). Parents and professionals have shown keen interest in data demonstrating that FFW-L intervention is associated with dramatic improvement in language test scores.
Two independent groups of researchers have evaluated the efficacy of FFW-L with children who have language impairments. Pokorni, Worthington, and Jamison (2004) randomly assigned a group of 54 children with language and reading impairments to one of three interventions: (a) FFW-L, (b) Earobics, or (c) the Lindamood Phoneme Sequencing Program. The children scored at least 1 SD below the mean on one of three subtests of the Clinical Evaluation of Language Fundamentals, Third Edition (CELF–3; Semel, Wiig, & Secord, 2000) — (a) Concepts and Directions, (b) Recalling Sentences, or (c) Listening to Paragraphs— and had reading skills that were at least 1 year below their grade level. The children were primarily African American (75%) and Caucasian (25%), and approximately 42% were living in families that were at or below the level of poverty. Children in all three groups received three 1-hr intervention sessions each day for 20 days. Testing occurred 4 to 6 weeks prior and 6 to 8 weeks after the completion of intervention. Pokorni et al. found no significant gains on the three subtests of the CELF–3 for any of the treatments.

Cohen et al. (2005) studied the effectiveness of FFW-L in a randomized controlled trial (RCT) that included 77 children with severe mixed receptive and expressive language impairment. On average, the children scored at least 2 SDs below the mean on both the Receptive and Expressive subtests of the CELF–3 (UK version; Semel, Wiig, & Secord, 2000). Children were randomly assigned to one of three groups: (a) one that received home intervention with FFW-L, (b) one that received home intervention with commercially available language and reading computer games that did not contain modified speech, or (c) one that received no home intervention. All the children received regular speech-language intervention services at school during the study. There were between 23 and 27 children in each group. The children in all three groups made significant gains on the CELF–3 UK at 9 weeks and at 6 months after intervention. However, there were no significant differences among the groups. Cohen et al. concluded that the FFW-L intervention did not provide any additional benefit for children with severe language impairments above the current language intervention services they were receiving at school. Thus, their findings call into question the utility of using acoustically modified speech to remediate severe, mixed language impairments.

There are some important limitations to the previous clinical trials that tested the language outcomes of FFW-L. Both trials included a relatively small subset of the children who typically receive treatment for language impairments. Cohen et al. (2005) studied only children with mixed receptive and expressive language impairments; Pokorni et al. (2004) studied children with language and reading disorders. In addition, none of the previous clinical trials measured changes in temporal auditory processing directly.

Our study was designed to overcome some of the limitations of the previous investigations. The study was undertaken over a period of 3 years and included a total of 216 children, almost three times the number of children in Cohen et al.’s (2005) study and four times the number of children in Pokorni et al.’s (2004) study. In our RCT, we compared FFW-L with two other computer instruction conditions: (a) a computer-assisted language intervention (CALI) condition, in which children played language intervention programs that specifically targeted cognitive, processing, and language skills similar to the ones targeted by FFW-L but without a modified speech signal and (b) an academic enrichment (AE) condition, in which children played educational computer games that were not specifically designed to improve language skills. The AE condition was designed to be an active comparison; that is, it shared a number of important features with the FFW-L and CALI conditions. However, the computer games in the AE condition were not specifically designed to improve language skills or auditory processing skills. We also included an individual language intervention (ILI) condition, which was delivered by a speech-language pathologist (SLP). Because we held frequency and duration of intervention constant, we were able to compare three critical dimensions of intervention: (a) computer-delivered versus human-delivered services, (b) modified speech versus unmodified speech, and (c) specific versus nonspecific intervention goals.

We had two primary research questions:

1. Would participants randomly assigned to the FFW-L condition show greater improvement on the CASL composite score than participants assigned to either of two specific intervention conditions (CALI or ILI) or to a nonspecific comparison condition (AE) immediately after intervention and on follow-up testing 3 and 6 months later?

2. Would participants randomly assigned to the FFW-L condition show greater improvement on a measure of backward masking than the participants assigned to the specific intervention condition (CALI or ILI) or to the nonspecific comparison condition (AE) immediately after intervention and on follow-up testing 3 and 6 months later?

We reasoned that if the auditory temporal processing deficit theory of language impairment is correct, then participants who received the FFW-L intervention should show more improvement on language and auditory processing measures than the children who were randomized to the other treatments or to an active comparison group that played academic enrichment computer
games. Following this logic, the FFW-L advantage should be maintained or strengthened 3 months and 6 months after treatment. The design of our study also enabled us to explore two secondary questions: (a) Is the modified speech in the FFW-L intervention necessary for language improvement to occur? and (b) are computer-delivered services that do not require a verbal response as efficacious as human-delivered services that include verbal interaction?

**Method**

**Participants**

Recruitment and identification. Two hundred sixteen children with language impairments participated in the study at three different sites. The children were recruited from nine school districts, with 96 children from northeast Kansas, 92 children from central Texas, and 28 children from north Texas. Between January and May of each year, children were recruited and tested to determine whether they qualified for inclusion in the study. The investigators met with SLPs, teachers, and special educators to explain the purpose of the investigation and the inclusion criteria. School district personnel gave brochures about the study to the parents of children whom the educators thought might meet the eligibility criteria. Parents contacted the research coordinators in each area if they were interested in having their child participate. After a brief conversation with the parents, the research coordinator scheduled an independent assessment for children who appeared to be potential participants.

Consistent with the EpiSLI model (Tomblin et al., 1997), children were determined to have a language impairment if they displayed a standard subtest score between 75 and 125 (±1.66 SD) on the Matrices subtest of the Kaufman Brief Intelligence Test (Kaufman & Kaufman, 1990) together with standard scores at or below 81 on two or more clusters of the Test of Language Development—Primary, Third Edition (TOLD–P;3; Newcomer & Hammill, 1997). To be eligible, participants could not present a hearing impairment, a visual impairment, gross neurological impairment, oral–structural anomalies, or emotional or social disorders. Vision and hearing status was confirmed either through school records of passing screenings within the past year or through live administration of vision and/or hearing screenings. When school vision screening records were unavailable, a vision screening was completed using Lea Symbols vision screening materials to ensure adequate vision in at least one eye with or without corrective lenses. Hearing screenings were administered at 20 dB HL at the frequencies of 1, 2, and 4 KHz in both ears. Oral structure and function screening with the Oral Speech Mechanism Screening Exam—Third Edition (St. Louis & Ruscello, 2000) ruled out the presence of oral–structural anomalies that could have interfered with normal language production.

Children were excluded if parent responses to questionnaires indicated three or more episodes of otitis media in the previous 12-month period, a history of focal brain lesions, traumatic brain injury, cerebral palsy, seizure disorders, symptoms of severely impaired reciprocal social interaction, or severely restricted activities listed in the Diagnostic and Statistical Manual of Mental Disorders (4th ed.; American Psychiatric Association, 1994) criteria for autism spectrum disorders. We also excluded children who had participated in 8 or more hours of language intervention or classroom activities using any of the Fast ForWord, Laureate, or Earobics speech-language or reading software, and/or the Lindamood–Bell auditory discrimination training; because this training may have confounded the results from the current study. In addition, parents agreed not to enroll their child in any other language intervention during the treatment phase of the study.

Randomization. After identification testing and pretesting, participants were randomly assigned to one of the four study conditions by the Biostatistics Center at the University of Iowa. Randomization of participants was stratified by treatment site (Austin, TX; Dallas, TX; and Lawrence, KS) and socioeconomic status (high or low). Randomization can minimize Type I error and improve statistical power for smaller RCTs (<400 participants; Kernan, Viscoli, Makuch, Brass, & Horwitz, 1999). In this study, stratified randomization helped by ensuring that the potential impact of otherwise uncontrollable factors associated with socioeconomic status were equally distributed across the treatment conditions at each site. The treatment assignments were sent to the research coordinators at each site via overnight express mail. The research coordinators were the only individuals who had a record of participant names, participant identification codes, and treatment assignments.

Participant characteristics. The preintervention means and standard deviations for all participants combined and for children in the four conditions are shown in Table 1. There were 54 children in each condition. Across the four conditions, there were more males (136) than females (80), a ratio of 1.7:1. The average age of the 216 participants was 7.6, with a range of 6.0 to 8.11. There were 58 six-year-olds, 78 seven-year-olds, and 80 eight-year-olds. The mean nonverbal intelligence score was within normal limits for each treatment group. Prior to intervention, group mean standard scores on the Spoken Language composite of the TOLD: P–3 ranged from 69.8 to 73.7, with an overall mean of 73.7. A preliminary analysis of variance (ANOVA)
indicated that none of the group differences reached significance on the identification measures. Overall, the children in the study fit the profile of having normal nonverbal intellectual abilities and language abilities that were consistently 1.2 SDs or more below the mean.

There were no differences in the gender distribution across the four conditions, $\chi^2(6, N = 216) = 6.13, p = .41$. Forty-six percent of the children who were enrolled in the study were White—not Hispanic. The next largest racial–ethnic group was Black/African American (29%), followed by children who were White–Latino/Hispanic (15%). The “Other” category (10%) included children who were Native American, Asian American, more than one race, and children whose parents elected not to report their race or ethnicity. There were no differences in the distribution of racial–ethnic groups across the four study conditions, $\chi^2(9, N = 216) = 7.75, p = .56$.

Parent education level was determined on the basis of self-report. There were three possible categories: (a) neither parent had attended college, (b) at least one parent had attended college (including community college), and (c) one or both parents had earned a college degree. In each treatment group, the majority of the children were from families in which at least one parent had attended a university, college, or community college. There was no significant difference between the number of children in the three parental groups across the four study conditions, $\chi^2(6, N = 215) = 2.45, p = .87$.

### Study Conditions

**FFW-L.** Children who were randomly assigned to the FFW-L condition played seven different computer games that targeted discrimination of tones (viz., Circus Sequence), detection of individual phoneme changes (viz., Old McDonald’s Flying Farm), matching phonemes to a target (viz., Phoneme Identification), identifying matched syllable pairs (viz., Phonic Match), discriminating between minimal pair words (viz., Phonic Words), recalling commands (viz., Block Commander), and comprehending grammatical morphemes and complex sentence structures (viz., Language Comprehension Builder).

The speech and nonspeech stimuli in the FFW-L computer games were modified by an algorithm that prolonged segments and differentially amplified particular frequencies (Nagarajan et al., 1998). The acoustic modifications were gradually decreased as children improved on each task. Children received trial-by-trial feedback for correct and incorrect responses, and their movement through the program was controlled automatically. When children responded incorrectly, the correct answer was provided prior to the next stimulus. Correct responses were rewarded by points, jingles, and extra animations on the computer screen.

The children played the games in a quiet area and wore Sony MDR-V-150 closed, supra-aural dynamic stereo headphones, which were used for all three computer conditions. A certified and licensed SLP supervised...
every session, and graduate student monitors were present in the room with the children and were encouraged to provide the child with positive nonverbal signals (i.e., thumbs up, smiles) to encourage continued work. Five of the seven games were presented each day for 20-min periods (a total of 1 hr and 40 min/day) until the child reached a criterion of 90% completion on any five games. Even though a few children reached this criterion, they continued playing the games at the highest level through the last day of the intervention session.

**CALI.** The CALI condition consisted of seven computerized instructional modules selected from the Earobics software (Cognitive Concepts, 2000a, 2000b) and Laureate Learning software (Semel, 2000; Wilson & Fox, 1997). Like the FFW-L exercises, the CALI modules targeted discrimination and memory of non-speech sounds (viz., Earobics Karloons’s Balloons), detection of individual phoneme changes (viz., Earobics C.C. Coal Car), phoneme discrimination (viz., Earobics Paint by Penguin), identifying matched syllable pairs (viz., Earobics Hippo Hoops), discriminating between minimal pair words (viz., Earobics Duck Luck), recalling commands (viz., Laureate Learning Systems Following Directions Series), and comprehending grammatical morphemes and complex sentence structures (Laureate Learning Systems Micro-LADS 1–6). None of the auditory stimuli in any of the CALI modules were modified acoustically. Similar to the FFW-L condition, each of the modules in the CALI arm were presented for 20-min periods (a total of 1 hr and 40 min/day), 5 days per week for 6 weeks via headphones in a quiet environment. Like the FFW-L treatment arm, a certified and licensed SLP supervised every session, and monitors, who were graduate students in speech-language pathology, observed the students and provided positive nonverbal feedback. Children started at the beginning of each computer game in CALI. When a child attained 90% correct at a particular level for 2 days in a row, the level of the exercise was considered to be mastered, and the next level was presented. If a child reached mastery level on a game before the end of the 6-week intervention session, he or she repeated that game until the last day of the intervention session.

**ILI.** ILI was delivered by certified and licensed SLPs who worked with each child in a quiet room. Like the FFW-L and CALI conditions, each of the intervention sessions in the ILI condition lasted 1 hr and 40 min per day, 5 days per week, for 6 weeks. Our approach to literature-based language intervention was influenced by Gillam and Ukrainetz (2006), Norris (1989), and Strong and Hoggan (1996), who have suggested that language therapy with school-age children should target specific language skills within activities that are related to the topic and content of children’s books. Opportunities for functional and interactive exchanges between clinicians and children (Gillam & Ukrainetz, 2005; Leonard, 1998; Nelson, Camarata, Welsh, & Butkovsky, 1996; Paul, 2001) were an integral part of therapeutic interactions.

The ILI activities were designed to target semantics, syntax (morphosyntax and clause structure), narration, and phonological awareness. The ILI units were developed around 13 picture books (see Appendix A) that were interesting to school-age children, that could be read in a short amount of time, and that contained a variety of vocabulary words that ranged in difficulty level for children who were 6 to 9 years of age. A minimum of six book units, usually one per week, were used with each child over the 6-week course of intervention. The SLP chose the book unit that she thought would be the most interesting for the child with whom she worked. The general outline that was used as the standard for each book unit is presented in Appendix B.

Language intervention activities were developed for three levels of difficulty within each of the target areas (i.e., semantics, grammatical morphology, clause structure, and narration and phonological processing). The language targets within the three levels of difficulty in each of the five areas are presented in Appendix C.

Throughout the ILI sessions, the SLP used a variety of language facilitation strategies that have been shown to be effective with children. These strategies included slower rate (Weismer, 1997), emphatic stress on target forms (Weismer, 1997), growth-relevant recasts (Camarata, Nelson, & Camarata, 1994; Nelson et al., 1996), focused stimulation (Cleave & Fey, 1997; Fey, Cleave, Long, & Hughes, 1993), incidental teaching (Kaiser, Yoder, & Keetz, 1992), scaffolding (Schneider & Watkins, 1996), and mediation (Miller, Gillam, & Peña, 2001).

The SLPs used a system developed by Miller et al. (2001) to track the children’s level of performance each day. After each activity, the clinicians judged the amount of teaching effort (low, medium, or high) and the amount of student responsiveness (low, medium, or high). Students advanced to the next level of difficulty when clinicians rated their teaching effort as low or moderate and rated student responsiveness as moderate or high for activities that were repeated in two consecutive sessions.

**AE.** The AE condition was designed to serve as an “active” comparison arm (Herbert & Gaudiano, 2005; Lohr, DeMaio, & McGlynn, 2003). AE shared common features with FFW-L and CALI, such as intervention setting, amount of contact with computers, type of child involvement, time spent on different computer games, reward system, and type of clinician involvement. AE was also designed to be a nonspecific intervention comparison. Even though many of the computer games in AE contained language, none of the software games were specifically designed to promote the development of
language or auditory processing skills; instead, the computer games in the AE condition had been designed to teach mathematics, science, and geography. The computer games included Magic School Bus Discovers Flight (Scholastic, Inc., 2001a), Magic School Bus Explores Dolphins and Whales (Scholastic, Inc., 2001b), Coin Critter (Nordic Software, Inc., 1999), Zurk’s Rainforest Adventure (Soleil Software, Inc., 1998), My Amazing World Explorer (Dorling Kindersley, 1999), Dinosaur 3D (Knowledge Adventure, Inc., 1999), and selected games from Arthur’s 1st Grade (TLC Educational Properties LLC, 1999a) and Arthur’s 2nd Grade (TLC Educational Properties LLC, 1999b). None of these games targeted language or reading directly.

**General Procedures**

Each child attended the summer intervention program for 3.5 hr daily. Of this time, 1 hr and 40 min were devoted to the intervention conditions. Children also received a 20-min snack break and attended a 1.5-hr group activity period in which they played board games, worked on informal arts and crafts (coloring, painting, cutting, etc.), and participated in general outdoor activities (recess). The activity groups were managed and monitored by a classroom teacher and two aides. Children were randomly assigned to one of two activity groups (A or B). The order of intervention and activity groups (first half or second half of the morning) was alternated each week.

**Attendance, Compliance, and Treatment Fidelity**

On rare occasions, an SLP in the ILI condition would be absent, or technical problems would lead to missed computer sessions. When an SLP was ill or when a computer was not usable because of technology problems, the child assigned to that condition would miss that day of treatment. The children’s average attendance rate overall was 28 of 30 days. Only 3 of 216 children withdrew from the study during the intervention period. There was no difference between the average number of days present for the FFW-L (27.31, SD = 3.30), CALI (27.80, SD = 2.75), ILI (27.27, SD = 3.30), and AE (28.31, SD = 2.13) conditions, F(3, 212) = 1.231, p = .299.

Clinicians documented compliance each day. Children received stickers after they completed each treatment session. At the end of the week, children selected a toy from a prize board if they had earned five stickers. There was a seven-step procedure for managing non-compliance. If a child refused to engage in an activity, the clinician verbally encouraged the child to resume her or his participation. If the first step was not successful, the clinician would offer the child an opportunity to win an extra sticker. In the third step, children could take a short break and get a drink of water. If that was not successful, children were offered a small food reward, such as a snack. In the fifth step, children were given a short (5-min) time-out period. If children were still non-compliant after a time-out period, we ended the session for the day. We planned to meet with the parents to discuss the child’s involvement with the study (Step 7) if a child reached Step 6 for 3 days in a row. The average numbers of days (out of 30) that any compliance step was recorded in the therapy logs are displayed in Table 2. Verbal encouragement was usually sufficient to obtain compliance. A multivariate comparison of the number of treatment days for which compliance steps were recorded yielded no significant group differences. In addition, there were no important adverse events or side effects for participants in any of the study conditions.

Clinicians were trained to administer the interventions. No clinicians provided treatment in more than one condition. The clinicians documented the number of games that the participants played (computer conditions) or the number and type of activities that were undertaken (ILI condition) each day. The research team assessed treatment fidelity each week during the summer program. The sessions in all four conditions were videotaped each day. At the end of each week, a randomly selected videotape of a session in each treatment condition was sent to another site for review. Research assistants reviewed the use of positive reinforcement, the presence of a quiet and nondistracting environment, use of facilitative talk (in the ILI units), wearing of headphones (for the computer conditions), and the extent and type of clinician assistance. The results of the fidelity reviews were faxed to the site coordinators each Monday. Any modifications or suggestions resulting from the fidelity review were implemented at the intervention site.

**Measurements**

Children received a battery of language, literacy, and auditory processing measures at four different times.
forced-choice paradigm

Monkeys have been known to use this technique to determine the location of objects. The task is widely referred to as a

backward masking task. Backward masking is widely considered to be a measure of temporal resolution because it requires listeners to distinguish between sounds in time (Hill, Hogben, & Bishop, 2005; Marler, Champlin, & Gillam, 2001). As with all other psychoacoustic measures, attention and memory contribute to performance on backward masking.

The backward masking task was delivered via a computer according to procedures that have been summarized by Marler et al. (2001); Thibodeau, Friel-Patti, and Britt (2002); and Wright et al. (1997). The task was administered in a quiet room with an ambient noise level of 45 dBA or less. Children looked at a computer screen showing three colorful pictures as they listened to three auditory intervals presented over TDH 39 headphones. The headphones were then positioned over the child with the test environment and procedures. Children who were seated at a computer looked at colorful pictures, one corresponding to each of the listening intervals that were displayed on the computer screen. The examiner explained that there were three sounds. After listening carefully, the child was instructed to indicate the sound that was different from the other two by clicking the mouse pointer on the appropriate picture. A colorful picture would appear on the screen after each correct response, but a solid horizontal green bar would appear after each incorrect response. After giving children the verbal instructions, the examiner provided a demonstration trial via a small loudspeaker. The signal was randomly presented in one of the three observation intervals. The observation intervals were 350 ms in duration and were separated by 500-ms pauses. The examiner indicated the interval that contained the signal by selecting one of the buttons. The child was given the opportunity to perform five practice trials independently. The headphones were then positioned over the child’s ears, and five practice trials were performed.

During testing, the examiner was seated behind the listener and assisted the child only to redirect his or her attention to the task when necessary. When ready, the child selected one of three buttons to begin the test. Three listening intervals were presented, and the child chose the one he or she thought contained the signal. If the response was incorrect, the first trial was repeated, systematically increasing the signal level, until the child responded correctly. After a correct response was obtained on the first trial, the signal levels for all subsequent trials were automatically adjusted on the basis of the listener’s response. The signal became louder (easier to hear) after each incorrect response and softer (more difficult to hear) after two consecutive correct responses. The initial step size was 5 dB, and the final step size was 2 dB.

A test run consisted of up to 60 trials and took 3 to 5 min to complete. The threshold estimate was based on the average of the last 10 of 12 reversals within a run. Threshold variability was based on the standard deviation from the same 10 reversals. If the threshold variability was greater than twice the step size, the run
was repeated. A maximum of three runs may have been attempted per test. The first run to satisfy the variability criterion was taken as the estimate of threshold, unless three runs were needed, after which the average of the three runs was taken. The child received verbal encouragement and a tangible reward on successful completion of each run. After the first threshold estimate, the child was given a break and returned later the same day for a second measure. The first threshold estimate was considered practice; only the second threshold estimate was used in the statistical analysis. The extensive training and practice that preceded the second threshold estimate were designed to minimize practice effects, which have been demonstrated in previous studies (Bishop, Carlyon, Deeks, & Bishop, 1999; Wright, 2001).

Secondary measures. The FFW-L and CALI arms primarily targeted sentence comprehension and phonological processing. These abilities were not specifically assessed by the subtests on the CASL or by the backward masking measure. We administered additional secondary measures that related more specifically to additional skills that were targeted in the FFW-L and CALI arms. The Token Test for Children (Disimoni, 1978) was administered to assess sentence comprehension. Because of the age and ability levels of the children in this study, there were ceiling effects on the first two sections of the Token Test. We report the combined raw score for Parts III, IV, and V, in which children manipulated small tiles in response to examiner directions, such as “Touch the red circle and the yellow square” (Part III), “Touch the small blue circle and the small red circle” (Part IV), and “Put the green square beside the yellow circle” (Part V).

We administered the Blending Words subtest of the Comprehensive Test of Phonological Processing (CTOPP; Wagner, Torgesen, & Rashotte, 1999) to assess phonological awareness. This required children to listen to audiorecorded sounds, presented one at a time, in sequence. Children combined the sounds mentally and said the complete word (e.g., “What word do these sounds make: j-u-m-p?”; children answer “jump”).

The primary and secondary measures were administered in day-long testing sessions that included snacks, craft breaks, and lunch. If a child was fatigued, we discontinued testing for a short period of time. There was no particular order for test administration. Once administered, language and literacy measures were rescored and checked for accuracy by three independent scorers.

Blinding

The principal investigators, data collectors, data analyzers, and members of a data safety and monitoring committee were blinded to the randomized treatment assignments. A research coordinator at each site was the only person who had a record of participant names, participant identification codes, and treatment assignments.1

Parents were not told which treatment their children received, and they were not allowed to observe therapy. It is possible that children told their parents enough about their experiences at school to enable parents to deduce the arm to which they had been randomized. To assess this possibility, 73 parents completed a questionnaire either in written format or via phone survey in which they were asked to guess which treatment their child had received. Only 20% of parents identified the correct condition to which their child had been randomized, which was slightly less than chance. These results suggest that the majority of the parents were blind to the treatment that their children had received. It is unlikely that potential parental expectations of the value of a particular treatment influenced the outcomes of the study.

Baseline Comparison

To check for selection bias, we tested for potential preintervention differences among the participants in the four study conditions (Berger, 2005). Cochran–Mantel–Haenszel tests were used to assess categorical variables, stratified by site and year. For continuous outcomes or covariates, mixed-model ANOVAs were used. No statistically significant differences were found among the four conditions for sex, age, socioeconomic status, or ethnicity. The mixed-model ANOVAs (with site, subsite, and classroom as random effects) were conducted to examine whether the three intervention and comparison groups differed at baseline on their pretest scores for each outcome. No statistically significant differences among the four conditions were identified.

Attrition

Three participants received only pretesting because they withdrew from the study before completing the treatment. Two additional participants missed two of the four testing sessions, and 8 participants missed one testing session (a total of 21 missed testing sessions of 648 possible sessions; 2.3% missing data). Preliminary analyses

---

1 There was a compromise to blinding at one site when some therapy files containing identifying information were reviewed. This occurred after the Year 2 posttest data had been analyzed and reported to the Data Management Center at the University of Iowa but before the Year 3 pretesting had begun. The primary and secondary outcome data were not observed or analyzed, and no person involved with testing participants, scoring tests, analyzing test results, or reporting data to the Data Management Center had any knowledge of the compromise to blinding. The individuals involved had no subsequent involvement in data collection or analysis. The incident was reported to the university’s institutional review board and the National Institutes of Health Data Safety Monitoring Board. Both committees determined that there were no adverse effects from this event.
with mixed-model ANOVAs (with site, subsite, and classroom as random effects) were conducted to compare pretest scores for participants who had completed the posttest versus participants with pretest scores who did not complete one or more posttests. Preliminary results indicated no significant differences in the complete cases versus incomplete cases on pretest scores. Following an intent to treat analysis strategy, we analyzed the data using all cases by imputing the missing values via an expectation—maximization algorithm.

**Results**

The overall goal of this RCT was to compare the language and auditory processing outcomes of three types of language intervention and an active, nonspecific intervention comparison group. We hypothesized that if the temporal processing deficit theory of language impairment is correct, then the participants who received the FFW-L intervention would show greater improvement on a standardized language test than children who were assigned to the other treatment arms or to the active, nonspecific comparison group at immediate posttest and at the 3- and 6-month follow-up periods. A mixed-model analysis of covariance was used to test the original hypotheses. The primary dependent variable was the CASL core composite score. Performance on the CASL composite pretest was significantly related to the posttest and follow-up outcomes, $F(1, 207) = 488.6, p < .001$, adjusted $R^2 = .62$. Because of this positive and significant correlation, the pretest CASL composite score was used as a covariate. Including the pretest as a predictor in the model effectively adjusted the posttest and follow-up tests for differences on the pretest scores. The fixed independent variables were the four conditions (AE, CALI, FFW-L, and ILLI) and three postintervention times (posttest, 3-month follow-up, and 6-month follow-up).

The mixed model used a multilevel model approach to avoid the possibility of statistically significant differences due to the decreased standard error resulting from nonindependence in participant scores within a particular sex, year (1, 2, or 3), site (Austin, Dallas, or Lawrence), subsite (particular school), or classroom group (A or B; Bryk & Raudenbush, 2002; Snijders & Bosker, 1999). If the variance in participant scores attributable to sex, year, site, or classroom was zero, that variance component was dropped from the model (Kreft & De Leeuw, 1998). The degrees of freedom for statistical significance tests were calculated using the Kenward–Roger adjustment for degrees of freedom in mixed models for repeated measures, because it reduces the potential for Type I error (Littell, 2002). For the within-subject effects, the unstructured variance–covariance matrix was based on a comparison of seven different structures using the Bayesian information criterion. An effect size was calculated using a pseudo $R^2$ based on the Level 1 residuals proposed by Snijders and Bosker (1999), which examined the proportional reduction in the residual variance. Standardized residuals from the hypothesis models were examined to identify any outliers and ensure that unusual values were accurate. The number of cases, means, and standard deviations for the CASL composite score and the backward masking score by treatment and time are presented in Table 3. Across time, there was a positive trend for the CASL composite score and a negative trend for the backward masking, meaning that performance on both tasks improved.

**Pre–Post Comparisons**

Because the statistical model that specifically addresses the hypotheses and research questions does not directly test whether change occurred from pretest to posttest, a preliminary mixed-model repeated measures analysis was performed to examine this issue of clinical importance. The predictors were condition (CALI, FFW-L, ILLI, and AE) and time (pre- and posttest) and the Condition × Time interaction. The dependent variable was the CASL core composite standard score. The

<table>
<thead>
<tr>
<th>Condition and time</th>
<th>CASL: Composite standard score</th>
<th>Backward masking (dB SPL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$N$</td>
<td>$M$</td>
</tr>
<tr>
<td><strong>CAU</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>54</td>
<td>79.0</td>
</tr>
<tr>
<td>Posttest</td>
<td>53</td>
<td>83.4</td>
</tr>
<tr>
<td>3 months</td>
<td>52</td>
<td>84.9</td>
</tr>
<tr>
<td>6 months</td>
<td>52</td>
<td>86.3</td>
</tr>
<tr>
<td><strong>FFW-L</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>54</td>
<td>78.0</td>
</tr>
<tr>
<td>Posttest</td>
<td>51</td>
<td>83.0</td>
</tr>
<tr>
<td>3 months</td>
<td>51</td>
<td>83.8</td>
</tr>
<tr>
<td>6 months</td>
<td>50</td>
<td>85.2</td>
</tr>
<tr>
<td><strong>ILLI</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>54</td>
<td>80.8</td>
</tr>
<tr>
<td>Posttest</td>
<td>52</td>
<td>85.7</td>
</tr>
<tr>
<td>3 months</td>
<td>51</td>
<td>87.8</td>
</tr>
<tr>
<td>6 months</td>
<td>50</td>
<td>89.9</td>
</tr>
<tr>
<td><strong>AE</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>54</td>
<td>78.2</td>
</tr>
<tr>
<td>Posttest</td>
<td>53</td>
<td>82.1</td>
</tr>
<tr>
<td>3 months</td>
<td>53</td>
<td>83.8</td>
</tr>
<tr>
<td>6 months</td>
<td>53</td>
<td>86.2</td>
</tr>
</tbody>
</table>
analysis used a mixed-model method that incorporated site and subsite random effects and an unstructured variance–covariance matrix for the within-subject residuals. There was a significant effect of time, \( F(1, 212) = 97.9, p < .001, d = 1.08 \), indicating significant improvement from pre- to posttest across the four study conditions. Neither the main effect of condition nor the Time \( \times \) Condition interaction was significant.

**Hypothesis Tests**

The primary research question was: Will participants randomly assigned to the FFW-L condition show greater improvement on the CASL composite score than the participants assigned to the treatment conditions (CALI or ILI) or to the active comparison (AE) condition immediately after intervention and at the 3- and 6-month follow-up periods? As discussed earlier, performance on the CASL composite pretest was significantly related to the posttest and follow-up outcomes. The main effect of condition was not significant, \( F(3, 206) = 1.9 \), indicating that the children in one condition of the study did not benefit from their experiences to a greater extent than the children in any other condition. The main effect of time was significant, \( F(2, 211) = 43.3, p \leq .001 \), adjusted \( R^2 = .05 \), indicating that the children in all four conditions demonstrated statistically significant change on the CASL from posttest to the 3- and 6-month follow-up periods. The least squares estimated adjusted means across time were as follows: 83.0 at posttest, 85.0 for the 3-month follow-up, and 86.8 for the 6-month follow-up. All three means were significantly different from each other using Tukey–Kramer adjusted \( p \) values ( \( p < .001 \)). The Condition \( \times \) Time interaction testing differential change across groups and time was not significant, \( F(6, 280) = 0.8 \), indicating that the lack of differences among study conditions did not vary across time.

It is possible that the core composite standard score on the CASL masked group differences among the language skills that were measured by the CASL subtests or that there were group differences in language abilities that were not measured by the CASL subtests. The unadjusted means and standard deviations for raw scores on four subtests on the CASL (Antonyms, Syntactic Construction, Paragraph Comprehension, and Pragmatic Judgment) are presented in Table 4. The raw score means and standard deviations for the sum of Parts III, IV, and V of the Token Test for Children, and the raw scores for the Blending Words subtest on the CTOPP, also are presented. Similar to the CASL analysis,

<table>
<thead>
<tr>
<th>Subtests</th>
<th>Antonyms</th>
<th>Syntactic Construction</th>
<th>Paragraph Comprehension</th>
<th>Pragmatic Judgment</th>
<th>Token Test</th>
<th>Blending Words</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Condition and time</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CALI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>13.5</td>
<td>5.5</td>
<td>18.3</td>
<td>6.7</td>
<td>24.2</td>
<td>11.4</td>
</tr>
<tr>
<td>Posttest</td>
<td>14.7</td>
<td>5.2</td>
<td>19.5</td>
<td>7.1</td>
<td>28.6</td>
<td>11.7</td>
</tr>
<tr>
<td>3 months</td>
<td>16.3</td>
<td>5.7</td>
<td>21.3</td>
<td>7.5</td>
<td>29.9</td>
<td>11.3</td>
</tr>
<tr>
<td>6 months</td>
<td>17.1</td>
<td>4.7</td>
<td>21.7</td>
<td>7.3</td>
<td>33.6</td>
<td>8.1</td>
</tr>
<tr>
<td><strong>FFW-L</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>13.7</td>
<td>5.4</td>
<td>18.3</td>
<td>6.5</td>
<td>22.6</td>
<td>10.9</td>
</tr>
<tr>
<td>Posttest</td>
<td>15.4</td>
<td>4.3</td>
<td>19.4</td>
<td>6.3</td>
<td>27.4</td>
<td>11.8</td>
</tr>
<tr>
<td>3 months</td>
<td>16.3</td>
<td>4.5</td>
<td>20.5</td>
<td>6.4</td>
<td>30.9</td>
<td>8.2</td>
</tr>
<tr>
<td>6 months</td>
<td>17.5</td>
<td>4.7</td>
<td>22.0</td>
<td>6.0</td>
<td>31.5</td>
<td>10.2</td>
</tr>
<tr>
<td><strong>ILI</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>15.4</td>
<td>6.3</td>
<td>19.7</td>
<td>5.4</td>
<td>22.0</td>
<td>15.5</td>
</tr>
<tr>
<td>Posttest</td>
<td>16.6</td>
<td>4.9</td>
<td>21.3</td>
<td>5.3</td>
<td>28.4</td>
<td>13.4</td>
</tr>
<tr>
<td>3 months</td>
<td>17.9</td>
<td>4.8</td>
<td>21.9</td>
<td>5.3</td>
<td>32.7</td>
<td>9.2</td>
</tr>
<tr>
<td>6 months</td>
<td>18.7</td>
<td>4.4</td>
<td>24.3</td>
<td>5.4</td>
<td>34.5</td>
<td>7.7</td>
</tr>
<tr>
<td><strong>AE</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pretest</td>
<td>14.1</td>
<td>6.1</td>
<td>17.9</td>
<td>6.5</td>
<td>22.9</td>
<td>12.3</td>
</tr>
<tr>
<td>Posttest</td>
<td>15.8</td>
<td>5.2</td>
<td>19.6</td>
<td>6.1</td>
<td>28.3</td>
<td>12.3</td>
</tr>
<tr>
<td>3 months</td>
<td>16.8</td>
<td>4.8</td>
<td>20.3</td>
<td>5.9</td>
<td>32.1</td>
<td>8.8</td>
</tr>
<tr>
<td>6 months</td>
<td>17.9</td>
<td>4.5</td>
<td>22.3</td>
<td>6.1</td>
<td>34.4</td>
<td>8.3</td>
</tr>
</tbody>
</table>

| Gillam et al.: Fast ForWord Language Intervention | 107 |
mixed-model analyses of covariance were performed on each measure, with condition (CALI, FFW-L, ILI, and AE) as the between-subjects factor and time (immediate posttest, 3-month follow-up, and 6-month follow-up) as the within-subject factor. The pretest score for each measure was entered as the covariate. The results for each CASL subtest and the Token Test for Children were consistent with the results for the CASL core composite: The main effects of time were significant, indicating change over time, but the Time × Condition interactions were not significant.

The Blending Words subtest of the CTOPP yielded a significant main effect of time, $F(2, 422) = 7.214, p < .01$, $\eta^2 = .034$, and a significant Time × Condition interaction, $F(6, 422) = 2.606, p < .05, \eta^2 = .036$. The means for the three assessment times differed significantly (Tukey, $p < .001$). At posttest, the means for the CALI, FFW-L, and ILI conditions were above the 95% confidence interval of the mean for the AE condition, indicating that the children in the three language treatment conditions outperformed the children in the active comparison condition on a measure of phonological awareness. There were no significant differences among the conditions at the 3-month follow-up; however, at the 6-month follow-up the means for the CALI and the FFW-L conditions exceeded the upper boundaries of the 95% confidence intervals for the AE and ILI conditions. Therefore, it appears that the intervention in the CALI and FFW-L conditions was better than the intervention in the AE and ILI conditions for improving phonological awareness.

Recall that the primary hypothesis (that children assigned to the FFW-L condition would have better outcomes than children in the other three conditions) was based on the temporal processing deficit hypothesis (Tallal, 2004). The results of the language testing were not consistent with the hypothesis. It is possible that the scores on the language measure were not reflective of changes in auditory temporal processing. Therefore, we applied the same statistical model, using backward masking as the dependent variable. If the acoustic modification that was built into the FFW-L software ameliorated a temporal processing deficit, then the backward masking scores of the children in the FFW-L condition should have improved more than the backward masking scores of the children in the other three conditions, which did not include modified acoustic input.

The pretest backward masking score was significantly related to the posttest and follow-up outcomes, $F(1, 193) = 214.9, p < .001$, adjusted $R^2 = .38$. When the posttreatment scores were adjusted to reflect differences in pretreatment performance levels, the main effect of condition was not significant, indicating no differences among the three treatments and the active comparison condition. The main effect of time was significant, $F(2, 211) = 20.9, p < .001$, adjusted $R^2 = .04$, indicating that the children in all four conditions demonstrated meaningful change on backward masking from posttest to the 3- and 6-month follow-up periods. The least squares estimated adjusted means across time were as follows: 55.7 dB at posttest, 53.5 dB at the 3-month follow-up, and 49.6 dB at the 6-month follow-up.

Another model tested the three-way interaction among backward masking, time, and condition to determine whether there were group differences depending on the initial backward masking score. This interaction was not statistically significant. The results of these analyses suggest that children’s auditory temporal processing skills improved regardless of whether the treatment they received contained modified speech.

We also conducted an analysis to examine whether the FFW-L group or other groups differed on the CASL composite score depending on both their backward masking score and time. The three-way interaction among condition, backward masking score, and time, and the two-way interaction between condition and backward masking score, were statistically nonsignificant. Participants with lower backward masking scores did not show greater language skills improvement in the FFW-L condition compared with the other conditions.

**Clinical Significance**

Statistical analyses suggest that the children in all four conditions made similar amounts of improvement on a global test of receptive and expressive language. Because this was an intervention study, we asked whether the changes demonstrated by the participants were clinically relevant. We used effect sizes, confidence intervals, and normalization percentages to assess the clinical significance of the results at posttest and at the 6-month follow-up testing.

The effect sizes for the intention-to-treat analysis and for the raw score changes on backward masking, the Token Test for Children, and the Blending Words subtest on the CTOPP are presented in Table 5. There are moderate CASL effect sizes from pretest to posttest across conditions, as reflected by Cohen’s $d$ values that ranged from 0.56 to 0.79. The gains from pretest to the 6-month follow-up test are uniformly large, as reflected by Cohen’s $d$ values that ranged from 0.93 to 1.34. The children in the CALI group obtained moderately large effect sizes on the backward masking measures from pretest to posttest and large effect sizes from pretest to the 6-month follow-up testing. The children in the other three conditions evidenced moderately small to moderate effect sizes on backward masking immediately after treatment and moderately large effect sizes from pretest to the 6-month follow-up. The effect sizes for performance on the Token Test were unremarkable at immediate posttest. At the 6-month follow-up period, the effect sizes
for the children in the CALI, FFW-L, and ILI conditions were in the moderate range (0.54 – 0.66), but there were relatively small effects (0.35) for children in the AE condition. The effect sizes for the Blending Words measure were noticeably different across the conditions of the study. Immediately after treatment, the participants in the AE condition evidenced a negative effect (a decrease in performance), and the children in the other three conditions evidenced a moderately small effect. At the 6-month follow-up testing, the Cohen’s d effect sizes for participants in the FFW-L condition were moderately high, and the effect sizes for the children in the other three conditions were in the moderate range. For this measure of phonological awareness, the effect size for the FFW-L group was 27% higher than the effect sizes for children in the CALI condition (0.79 vs. 0.62), 49% higher than those for the children in the ILI condition (0.79 vs. 0.53), and 72% higher than those for the children in the AE condition (0.79 vs. 0.46).

Another indicator of clinical significance is the participants in each condition whose posttest and 6-month follow-up scores were outside the upper limits of the 68% and 95% confidence intervals (1 SE and 1.96 SE) for the CASL core composite score. As shown in Table 6, more than half of the children in each condition earned scores that were above the upper limit of the 68% confidence interval immediately after the intervention period. Between 63% and 80% of the children in each condition earned 6-month follow-up scores that were greater than 1 standard error of measurement on the CASL core composite. The 95% confidence interval is a more conservative indication of clinically significant change. More than half of the children in each condition earned scores that were above the upper limit of the 95% confidence interval for their pretest score by the 6-month follow-up testing period.

The final indicator of clinical significance was the percentage of the children in each condition whose scores moved into the normal range on the CASL over the course of the study. Following Tomblin et al. (2003), the normalization criterion was a standard score of 83 or above. According to this criterion, the percentages of children who normalized over the course of the study were 64.8% for CALI, 59.3% for FFW-L, 75.9% for ILI, and 70.4% for AE.

### Discussion

The primary purpose of this RCT was to determine whether Fast ForWord Language, a computer instructional program designed to improve auditory temporal processing skills in children with language impairments, was more effective than other types of interventions for improving language and auditory processing. If language impairments primarily reflect a deficit in auditory temporal processing skills, then children randomized to the FFW-L condition should have presented greater pretest-to-posttest improvement on a standardized measure of language and on a measure of backward masking in comparison to children in the other conditions. Furthermore, if the auditory processing deficit was remediated, then the FFW-L advantage should have been maintained over time.

The children with language impairments in our sample who received FFW-L did not fare better than children in other language interventions of equal intensity on our primary outcome measure of language, the CASL core composite score; on the individual CASL subtests; on the Token Test for Children; or on a measure of temporal auditory processing, the backward masking threshold. With one exception, the children in all four conditions made significant improvement on the language and the

### Table 5

<table>
<thead>
<tr>
<th>Condition</th>
<th>CASL Posttest</th>
<th>CASL 6 months</th>
<th>Backward masking Posttest</th>
<th>Backward masking 6 months</th>
<th>Token Test Posttest</th>
<th>Token Test 6 months</th>
<th>Blending Words Posttest</th>
<th>Blending Words 6 months</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALI</td>
<td>0.56</td>
<td>1.02</td>
<td>0.71</td>
<td>1.14</td>
<td>0.12</td>
<td>0.54</td>
<td>0.26</td>
<td>0.62</td>
</tr>
<tr>
<td>FFW-L</td>
<td>0.71</td>
<td>0.93</td>
<td>0.29</td>
<td>0.65</td>
<td>0.29</td>
<td>0.66</td>
<td>0.32</td>
<td>0.79</td>
</tr>
<tr>
<td>ILI</td>
<td>0.79</td>
<td>1.33</td>
<td>0.51</td>
<td>0.79</td>
<td>0.11</td>
<td>0.56</td>
<td>0.27</td>
<td>0.53</td>
</tr>
<tr>
<td>AE</td>
<td>0.61</td>
<td>1.34</td>
<td>0.39</td>
<td>0.76</td>
<td>0.02</td>
<td>0.35</td>
<td>-0.11</td>
<td>0.46</td>
</tr>
</tbody>
</table>

### Table 6

<table>
<thead>
<tr>
<th>Condition</th>
<th>68% Confidence interval Immediate posttest</th>
<th>68% Confidence interval 6-month follow-up</th>
<th>95% Confidence interval Immediate posttest</th>
<th>95% Confidence interval 6-month follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>CALI</td>
<td>55.6</td>
<td>63.0</td>
<td>31</td>
<td>54</td>
</tr>
<tr>
<td>FFW-L</td>
<td>53.7</td>
<td>74.1</td>
<td>33</td>
<td>59</td>
</tr>
<tr>
<td>ILI</td>
<td>59.3</td>
<td>79.6</td>
<td>43</td>
<td>69</td>
</tr>
<tr>
<td>AE</td>
<td>53.7</td>
<td>68.5</td>
<td>31</td>
<td>54</td>
</tr>
</tbody>
</table>
auditory processing measures from pre- to posttest as well as at the follow-up testing 3 and 6 months later. The exception was the Blending Words subtest of the CTOPP, which is a measure of phonological awareness. On that measure, children in the CALI and FFW-L conditions evidenced significantly better performance than the children in the AE and ILI conditions.

One important finding is that the primary measures of language and auditory processing improved significantly across groups. One possible conclusion from this pattern of results is that all four conditions were equally ineffective; that is, the gains in language and auditory processing that were observed across the four conditions of the study may have arisen as a result of practice effects, maturation, regression to the mean, and/or spontaneous recovery. If we had used a no-treatment comparison group in the design of the study, then an interpretation of equally ineffective intervention practices would be entirely appropriate. However, this study was designed with an active comparison group that received daily computer-assisted instruction on academic subjects. Without a no-treatment control group, we have no way to assess the relative extent to which intervention and changes over time contributed to changes in the primary outcome variables. Most of the children in this study had moderate to severe language impairments. We felt an obligation to their parents to provide each child with some sort of positive experience. By including an active control group and an ILI group, we controlled the variables (attention by an interested clinician, engagement in activities designed to promote thinking and problem solving, and opportunities to socialize with peers) that would have been uncontrolled if we had included a no-treatment group.

**Language Measures**

Careful consideration of the pattern of language results in light of what is known about the longitudinal course of language impairment supports the validity of the statistical results demonstrating significant improvements in language and auditory processing skills for participants in all four conditions. A number of longitudinal studies have shown that a high percentage of school-age children with language impairments do not make gains beyond 1 standard error of measurement on formal tests administered 1, 3, or even as many as 10 years after they were identified (Aram & Hall, 1989; Bernstein & Stark, 1985; Conti-Ramsden & Botting, 1999, 2004; Fazio et al., 1996; Silva et al., 1987; Snowling & Hulme, 1989; Tallal et al., 1996). In this study, the 6-month follow-up scores of 63% to 80% of the participants were more than 1 standard error of measurement above their pretest scores. In addition, more than 50% of the children had 6-month follow-up scores that were outside the 95% confidence interval of their pretest scores, yielding large pretest-to-posttest effect sizes across the groups (range: 0.93–1.34). These findings support the conclusion that children in all four conditions made clinically relevant improvements on the CASL that are greater than would be expected for children with language impairments who have participated in longitudinal studies.

An additional measure of clinical significance adds more support to our conclusion of improved language skills across the treatment and comparison conditions. At the time of the 6-month follow-up evaluation, between 59% and 76% of the participants had good outcomes, as indicated by CASL scores that were above Tomblin et al.’s (2003) suggested cutoff score of 83. In many previous studies, between 55% and 80% of children with language impairments evidenced follow-up scores on global language tests that remained in the impaired range (Beitchman, Wilson, Brownlie, Walters, & Lancee, 1996; Conti-Ramsden & Botting, 1999; Tomblin et al., 2003). It is important to remember that in the current study, children were identified as language impaired with one measure (the TOLD–P.3) that differed from the primary outcome measure (the CASL). This procedure was designed to control for regression effects by minimizing regression to the mean after pretesting (Tomblin et al., 2003). In comparison to other studies of similar age children with similar types and degrees of language impairments, the treatments used in this study reduced the persistence of language impairment by half or more.

The children made less improvement on a test of sentence comprehension than they did on the primary outcome measure, which was a global language test. The immediate posttest and 6-month follow-up effect sizes for the Token Test for Children were approximately half as large as the effect sizes for the CASL. Performance on all the subtests on the CASL requires some degree of language comprehension; however, only one subtest—Paragraph Comprehension—primarily assesses language comprehension. The rest of the CASL subtests involve language production. The difference in the effect sizes of the gains on the CASL and the Token Test for Children suggests that children’s expressive language skills improved more than their receptive language skills. This occurred despite the fact that three of the four conditions of the study primarily focused on listening. These results are consistent with the outcomes of a recent study that compared computerized intervention using modified or unmodified speech (Bishop, Adams, & Rosen, 2006). In that investigation, the children in both intervention groups and a no-treatment control group made relatively small gains on a test of sentence comprehension.
One language measure yielded significant group differences. At the immediate posttest, participants in the FFW-L and CALI conditions earned higher scores on the Blending Words subtest of the CTOPP than the participants in the AE condition. At the 6-month follow-up, scores on the Blending Words subtest for participants in the FFW-L and CALI conditions were still significantly higher than the scores for participants in the AE and ILI conditions. It appears that the two computerized instructional programs that focused primarily on auditory discrimination of sounds, syllables, and words yielded better phonological awareness results than a computerized treatment that focused on general academic skills or a clinician-directed language treatment that emphasized phonological awareness to the same extent as other language skills (vocabulary, syntax, grammatical morphology, and narration). Finally, similar positive outcomes for phonological awareness for the FFW-L and CALI conditions suggest that the inclusion of modified speech in the FFW-L program does not appear to be necessary for improvements in phonological awareness.

Auditory Processing

We used a backward masking task to assess aspects of temporal processing. In backward masking, a brief tonal signal precedes a long-duration noise masker. We patterned this listening task after one developed by Wright et al. (1997), who reported deficits in temporal resolution in children with language impairment. Like Wright et al., we found that some children with language impairment had difficulty distinguishing the signal from the masker that followed. However, some children with language impairments performed relatively well on the backward masking measure at pretest. In our study, children in all four groups evidenced improvement on the backward masking measure over time regardless of the level of their pretest performance. These results suggest that individual instruction that involves intensive daily experiences in listening and responding to verbal input had a positive effect on the ability to perform a backward masking task, whether that intervention was delivered by a human agent or by a computer software program.

What mechanism might be responsible for deficits and improvement on our temporal processing measure? There appear to be two schools of thought on the topic. One view is that difficulties with temporal processing occur because the incoming stream of information is altered so that fluctuations in the waveform are smeared in the time domain. If this were the case, then a child with language impairment would have difficulty encoding auditory input, which would yield a poor-quality template for learning language. Although there is no direct evidence that this actually occurs, physiologic data that indicate auditory brainstem responses recorded from children with specific language impairment (Marler & Champlin, 2005) or more generalized learning impairment (e.g., Kraus et al., 1996) are significantly different from those obtained from children with typically developing language who were matched by age and nonverbal intelligence. Because the auditory brainstem response is not influenced by the participant’s state of consciousness, cognitive factors such as attention, memory, motivation, and so on, are not believed to affect the magnitude of this obligatory physiologic response. Therefore, some of the children in the present study may have been afflicted by suboptimal neural encoding, and their participation in generalized listening experiences may have changed the functional anatomy and physiology of brain areas that were used for creating auditory templates.

Alternatively, it is possible that general cognitive abilities related to attention, memory, and perception play a similar role in detecting a signal in the presence of backward masking. Children with language impairments may have generalized information-processing difficulties (Bishop et al., 2006; Ellis Weismer & Evans, 2002; Hoffman & Gillam, 2004; Johnston, 2006). If this were the case, then detecting the signal in the presence of backward masking would be a relatively difficult task for a number of reasons. It has been suggested that, with appropriate training, signal thresholds in backward masking can be lowered. For example, Marler et al. (2001) reported that computer-based training with FFW-L over a 5-week period yielded progressively lower signal thresholds in a small group of children with language impairment. In addition, providing a brief sound in the opposite ear to cue the presence of the signal has long been known to reduce backward masking in adults (e.g., Puelo & Pastore, 1980). Introducing a contralateral marker may have helped children learn how to listen for the signal, which may have led to lower thresholds in backward masking. It is reasonable to suspect that both sensory and nonsensory factors contribute to a person’s temporal processing ability. The three treatment conditions and the active comparison condition all impacted the cognitive functions related to the information-processing skills required for simultaneously listening to verbal information and looking at related visual stimuli for long periods of time. Even though children with language impairments are likely to represent a heterogeneous population in terms of the source and the magnitude of the cognitive processing deficits that contribute to their language impairments, it appears that most of these children can profit from instructional programs that stimulate multiple cognitive functions.

Critical Components of Language Intervention

Fast-ForWord Language was designed to capitalize on the cerebral cortex’s demonstrated ability to reorganize
in response to focused, intensive stimulation. The core principles of the FFW-L program included heavy repetition, intensive intervention across successive days (1-hr and 40-min sessions, 5 days/week for 4–6 weeks), active attention to each trial, motivation through behavioral reinforcement, immediate feedback about response correctness, and programs that adapt to increased skill by increasing difficulty as a function of improvement (Agocs et al., 2006; Tallal, 2004). These principles were based on studies of training paradigms that led to neural change in animal models (Merzenich et al., 1996). The outcomes of this RCT inform our understanding of the necessity of the core principles of neuroplasticity-based training for language intervention with children. Many of the games in the AE condition were not adaptive, involved minimal repetition, and offered no feedback on accuracy. Because the children in the AE condition had language outcomes that were very similar to those of the children in the FFW-L condition, some basic behavioral teaching principles that lead to behavioral and neurophysiological change in animal models may not be necessary components of protocols for teaching language to human beings. We did not obtain objective measures of neurophysiological change; however, a clear relationship between brain and behavioral changes has been established for children who received FFW-L training (Tallal & Gaab, 2006). One can reasonably assume that any kind of teaching or intervention leads to neural reorganization. Adaptive responses, repetition, and feedback on accuracy probably play a role in human learning, but their roles may be mitigated by higher cognitive and linguistic functions.

**Modified speech stimuli.** The publishers of FFW-L have claimed that an additional important component of this instructional software is the adaptive modification of auditory stimuli that promotes rapid auditory processing (Merzenich et al., 1996; Tallal et al., 1996). We wanted to know whether modified speech stimuli were better for promoting language development in children with language impairment compared with unmodified speech (as in our CALI condition). Consistent with earlier findings by Gillam, Crofford, Gale, and Hoffman (2001), our results indicate that the modified speech in FFW-L was not a necessary component of interventions that positively affect immediate or long-term language or auditory processing outcomes.

Children in the FFW-L and CALI conditions had better outcomes on a measure of phonological awareness than children in the AE and ILI conditions. We did much to equate FFW-L and CALI, but there were many factors other than modified speech that distinguished these programs. A strict comparison would have required us to use FFW-L stimuli exactly, but without using the algorithm for modifying the speech children hear. We were not able to make these changes in the FFW-L program. The exercises in our CALI condition had natural speech, but they had a number of other differences as well. We really compared two different commercially available packages, both of which were delivered via computer and both of which were very intensive. This made an interesting and important contrast, but it is not a contrast that was based solely on modified speech. The critical point is that a computer program that did not contain modified speech yielded the same improvements in phonological awareness as a program that did.

**Intervention agent.** We found that children who received one-to-one language intervention with a certified SLP fared as well as children who interacted with computers during their intervention time. However, one should consider that the time investment of the SLP is greater than that of the computerized interventions. In the three computer conditions, 6 children received the computerized interventions at one time, with two or three monitors to oversee the children's play and to encourage the children when needed. In contrast, in the ILI condition children worked with SLPs in one-to-one sessions. It is possible that the ILI condition could be adjusted to a group setting, and we are currently conducting studies to evaluate this possibility.

With respect to the two computer interventions versus the human (ILI) intervention, we believe we designed our study to take advantage of the best that each modality had to offer. The computer-directed intervention provided many discrete trials in a limited period of time, and the computer control allowed stimuli to vary slightly along specified dimensions. On the other hand, human (clinician) interaction afforded opportunities for the clinician to respond directly to the child's intended meaning within the context of goal-directed social interaction. It does not appear that differences related to discrete-trial learning and goal-directed social interaction played a key role in the outcomes of this study. Our results suggest that both types of intervention can lead to clinically relevant changes on a global language test.

**Comparisons With Previous RCTs**

Our results are consistent with Cohen et al.'s (2005) RCT, which was conducted with a smaller group of children who had severe mixed language impairment in that there was no additional benefit of FFW-L compared with another computerized intervention without modified speech. However, it is important to note that the children in our RCT made larger gains on a global language test than the children in Cohen et al.'s (2005) RCT. This may have occurred because the children in their study, as a group, had more severe language impairments than the children in our study.

Our results were not consistent with those of Pokorni et al. (2004). They did not find significant improvement in
language skills following FFW-L intervention, yet we did. This difference in results may be due to participant characteristics as well as the statistical power of the studies. Their participants had reading disorders, and almost half of the children were at risk because of poverty. The inclusion of children at risk for language difficulties may have meant that the participants included children whose language weaknesses were due to reduced exposure to a language-rich input and/or reduced opportunities to have meant that the participants included children whose language weaknesses were due to reduced exposure to a language-rich input and/or reduced opportunities to engage in language-stimulating interactions and not due to the kinds of underlying factors that lead children to have problems learning language in the presence of ample input and interactions that would normally foster language development. In addition, Pokorni et al. had 16 to 20 children in their intervention groups, whereas we had 54 children per intervention group. Such factors may have limited the benefits of FFW-L intervention in their study. Future studies with children who have language impairments and who are at risk because of poverty will shed light on this finding.

**The AE Active Comparison Condition**

A final issue is our finding that children who were randomly assigned to an active comparison condition made improvements on our language and auditory processing measures that were similar to the improvements by participants in the three treatment conditions. Recall that we did not hypothesize that the children in the AE condition would make significant gains on our outcome measures. We are somewhat limited in our ability to answer why the AE comparison condition was effective for some children. The computer games that we used varied in the amount of oral language that was heard by the children. For example, in the computer game Coin Critters, children received very little verbal input. They primarily added coin values and heard “Okay” or “Oops” in response to a correct or incorrect answer. In contrast, computer games such as Magic School Bus Discovers Flight contained some areas in the program in which new vocabulary was introduced and defined, whereas other areas of the same program had an art activity with little verbal input. Although we could attempt to discern which aspects of each software package may have facilitated language, the children were free to choose whichever parts of the academic enrichment games that they wanted to pursue each day. Thus, some children may have spent more times in the picture gallery of Zurk’s Rainforest snapping pictures of animals, whereas others may have spent more time categorizing and hearing labels of rainforest animals.

We evaluated the commonalities across the conditions, and there appear to be three possible explanations for the language improvements that were seen in the children who were randomized to the AE condition. The first is related to our overall program. The children attended school for nearly 4 hr each day, of which only 1 hr and 40 min was spent in their assigned condition. During the remainder of the day, the children participated in crafts, recess, snack, and other activities. There was no structured language intervention during this time or a trained professional targeting language skills. However, many parents reported that their child developed friendships in the classroom setting. Children with language impairments typically are the least likely to be selected as a friend in the classroom (Gertner, Rice, & Hadley, 1994). Because the children were grouped with other children with language impairments, friendships may have been made more easily in comparison to classrooms in which only one or two children have language impairments. As a result of new bonds, the children may have had a very socially oriented classroom environment that facilitated their language skills. We are not able to evaluate this hypothesis with data from our study, but further study of classrooms with children with language impairments over short periods of time, like that in this study, may lead to some important insights on the value of productive and mutually satisfactory social interactions among peers, particularly among and between children who have language impairments.

Another possible explanation for the positive outcomes for the children in the AE condition may be related to improvements in general information-processing skills. The AE computer games were not devoid of language; they contained words, sentences, and music that may have been facilitative to the children. Listening under headphones for 1.5 hr per day for 30 days may have tuned the children’s listening skills. In addition, the children were receiving visual stimulation via the computer games. Intensive and simultaneous auditory and visual experiences may have improved the children’s visual organization skills. Hoffman and Gillam (2004) found that children with specific language impairment had a smaller overall capacity for visual–verbal and visual–spatial information along with significant difficulty processing visual information when both verbal and spatial domains were activated in dual-processing tasks. The possibility exists that children who participated in the AE condition benefited from intense practice with visual and auditory stimuli as provided by the academic computer games and that their information processing systems were subsequently modified in a way that was beneficial for test-taking skills. Perhaps the instruction in all four conditions facilitated the ability to focus, execute, sustain, and shift attention, as well as the ability to encode information. If so, improved attention and information-processing abilities might be shared across the treatments, including stimulation with academic enrichment software. Additional research is required to examine the role of attention in language impairments.
and the potential impact of general facilitation of attention and information-processing abilities.

A third possibility is that the similar results for the active comparison condition and the three treatment conditions reflected the value of careful attention. All the clinicians who worked in this research study, regardless of the condition to which they were assigned, believed that they were participating in something important that was likely to be beneficial to children with language impairments. It is likely that this positive attitude was conveyed to parents during the various interactions they had with project staff. Staff members also made children feel special because they were involved in this study. It was apparent to us that parents expected good results from their children’s involvement in this study, and children were excited and happy about their participation. It is possible that the core variable in our results was not changes in children’s attention regulation; instead, it may have been having the right kind of attention paid to them.

Summary

This RCT was conducted to determine whether children with language impairments who received FFW-L intervention made more improvements on measures of language and auditory processing than children who received other types of interventions. Children who were randomized to the FFW-L condition did not present greater pretest-to-posttest improvement on a standardized measure of language and on a measure of backward masking in comparison to children who received computer-assisted instruction that taught auditory processing, memory, and language without modified speech stimuli (the CALI condition); children who received individual language intervention with a certified SLP (the ILLI condition); or children in a active comparison group who received academically oriented computer-assisted instruction that was not specifically designed to improve language, reading, or auditory processing skills (the AE condition). The children in all four conditions made similar amounts of improvement during intervention as measured by a global test of receptive and expressive language and a test of backward masking, and improvements on these measures continued for the next 6 months. Children in the FFW-L and CALI conditions made greater improvements on a measure of phonological awareness than children in the ILLI and AE conditions. Even though these results were not consistent with the temporal auditory processing hypothesis, they do not mean that auditory processing skills are not important for language development. As Johnston (2006) pointed out, language development requires listening, and temporal processing is necessarily part of listening to speech. However, mechanisms such as auditory selectivity and volitional mental focus may be as important for processing and comprehending speech as the specific auditory abilities that are related to temporal processing. The results of this study suggest that intensive language intervention experiences that require close attending and immediate responding to auditory and visual stimuli in combination with opportunities for socialization with same-ability peers and a great deal of positive attention from caring and interested adults should result in clinically relevant improvements in language and auditory processing skills in children with language impairments.

Acknowledgments

This research was funded by National Institutes of Health (NIH) Grant U01 DC04560, awarded to the first two authors. It also was supported by resources from NIH Grant P30 HD02528, Grant BNCID P30 DC005803 from The Kansas Mental Retardation and Developmental Disabilities Research Center at the University of Kansas, and the Lillywhite endowment at Utah State University.

We acknowledge the dedication and assistance of Emily Tobey, Lori Betourne, and Alicia Wanek. We also acknowledge the guidance of our Project Officer, Julia Gulya. We thank Steven Camarata, Chris Doolough, Judith Gravel, Fred Gruber, and Mark Espland, who served on our Data Safety Monitoring Board. Jack Fletcher, Judith Johnston, and Karen Rascati, who served on our Advisory Committee, provided important suggestions throughout the course of the project. Diane Anderson and William Clarke, at the University of Iowa, oversaw the data management with great diligence. This study could not have been conducted without the very able assistance of numerous speech-language pathologists, graduate students, and undergraduate students. We are grateful to the school districts that participated: Blue Valley Schools, Dallas Independent School District, Kansas City (Kansas) Public Schools, Lawrence Public Schools, Leander Independent School District, Pflugerville Independent School District, and Plano Independent School District. Last, but not least, we thank the children and their families for their long-term commitment to this study.

References


Received November 9, 2006
Revision received April 24, 2007
Accepted July 2, 2007
DOI: 10.1044/1092-4388(2008/007)
Contact author: Ronald B. Gillam, Department of Communicative Disorders and Deaf Education, 1000 Old Main Hill, Logan, UT 84322. E-mail: ron.gillam@usu.edu.
Appendix A. Book list for the Individualized Language Intervention (ILI) units.

Amazing Grace
Mary Hoffman
Dial Books for Young Readers (1991)

How Do Dinosaurs Say Good Night?
Jane Yolen

Saturday Night at the Dinosaur Stomp
Carol Diggory Shields, Scott Nash (Illustrator)
Candlewick Press (1997)

Little Grunt and the Big Egg
Tomie dePaola
Holiday House (1993)

Taps and Bottoms
Janet Stevens
Harcourt (1995)

Tacky the Penguin
Helen Lester
Houghton Mifflin (1988)

Strega Nona
Tomie De Paola

Doctor Me Di Cin
Roberto Piumini
Front Street/Lemniscaat (2001)

Can I Have a Tyrannosaurus Rex, Dad? Can I? Please?
Lois G. Grambling
Bridgewater Paperback (2000)

Miss Nelson Is Missing
Harry Allard and James Marshall
Houghton Mifflin (1977)

Alexander and the Terrible, Horrible, No Good, Very Bad Day
Judith Viorst
Aladdin Paperbacks (1972)
Appendix B. ILI literature-based language unit outline.

<table>
<thead>
<tr>
<th>Day</th>
<th>Minimum time (in minutes)</th>
<th>Activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>5</td>
<td>Prestory presentation: Semantic map</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Prestory presentation: Preparatory set</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Read entire story</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Narrative activity: Poststory presentation—discussion questions</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Semantic activity: Add vocabulary targets to New Word Book</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Semantic activity choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Syntax activity choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Syntax activity choice</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Narrative activity: Narrative retelling through pictography</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Phonological awareness choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Phonological awareness choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Semantics activity choice</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>Semantics activity: Review target vocabulary in New Word Book</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>Reread the entire story</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>Narrative activity: Create a retold book: Dictation—use target vocabulary</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Semantics activity choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Syntax activity choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Syntax activity choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Phonological awareness choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Phonological awareness choice</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Narrative activity: Audiotape of retelling</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If there is additional time: Repeat any activities that the child had difficulty with in previous day(s) or give child the choice of prior activities</td>
</tr>
<tr>
<td>3</td>
<td>10</td>
<td>Reread the entire story</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Semantics activity: Review the target vocabulary in New Word Book</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>Read the child’s retelling</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Narrative activity: Create a parallel book—include target vocabulary</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Semantics activity choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Syntax activity choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Syntax activity choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Phonological awareness choice</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>Phonological awareness choice</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>Narrative activity: Audiotape of a parallel story</td>
</tr>
<tr>
<td></td>
<td></td>
<td>If there is additional time, repeat any activities that the child had difficulty with in previous day(s) or give the child the choice of prior activities</td>
</tr>
</tbody>
</table>

Appendix C. Levels of language targets for the ILI units.

<table>
<thead>
<tr>
<th>Language target</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Narratives</td>
<td>1. Setting</td>
<td>1. Basic episode</td>
<td>1. Multiple episodes</td>
</tr>
<tr>
<td></td>
<td>2. Abbreviated setting</td>
<td>2. Added elements</td>
<td>2. Complex episode</td>
</tr>
<tr>
<td>Semantics</td>
<td>Five concrete nouns</td>
<td>Five concrete nouns</td>
<td>Four verbs (from Level 2) + 8 adjectives,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(from Level 1) + 4 verbs</td>
<td>nouns, or adverbs</td>
</tr>
<tr>
<td>Syntax: Grammatical morphology</td>
<td>1. Copula be</td>
<td>1. Auxiliary inversion</td>
<td>1. Wh-questions with modals</td>
</tr>
<tr>
<td>Syntax: Clause structure</td>
<td>1. Coordination and</td>
<td>1. Coordination but, or, so</td>
<td>1. Subject relative clauses</td>
</tr>
<tr>
<td></td>
<td>2. Simple infinitive complements</td>
<td>2. Subordinating conjunctions</td>
<td>2. Objective relative clauses</td>
</tr>
<tr>
<td>Phonological awareness</td>
<td>1. Rhyming</td>
<td>1. Initial and final sound identification</td>
<td>1. Blending and segmenting nonwords</td>
</tr>
<tr>
<td></td>
<td>2. Sound matching</td>
<td>2. Blending and segmenting words</td>
<td>2. Making words</td>
</tr>
</tbody>
</table>