Supplement

Language Intervention Practices for School-Age Children With Spoken Language Disorders: A Systematic Review

Frank M. Cirrin
Minneapolis Public Schools, MN

Ronald B. Gillam
Utah State University, Logan

As part of the current climate of accountability in our nation’s schools, speech-language pathologists (SLPs) face mandates to use instructional methods that have demonstrated effectiveness and efficiency and to apply evidence-based practice (EBP) when making assessment and intervention decisions (Individuals With Disabilities Education Act [IDEA], 2004). New graduates from personnel preparation programs who serve school-age children must be well prepared to apply EBP to clinical decision making when they are employed in schools. This requires that they begin their careers with knowledge of evaluation and intervention techniques that are supported by scientific evidence. EBP has been defined as “the conscientious, explicit, and judicious use of current best evidence in making decisions about the care of individual patients...[by] integrating individual expertise with the best available external clinical evidence from systematic research” (Sackett, Rosenberg, Gray, Haynes, & Richardson, 1996, p. 71). The American Speech-Language-Hearing Association (ASHA) has conceptualized the goal of EBP as the integration of (a) clinical expertise, (b) best current external evidence, and (c) client values to provide high-quality services that reflect the values, needs, and choices of the students and families who are served by SLPs (ASHA, 2004a).

ABSTRACT: Purpose: This systematic review focuses on peer-reviewed articles published since 1985 that assess the outcomes of language intervention practices for school-age students with spoken language disorders.

Method: We conducted computer searches of electronic databases and hand searches of other sources for studies that used experimental designs that were considered to be reliable and valid: randomized clinical trials, nonrandomized comparison studies, and multiple-baseline single-subject design studies.

Results: The review yielded 21 studies concerning the efficacy or effectiveness of language intervention practices with school-age children since 1985. Eleven of the studies limited participants to children in kindergarten and first grade, and no studies were located that focused on students in middle grades or high school. The relatively high quality of the studies that met our criteria, and the moderate-to-high effect sizes we calculated for the majority of studies, suggests that clinicians can have some confidence in the specific language intervention practices examined.

Conclusion: The fact that only 21 studies met our criteria means that there is relatively little evidence supporting the language intervention practices that are currently being used with school-age children with language disorders. We outline significant gaps in the evidence and discuss the implications for clinical practice in schools.

KEY WORDS: evidence-based practice, language disorder, school-age children
Readers are referred to ASHA’s position paper and technical report on EBP (ASHA, 2005a, 2005b) for further information on EBP in the discipline of speech-language pathology.

SLPs working in the schools provide services to students from every disability category in IDEA (2004), and these students present with a wide range of speech and language disorders (ASHA, 2004b). This review focuses on experimental research concerning the efficacy and effectiveness of language intervention provided to school-age students with spoken language disorders. These children make up the largest subgroup of students with communication disorders who receive intervention services from SLPs in schools (ASHA, 2004b). Children with language disorders as a primary disability, often referred to as having a specific language impairment (SLI) or a language learning disability (LLD) (Paul, 2001; Wallach & Liebergott, 1984), are a heterogeneous group who have difficulty acquiring, comprehending, and/or expressing themselves with spoken and/or written language. These children may have difficulties in one or more aspects of language, including grammatical morphology, complex syntax, semantics, and/or pragmatics. Many of these children present problems in language processing skills related to attention, speech perception, working memory, and phonological awareness. Their difficulties with learning and using language place them at significant risk for social and academic problems throughout the school-age years and into young adulthood.

Guidelines for clinical practice that have been developed through comprehensive reviews of the experimental evidence on the effectiveness of language intervention for school-age children are needed if SLPs are to make intervention decisions that are consistent with EBP. Unfortunately, published guidelines pertaining to language intervention with school-age children have yet to be developed. Treatment efficacy and effectiveness for children with specific language disorders as a primary disability has been the focus of a number of narrative reviews of the literature (Gallagher, 1998; Goldstein & Hockenburger, 1991; Law, 1997; Leonard, 1998; McClean & Woods Cripe, 1997; Yoder & McDuffie, 2002). These reviews provide some support that speech and language therapy is beneficial for young children, but there are a number of concerns with these reviews related to the quality of studies that were included and the resulting potential exaggeration of positive effect sizes (Law, Garrett, & Nye, 2004). Another factor that affects the generalizability of the findings is that the majority of the studies described in these reviews limit participants to children of preschool age (i.e., between the ages of 2 and 5 years).

Law et al. (2004) recently completed a meta-analysis of language intervention studies. A meta-analysis applies specific statistical procedures to a group of studies in order to compute an average effect size as well as other summary statistics of the direction, size, and precision of findings from individual studies. The results of this meta-analysis revealed a significant positive effect of intervention when children had phonological difficulties or expressive vocabulary difficulties. There was mixed evidence on the effects of intervention for children with expressive syntax difficulties, and the results for children with receptive language problems were inconclusive due to the limited number of studies that had been carried out.

Although this meta-analysis represents an important contribution to the literature on the efficacy and effectiveness of language intervention, the results do not address at least two issues of importance to SLPs who work with school-age children with language disorders. First, all but 2 of the 36 articles reviewed by Law et al. (2004) limited participants to preschool children under the age of 5 years, and the majority of those were toddlers between the ages of 2 and 3 years. No studies were included in the meta-analysis that had children over the age of 6 as participants. Second, the results of this meta-analysis did not address the efficacy or effectiveness of specific language intervention procedures. To determine effect sizes, intervention methods were collapsed across treatment target areas (i.e., expressive phonology, expressive vocabulary, expressive syntax, receptive phonology, receptive vocabulary, and receptive syntax). Thus, the results of this meta-analysis would not help clinicians select a particular treatment strategy.

To provide effective services to school-age children with language disorders, practicing clinicians and students in university training programs need information on language intervention strategies that have been demonstrated to be beneficial. This article attempts to help meet this need by systematically reviewing the available evidence on language intervention practices with school-age children with LLDs. A systematic review is a comprehensive overview of the scientific literature on a specific clinical question, which is developed using systematic procedures and explicit criteria to reduce bias. It describes the extent to which various treatment approaches are supported by the evidence (ASHA, 2004a). A systematic review identifies common themes and gaps in the literature without using the statistical techniques that are employed in a meta-analysis.

The clinical question that this review attempts to answer is: Are there high-quality studies (using experimental designs generally considered to be reliable and valid) in the past 20 years that document the efficacy or effectiveness of language intervention practices for school-age students with spoken language disorders? In addition, this article identifies gaps in the evidence base and areas in need of further research, and discusses implications for clinical practice in schools and for personnel preparation of SLPs.

METHOD

Computer searches of electronic databases were conducted to locate appropriate studies. A list of these databases and the search terms used are presented in Appendix A. In addition, the Bamford-Lahaye Children’s Foundation (2004) database on EBP in child language disorders and the following journals were searched to locate relevant literature: Language, Speech, and Hearing Services in Schools; Journal of Speech, Language, and Hearing Research; and American Journal of Speech-Language Pathology. We also examined the reference lists in the studies that were identified in our search for articles that we may have missed. One additional article accepted for publication (Gillam et al., in press) was also included. We did not search sources for articles that had not been peer reviewed (e.g., Dissertation Abstracts International, presentations, and textbooks) or publications in languages other than English; thus, the conclusions based on this review should be considered tentative. We reviewed all studies that met the following selection criteria:

- Studies had to focus on experimental measures of the outcomes of language intervention practices for students with spoken language disorders. Descriptions of language intervention programs without objective measures of treatment outcomes were not included.
- By language intervention practices, we meant that our review would consider a broad range of experimental studies of specific language intervention methods (e.g., modeling), service

Cirrin & Gillam: Language Intervention: Systematic Review

S111

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
delivery models (e.g., classroom-based, individual, group), dosage of intervention (e.g., amount and frequency of therapy), and so on. We also wanted to include both efficacy and effectiveness studies in our review. Efficacy studies examine possible cause–effect relationships between intervention variables and outcomes from therapy under ideal, laboratory conditions; effectiveness studies follow efficacy studies to determine whether the same or similar effects are observed under more typical clinical conditions (Fey, 2006; Pring, 2004; Robey & Schultz, 1998).

By spoken language disorders, we meant that our review would concentrate on students who had primary disabilities in the area of oral language, both expressive and receptive. We recognized that students with LLDs often have difficulties with written language but made the decision not to include studies that focused on literacy intervention (reading/writing skills or the comprehension and production of read/written vocabulary) because of constraints on time and resources, and because comprehensive reviews of studies with students with learning disabilities on intervention for building reading fluency (Chard, Vaughn, & Tyler, 2002) and for vocabulary instruction and word learning (Jitendra, Edwards, Sacks, & Jacobson, 2004) have been published recently. We understand that the scope of practice for SLPs includes literacy assessment and intervention for children, adolescents, and adults (ASHA, 2001), and SLP involvement in these activities is critical to the academic and social success of students with LLDs.

We did include studies on language intervention for problems with phonological awareness and metalinguistics in our review. Even though these areas of language are tied closely to reading and writing, experts in the field consider problems with phonological awareness as part of the oral language deficits that are experienced by many children with LLDs (e.g., Gillon, 2000).

* Studies had to include school-age students with language disorders as a primary disability. By school-age, we meant that students had to be in kindergarten through 12th grade, and/or over the typical age for kindergarten entry (>5 years). Studies on language intervention with preschool children (under the age of 5) with LLDs were not included in this review because of our focus on school-age students and constraints on time and resources. Further discussion of the implications of this decision appears below. Studies of language intervention (e.g., vocabulary, linguistic concepts) with “at risk” students (without identified language disabilities) in general education settings were also not included.

Studies of students with secondary language and communication problems related to autism spectrum disorder (ASD) and to cognitive disabilities or general developmental delays were not included in this review because comprehensive reviews of this literature have already been published. Readers are referred to Goldstein (2002) for a systematic review of language intervention studies for students with ASD and to Sigafos and Dragow (2003) for a review of studies of language intervention for students with cognitive disabilities.

* Studies that were included in the review had to be one of the following design types: randomized clinical trials (RCTs), meta-analyses of RCTs, and systematic reviews of RCTs (sometimes referred to as Level 1 evidence); nonrandomized comparison studies; or multiple-baseline single-subject design studies (sometimes referred to as Level 2 evidence). In RCTs, the investigator actively compares treatment and control groups that have been created by randomly assigning participants to the groups. Nonrandomized comparison studies contrast the outcomes of treatment groups that were comprised of students who were matched to each other on a particular variable or who were assigned to groups on any basis other than randomization. We also included multiple-baseline single-subject design studies because they demonstrate causality in a manner that is generally considered to be reliable and valid. An overview of levels of an evidence system for evaluating individual studies of treatment is presented in Appendix B.

* Studies had to be published since 1985 in peer-reviewed journals. We reasoned that our 20-year timeframe was sufficient because in the meta-analysis by Law et al. (2004), only 1 of 7 articles on expressive language intervention, none of the 10 articles on vocabulary intervention, and none of the 2 articles on receptive language intervention were published before 1985. In addition, it was our impression that articles older than 20 years generally do not include enough methodological detail for the reader to be able to replicate the language intervention. However, it is possible that some bias was introduced into our review by not including studies published before 1985. Except for the Gillam et al. (in press) study, all studies were published between 1985 and 2006.

We calculated and reported effect sizes when sufficient data (i.e., pre- and posttest means and standard deviations) were provided. Effect size is a method of quantifying the effectiveness of a particular intervention relative to some comparison intervention (i.e., quantifies the size of the difference between two groups). See Schuere and Justice (2006) for a tutorial on the interpretation of effect size. When no-treatment control groups were used, group effects were calculated using Cohen’s $d$, in which effect size = $M_1$ (posttest score of the experimental group) $- M_2$ (posttest score of the control group)/pooled standard deviation. When the design did not include a no-treatment control group, effect sizes were calculated for pretest to posttest differences for each of the treatments. When possible, we provided the effect sizes that were reported in the articles. When effect sizes were not listed in the article, we computed them using the following on-line effect size calculators: www.cemcentre.org/ebeuk/research/effectsize/Calculator.htm and web.uccs.edu/~becker/Psy590/escalc3.htm. With this method, an effect size of .2 is considered to be small, an effect size of .5 is medium, and an effect size of .8 or greater is large (Cohen, 1988).

The treatment effect size formula could not be applied to 2 studies that taught invented language forms (Connell & Stone, 1992; Weismer & Hesketh, 1993). Because the children in the 2 studies learned morphemes or words that they only heard during the treatment, all learning could be attributed to the intervention. The true effect size is the amount of learning divided by the standard deviation, and that was what we computed in these 2 cases.

Effect sizes for single-subject designs (in which a single child’s performance is compared over one or more baseline periods with treatment periods) were calculated by computing the percentage of nonoverlapping data (PND). This value indicates the percentage
of data points during the treatment phase that exceed the most extreme data point in the baseline phase (Scruggs, Mastropieri, & Casto, 1987). With this method, higher PNDs indicate stronger effects (90% = very effective treatments; 70%–90% = effective treatments; 50%–70% = questionable; <50% = ineffective) (Scruggs & Mastropieri, 2001).

RESULTS

The initial computer search yielded 593 published reports. Of these, 495 articles were rejected because they failed to meet at least one of our criteria based on the information available in the abstract, and 98 were accepted for a review of the full text of the article. Our hand search yielded 36 additional publications that were accepted for a review of the full text, for a total of 134 articles that we reviewed against our criteria. After eliminating articles that were nonexperimental descriptions of language intervention methods, that focused on speech sound disorders, or that studied populations other than children with LLDs (e.g., students with autism, cognitive disabilities, dyslexia, learning disabilities, and children with typically developing language), a total of 21 studies met all four selection criteria.

Not all studies that were included in the final review were designed and conducted equally well. An assessment of the quality of a design is important to SLPs as they review published research because certain aspects of study design influence the reliability and generalizability of results. As proposed by Gillam and Gillam (2006), each of the articles was coded for eight appraisal questions that can be used to help evaluate the quality of treatment research (see Table 1). We evaluated each appraisal point for all 21 studies. In general, a clinician can have more confidence in the results of a study that receives mostly yes answers to the eight questions in comparison to a study that receives few yes answers to the questions. These appraisal questions are more appropriate for group design studies than single-subject design studies. For single-subject design studies, we noted NA for those questions that did not apply to single-subject designs (such as, "Was there a comparison and/or control group?").

A summary of appraisal points for all 21 studies reviewed in this article is presented in Table 2. The group studies received an average of five of eight possible appraisal points (range = 4–6). The single-subject design studies received an average of 3.3 of 5 possible appraisal points (range = 2–5). With a few exceptions noted below, clinicians can have a moderate degree of confidence in the results of this set of studies.

Tables 3 through 7 provide a summary of the language intervention studies reviewed in this article grouped according to the specific aspects of language that were the target of the intervention (as recommended by Law, Garrett, & Nye, 2005). Even though only 21 studies were reviewed, we grouped the studies into the general domains of syntax/morphology (3 studies), semantics/vocabulary (6 studies), phonological awareness/metalinguistics (5 studies), language processing (4 studies), and pragmatics/discourse (2 studies) because they represent major categories of language treatment targets for intervention that is frequently provided to school-age students by SLPs in schools (ASHA, 2004b). We have attempted to offset the heterogeneity of the treatment targets in each category and, in some cases, the overlap in targets across categories, by describing the actual language targets in the tables.

Syntax and Morphology

Our search yielded only 3 Level 2 studies of interventions designed to treat aspects of syntax and morphology in school-age children. Two studies were nonrandomized comparisons of experimental and control groups (Bishop, Adams, & Rosen, 2006; Connell & Stone, 1992), and 1 study was a multiple-baseline single-subject design (Weismer & Murray-Branch, 1989). A total of 72 students with language impairments between the ages of 6,11 (years;months) and 8;11 participated in the studies. Three of four participants in the single-subject study had expressive disorders, and the fourth had both receptive and expressive language disorders. The type of language disorder was not described for the 32 participants in the Connell and Stone study. All 36 students in the Bishop et al. study had receptive language impairments.

The single-subject study (Weismer & Murray-Branch, 1989) assessed participants’ production of individualized grammatical

<table>
<thead>
<tr>
<th>Table 1. Critical appraisal questions for research studies.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Comparison control group</strong></td>
</tr>
<tr>
<td>Random assignment Participants</td>
</tr>
<tr>
<td>Initial group similarity Blinding</td>
</tr>
<tr>
<td>Measures Statistical significance</td>
</tr>
<tr>
<td>Practical significance</td>
</tr>
<tr>
<td>Did the individuals who conducted the assessments and analyzed the data know which groups (treatment, comparison, or control) the participants were assigned to?</td>
</tr>
<tr>
<td>Did the authors report p values (the probability that a difference between the mean scores of two or more groups would not occur by chance alone) that were less than .05?</td>
</tr>
<tr>
<td>Did the authors report ETA squared values (percent of variance accounted for) or standardized d values (number of standard deviations of difference between pre- and post testing or between groups) that were at least moderately large? If not, could moderately large values be calculated from the data that were provided in the article?</td>
</tr>
</tbody>
</table>

*Note. The appraisal system was developed for assessing group studies; therefore, single-subject design studies are at a disadvantage when they are appraised according to this system. From "Making Evidence-Based Decisions About Child Language Intervention in Schools," by S. Gillam and R. Gillam, 2006, Language, Speech, and Hearing Services in Schools, 37. Copyright 2006 by the American Speech-Language-Hearing Association. Adapted with permission.*

Cirrin & Gillam: Language Intervention: Systematic Review S113
<table>
<thead>
<tr>
<th></th>
<th>Comparison/control group</th>
<th>Random assignment</th>
<th>Participants</th>
<th>Initial group similarity</th>
<th>Blinding</th>
<th>Measures</th>
<th>Statistical significance</th>
<th>Practical significance</th>
<th>Number of appraisal points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Syntax</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Weisner &amp; Murray-Branch, 1989</td>
<td>NA (Single-subject multiple baseline)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Connell &amp; Stone, 1992</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Bishop, Adams, &amp; Rosen, 2006</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Semantics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>McGregor &amp; Leonard, 1989</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Wing, 1990</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Masterson &amp; Perrey, 1999</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Crowe, 2003</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Throneburg, Calvert, Sturm, Parmboukas, &amp; Paul, 2000</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Unclear</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Weisner &amp; Keshteh, 1993</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Phonological awareness</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>van Kleeck, Gillam, &amp; McFadden, 1998</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Gilion, 2000</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Gilion, 2002</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Blischak, Shah, Lombardino, &amp; Chiarella, 2004</td>
<td>NA (Single-subject design)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Seger &amp; Voeroheoven, 2005</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td></td>
<td>4</td>
</tr>
<tr>
<td>Language processing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Merzenich et al., 1996</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Tallal et al., 1996</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>Gillam, Crofford, Gale, &amp; Hoffman, 2001</td>
<td>NA (Single-subject multiple baseline)</td>
<td>Yes</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>Yes</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>Yes</td>
<td>5</td>
</tr>
<tr>
<td>Cohen et al., 2005</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
<td></td>
<td>6</td>
</tr>
<tr>
<td>Gillam et al., in press</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes (able to be computed)</td>
<td></td>
<td>8</td>
</tr>
<tr>
<td>Pragmatics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bellinson &amp; Oswald, 2003</td>
<td>NA (Single-subject multiple baseline)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Bedosian &amp; Willis, 1987</td>
<td>NA (Single-subject multiple baseline)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>Yes</td>
<td>3</td>
</tr>
<tr>
<td>Bellinson &amp; Oswald, 2003</td>
<td>NA (Single-subject multiple baseline)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>No</td>
<td>Yes</td>
<td>NA (Single-subject design)</td>
<td>Yes</td>
<td>3</td>
</tr>
</tbody>
</table>
Table 3 (p. 1 of 2). Language intervention studies focusing on syntax and morphology.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Participants</th>
<th>Research design</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weismer &amp; Murray-Branch (1989). Modeling versus modeling plus evoked production training: A comparison of two language intervention methods. <em>Journal of Speech and Hearing Disorders, 54</em>, 269–281.</td>
<td><strong>Number of participants:</strong> 4</td>
<td>Level 2: Single-subject (alternating treatment design with baseline)</td>
<td>Dependent variables were number of correct productions of target grammatical form and total attempts at target form.</td>
</tr>
<tr>
<td></td>
<td><strong>Age/grade:</strong> 5½–6½;11 (years;months)</td>
<td><strong>Treatment targets</strong> Each participant had different grammatical targets as per a language evaluation.</td>
<td>No marked difference in outcome between modeling and modeling plus evoked production training for any participant.</td>
</tr>
<tr>
<td></td>
<td><strong>Disability:</strong> Specific language impairment (SLI)</td>
<td>Participant A: present progressive –ing, 3rd person singular –s, auxiliary “be” forms, copula “be” forms</td>
<td>For the 3 participants with only production delay, both treatments were equally effective; experimental control demonstrated across participants from baseline to treatment phase (increase in use of specific targets).</td>
</tr>
<tr>
<td></td>
<td>Three participants (A, B, D) were –1 SD on production (mean length of utterance [MLU]; normed test); 1 participant (C) was –1 SD in both production and comprehension</td>
<td>B: regular &amp; irregular past tense, nominative case pronouns, auxiliary inversion for questions</td>
<td>For the 1 participant with delay in both production and comprehension, neither treatment led to an increase from baseline over 15 treatment sessions.</td>
</tr>
<tr>
<td></td>
<td><strong>Class placement:</strong> Students in public schools, also receiving service at university clinic</td>
<td>C: articles, present progressive –ing, 3rd person singular –s, auxiliary “be” forms, copula “be” forms</td>
<td><strong>Effect sizes (Percentage of non-overlapping data [PND] for either treatment over baseline):</strong></td>
</tr>
<tr>
<td></td>
<td><strong>Service provider:</strong> Investigator provided intervention</td>
<td>D: articles, present progressive –ing, auxiliary “be” forms, copula “be” forms, prepositions (in, to, at)</td>
<td>Participant A PND = .83 (moderate)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Therapy methods</td>
<td>Participant B PND = 1.00 (very high)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Treatment 1 (T1): Modeling (M) — focused models of target form provided by clinician; no verbal responses required from child; activities included storytelling, art construction activities, puppet play.</td>
<td>Participant C PND = .20 (ineffective)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Treatment 2 (T2): Modeling plus evoked production (MEP) — focused models of target form with intermittent opportunities for participants to produce the form and receive feedback regarding correctness of production; training environment structured to evoke spontaneous (non-imitative) productions of target form.</td>
<td>Participant D PND = .64 (mildly effective)</td>
</tr>
<tr>
<td>Connell &amp; Stone (1992). Morpheme learning of children with specific language impairment under controlled instructional conditions. <em>Journal of Speech and Hearing Research, 35</em>, 844–852.</td>
<td><strong>Number of participants:</strong> 32 SLI; 24 age-matched and 20 language-matched controls</td>
<td>Level 2: split-plot factorial</td>
<td>Dependent variables were pre- and posttest performance on both production and comprehension probes of invented morphemes.</td>
</tr>
<tr>
<td></td>
<td><strong>Age/grade:</strong> 5½–6½;11</td>
<td><strong>Treatment targets</strong> Investigator-designed invented morphemes: one of four suffixes attached to concrete nouns capable of being represented by pictures, to indicate one of the following meanings (large/small, whole/part, whole/broken) (e.g., “TVum” to indicate a broken TV).</td>
<td>Modeling alone did not significantly increase SLI participants’ use of morphemes; imitation training did result in SLI participants’ increased performance for morpheme production. Both modeling and imitation training appeared equally effective for increasing SLI participants’ performance on morpheme comprehension probes.</td>
</tr>
<tr>
<td></td>
<td><strong>Disability:</strong> SLI; –1 SD on either MLU or normed language test</td>
<td></td>
<td>However, results from the study were confounded by large effects of order of treatment procedure (whichever treatment strategy administered first overrode any differential effect within the counterbalanced design).</td>
</tr>
<tr>
<td>Citation</td>
<td>Participants</td>
<td>Research design</td>
<td>Major findings</td>
</tr>
<tr>
<td>----------</td>
<td>--------------</td>
<td>----------------</td>
<td>---------------</td>
</tr>
<tr>
<td><strong>Classroom placement:</strong> Participants selected from caseloads of SLPs in schools and clinics, but no other breakdown described</td>
<td>Therapy methods</td>
<td>For SLI students in modeling condition for: a) Production probes .78 (moderate) b) Comprehension probes 1.32 (large)</td>
<td><strong>Effect sizes</strong> For SLI students in imitation condition for: a) Production probes .50 (moderate) b) Comprehension probes 1.30 (large)</td>
</tr>
<tr>
<td><strong>Service provider:</strong> Assumed to be investigators, but not described</td>
<td>Participants taught set of invented morphemes by computerized language teaching program (to standardize teaching procedures and minimize unintentional examiner bias). Modeling treatment: Participant heard morpheme being used in meaningful way but was not asked to repeat or use the morpheme in any way. Imitation treatment: Same as modeling with additional step of instructing participant to give a direction to a computer cartoon figure that contains the target morpheme.</td>
<td><strong>Primary outcome variable:</strong> Test of Reception of Grammar—2 (TROG—2). <strong>Secondary outcome variables:</strong> Expression, Reception, and Recall of Narrative Instrument (ERRNI) and a variety of speech and nonspeech auditory processing measures. Children in the slow speech condition, the modified speech condition, and the control condition did not differ on any measure. Potential improvements from time 1 to time 2 were not assessed.</td>
<td></td>
</tr>
<tr>
<td><strong>Bishop, Adams, &amp; Rosen (2006).</strong> Resistance of grammatical impairment to computerized comprehension training in children with specific and non-specific language impairments. <em>International Journal of Language and Communication Disorders, 41</em>(1), 19–40.</td>
<td><strong>Number of participants:</strong> 36 Level 2: nonrandomized comparison</td>
<td><strong>Effect sizes (Pre-post)</strong> Slow speech condition TROG—2: .77 ERRNI comprehension: -.17 ERRNI MLU: -.84 Modified speech condition TROG—2: .37 ERRNI comprehension: −.03 ERRNI MLU: −.44 Untrained condition TROG—2: .53 ERRNI comprehension: −.74 ERRNI MLU: −.68</td>
<td></td>
</tr>
<tr>
<td><strong>Age/grade:</strong> 8–13 years Receptive training of grammatical skills in order to make sentence interpretation more fluent, accurate, and automatic.</td>
<td><strong>Disability:</strong> Receptive language impairments</td>
<td><strong>Therapy methods</strong> Children moved or activated objects on a computer screen to match spoken sentences that increased in syntactic complexity.</td>
<td></td>
</tr>
<tr>
<td><strong>Classroom placement:</strong> Recruited from four special language education schools and three special classes within mainstream schools in England</td>
<td><strong>Service provider:</strong> The computer intervention was provided in schools by school staff</td>
<td>Slow speech condition: there was a 1.2 s delay between the end of one phrase within a sentence and the start of the next. Modified speech condition: same algorithm for modifying rate and amplification of frequencies as used in Fast ForWord—Language. Untrained: children did not receive the computerized training. The authors do not state whether the children continued to receive language services in school or not. Participants received 20 computerized language-teaching sessions, 15 min each, administered on consecutive school days.</td>
<td></td>
</tr>
<tr>
<td>Citation</td>
<td>Participants</td>
<td>Research design</td>
<td>Major findings</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>
Age/grade: 9;1-10;5  
Disability: All diagnosed as SLI; all students scored more than -1 SD on at least two normed tests of language and word finding  
Class placement: All enrolled in class for SLI students at their school  
Service provider: Not specified but assumed to be investigator | Level 2: nonrandomized comparison  
Two participants participated in 6 weeks of language therapy with word-finding focus; other 2 served as controls and participated in narrative and syntax activities. Nonrandom assignment of participants to treatment and control groups. Controls matched on age and word-finding abilities.  
Treatment targets  
Investigator-designed elaboration strategies (to provide child with a richer knowledge of target words) and retrieval strategies (teach child to use information already known about target words) thought to aid word retrieval taught via a set of 120 concrete nouns capable of being represented by pictures. | Dependent variables were number of errors on naming task; number of words recalled in free recall task.  
On naming and recall tasks, treatment participants improved from pre- to posttest, while controls either made smaller gains or no gains. Performance on maintenance tasks, however, was sometimes poorer than posttest for all groups. Data did not clearly indicate whether elaboration or retrieval training was more effective.  
Effect sizes cannot be computed from data presented in the article. Difference scores for each participant are presented from pre- to posttest on the number of errors on the naming task and the number of words recalled in the free recall task. |
Age/grade: Range = 71-85 months, mean = 77 months; first grade  
Disability: Identified by school as having severe language impairment  
Class placement: All children were in a self-contained class for language disabilities; intervention was provided in the school | Level 2: nonrandomized comparison  
Students divided into two groups: one received semantic treatment; the other received phonological treatment; not random assignment.  
Each group received 30 group therapy sessions, each 25 min long, over 2.5 months.  
Treatment targets  
Investigator-designed semantic or phonological strategies thought to aid word retrieval taught via set of vocabulary items similar to but not duplicating items on the Test of Word Finding (TWF). | Dependent variables were pre- and posttest scores on the TWF  
Significant gains on TWF scores for phonological treatment group; nonsignificant gain for semantic treatment group, even though all participants but one (who was in semantic group) improved TWF scores from pre-to posttest.  
Effect sizes  
Semantic treatment group:  
$d = .6$ (moderate)  
Phonological treatment group:  
$d = .7$ (moderate) |
<table>
<thead>
<tr>
<th>Citation</th>
<th>Participants</th>
<th>Research design</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Masterson &amp; Perrey (1999). Training analogical reasoning skills in children with language disorders. <em>American Journal of Speech-Language Pathology</em>, 8, 53–61.</td>
<td>Service provider: Investigator was students’ regular SLP</td>
<td>Therapy methods: Semantic treatment: Activities to improve elaboration and organization of semantic storage (put pictures/words in categories; supply attributes; use categories and attributes to define words; explain function and attributes of objects). Phonological treatment: Activities for phonological segmentation (count syllables and phonemes, match and supply rhyming words) and activities for auditory imagery (hear picture name in their mind, silent verbal rehearsal).</td>
<td>Dependent variable was performance on 100-item verbal analogies test (20 from each of the following categories: synonym, antonym, functional, linear order, category membership). None of specific analogies in pre- and posttest were used in treatment.</td>
</tr>
<tr>
<td></td>
<td>Number of participants: 12</td>
<td>Treatment targets: Investigator-designed mediated learning and bridging strategies thought to aid in analogical reasoning (verbal analogy performance) taught via a set of verbal analogies from five categories (synonyms, antonyms, function, linear order, and category membership).</td>
<td>Average gain made by students receiving treatment was 2.37 SDs more than participants who did not receive treatment (large effect size). Performance appeared to be consistent across all five semantic analogy categories.</td>
</tr>
<tr>
<td></td>
<td>Matched 6 treatment to 6 controls (not randomly assigned; treatment volunteers expressed willingness to attend extra treatment sessions)</td>
<td>Therapy method: Phase 1 treatment: Mediated learning, which included direct instruction in component processes of analogical thinking (encoding, inferring, mappings, applying); Phase 1 treatment consisted of 8 sessions over 2 weeks. Phase 2 treatment: Bridging, which included exercises designed to help participants incorporate the component processes of analogical thinking into everyday activities; Phase 2 treatment consisted of 8 sessions over 2 weeks.</td>
<td>Direct instruction in analogical reasoning has positive effects in children with language disorders. No measure of generalization of trained strategies to classroom curricular activities.</td>
</tr>
<tr>
<td></td>
<td>Age/grade: Between 9 and 14 years (mean = 11.1)</td>
<td></td>
<td>Effect sizes</td>
</tr>
<tr>
<td></td>
<td>Disability: SLI</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Class placement: All receiving services for language in public schools or private schools</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Participants received treatment in small groups; assume in pull-out school setting but not described</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Service provider: Training administered by one of investigators</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Crowe (2003). Comparison of two reading feedback strategies in improving the oral and written language performance of children with language-learning disabilities. <em>American Journal of Speech-Language Pathology</em>, 12, 16–27.</td>
<td>Number of participants: 12</td>
<td>Level 2: nonrandomized controlled trial</td>
<td>Dependent variables were pre- and posttest difference scores on the Gray Oral Reading Test—Revised (reading comprehension questions, oral reading rate/accuracy, and combined score) and on the Comprehensive Receptive and Expressive Vocabulary Test (CREVT) (normed test of expressive and receptive vocabulary; participants had to point to pictures and verbally define vocabulary items).</td>
</tr>
<tr>
<td>Citation</td>
<td>Participants</td>
<td>Research design</td>
<td>Major findings</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>Throneburg, Calvert, Stum, Parmboukas, &amp; Paul (2000). A comparison of service delivery models: Effects on curricular vocabulary skills in the school setting. <em>American Journal of Speech-Language Pathology</em>, 9, 10–20.</td>
<td><strong>Number of participants</strong>: 32 (who were eligible for S/L service out of 177 total students in 12 classrooms)  <strong>Age/grade</strong>: Grades K–3  <strong>Disability</strong>: 13 eligible for speech services; 19 eligible for language services; criteria for placement was a score of −1.5 SD or greater on two normed tests of language or one normed test of articulation  <strong>Class placement</strong>: Regular education classrooms  <strong>Service provider</strong>: Students’ regular school SLP; classroom teachers; graduate students</td>
<td><strong>Treatment targets</strong>  Investigator-designed traditional reading decoding strategies or meaning-based communicative reading strategy (CSR) thought to aid oral language performance (e.g., receptive and expressive vocabulary performance).  <strong>Therapy methods</strong>  <strong>(a)</strong> Collaborative: SLP and classroom teacher collaboratively planned and implemented activities to target curriculum vocabulary words in the classroom (large group vocabulary instruction and hands-on activities for topic units with embedded vocabulary words).  <strong>(b)</strong> Classroom-based: Teacher and SLP independently planned and implemented vocabulary activities, similar to (a) above; SLP provided independent classroom lesson.  <strong>(c)</strong> Traditional pull-out: SLP provided vocabulary instruction for S/L impaired students in traditional pull-out sessions averaging 50 min per week. SLP used same materials that were used in collaborative and classroom-based settings.</td>
<td>Standardized measures of receptive and expressive vocabulary indicated that children in the CRS group performed better than children in either the T1 or control group, but group differences were not statistically significant. All groups performed worse on posttest receptive and general vocabulary measures (an alternate form of the CREVT given as pretest); CRS group increased performance on alternate form expressive posttest; T1 and control groups performed worse.  <strong>Effect size</strong>  For the CRS treatment group’s increase on expressive posttest: ( d = .5 ) (moderate)</td>
</tr>
</tbody>
</table>
**Table 4 (p. 4 of 4).** Language intervention studies focusing on semantics and vocabulary.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Participants</th>
<th>Research design</th>
<th>Major findings</th>
</tr>
</thead>
</table>
| Weismer & Hesketh (1993). The influences of prosodic and gestural cues on novel word acquisition by children with specific language impairment. *Journal of Speech and Hearing Research*, 36, 1013–1025. | **Number of participants:** 16 (8 SLI and 8 normal language)  
**Age/grade:** 5;1–6;7/kindergarten  
**Disability:** Treatment group: SLI; controls: normal language  
**Class placement:** Treatment group: Self-contained classroom for children with severe language disorders in a public school kindergarten  
**Controls:** Regular kindergarten class  
**Service provider:** Assume investigator, but not stated | **Level 2:** nonrandomized comparison  
**Treatment targets**  
Meanings of a set of 9 investigator-designed novel words (one-syllable, consonant/vowel/consonant forms with early developing sounds) representing either object labels or locatives.  
**Therapy methods**  
Investigators varied the presentation of target words in 3 experimental treatment conditions:  
**Rate:** Stimulus sentences with target words presented at 3 rates (slow, normal, fast)  
**Stress:** Stimulus sentences with target words presented with and without emphatic stress  
**Visual:** Stimulus sentences with target words presented verbally or verbally with an accompanying iconic gesture | Dependent variables were number of novel words comprehended and produced.  
For therapy group and controls, acquisition of novel words was affected by alterations in speaking rate and by the use of gestures accompanying spoken language, but not by stress manipulations. Slower presentation rate helped SLI performance as did the addition of gesture cues.  
**Effect sizes:** For SLI students, effect sizes calculated from individual subject difference scores for comparisons:  
Slow-Fast rate for comprehension probes: 1.1 (large)  
Slow-Fast rate for production probes: 1.1 (large)  
Emphatic-Neutral stress for comprehension probes: .12 (no effect)  
Emphatic-Neutral stress for production probes: .74 (moderate)  
Gesture-No gesture for comprehension probes: .57 (moderate)  
Gesture-No gesture for production probes: .33 (small) |
Table 5 (p. 1 of 2). Language intervention studies focusing on phonological awareness and metalinguistics.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Participants</th>
<th>Research design</th>
<th>Major findings</th>
</tr>
</thead>
</table>
Disability: Speech and/or language disorder  
Age/grade: Preschool-age mean = 48 months and prekindergarten mean = 60 months  
Class placement: Preschool and prekindergarten classrooms in a private school for children with language disorders  
Service provider: SLPs and graduate students in speech-language pathology | Level 2: nonrandomized comparison of treatment groups with a control group of older children who had attended the same classrooms.  
Treatment targets: Rhyming: rhyme identification, rhyme judgment, and rhyme generation.  
Phoneme awareness: matching and identifying initial sounds, generating words, phoneme blending, and phoneme segmentation.  
Therapy methods: Children in the two treatment groups received rhyming instruction during the fall semester and phoneme awareness instruction during the spring semester. | Both treatment groups made significant improvements in rhyming and phoneme awareness. Gains in rhyming fell below the lower boundary of the 95% confidence interval of the control group, suggesting that the development in rhyming was not dependent on treatment. However, gains in phoneme awareness were above the upper limits of the control group's confidence interval, suggesting that training contributed to improvements in phonemic awareness. The children in the treatment groups performed better on the phonological awareness tasks than did older kindergartners and first graders who had previously attended the pre-K class.  

Effect sizes  
d (preschool vs. control) = 1.58 (large)  
d (pre-K vs. control) = 1.76 (large) |
Age/grade: 5:6–7:6  
Disability: Speech and language disorder  
Class placement: Pull-out services in clinic and school settings compared to classroom consultation or classroom instruction  
Service provider: SLPs and graduate students in SLP | Level 2: nonrandomized (matched) assignment to treatment and control groups compared with a cohort of typically achieving children.  
Treatment targets: Experimental group: identification of phonological similarities, phoneme manipulation, sound identification, phoneme segmentation, grapheme-phoneme correspondence, phoneme production.  
Traditional control: phoneme production in isolation, syllables, words, and phrases.  
Therapy methods: 4 groups:  
1. Experimental Intervention: Gillon phonological awareness training program  
2. Traditional Control: Van Riper speech therapy  
3. Minimal Intervention  
4. Classroom Consultation: normal comparison | Children in Group 1 made more improvement on tests of phonological awareness and reading than did children in Groups 2 and 3. At the end of the study, the phonological awareness performance of the children in Group 1 was similar to the normal controls.  

Effect sizes  
Group 1 vs. 2  
Phoneme awareness 2.58 (large)  
Rhyming .48  
Reading accuracy .59  
Reading comprehension .67 (moderate)  

Group 1 vs. 3  
Phoneme awareness 1.77 (large)  
Rhyming .67 (moderate)  
Reading accuracy .34  
Reading comprehension .61 |
<table>
<thead>
<tr>
<th>Citation</th>
<th>Participants</th>
<th>Research design</th>
<th>Major findings</th>
</tr>
</thead>
</table>
| Gillon (2002). Follow-up study investigating the benefits of phoneme awareness intervention for children with spoken language impairment. *International Journal of Language and Communication Disorders, 37*, 381–400. | **Number of participants:** 20 of the original 23 SLI and all of the normal controls  
**Age/grade:** 5;6–7;6  
**Disability:** Speech and language disorder  
**Class placement:** Pull-out services in clinic and school settings compared to classroom consultation or classroom instruction  
**Service provider:** SLPs and graduate students in speech-language pathology | **Level 2:** nonrandomized (matched) assignment to treatment and control groups compared with a cohort of typically achieving children  
**Treatment targets**  
Same as Gillon, 2000  
**Therapy methods**  
4 groups: Same as Gillon (2000) | Treatment led to sustained growth in phoneme awareness and word recognition. The majority of the children were reading at or above age-level expectations and improved on nonword spelling.  
**Effect sizes**  
Means and standard deviations were not provided for the phoneme awareness measure. They were provided for the word recognition measure.  
$d$ (SLI, pre-to follow-up) = 2.42 (large)  
$d$ (Controls, pre-to follow-up) = 1.52 (large)  
$d$ (SLI vs. controls at follow-up) = .64 (moderate) |
| Blischak, Shah, Lombardino, & Chiarella (2004). Effects of phonemic awareness instruction on the encoding skills of children with severe speech impairment. *Disability and Rehabilitation, 26*, 1295–1304. | **Number of participants:** 3  
**Age/grade:** 5;6–7;6  
**Disability:** Severe speech impairment with coexistent language disorder  
**Class placement:** Not stated  
**Service provider:** SLPs | **Multiple-baseline single-subject design across behaviors and participants**  
**Treatment targets**  
Phoneme–grapheme correspondence, phoneme segmentation, phoneme manipulation, pseudo word spelling.  
**Therapy methods**  
Treatment: Phoneme–grapheme instruction and phonemic awareness instruction. | Phoneme–grapheme correspondence and phonemic awareness instruction increased the encoding skills of 2 of the 3 participants and generalized these skills to untrained pseudo- and real words.  
**Effect sizes (PND for either treatment over baseline):**  
Phoneme manipulation: moderate PND, 85% of probes in instruction phase were above highest baseline level.  
Encoding: large PND, 98% of probes in instruction phase were above highest baseline level. |
**Disability:** SLI  
**Age/grade:** 4;6–6;11/kindergarten  
**Class placement:** Special schools for children with SLI in the Netherlands  
**Service provider:** Computer games | **Level 2:** nonrandomized (matched) comparison of 2 experimental groups and a control group  
**Treatment targets**  
Rhyming and phoneme synthesis.  
**Therapy methods**  
Group 1: 10 rhyming and sound synthesis computer games with normal speech.  
Group 2: 10 rhyming and sound synthesis computer games with modified speech (slower and amplified formant transitions).  
**Control group:** Vocabulary computer games.  
All children received 14 sessions, 15 min long, across 5 weeks. | Positive treatment results were found for Experimental Group 1 when the phonological awareness task results were combined into difference z scores. The positive results were no longer significant 18 weeks after completion of the intervention.  
**Effect size**  
The effect size of treatment for Experimental Group 1 compared with the control group: $d = 0.29$ (small). |
<table>
<thead>
<tr>
<th>Citation</th>
<th>Participants</th>
<th>Research design</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Merzenich, Jenkins, Johnston, Schreiner, Miller, &amp; Tallal (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. Science, 271, 77–80.</td>
<td>Number of participants: 22 Age/grade: 5.2–10.0 Disability: mixed (receptive and expressive) language impairment Class placement: Unknown Service provider: Computer games</td>
<td>Level 2: nonrandomized (matched) comparison of experimental and control groups Treatment targets: Perceptual identification of tone sequences and phoneme recognition. Therapy methods: Experimental group received FFW–L games: Circle Sequence and Phoneme Identification.</td>
<td>The students in the experimental group improved on the Tallal Repetition Test (a measure of auditory temporal processing), but the students in the control group did not. Effect sizes were not reported and cannot be calculated because means and standard deviations were not provided.</td>
</tr>
<tr>
<td>Tallal, Miller, Bedi, Byma, Wang, Nagarajan, et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. Science, 271, 81–84.</td>
<td>Number of participants: 22 Age/grade: 5.2–10.0 Disability: mixed (receptive/expressive) language impairments Class placement: Unknown Service provider: Computer games and trained clinicians</td>
<td>Level 2: nonrandomized (matched) comparison of treatment and control groups Treatment targets: Speech discrimination and on-line language comprehension. Therapy methods: Treatment group: early versions of FFW–L games (block commander, phonic match, phonic word, and language comprehension builder with modified speech stimuli). Control group: Computer versions of these same tasks without modified speech stimuli.</td>
<td>Participants in the experimental group showed significantly larger improvements on measures of speech discrimination, language processing, and grammatical comprehension. Effect sizes were not reported and cannot be calculated because means and standard deviations were not provided.</td>
</tr>
<tr>
<td>Gillam, Crofford, Gale, &amp; Hoffman (2001). Language change following computer-assisted language instruction with Fast ForWord or Laureate Learning Systems software. American Journal of Speech-Language Pathology, 10, 231–247.</td>
<td>Number of participants: 4 Age/grade: 6–8 years Disability: SLI Class placement: Regular classroom plus pull-out SLP services Service provider: Research assistants</td>
<td>Level 2: multiple-baseline single-subject design Treatment Targets: FFW–L: discrimination of tones, phonemes, syllables, and words; memory for commands; comprehension of grammatical morphology and complex sentences. Therapy methods: Laureate Learning software (LLS): comprehension and memory of words, grammatical morphemes, sentences and stories. Two children received FFW–L, 2 other children received a group of language games produced by Laureate Learning Systems (LLS).</td>
<td>All 4 children made clinically significant gains on the Oral and Written Language Scales. Both children who received the LLS games and 1 of the 2 children who received FFW–L made clinically significant gains on MLU computed from language samples. Effect sizes (PND for MLU in the follow-up phase vs. the baseline phase): LLS – Immediate 100% LLS – Delayed 100% FFW – Immediate 0% FFW – Delayed 100%</td>
</tr>
</tbody>
</table>
Table 6 (p. 2 of 2). Language intervention studies focusing on language processing.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Participants</th>
<th>Research design</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age/grade: 6–10 years</td>
<td><em>Treatment targets</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disability: mixed (receptive/expressive) language impairment</td>
<td>FFW–L: discrimination of tones, phonemes, syllables, and words; memory for commands; comprehension of grammatical morphology and complex sentences.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Class placement: Regular classroom plus pull-out for language therapy</td>
<td>Other computer programs: listening, spelling, phonological awareness, reading, writing, vocabulary, problem solving, narration, syntax, and morphology.</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Service provider: Computer plus school clinicians</td>
<td>Therapy methods</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Three groups:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>1. FFW–L</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>2. Other computer programs</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Control (school therapy services only)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Similar gains on the Clinical Evaluation of Language Fundamentals, Third Edition for the treatment and control groups at 9-week and 6-month follow-up periods. Suggests that the computer intervention plus school therapy was more effective than school therapy alone.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Effect sizes</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9 weeks:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .09$ for FFW vs. Control (no effect)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .27$ for FFW vs. Computer (small)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 months:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .05$ for FFW vs. Control (no effect)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .27$ for FFW vs. Computer (small)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Age/grade: 6–9 years</td>
<td><em>Therapy methods</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Disability: specific language impairment</td>
<td>4 groups:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Class placement: conducted in special summer school sessions with socialization groups plus a computer lab or an individual intervention session</td>
<td>1. FFW–L</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Service provider: Computer or SLP</td>
<td>2. Computer Assisted Language Intervention (CALI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>3. Individual Language Intervention (ILI)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>4. Academic Enrichment (AE)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td><em>Treatment targets</em></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>FFW–L: discrimination of tones, phonemes, syllables, and words; memory for commands; comprehension of grammatical morphology and complex sentences.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>CALI: discrimination of tones, phonemes, syllables, and words; memory for commands; comprehension of grammatical morphology and complex sentences.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ILI: vocabulary, grammatical morphology, complex syntax, narration, phonological awareness.</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>AE: listening, mathematical calculation, non-verbal problem solving, science, geography.</td>
<td>Children in all 4 treatment conditions improved significantly on a global language test, a test of backward masking, and a test of language comprehension. Children in the FFW–L and CALI groups made significantly more progress on a test of phonological awareness than did children in the AE and ILI conditions.</td>
</tr>
<tr>
<td></td>
<td></td>
<td><strong>Effect sizes (pre-post comparisons on the CASL):</strong></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>9 weeks:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .56$ for CALI (moderate)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .71$ for FFW–L (moderate)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .79$ for ILI (moderate)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .61$ for AE (moderate)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>6 months:</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .10$ for CALI (large)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .93$ for FFW–L (large)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = 1.33$ for ILI (large)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>$d = .34$ for AE (large)</td>
<td></td>
</tr>
</tbody>
</table>
Table 7. Language intervention studies focusing on pragmatics and discourse.

<table>
<thead>
<tr>
<th>Citation</th>
<th>Number of participants</th>
<th>Research design</th>
<th>Major findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beilinson &amp; Olswang (2003). Facilitating peer-group entry in kindergartners with impairments in social communication. Language, Speech, and Hearing Services in Schools, 34, 154-166.</td>
<td>3</td>
<td>Level 2: Single-subject (multiple baseline across subjects)</td>
<td>Dependent variables were the frequency of low- and high-risk entry behaviors and the frequency of prop use; investigators also measured the frequency of solitary vs. cooperative play.</td>
</tr>
<tr>
<td></td>
<td>Age/grade: 5.6–6.3/kindergarten</td>
<td>Treatment targets</td>
<td>Experimental control demonstrated for the intervention across all participants between baseline, treatment, and withdrawal phases.</td>
</tr>
<tr>
<td></td>
<td>Disability: “Social communication deficits” identified by SLP and teachers based on observation and comparison with normal peers; also, scores on normed language tests (range from -2 SD to “average” on Peabody Picture Vocabulary Test, Test of Language Development)</td>
<td>Investigator-designed peer-group entry behaviors (general and specific statements to peers, and use of desirable toys as entry props), and investigator-designed steps in a peer-group entry sequence taught via a set of Meyer-Johnson picture symbols.</td>
<td>Effect sizes: For participants between baseline and treatment and between treatment and withdrawal</td>
</tr>
<tr>
<td></td>
<td>Class placement: Full-day integrated classroom in a university lab school</td>
<td>Therapy methods</td>
<td>High-risk entry: All 3 students increased frequency of high-risk entry behaviors: d from 2.2 to 4.5 (large)</td>
</tr>
<tr>
<td></td>
<td>Service provider: Second-year master’s student in speech-language pathology and teacher in classroom</td>
<td>Direct instruction in the use of high-risk entry behaviors (general and specific statements to peers) and the use of props (desirable toy) to gain entry.</td>
<td>Prop use: All 3 showed increase in prop use: d from 2.5 to 10.2 (large)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Direct instruction included the use of Meyer-Johnson symbols for steps in an entry sequence (e.g., watch your friend, get a toy like your friend is using, do the same thing as your friend, tell an idea). Teachers were instructed to prompt students to use an entry sequence in the classroom.</td>
<td>Cooperative play: All 3 showed increase in frequency of cooperative play: d from 1.5 to 13.2 (large)</td>
</tr>
<tr>
<td></td>
<td>Age/grade: 60 months/kindergarten</td>
<td>Treatment targets</td>
<td>Experimental control and effectiveness of treatment demonstrated across behaviors.</td>
</tr>
<tr>
<td></td>
<td>Disability: SLI; expressive syntax at Brown’s Stage 4; only 2 of Brown’s 14 morphemes mastered (ing, on); few topics initiated spontaneously, and limited to “here and now”</td>
<td>Increase in frequency of topic initiations in child/clinician conversation related to memory (past events) and topic initiations related to future events.</td>
<td>Increase in the frequency of past topic initiations, even though the absence of toys was related to the number of past topics, thus confounding the interpretation.</td>
</tr>
<tr>
<td></td>
<td>Class placement: Regular kindergarten class</td>
<td>Therapy methods</td>
<td>Increase in the frequency of future topic initiations.</td>
</tr>
<tr>
<td></td>
<td>Service provider: Not specified</td>
<td>Two 30-min therapy sessions per week for 6 months.</td>
<td>Also, increase in participant’s use of appropriate syntactic forms to mark past and future topics, as well as a general increase in syntax level and the use of Brown’s morphemes.</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Treatment: Direct instruction, modeling and feedback within a communicative context; for past topics, clinician modeled using appropriate grammatical markers (e.g., what we did yesterday), direct/indirect requests to elicit comments from child, and visual sequence cards. Similar procedures used to increase frequency of future topics.</td>
<td>Effect sizes</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Past topics d = .6 (moderate)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Future topics d = 1.5 (large)</td>
</tr>
</tbody>
</table>
targets after intervention consisting of adult models of targets and intervention that included models plus training stimuli structured to evoke child productions of target forms. In the modeling procedure, the clinician provided models of the target form without requiring a response from the child. In the modeling plus evoked production procedure, the clinician modeled the use of the target form but then provided opportunities for the child to respond and/or receive feedback about the accuracy of his or her utterances. Both intervention procedures were provided within the context of play activities. Results indicated that for three children with expressive language problems, both modeling and modeling plus evoked production treatment strategies worked equally well (PND between .64 and 1.00, for a moderate to large effect size). For the child with a delay in both expressive and receptive language, neither treatment was effective (PND of .20).

Connell and Stone (1992) assessed students’ production and comprehension of novel morphemes after treatment that presented adult models and treatment that included models plus a direction for the participant to imitate the target morpheme. In the modeling condition, learners heard a morpheme being used in a meaningful way but were not asked to repeat or use the morpheme in any way. The learner did receive an instruction to give a command to a cartoon computer character after hearing the morpheme but was not specifically instructed to use the target form. The imitation condition was the same as for the modeling trials, with the exception that the learner received instructions from the cartoon character to specifically imitate the target morpheme. Results indicated that typically developing children produced the new morphemes with relatively equal ease following either the modeling or imitation instructions. The children with LLDs showed very little use of the morphemes following the modeling instruction alone, and significantly more use following the imitation instruction. For students with language impairments, we calculated effect sizes of .78 for morpheme production probes and 1.3 for morpheme comprehension probes following imitation instructions. Both were considered to be large effects (in this case, the effect size is the amount of learning divided by the standard deviation because the study used invented morphemes that were only heard during the treatment).

Bishop et al. (2006) evaluated computerized language instruction that was designed to improve sentence (grammar) and story comprehension. Thirty-six students were assigned to one of three conditions: a computer intervention with slowed speech, a computer intervention with modified speech that was designed to be similar to the auditory stimuli used in Fast ForWord—Language (FFW–L; Scientific Learning Corporation, 1998), or an untrained condition. The treatments were provided in 15-min sessions across a 20-day period. The students were recruited from language classrooms in mainstream schools in England. They received the treatment in the school setting, so it is reasonable to assume that the students in the treatment and control conditions continued to receive school language instruction during the treatment phase of the study. Results indicated that there were no group differences in the posttest scores on any of the language comprehension measures. Change from pretest to posttest was not assessed statistically. However, the means and standard deviations for the pretest and posttest scores were available, so we calculated pre–post effect sizes for the receptive grammar, story comprehension, and mean length of utterance (MLU) measures. The children in all three groups had moderate improvements on the sentence comprehension measure (effect sizes between .37 and .77) and got worse on the story comprehension and MLU measures (effect sizes between -.03 and -.84). Thus, it would appear that 20 days of a 15-min computer intervention with either slow speech or modified speech stimuli had no additional effect on the language services that the children were already receiving in school. The authors did not offer an explanation for the poorer scores on narrative comprehension measures and MLU at posttest.

It is not possible to draw compelling conclusions from 3 Level 2 studies. Clinicians who wish to improve children’s expressive syntax can have some confidence that they will obtain moderately large to large effects from procedures that employ imitation, modeling, or modeling plus evoked production strategies. The results of 1 study (Bishop et al., 2006) indicated that computerized intervention treatments that use slowed speech or modified speech input did not yield outcomes that were appreciably greater than regular school services. Unfortunately, clinicians who work on syntax and grammatical morphology with school-age children have very little evidence available to draw on when selecting an intervention approach.

Semantics and Vocabulary

Our study yielded 6 studies of interventions designed to treat aspects of semantics, vocabulary, concepts, and word finding in school-age children. All 6 studies were Level 2 nonrandomized comparison studies. A total of 65 students classified as having SLI participated; 28 of these students were between the ages of 8 and 14; the remainder were under 7 years of age. The grade levels of the participants were not specified in most studies.

To improve students’ word retrieval skills, McGregor and Leonard (1989) taught participants two investigator-designed word retrieval strategies. Elaboration strategies were designed to provide the child with a richer knowledge of target words. These included both phonemic strategies (SLP presented pictures for words that rhymed with the target words) and semantic strategies (SLP presented pictures of words that were similar to and different than target words). Retrieval strategies were designed to teach the child to use information that was already known about target words (SLP presented a series of retrieval cues related to target words, such as the name of the category and the sound the word begins with). Both elaboration and retrieval strategies improved the naming and recall performance of students with SLI, but data did not indicate which of the two treatment strategies was more effective. Effect sizes could not be calculated for McGregor and Leonard.

Similarly, Wing (1990) sought to increase children’s word retrieval performance by teaching investigator-designed semantic or phonological strategies. For the semantic strategy, the investigator had students engage in activities such as putting pictures/words into categories, defining words, and explaining the function of objects. The phonological treatment consisted of activities such as counting the number of sounds and syllables in words, supplying rhyming words, and silent verbal rehearsal. Scores of participants on the Test of Word Finding (TWF; German, 1986) increased significantly after the phonologically based treatment strategy but not after the semantic-based treatment. However, based on data presented in this article, we calculated moderately large pre–post effect sizes for both the semantic (d = .6) and phonological (d = .7) treatment groups, suggesting that both intervention strategies had a positive effect.

Masterson and Perrey (1999) taught students mediated learning and bridging strategies that are thought to aid in analogical reasoning (verbal analogy performance). In the mediated learning
condition, the SLP provided students with a rationale for the lesson’s objectives and activities and guided them to comprehend the meaning underlying the lesson. In this condition, students were taught several components of analogical thinking. For example, they were taught mapping, which is using the relationship between one pair of concepts to identify a similar relationship for a second pair. In the bridging condition, students were given exercises to help them incorporate previously taught components of analogical reasoning. For example, the SLP helped students use mapping to solve real-life problems of increasing a recipe to feed more people and then applying the strategy to an analogous situation of increasing a budget to cover expenses for more people. The results indicated that direct instruction on mediated learning and bridging strategies had large positive effects on students’ performance on a verbal analogies test over a control group.

Crowe (2003) compared investigator-designed traditional reading decoding strategies to meaning-based communicative reading strategies (CRS) on student’s receptive and expressive vocabulary performance. In the traditional decoding treatment, students read passages, sounded out words, answered comprehension questions, and retold what was read while the SLP provided cues and feedback. In the CRS treatment, small-group reading time was turned into an interactive conversational activity in which students asked each other questions, commented, summarized, and reacted to story events with the interventionist providing cues and feedback on linguistic features of the read material (e.g., pronoun reference, cohesive ties). Results indicated that students with LLD in the CRS treatment group improved their expressive scores on the Comprehensive Receptive and Expressive Vocabulary Test (CREVT; Wallace & Hammill, 1994) more than those in the control and traditional decoding feedback groups (for a moderate treatment effect size of .5). However, posttest performance of all groups on receptive vocabulary measures was worse than pretest performance, which the author hypothesized was due to differences in the difficulty of vocabulary items across alternate forms of the standardized vocabulary test.

In one of the few direct comparisons of service delivery models for language intervention, Throneburg, Calvert, Sturm, Paramboukas, and Paul (2000) compared the effects of vocabulary intervention provided through collaborative, classroom-based, and traditional pull-out service delivery models. In the collaborative model, the SLP and classroom teacher collaboratively planned and implemented large-group vocabulary instruction and hands-on activities for topic units with 60 embedded vocabulary words. The classroom-based instruction condition was similar except the SLP and classroom teacher independently planned and implemented their own classroom vocabulary lessons. In the pull-out condition, the SLP provided vocabulary instruction to students with LLDs in one 50-min session each week, using the same words and materials that were used in the other two conditions. Results showed similar improvements (large effect sizes between 1.2 and 3.5) on the language measures for children who received vocabulary intervention through collaborative, classroom-based, and traditional pull-out service delivery models, even though treatment effects were largest for collaborative and classroom-based models. It is noted that intervention effect sizes for Throneberg et al. must be interpreted cautiously because no nontreatment control group was used. Without this control, confounding factors that could influence word learning scores cannot be ruled out. It is well known that children learn words in and out of school and that their word learning ability increases with intellectual maturation.

Weismer and Hesketh (1993) studied the effects of modifying prosody during intervention (rate of stimulus presentation and use of stress) and the use of supplemental visual cues on the ability of students with LLDs to learn the meanings of a set of nine investigator-designed novel words. In the rate condition, stimulus sentences with target words were presented at three rates (slow, normal, fast). In the stress condition, target words were presented in sentence contexts with and without emphatic stress. In the visual condition, the invented lexical items were presented using verbal input only or with accompanying gestures associated with each target word. As with Connell and Stone (1992), the effect size is the amount of learning divided by the standard deviation because the study used invented words that were only heard during the treatment. We calculated large effect sizes (d = 1.1) for students with LLDs for both production and comprehension probes when stimuli sentences were presented at slow rates. There was a moderate effect (d = .74) for lexical production when target words were presented with emphatic stress, but no effect for lexical comprehension. There was a moderate effect for comprehension of novel words when target items were presented with accompanying iconic gestures, but only a small effect for production of novel words with gestures.

The results of these studies give some direction to clinicians searching for evidence about methods for improving semantic processing and vocabulary. Clinicians can have some confidence that collaborating with teachers on large-group instruction (Throneburg et al., 2000) and slowed presentation rate can impact vocabulary development. In addition, interactive conversational reading strategies may be somewhat helpful for improving receptive and expressive vocabulary (Crowe, 2003). Direct instruction about the processes involved in analogical thinking has yielded a large effect in comparison to no treatment (Masterson & Perrey, 1999). For assisting children with word finding problems, there does not appear to be clear differences in outcomes between semantic elaboration, semantic retrieval, or phonologically based treatments. Without evidence from multiple, well-conducted clinical trials, clinicians can only have some confidence in the evidence base for procedures that are designed to improve semantic processing and vocabulary development with school-age children.

**Phonological Awareness and Metalinguistics**

Our search yielded 5 studies of the effects of phonological awareness instruction with school-age children with language impairments. All 5 studies were Level 2 nonrandomized comparisons of experimental and control groups or cohort studies. The experimental groups received training in various combinations of rhyming, phoneme identification, phonological segmentation, phoneme blending, and sound–symbol correspondence. The treatment was provided within classrooms in 1 study (van Kleeck, Gillam, & McFadden, 1998). The other 4 studies provided treatment in pull-out sessions. van Kleeck et al. (1998) provided instruction on rhyming and phoneme awareness to 24 children in prekindergarten and kindergarten classrooms in a private school. They worked with the children in small groups twice each week for two semesters. Instruction focused on rhyme recognition, rhyme identification, rhyme judgment, and rhyme generation during the fall semester. During the spring semester, instruction focused on teaching children to match initial sounds, identify initial sounds, generate words that begin with target sounds, blend sounds together, and analyze the sounds that make up words. van Kleeck et al. compared children in

---

Cirris & Gillam: Language Intervention: Systematic Review  S127

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.
preschool and prekindergarten classes with a cohort of children who had attended those classes the previous year. The children who received treatment performed more than 1.5 SD better on their phoneme awareness measures (a large effect) than the older children who had attended the same classes a year earlier. Even the children in the prekindergarten class outperformed the control children, who were more than 2 years older ($d = 1.78$, which is also a large effect).

Gillon (2000, 2002) studied phonological awareness outcomes in three intervention conditions. The phonological awareness intervention targeted phonological identification, phoneme manipulation, phoneme segmentation, grapheme–phoneme correspondence, and phoneme production. The traditional intervention targeted phoneme production in isolation, syllables, words, and phrases. The minimal intervention focused on phoneme production through monthly consultation between SLPs, teachers, and parents. Gillon compared the outcomes of children ages $5.6$ to $7.6$ with and without language impairment immediately after treatment (2000) and 1 year after treatment (2002). Her results suggest that children with speech and language impairments normalized on the phonological awareness and word recognition measures. In these investigations, children with speech and language impairments who received the phonological awareness training performed significantly better on formal measures of phonological awareness and reading than did the children who received traditional speech therapy and the children who received minimal intervention (consultation only). The effect sizes for the various measures ranged between moderately high ($d = .67$) and high ($d = 1.77$) immediately after treatment and 1 year later.

Blisckach, Shah, Lombardino, and Chiarella (2004) conducted a multiple-baseline study with three children between the ages of $5.6$ and $7.6$. They taught phoneme–grapheme correspondence, phoneme segmentation, phoneme manipulation, and pseudo word spelling. Blisckach et al. found that all participants improved on phoneme manipulation and encoding probes. Across the three children who participated in their study, 85% of the phoneme manipulation probes in the instruction phase were above baseline levels, and 98% of the encoding probes were above baseline levels.

Finally, Seger and Voerhoeven (2005) assessed phonological awareness outcomes in groups of children who played phonological awareness computer games containing normal speech, phonological awareness computer games with modified speech input, or vocabulary computer games with normal speech. The children who received the phonological awareness computer games with modified speech made significantly greater improvements than did the children who received vocabulary instruction on measures of phonological awareness immediately after treatment. The effect sizes were small ($d = .29$), however, and the group differences were no longer significant 18 weeks after intervention ended.

Clinicians can have a moderate degree of confidence in techniques designed to improve phonological awareness in school-age children. The results of 5 Level 2 studies suggest that phonological awareness interventions that include tasks designed to improve rhyming, sound identification, phoneme segmentation, phoneme manipulation, and grapheme–phoneme correspondence consistently yield moderately large to large effects. The results of 1 study (van Kleek et al., 1998) suggest that similar effects can be obtained with a classroom collaboration approach. Further, these effects can be maintained over time (Gillon, 2002).

**Language Processing**

Our search yielded 5 studies of language processing. These studies assessed intervention programs that were designed to improve the way children attend to, perceive, discriminate, and recall sounds, syllables, words, and sentences. Two studies were RCTs (Cohen et al., 2005; Gillam et al., in press), 2 studies were Level 2 nonrandomized comparisons of experimental and control groups (Merzenich et al., 1996; Tallal et al., 1996), and 1 study was a Level 2 multiple-baseline single-subject design (Gillam, Crofford, Gale, & Hoffman, 2001). All 5 studies assessed language or auditory processing outcomes after children played computer games associated with FFW–L.

Merzenich et al. (1996) studied perceptual identification of tone sequences and phoneme recognition in children who did or did not receive early versions of some of the games that later became FFW–L. The children who received training with modified speech software outperformed the control children on measures of auditory temporal processing. Tallal et al. (1996) studied speech discrimination and on-line language comprehension in children between the ages of $5.2$ and $10.0$ who received an early version of FFW–L games or computer games without modified speech stimuli. Children who played early versions of the FFW–L games made significantly larger improvement on speech discrimination, language processing, and grammatical comprehension than did children who played computer games without modified speech. Effect sizes cannot be computed for either of these studies because standard deviations were not reported in the articles.

Gillam et al. (2001) compared the general language outcomes of children who received FFW–L software and children who received Laureate Learning software (Wilson & Fox, 1997a, 1997b). The language targets for the FFW–L condition were the same as those in the Cohen et al. (2005) study. The Laureate Learning software targeted discrimination and memory of words, grammatical morphemes, and sentences. The participants were four children (two of whom were identical twins) with moderate to severe language impairments. In a multiple-baseline design, two children were randomly assigned to immediate treatment conditions, and two children received delayed treatment (as one form of control). All four children made clinically significant gains on a formal language test, indicating that the Laureate Learning software (without modified speech stimuli) was as effective as FFW–L (with modified speech stimuli). Effect sizes could only be computed for the language sample measures. Both students who received Laureate Learning games, but only one student who received FFW–L, demonstrated improvements (large effects; PND of 100%) on MLU in the follow-up phase.

Cohen et al. (2005) compared language and literacy gains in children who received regular school services only, children who received regular school services plus FFW–L, and children who received regular school services plus computer software without a modified speech component. The treatment targets for the individual speech and language services were not reported. The treatment targets for FFW–L included discrimination of tones and minimal pair words, detection of phoneme changes, phoneme and syllable matching, recalling commands, and comprehending grammatical morphemes and complex sentence structures. The computer software condition targeted listening, spelling, phonological awareness, reading, writing, vocabulary, problem solving, narration, syntax, and morphology. The Cohen et al. RCT yielded similar improvements.
on the language measures for children who received FFW–L (with modified speech) and children who received another type of computer intervention (without modified speech) for the same amount of time. The outcomes for the FFW–L group and the control group did not differ immediately after intervention or 6 months later. The effect sizes of the comparisons between the three FFW–L groups in the Cohen et al. study varied between small ($d = .27$) and trivial ($d = .09$), suggesting that computer intervention plus school therapy was no more effective than school therapy alone.

Gillam et al. (in press) compared language and auditory processing gains in children who attended specialized research summer school. Children were randomly assigned to one of four groups: FFW–L, computer-assisted language intervention (CALI) without modified speech stimuli, individual language intervention (ILI), or an “active” comparison group that received computerized academic enrichment (AE). The treatment targets for FFW–L were the same as those reported in the Cohen et al. (2005) study. The computerized language intervention condition targeted similar skills as FFW–L, but the auditory stimuli was not modified in any way. Individualized language intervention with clinicians targeted vocabulary, grammatical morphology, complex syntax, narration, and phonological awareness. Finally, the computerized AE condition targeted mathematical computation, nonverbal problem solving, science, and geography. The Gillam et al. RCT yielded similar improvements on language measures and auditory processing measures for children in all four groups. There were moderate effects ($d$ values between .56 and .79) for language gains immediately after 6 weeks of intervention and large effects ($d$ values between .93 and 1.34) for language gains at the 6-month follow-up testing. The results of this study suggest that a variety of 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.

Clinicians can have a moderately high degree of confidence in the quality of most of the language processing studies reviewed in this article. The results of 2 RCTs and a multiple-baseline single-subject design study appear to show that the effects associated with FFW–L were indistinguishable from the control conditions, which included computer-based interventions without modified speech and other specific and nonspecific language interventions (including school services). Bishop et al. (2006; see previous section on syntax and morphology) also found that computer intervention using modified speech stimuli did not improve performance on any language measures over a control condition. Collectively, these results suggest that targeting children’s language processing with proprietary computer programs is neither necessary nor sufficient to induce clinically significant changes in language processing or expressive and receptive language skills.

**Pragmatics and Discourse**

Our search yielded only 2 studies of interventions designed to treat aspects of pragmatics, conversation, discourse, and narratives in school-age children that met our criteria. Both studies employed multiple-baseline single-subject designs (Bedrosian & Willis, 1987; Beilinson & Olswang, 2003). A total of four students with SLI participated. Participants ranged in age from 5 to 6 years and were all kindergarteners. One participant was classified as specific language impaired, and three were identified as having social communication deficits. We excluded one pre- and posttest comparison study (Swanson, Fey, Mills, & Hood, 2005) that assessed the feasibility of an intervention designed to increase the quality and quantity of students’ narratives because it did not employ a research design that met the Level 1 or 2 criteria (Note: It was not these authors’ intent to evaluate the efficacy of the intervention approach).

Beilinson and Olswang (2003) assessed participants’ use of high-risk peer-group entry behaviors after children received direct instruction in the steps in a group entry sequence and the use of desirable props/toys. Direct instruction from the SLP (in the hall outside the classroom) included the use of Meyer-Johnson symbols for steps in an entry sequence (e.g., watch your friend get a toy like your friend is doing, tell an idea to your friends). Once students returned to their classrooms, their teachers were instructed to prompt them to use the trained entry sequence during a free-play period with peers. Results indicated that the direct instruction was effective for all three students, who increased the frequency of group entry behaviors ($d$ values between 2.2 and 4.5), the use of desirable toys as props to gain group entry ($d$ values between 2.5 and 10.2), and the use of cooperative play with peers ($d$ values between 1.5 and 13.2).

Bedrosian and Willis (1987) used a combination of direct instruction, modeling, and feedback to increase the frequency of topic initiations in child/clinician conversation related to memory (past events) and topic initiations related to future events. For past topics, clinicians modeled using appropriate grammatical markers (e.g., Remember what we did yesterday?), visual sequence cards and toys, and direct and indirect requests to encourage the child to talk about feelings he was having or activities he was engaged in before the treatment session. Intervention procedures to increase the frequency of future-related topic initiations were similar; the child was encouraged to talk about activities he would be engaged in at any point in time following the session with the SLP modeling appropriate grammatical forms (e.g., Then what will we do?). For this single subject, the intervention had a moderate effect ($d = .6$) on increasing the number of past topic initiations and a large effect ($d = 1.5$) on the number of future topic initiations.

Unfortunately, like the situation with syntax and grammatical morphology, clinicians can only draw on the results of 2 Level 2 studies for evidence pertaining to interventions designed to improve pragmatics and discourse in school-age children. Direct instruction on topic initiation (Bedrosian & Willis, 1987) and group entry behaviors (Beilinson & Olswang, 2003) can yield moderately large to large effects with students who present social communication deficits. More research that compares the effects of a variety of treatments is needed before clinicians will have an evidence base on pragmatic and discourse treatment in which they can have confidence.

**DISCUSSION**

**Implications for Clinical Practice**

This article presents a systematic review of research articles that evaluated language intervention practices for school-age children.
with spoken language disabilities. SLPs who provide intervention to students with language disorders in schools may find the results somewhat disconcerting. The fact that our systematic search for experimental evidence revealed only 21 studies that met our criteria means that clinicians who work with these children have relatively little research evidence to guide evidence-based decisions about treatment options. Further, as our university training programs train graduate students to critically analyze efficacy and effectiveness research and to incorporate EBP methods into their own clinical practices, the reality of this difficult clinical situation should not be minimized. Newly graduated SLPs can be trained on EBP methods, but they need to be aware that there are relatively few valid and reliable treatment studies to guide their search for effective clinical methods.

Clinicians committed to EBP must consider the quantity and level of research evidence supporting a candidate intervention relative to the individual needs of each client and the practical constraints of practice in school settings. EBP does not require clinicians to use research evidence as the only basis for clinical decisions. When making decisions about the value and desirability of an intervention, clinicians should consider the research evidence in light of factors related to student, parent, and clinician beliefs and opinions as well as the resources and policies of the school district that the clinician works for. Gillam and Gillam (2006) present a decision-making process for clinicians to systematically integrate research evidence with their clinical knowledge and experience, parent–client preferences, and school district guidelines (see Table 8).

A detailed elaboration of each of the steps displayed in Table 8 is beyond the scope of this article, but we wish to make two crucial points here. First, systematic reviews of the experimental literature on language intervention, like the present article, will help SLPs in this EBP decision process as they attempt to find, synthesize, and evaluate the external evidence (e.g., steps 2 and 3). Second, given the paucity of evidence on beneficial language intervention practices for school-age children with language disorders, SLPs need to continue to do what they have always done; that is, weigh internal and client factors more heavily than external research results in their clinical decision-making process (e.g., steps 4, 5 and 7). Thus, as described by Gillam and Gillam (2006), SLPs should evaluate the evidence related to student–parent factors, giving particular attention to maximizing participation in ecologically relevant intervention activities (e.g., step 4). Clinicians should select interventions that are associated with the kinds of activities that children find motivating and relevant to the types of classroom activities and curricula that children typically experience in school settings. For example, Gillam and Ukrainetz (2006) and Swanson et al. (2005) have used a therapy approach known as literature-based language intervention in which language lessons are tied to the content of children’s books.

Further, SLPs should evaluate the internal evidence related to clinician-agency factors, specifically, the data that individual clinicians collect on the children they treat (e.g., step 5). A clinician may collect data on a number of children who received the same kind of treatment over the course of several semesters, and these data would suggest the range of outcomes that the SLP could expect for a specific language intervention method. Clinicians should compare their own outcome data to published outcomes to determine whether their intervention routinely yields results that are poorer than, similar to, or better than the results that were reported in a journal article. Finally, SLPs should evaluate the functional outcomes of the language intervention procedures they choose by documenting instances in which their students demonstrate use of individualized language targets correctly (without prompts or cues) in natural speaking situations (e.g., step 7). In addition, clinicians should obtain information on changes in the type and amount of students’ functional language abilities in daily classroom activities from parents, teachers, and other professionals who work with the children on their caseloads (Gillam & Gillam, 2006). In sum, until the research base expands and confirms the efficacy and effectiveness of specific intervention practices for older students with language problems, clinicians working in school settings will need to select intervention approaches carefully, monitor students’ progress on a regular and frequent basis, and validate the effectiveness of specific interventions for each student to whom they are applied.

Clinicians can also use scientific thinking, theory, and reason to back their instructional procedures when strong external evidence is not available. Writing for the field of education, Stanovich and Stanovich (2003) recommend that teachers go through the following thought process when considering new or untested instructional methods:

Even though there is no direct evidence for this [teaching] method, how is the theory behind it (e.g., the causal model of the effects it has) connected to the research consensus in the literature surrounding this [curriculum] area? Even in the absence of direct empirical evidence on a particular method, there could be a theoretical link to the consensus in the existing literature that would support the method. A lack of direct empirical evidence should be grounds for suspicion, but should not immediately rule out a new teaching method. (p. 19)

An example of teaching procedures that align well with the theoretical literature but have not yet been tested empirically with Level 1 or Level 2 experimental designs are what have been called narrative-teaching strategies (Hoggan & Strong, 1994). These include episode and story maps to support story comprehension, and involve teaching the story elements and the way in which the elements fit together. Mapping focuses on both the overall structure and theme of the story as well as the story’s individual episodes (e.g., setting, problem/goal, major episodes, theme, and resolution). The theory of text structure (Kintsch & van Dijk, 1978) is one of the theoretical bases for this type of narrative intervention and includes descriptions of text microstructure (semantic representations of propositions), macrostructure (overall structure or theme of the text, specific to different genres), and cohesion (devices used to relate microstructures to each other and to the macrostructure).
Another theoretical base for narrative teaching are schema theories used to explain the organization of narratives in memory and how individuals use schema for both text comprehension and production (Rumelhart, 1975; Stein & Glenn, 1979). Narrative schemas are frequently called story grammars (Westby, 1984) and include a setting and a goal that is met in an episode system. Setting introduces main characters and describes the social, physical, or temporal context of the story. Episodes have an initiating event, a reaction of the characters to the initiating event, an action or attempt by characters to deal with the initiating event, an outcome or consequence of the attempt, and an ending. In addition to theoretical support for narrative teaching strategies, there is accumulating evidence that children's narrative abilities continue to develop during adolescence (Nippold, 1998). Finally, current best practices in school speech-language pathology include an SLP role in intervention for students with LLDs to support literacy development (ASHA, 2001). Specifically, helping students gain explicit knowledge of text structures and linguistic cohesion devices may help them to improve their story and reading comprehension and oral and written discourse structures.

Gaps in the Literature

This systematic review revealed relatively few studies concerning the efficacy or effectiveness of language intervention practices with school-age children since 1985. However, a wide variety of language intervention practices are being used in public school settings, with a large number of students with LLDs. As a result, clinicians have relatively little research evidence on which to base their practices. SLPs should carefully distinguish between situations in which there is insufficient high-quality evidence to support strong conclusions (i.e., high-quality studies such asRCTs have not been done, thus, we do not yet know whether such interventions work) and those in which there is high-quality evidence that interventions are not effective.

In conducting the literature search for this article, it was apparent that the majority of studies on language intervention practices for children with language disorders have been carried out with preschool children under the age of 5 years. This includes studies on intervention for targets in the general areas of syntax morphology (e.g., Cole & Dale, 1986; Fey, Cleave, & Long, 1997; Fey, Cleave, Long, & Hughes, 1993; Fey & Loeb, 2002; Kaiser & Hester, 1994; Schwartz, Chapman, Terrell, Prelock, & Rowen, 1985; Tyler, Lewis, Haskill, & Tolbert, 2002, 2003), semantics/vocabulary/concepts (e.g., Girolametto, Pearce, & Weitzman, 1996; Robertson & Weismer, 1997; Wilcox, Kouri, & Caswell, 1991), phonological awareness/metalinguistics (e.g., Justice, Chow, Capellini, Flanigan, & Colton, 2003; Justice & Ezell, 2000), and pragmatic/disourse/conversation/narratives (e.g., Bradshaw, Hoffman, & Norris, 1998; Hayward & Schneider, 2000). In addition, a meta-analysis of the effectiveness of speech and language intervention for children has been carried out on studies with preschool children (Law, Garrett, & Nye, 2003, 2004). Eleven of the 21 studies we found limited participants to children in kindergarten and first grade. No studies were located that examined the efficacy of language intervention with students with language disorders in middle grades or in high school. This is a major gap in the language intervention evidence base and is especially problematic for SLPs in school settings. It is possible that older children might respond differently to various language intervention methods than young children would. The efficacy and effectiveness of many language intervention practices that are currently used with school-age children (and apparently adapted from preschool language intervention studies) has not been tested directly. There is a strong need for language intervention studies to be carried out directly with school-age children with language disorders.

SLPs who work in schools need information from intervention studies that examine the outcomes of a number of specific language intervention procedures. For example, a great deal of evidence exists supporting the use of contingent language facilitation procedures (e.g., modeling, imitation, recasts, focused stimulation) on the grammatical development of preschoolers (Fey, 1986; Leonard, 1998; McCauley & Fey, 2006). However, there is no Level 1 or Level 2 research that examines the use of these, or other, language facilitation strategies with school-age children. The few studies we found on syntax were limited to intervention programs designed to increase children's use of grammatical morphemes. A major gap in the evidence is the lack of research on interventions for "complex syntax" (e.g., complex sentences, elaborated noun phrases, elaborated verb phrases, interrogatives, etc.) necessary for students with language impairments to function in school settings. Further, we found no research examining the effectiveness of various language goal attack strategies with school-age children with LLDs (i.e., Fey, 1986; Tyler et al., 2003). A goal attack strategy refers to the way in which multiple goals are approached or scheduled, and may include (a) a vertical strategy in which one goal at a time is focused on until some predetermined level of accuracy is achieved, or (b) a horizontal strategy in which several goals are repeatedly targeted within each session, and (c) a cycle strategy in which several goals are targeted, each for a specified time period independent of accuracy, and the sequence is repeated (Tyler et al., 2003). Studies with preschool children have demonstrated differential effects of goal attack strategies on changes in the children's use of morphosyntactic forms (Tyler et al., 2003), and it is reasonable to expect similar effects with older students with LLDs.

Only 2 of the 21 studies we reviewed (Cohen et al., 2005; Gilliam et al., in press) examined the maintenance of treatment effects, and the Cohen et al. study found little or no effect of a computer-based language intervention program on student performance after 6 months. The lack of research on whether various language interventions produce lasting positive effects in school-age children is a major gap in the evidence. Proof that our interventions lead to lasting improvements in the functional language abilities of students with LLDs is especially critical as SLPs face increasing mandates to demonstrate their effectiveness in the current climate of accountability in our nation's schools.

Clinical outcomes of specific intervention procedures designed to improve semantics and vocabulary/concepts need to be examined systematically. Specifically, research on treatment strategies is needed for curriculum vocabulary use and understanding (in conjunction with reading), the use of storybooks as a tool for fostering vocabulary development (i.e., Justice, Meier, & Walpole, 2005), and facilitating higher level language in school-age students with language disabilities (e.g., figurative language, multiple meaning words, idioms, metaphor). In a related area, we found only 4 studies, 1 concerning syntax and 3 in the domain of semantics/vocabulary (Crowe, 2003; Throneburg et al., 2000; Weismer & Hesketh, 1993), that included receptive language outcome measures, and 1 of those found a negative effect of intervention on receptive vocabulary (Crowe, 2003). Law et al.'s (2003, 2004) meta-analyses showed
inconclusive effects for intervention on improving preschool students’ receptive language skills. Research that examines the efficacy and effectiveness of language intervention on the receptive language abilities of school-age students with language disorders is needed, especially given the large number of SLPs in schools who report serving students in this area (ASHA, 2004b).

The current research on language processing intervention profits from the presence of 2 RCTs that focus on children with language impairments (Cohen et al., 2005; Gillam et al., in press). The results of these (and other) studies did not provide evidence that computer-assisted language intervention procedures, including FFW–L, which uses modified speech stimuli, were superior to traditional school therapy services for expressive or receptive language processing targets. Studies are needed that examine the efficacy and effectiveness of the many other computer-based language intervention programs that are being used in school settings.

In the recent ASHA Schools Survey (2004b), 75% of SLPs in schools reported that they regularly serve students with pragmatic and social communication problems. Yet we were able to find only 2 single-subject Level 2 studies on pragmatics and conversational discourse with school-age children with language disorders, and no studies with children over the age of 6 years that met our criteria. SLPs who serve students with these language problems need efficacy and effectiveness studies on intervention methods designed to teach topic initiation, topic maintenance and relevance, and other conversational discourse skills, including the use of language to form and enhance peer social interactions within and outside the classroom.

Another major gap in language intervention outcome research for school-age children is in the area of narrative treatment strategies. Given the theoretical and practical importance of students’ narrative skills to literacy (ASHA, 2001), it is surprising that we found no Level 1 or 2 studies that investigated the effects of narrative-based intervention with school-age students. There are many more articles and reports available regarding how it is done, or should be done (e.g., Hoggan & Strong, 1994). SLPs who work in schools would benefit from studies on narrative outcomes for (a) prestory intervention methods (e.g., preparatory sets, summarizing, semantic word mapping, think-aloud, directed reading/thinking activities), (b) during-story intervention methods (e.g., extensions, questioning, episode/story mapping), (c) post-story intervention methods (e.g., question-answer relationships, internal states, word substitutions, discussion web, flow charting, story retelling, story grammar cueing, journal, dramatic play, story generation), and (d) interventions for narrative and expository text comprehension questions. Swanson et al. (2005) conducted a study on narration that we could not include in the systematic review because there was no control group. Swanson et al. assessed the quality and quantity of students’ narratives after treatment that included story imitations and retells, story generation tasks, and repeated retellings of stories at home. Results indicated an increase in the narrative quality index for most participants (a rating of the characters, physical setting, plot, ending, and language sophistication of a child’s narrative production), but not for the narrative quantity index (number of different words used). We hope to see group design studies that compare the outcomes for a group of children who receive the type of intervention that Swanson et al. used to a no-treatment control group.

Research is also needed on the effects of various service delivery models, including the effectiveness of different service delivery models on outcomes for different language targets (e.g., grammar, vocabulary, conversational discourse, or narrative). A service delivery model is an organized configuration of resources that is aimed at achieving a particular educational goal and includes personnel, materials, specific instructional or intervention procedures, the schedule for provision of services, settings in which intervention services will be delivered, and the direct and indirect roles that service providers assume as they deliver language intervention to students with language impairments (Cirrin & Penner, 1995). Gaps in the evidence for service delivery models include research on the effectiveness of classroom-based and collaborative language interventions (e.g., Farber & Klein, 1999), collaborative consultation (e.g., Ellis, Schlaudecker, & Regimbal, 1995), classroom versus individual treatment for language disorders (e.g., Wilcox et al., 1991), and integrated indirect services for increasing the language abilities of students with severe disabilities (e.g., Giangreco, 2000). We were surprised to find that even the basic question of treatment efficacy for students in group therapy versus individual therapy remains unanswered for school-age children with language disorders.

The lack of evidence that relates to the use of curriculum-relevant materials and general education standards in language intervention (i.e., ecologically relevant therapy), and on the effects that language therapy has on students’ progress in the general education curriculum (reading, writing, math), is especially problematic for SLPs who work in schools. IDEA requires that special education services must relate to students’ progress in the general education curriculum, yet the positive effects of language intervention on students’ classroom language performance remain an untested hypothesis.

Another major gap in the evidence is that no studies were found that examined the amount and frequency of intervention required to make significant progress on language targets for children in schools. McCauley and Fey (2006) describe dosage as both the frequency with which treatment sessions are held as well as the frequency with which targeted goals are addressed within a given treatment session. A study by Jacoby, Lee, Kummer, Levin, and Creaghead (2002) determined the average number of treatment units needed to achieve improvements in functional communication for preschool children receiving individual therapy in a hospital setting. This study also provided some preliminary data on which children showed differential gains and needed more treatment units. For example, Jacoby et al. found that younger children received the greatest benefit per units of therapy provided, and that children with lower initial functional communication abilities required more units of therapy to demonstrate improvement than did children with higher initial ability levels. Research on the basic clinical question of the effects of intervention dosage on students’ functional language abilities is desperately needed by SLPs who work in school settings with older children in grades K through 12.

Caveats

There are a number of caveats concerning the present review of language intervention practices in schools. First, there were several potential sources of bias in the selection criteria that we used. By limiting our review to studies using RCTs or other experimental designs (e.g., Level 1 and Level 2), we may have excluded sectors of the practice of speech-language intervention (Johnston, 2005). Further, it has not been demonstrated that the evidence standards developed for health care clinical questions
(e.g., the safety of new drugs or medical procedures) may be the most appropriate ones to use for nonmedical or behavioral therapy methods. RCT studies may not prove to be the design of choice for studies on effective speech-language intervention methods (Johnston, 2005). The design requirements of these studies (group comparisons, random assignment, blind assessment) make it nearly impossible for researchers to individualize instruction in ways that are similar to what routinely happens in language intervention in school settings. We excluded evidence in the form of case studies and pre-post single group studies, and we did not search for book chapters or presentations. Thus, it is possible that we missed some evidence that would be quite useful to school clinicians. However, we felt that we could have more confidence in the reports of studies that directly assessed causal relationships and that had made it through the peer review process.

Second, we did not include studies of language intervention (e.g., vocabulary, linguistic concepts) with "at risk" students (without identified language disabilities) in general education settings. Thus, it is likely that a number of promising language intervention practices with potential benefit for school-age students with LLDs were not reviewed in this manuscript (e.g., Justice et al., 2003; Justice et al., 2005).

Third, we opted to organize the results of our review in terms of the language outcomes used in the studies rather than intervention type because of the heterogeneity of treatments used, as recommended by Law et al. (2004). There was an insufficient number of studies that used any one treatment to organize our results by intervention method, which would have been beneficial for clinicians seeking external evidence about specific language intervention methods. In this sense, the value of our review may be compromised, as are results of meta-analyses that fail to distinguish adequately among different types of therapies, and that must group therapies of different types and durations in order to apply statistical measures of effect size (Johnston, 2005).

In conclusion, this systematic review of EBP of language intervention for school-age students with language disorders is a narrow sample of what needs to be done given the extensive scope of practice of communication disorders in schools. This review has identified a number of critical gaps in the evidence base for language intervention practices. In general, the quantity and quality of research for informing EBP optimally in schools must be enhanced. Specifically, resources are needed for studies on effective language intervention practices for school-age children (i.e., in Grades K–12). Finally, we hope that funding agencies such as the Institute for Educational Science and the National Institute on Deafness and Other Communication Disorders will give priority to optimizing the research base supporting EBP for speech-language pathology conducted in schools.

REFERENCES


ACKNOWLEDGMENTS

The authors gratefully acknowledge the contributions of our colleagues, without whose support this project would not have been possible: Kathleen Whitmire, Director of School Services, American Speech-Language-Hearing Association; and Rob Mullen, Gretchen Gould, Andrea Castrogiovanni, Tobi Frymark, and Tracy Schooling, National Center for Evidence-Based Practice in Communication Disorders, American Speech-Language-Hearing Association.

The development of this manuscript was supported in part by the Center for Personnel Studies in Special Education (COPSSSE) funded by the Office of Special Education Programs (OSEP, U.S. Department of Education) and located at the College of Education, University of Florida, Gainesville, FL. Additional information on COPSSSE and work by the Center can be found at www.copssse.org.

Cirrin & Gillam: Language Intervention: Systematic Review 5133


International Journal of Language and Communication Disorders, 37, 381–400.


Received February 8, 2007
Revision received May 30, 2007
Accepted August 13, 2007
DOI: 10.1044/0161-1461(2008/012)

Contact author: Frank M. Cirrin, Minneapolis Public Schools/Special Education, 425 5th St. NE, Minneapolis, MN 55413. E-mail: fcirrin@mpls.k12.mn.us.
APPENDIX A. DATABASES SEARCHED AND SEARCH TERMS USED

Databases
Campbell Collaboration
Cochrane Database of Systematic Reviews
CINAHL
Education Abstracts
ERIC
Exceptional Child Education Resources
Health Source: Nursing
Linguistics and Language Behaviour Abstracts
Medline (CSA)
Medline (EBSCO)
Medline (PubMed)
Pre-CINAHL
PsycARTICLES
PsycINFO
Science Citation Index
ScienceDirect
Social Science Citation Index
What Works Clearinghouse

Search Terms
* Language AND child*
* Language AND school-based and child*
* Language AND treatment AND child*
* Language AND (Specific language impairment OR language disorders in children OR aphasic children OR specific language impairment in children OR specific language disability in children)
* Language AND (syntax OR (syntax AND semantics)) OR (grammatical morphemes OR morphemes OR agrammatism OR grammatical speech disorders OR speech disorder, grammatical) OR vocabulary/semantics OR pragmatics OR (narration AND expository text) OR phonological awareness OR comprehension
* (“Child language” OR “language research” OR “phonological awareness” OR “speech pathology”) AND (“control groups” OR “case studies”)
* (“Speech” OR “speech pathology” OR “language research”) AND “school”
* (“Language disorders in children” OR “phonological awareness”)
* (Speech-language pathology OR child language OR language development/disorders) AND (schools OR students)
* Language OR language development disorders OR language tests OR language therapy OR speech-language pathology OR language development OR rehabilitation of speech and language disorders OR child language

AND

* Randomized controlled trials OR cohort studies OR meta-analysis

AND

* School health services OR schools OR Students
* “Phonological Awareness” AND “Language” OR “Figurative Language” OR “Form Classes (Language)” OR “Phrases” OR “Sentences” OR “Vocabulary” OR “Language Delay” OR “Language Disorders”

AND

* “School Environment” OR “School Facilities” OR “Classrooms” OR “Educational Laboratories” OR “Learning Centers (Educational)” OR “Classroom Environment”

AND

* “Cohort Analysis” OR “Randomized Controlled Trials” OR “Single-Subject Designs”
* Language and education OR language awareness OR language attrition OR language awareness in children OR child language OR language disorders OR language disorders in children

OR

* Grammar OR comparative and general OR phonological awareness

Note. The literature search was conducted from October, 2005 to November, 2005 by Gretchen Gould and Andrea Castrogiovanni. References were managed using the bibliographic database EndNote.
## APPENDIX B. SOME KEY FEATURES TO CONSIDER FOR EVALUATING THE QUALITY OF EVIDENCE FROM INDIVIDUAL STUDIES OF TREATMENT (ASHA, 2004A)

<table>
<thead>
<tr>
<th>Level</th>
<th>Design features</th>
<th>Bias control</th>
<th>Importance</th>
<th>Precision</th>
<th>Harm potential</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>• Data-based, controlled, experimental, randomized, prospective (RCT)</td>
<td>Selection/group, measurement, analysis</td>
<td>Large effect size (ES), in and outside of therapy context</td>
<td>Narrow confidence interval (CI)</td>
<td>Minimal potential for any harm</td>
</tr>
<tr>
<td></td>
<td>• Systematic reviews</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>• Data-based, controlled, nonexperimental, nonrandomized, prospective (cohort; cross-sectional)</td>
<td>Selection/group, measurement</td>
<td>Large ES in therapy context</td>
<td>Moderately narrow CI</td>
<td>Low potential for any harm</td>
</tr>
<tr>
<td></td>
<td>• Well-designed single-subject designs with evidence of experimental control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>• Data-based, controlled, nonexperimental, nonrandomized, retrospective (case-control)</td>
<td>Selection/group, high interrater reliability</td>
<td>Medium ES in therapy context</td>
<td>Wide CI</td>
<td>Moderate potential for mild, short-term harm is outweighed by potential benefit</td>
</tr>
<tr>
<td>4</td>
<td>• Data-based, uncontrolled, nonexperimental, nonrandomized (case series/study)</td>
<td>No blinding; prospect of direct benefit to principal investigator</td>
<td>Small ES</td>
<td>No CI reported or calculable</td>
<td>Potential for serious, short-term harm is outweighed by potential benefit</td>
</tr>
<tr>
<td>5</td>
<td>• No empirical data (opinion, belief, inductive logic)</td>
<td>No independent evaluation; conclusions may benefit those making assertions</td>
<td>No ES reported or calculable</td>
<td>No CI reported or calculable</td>
<td>Potential for serious, lasting harm is not clearly outweighed by potential benefit</td>
</tr>
</tbody>
</table>

*Note.* Only studies categorized as Level 1 or 2 were included for review in the present article.