Screening for Speech and Language Delay in Preschool Children: Systematic Evidence Review for the US Preventive Services Task Force

Heidi D. Nelson, MD, MPH (1, 2, 4)

Peggy Nygren, MA (1, 4)

Miranda Walker, BA (1, 4)

Rita Panoscha, MD (1, 3, 4)

Departments of (1)Medical Informatics and Clinical Epidemiology, (2)Medicine, and (3)Pediatrics, and (4) the Oregon Evidence-based Practice Center, Oregon Health and Science University, Portland, Oregon.

Corresponding Author: Heidi D. Nelson, MD, MPH, Oregon Health & Science University, Mail Code BICC 504, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239; telephone (503) 494-1566, FAX: (503) 418-3332, e-mail: nelsonh@ohsu.edu.

Support: This study was conducted by the Oregon Evidence-based Practice Center under contract to the Agency for Healthcare Research and Quality Contract #290-02-0024, Task Order Number 2, Rockville, MD.

Running head: Screening for Speech and Language Delay in Preschool Children

Word Count: 5,214; Abstract: 312; Number of tables: 7; Figures: 1; Appendices: 1

ABSTRACT

Background. Speech and language development is a useful indicator of a child's overall development and cognitive ability and is related to school success. Identification of children at risk for developmental delay or related problems may lead to intervention services and family assistance at a young age, when the chances for improvement are best. However, optimal methods for screening for speech and language delay have not been identified, and screening is inconsistently practiced in primary care.

Purpose. We sought to evaluate the strengths and limits of evidence about the effectiveness of screening and interventions for speech and language delay in preschool-aged children to determine the balance of benefits and adverse effects of routine screening in primary care for the development of guidelines by the U.S. Preventive Services Task Force. The target population includes all children up to 5 years old without previously known conditions associated with speech and language delay, such as hearing and neurologic impairments.

Methods. Studies were identified from MEDLINE, PsycINFO, and CINAHL databases (1966 to November 19, 2004), systematic reviews, reference lists, and experts. The evidence review included only English-language, published articles that are available through libraries. Only randomized, controlled trials were considered for examining the effectiveness of interventions. Outcome measures were considered if they were obtained at any time or age after screening and/or intervention as long as the initial assessment occurred while the child was ≤ 5 years old. Outcomes included speech and language measures and other functional and health outcomes such as social behavior. A total of 745 full-text articles met our eligibility criteria and were reviewed. Data were extracted from each included study, summarized descriptively, and rated for

quality by using criteria specific to different study designs developed by the U.S. Preventive Services Task Force.

Results. The use of risk factors for selective screening has not been evaluated, and a list of specific risk factors to guide primary care physicians has not been developed or tested. Sixteen studies about potential risk factors for speech and language delay in children enrolled heterogeneous populations, had dissimilar inclusion and exclusion criteria, and measured different risk factors and outcomes. The most consistently reported risk factors included a family history of speech and language delay, male gender, and perinatal factors. Other risk factors reported less consistently included educational levels of the mother and father, childhood illnesses, birth order, and family size.

The performance characteristics of evaluation techniques that take ≤ 10 minutes to administer were described in 24 studies relevant to screening. Studies that were rated good-to-fair quality reported wide ranges of sensitivity and specificity when compared with reference (sensitivity 17%-100%; specificity 45%-100%). Most of the evaluations, however, were not designed for screening purposes, the instruments measured different domains, and the study populations and setting were often outside of primary care. No "gold standard" has been developed and tested for screening, reference standards varied across studies, few studies compared the performance of ≥ 2 screening techniques in 1 population, and comparisons of a single screening technique across different populations are lacking.

Fourteen good- and fair-quality randomized, controlled trials of interventions reported significantly improved speech and language outcomes compared with control groups. Improvement was demonstrated in several domains including articulation, phonology, expressive language, receptive language, lexical acquisition, and syntax among children in all age groups studies and across multiple therapeutic settings. Improvement in other functional outcomes such

as socialization skills, self-esteem, and improved play themes were demonstrated in some, but not all, of the 4 studies that measured them. In general, studies of interventions were small and heterogeneous, may be subject to plateau effects, and reported short-term outcomes based on various instruments and measures. As a result, long-term outcomes are not known, interventions could not be compared directly, and generalization is questionable.

Conclusions. Use of risk factors to guide selective screening is not supported by studies. Several aspects of screening have been inadequately studied to determine optimal methods, including which instrument to use, the age at which to screen, and which interval is most useful. Trials of interventions demonstrate improvement in some outcome measures, but conclusions and generalizability are limited. Data are not available addressing the effectiveness of screening in primary care, role of enhanced surveillance by primary care physiciansbefore referral fro diagnostic evaluation, non-speech and language and long-term benefits of interventions, and adverse effects of screening and interventions.

Keywords: speech and language delay and disorders, preschool children, screening, interventions

Speech and language development is considered by experts to be a useful indicator of a child's overall development and cognitive ability[1] and is related to school success.[2-7] Identification of children at risk for developmental delay or related problems may lead to intervention services and family assistance at a young age when chances for improvement are best.[1] This rationale supports preschool screening for speech and language delay, or primary language impairment/disorder, as a part of routine well-child care.

Several types of speech and language delay and disorders have been described,[8] although terminology varies (Table 1). Expressive language delay may exist without receptive language delay but often they occur together in children as a mixed expressive/receptive language delay. Some children also have disordered language. Language problems can involve difficulty with grammar (syntax), words or vocabulary (semantics), the rules and system for speech sound production (phonology), units of word meaning (morphology) and the use of language particularly in social contexts (pragmatics). Speech problems may include stuttering or dysfluency, articulation disorders, or unusual voice quality. Language and speech problems can exist together or by themselves.

Prevalence rates for speech and language delay have been reported across wide ranges. A recent Cochrane review summarized prevalence data on speech delay, language delay, and combined delay in preschool and school-aged children.[9] For preschool-aged children, 2 to 4.5 years old, studies that evaluated combined speech and language delay reported prevalence rates ranging from 5% to 8%,[10, 11] and studies of language delay have reported prevalence rates from 2.3% to 19 %.[9, 12-15] Untreated speech and language delay in preschool children has shown variable persistence rates (from 0% to 100%), with most studies reporting 40% to 60%.[9] In 1

study, two-thirds of preschool children who were referred for speech and language therapy and given no direct intervention proved eligible for therapy 12 months later.[16]

Preschool-aged children with speech and language delay may be at increased risk for learning disabilities once they reach school age [17]. They may have difficulty reading in grade school,[2] exhibit poor reading skills at age 7 or 8,[3-5] and have difficulty with written language,[6] in particular. This may lead to overall academic underachievement[7] and, in some cases, lower IQ scores[13] that may persist into young adulthood.[18] As adults, children with phonological difficulties may hold lower skilled jobs than their non-language-impaired siblings.[19] In addition to persisting speech- and language-related underachievement (verbal, reading, spelling), language-delayed children have also shown more behavior problems and impaired psychosocial adjustment.[20, 21]

Assessing children for speech and language delay and disorders can involve a number of approaches, although there is no uniformly accepted screening technique for use in the primary care setting. Milestones for speech and language development in young children are generally acknowledged.[22] Concerns for delay arise if there are no verbalizations by the age of 1 year, if speech is not clear, or if speech or language is different from that of other children of the same age. Parent questionnaires and parent concern are often used to detect delay.[23] Most formal instruments were designed for diagnostic purposes and have not been widely evaluated for screening. Instruments constructed to assess multiple developmental components, such as the Ages and Stages Questionnaire,[24] Clinical Adaptive Test/Clinical Linguistic and Auditory Milestone Scale,[25] and Denver Developmental Screening Test,[26] include speech and language components. Instruments designed for specific communication domains include the McArthur Communicative Development Inventory,[27] Ward Infant Language Screening Test, Assessment,

Acceleration, and Remediation (WILSTAAR),[28] Fluharty Preschool Speech and Language Screening Test,[29] Early Language Milestone Scale,[30] and several others.

A specific diagnosis is most often made by a speech and language specialist using a battery of instruments. Once a child has been diagnosed with a speech and/or language delay, interventions may be prescribed. Therapy takes place in various settings including speech and language specialty clinics, home, and schools or classrooms. Direct therapy or group therapy provided by a clinician, caretaker, or teacher can be child centered and/or include peer and family components. The duration of the intervention varies. Intervention strategies focus on 1 or more domains depending on individual needs, such as expressive language, receptive language, phonology, syntax, and lexical acquisition. Therapies can include naming objects, modeling and prompting, individual or group play, discrimination tasks, reading, and conversation.

It is not clear how consistently clinicians screen for speech and language delay in primary care practice. In 1 study, 43% of parents reported that their young child (aged 10 to 35 months) did not receive any type of developmental assessment at their well-child visit, and 30% of parents reported that their child's physician had not discussed how the child communicates.[31] Potential barriers to screening include lack of time, no clear protocols, and the competing demands of the primary care visit.

This evidence review focuses on the strengths and limits of evidence about the effectiveness of screening and interventions for speech and language delay in preschool age children. Its objective is to determine the balance of benefits and adverse effects of routine screening in primary care for the development of guidelines by the U.S. Preventive Services Task Force (USPSTF). The target population includes all children up to age 5 years without previously known conditions associated with speech and language delay, such as hearing and neurological impairments. The evidence synthesis emphasizes the patient's perspective in the

choice of tests, interventions, outcome measures, and potential adverse effects, and focuses on those that are available and easily interpreted in the context of primary care. It also considers the generalizability of efficacy studies performed in controlled or academic settings and interprets the use of the tests and interventions in community-based populations seeking primary health care.

METHODS

Analytic Framework and Key Questions

Evidence reviews for the USPSTF follow a specific methodology[32] beginning with the development of an analytic framework and key questions in collaboration with members of the USPSTF. The analytic framework represents an outline of the evidence review and includes the patient population, interventions, outcomes, and adverse effects of the screening process (Figure 1). Corresponding key questions examine a chain of evidence about the effectiveness, accuracy, and feasibility of screening children age 5 years and younger for speech and language delay in primary care settings (key questions 1 and 2), adverse effects of screening (key question 3), the role of enhanced surveillance in primary care (key question 4), effectiveness of interventions for children identified with delay (key questions 5, 6, and 7), and adverse effects of interventions (key question 8).

Studies addressing key question 1, corresponding to the overarching arrow in the analytic framework, would include all components in the continuum of the screening process, including the screening evaluation, diagnostic evaluation for children identified with delay by the screening evaluation, interventions for children diagnosed with delay, and outcome measures allowing determination of the effectiveness of the overall screening process. Enhanced

surveillance in primary care relates to the practice of closely observing children who may have clinical concern for delay but not of the degree warranting a referral ("watchful waiting"). Outcome measures in this review include speech and language specific outcomes as well as non-speech and language health and functional outcomes such as social behavior, self-esteem, family function, peer interaction, and school performance. Key questions 5 examines whether speech and language interventions lead to improved speech and language outcomes. Key question 6 examines whether speech and language interventions lead to improved non-speech and language outcomes. Key question 7 evaluates the subsequent effects of improved speech and language, such as improved school performance at a later age.

Literature Search and Selection

Relevant studies were identified from multiple searches of MEDLINE, PsycINFO, and CINAHL databases (1966 to November 19, 2004). Search terms were determined by investigators and a research librarian and are described elsewhere.[33] Articles were also obtained from recent systematic reviews,[34, 35] reference lists of pertinent studies, reviews, editorials, and websites, and by consulting experts. In addition, investigators attempted to collect instruments and accompanying manuals, however, these materials are not generally available and must be purchased, which limited the evidence review to published articles.

Investigators reviewed all abstracts identified by the searches and determined eligibility of full-text articles based on several criteria. Eligible articles had English-language abstracts, were applicable to U.S. clinical practice, and provided primary data relevant to key questions. Studies of children with previously diagnosed conditions known to cause speech and language delay (e.g., autism, mental retardation, Fragile X, hearing loss, degenerative and other

neurological disorders) were not included because the scope of this review is screening children without known diagnoses.

Studies of risk factors were included if they focused on children age 5 years or younger, reported associations between predictor variables and speech and language outcomes, and were relevant to selecting candidates for screening. Otitis media as a risk factor for speech and language delay is a complex and controversial area and was not included in this review.

Studies of techniques to assess speech and language were included if they focused on children aged 5 years and younger, could be applied to a primary care setting, used clearly defined measures, compared the screening technique to an acceptable reference standard, and reported data allowing calculation of sensitivity and specificity. Techniques that take 10 minutes or less to complete that could be administered in a primary care setting by nonspecialists are most relevant to screening and are described in this report. Instruments taking more than 10 minutes and up to 30 minutes or for which administration time was not reported are described elsewhere.[33] In general, if the instrument was administered by primary care physicians, nurses, research associates, or other nonspecialists for the study, it was assumed that it could be administered by nonspecialists in a clinic. For questionable cases, experts in the field were consulted to help determine appropriateness for primary care. Studies of broader developmental screening instruments, such as the Ages and Stages Questionnaire and Denver Developmental Screening Test, were included if they provided outcomes related to speech and language delay specifically.

Only RCTs were considered for examining the effectiveness of interventions. Outcome measures were considered if they were obtained at any time or age after screening and/or intervention as long as the initial assessment occurred while the child was aged 5 years or

younger. Outcomes included speech and language measures as well as other functional and health outcomes as previously described.

Data Extraction and Synthesis

Investigators reviewed 5,377 abstracts identified by the searches. A total of 690 full-text articles from searches and an additional 55 non-duplicate articles from reference lists and experts met eligibility criteria and were reviewed. Data were extracted from each study, entered into evidence tables, and summarized by descriptive methods. For some studies of screening instruments, sensitivity and specificity were calculated by the investigators if adequate data were presented in the paper. No statistical analyses were performed because of heterogeneity of studies. Investigators independently rated the quality of studies using criteria specific to different study designs developed by the USPSTF (Appendix).[32] The quality of the study does not necessarily indicate the quality of an instrument or intervention but may influence interpretation of the results of the study.

RESULTS

Key Question 1. Does Screening for Speech and Language Delay Result in Improved Speech and Language as well as Improved Other Non-speech and Language Outcomes?

No studies directly addressed this question.

Key Question 2. Do Screening Evaluations in the Primary Care Setting Accurately Identify Children for Diagnostic Evaluation and Interventions?

Key Question 2a. Does Identification of Risk Factors Improve Screening?

Nine studies conducted in English speaking populations,[36-44] and 7 studies from non-English speaking populations[45-51] met inclusion criteria (Table 2). The most consistently reported risk factors include a family history of speech and language delay, male gender, and perinatal risk factors; however, their role in screening is unclear. A list of specific risk factors to guide primary care physicians in selective screening has not been developed or tested.

English-language studies include case control,[37, 39-41, 43] cross sectional,[36, 38, 42] and prospective cohort[44] designs. Most studies evaluated risk for language delay with or without speech delay, and 1 restricted the evaluation to expressive language only.[44] Family history was the most consistent significantly associated risk factor in 5 of 7 studies that examined it.[37, 39, 41-43] Family history was defined as family members who were late talking or had language disorders, speech problems, or learning problems. Male gender was a significant factor in all 3 of the studies examining it.[37, 39, 42] Three[37, 41, 43] of 5 studies reported an association between lower maternal education level and language delay, while 3[41-43] of 4 studies evaluating paternal education level reported a similar relationship. Other associated risk factors that were reported less consistently included childhood illnesses,[36, 40] born late in the family birth order,[42] family size,[39] older parents[39] or younger mother[43] at birth, and low socioeconomic status or minority race.[40] One study that evaluated history of asthma found no association with speech and language delay.[39]

The 7 studies assessing risk in non-English speaking populations included case-control,[47] cross-sectional,[45] prospective-cohort,[48-51] and concurrent-comparison[46] designs. Studies evaluated several types of delay including vocabulary,[46] speech,[45] stuttering,[47] language,[48-51] and learning.[49-51] Significant associations were reported in the 2 studies evaluating family history,[45, 48] and 1 of 2 studies evaluating male gender.[51] Three of 4 non-English language studies,

including a cohort of more than 8,000 children in Finland,[51] reported significant associations with perinatal risk factors such as prematurity,[50, 51] birth difficulties,[45] low birth weight,[50, 51] and sucking habits.[45] An association with perinatal risk factors was not found in the 1 English language study that examined low birth weight.[43] Other associated risk factors reported less consistently include parental education level,[49, 50] and family factors such as size and overcrowding.[50, 51] These studies did not find associations with mother's stuttering or speaking style or rate,[47] mother's age,[51] or child temperament.[46]

Key Questions 2b & 2c. What Are Screening Techniques and How Do They Differ by Age? What Is the Accuracy of Screening Techniques and How Does It Vary by Age?

A total of 22 articles reporting performance characteristics of 24 evaluations met inclusion criteria.[33] Studies utilized several different standardized and nonstandardized instruments (Table 3), although many were not designed specifically for screening purposes. Results of instruments were compared with those of a variety of reference standards and no gold standard was acknowledged or used across studies, which limited comparisons between them.

The studies provided limited demographic details of subjects, and most included predominantly white children with similar proportions of boys and girls. One study enrolled predominantly black children[52] and another, children from rural areas.[53] Study sizes ranged from 25[54] to 2,590[11] subjects. Testing was conducted in general health clinics, specialty clinics, day care centers, schools, and homes by pediatricians, nurses, speech and language specialists, psychologists, health visitors, medical or graduate students, teachers, parents, and research assistants. Studies are summarized below by age categories according to the youngest ages included, although many studies included children in overlapping categories.

Ages 0 to 2 years. Eleven studies from 10 publications utilized instruments taking 10 minutes or less to administer for children up to 2 years old including the Early Language Milestone Scale,[30, 55] Parent Evaluation of Developmental Status,[56] Denver Developmental Screening Test II (language component),[57] Pediatric Language Acquisition Screening Tool for Early Referral,[52] Clinical Linguistic and Auditory Milestone Scale,[58] Language Development Survey,[59-61] Development Profile II,[57] and the Bayley Infant Neurodevelopmental Screener[62] (Table 4). Of these studies, 6 tested expressive and/or receptive language,[30, 52, 55, 57, 62] 3 expressive vocabulary,[59-61] 1, expressive language and articulation,[56] and 1, syntax and pragmatics.[58]

For the 10 fair- and good-quality studies that provided data to determine sensitivity and specificity, sensitivity ranged from 22% to 97% and specificity from 66% to 97%.[30, 52, 56-62] Four studies reported sensitivity and specificity of 80% or more using the Early Language Milestone Scale,[30] the Language Development Survey,[59,60] and the Clinical Linguistic and Auditory Milestone Scale.[58] The study of the Clinical Linguistic and Auditory Milestone Scale also determined sensitivity and specificity by age, and reported higher sensitivity/specificity at age 14 to 24 months (83%/93%) than 25 to 36 months (68%/89%) for receptive function, but lower sensitivity/specificity at age 14 to 24 months (50%/91%) than 25 to 36 months (88%/98%) for expressive function.[58] A study testing expressive vocabulary using the Language Development Survey indicated higher sensitivity/specificity at age 2 years (83%/97%) than at age 3 years (67%/93%).[60]

Ages 2 to 3 years. Ten studies in 9 publications used instruments taking 10 minutes or less to administer for children aged 2 to 3 years including the Parent Language Checklist,[11] Structured Screening Test,[63] Levett-Muir Language Screening Test,[64] Fluharty Preschool Speech and Language Screening Test,[53, 65] Screening Kit of Language Development,[66] Hackney Early

Language Screening Test,[54, 67] and Early Language Milestone Scale[68] (Table 5). All studies tested expressive and/or receptive language.[11, 53, 54, 63-68] In addition, 3 studies tested articulation[53, 65] and 1 tested syntax and phonology.[64]

For the 8 fair and good-quality studies providing data to determine sensitivity and specificity, sensitivity ranged from 17% to 100% and specificity from 45% to 100%. Two studies reported sensitivity and specificity of 80% or better using the Levett-Muir Language Screening Test[64] and the Screening Kit of Language Development.[66] The study of the Screening Kit of Language Development reported comparable sensitivity/specificity at ages 30 to 36 months (100%/98%), 37 to 42 months (100%/91%), and 43 to 48 months (100%/93%).[66]

Ages 3 to 5 years. Three studies used instruments taking 10 minutes or less to administer including the Fluharty Preschool Speech and Language Screening Test,[69] Test for Examining Expressive Morphology,[70] and the Sentence Repetition Screening Test[71] (Table 6). Of these, 2 studies tested expressive and receptive language and articulation,[69, 71] and 1 tested expressive vocabulary and syntax.[70] The 2 fair-quality studies reported sensitivity ranging from 57% to 62% and specificity from 80% to 95%.[66, 69, 71]

Systematic review. A Cochrane systematic review of 45 studies, including most of the studies cited above, summarized the sensitivity and specificity of instruments taking 30 minutes or less to administer.[34] Sensitivity of instruments for normally developing children ranged from 17% to 100%, and for children from clinical settings it ranged from 30% to 100%. Specificity ranged from 43% to 100%, and 14% to 100% respectively. Studies considered to be of higher quality tended to have higher specificity than sensitivity (t=4.41, P<0.001), however, high false-positive and false-negative rates were reported often.[34]

Key Question 2d. What Are the Optimal Ages and Frequency for Screening?

No studies addressed this question.

Key Question 3. What Are the Adverse Effects of Screening?

No studies addressed this question. Potential adverse effects include false-positive and false-negative results. False-positive results can erroneously label children with normal speech and language as impaired, potentially leading to anxiety for children and families and further testing and interventions. False-negative results would miss identifying children with impairment, potentially leading to progressive speech and language delay and other long-term effects including communication, social, and academic problems. In addition, once delay is identified, children may be unable to access services because of unavailability or lack of insurance coverage.

Key Question 4. What Is the Role of Enhanced Surveillance by Primary Care Clinicians?

No studies addressed this question.

Key Question 5. Do Interventions for Speech and Language Delay Improve Speech and Language Outcomes?

Twenty-five RCTs in 24 publications met inclusion criteria including 1 rated good,[72] 13 rated fair,[73-85] and 11 rated poor quality[77, 86-95] (Table 7). Studies were considered poor quality if they reported important differences between intervention and comparison groups at baseline, did not use intention-to-treat analysis, no method of randomization was reported, and there were fewer than 10 subjects in intervention or comparison groups. Limitations of studies, in general,

include small numbers of participants (only 4 studies enrolled more than 50 subjects), lack of consideration of potential confounders, and disparate methods of assessment, intervention, and outcome measurement. As a result, conclusions about effectiveness are limited. Although children in the studies ranged from 18 to 75 months old, most studies included children age 2 to 4 years old and results do not allow for determination of optimal ages of intervention.

Studies evaluated the effects of individual or group therapy directed by clinicians and/or parents that focused on specific speech and lanugage domains. These include expressive and receptive language, articulation, phonology, lexical acquisition, and syntax. Several studies used established approaches to therapy, such as the WILSTAAR program[96] and the HANEN principles.[78, 79, 85, 93] Others used more theoretical approaches, such as focused stimulation,[78, 79, 86, 87, 93] auditory discrimination,[83, 90] imitation or modeling procedures,[76, 92] auditory processing or work mapping,[85] and play narrative language.[80, 81] Some interventions focused on specific words and sounds, used unconventional methods, or targeted a specific deficit.

Outcomes were measured by subjective reports from parents,[77, 78, 80, 85] and by scores on standardized instruments, such as the Reynell Expressive and Receptive Scales,[74, 77] the Preschool Language Scale,[72, 75, 85] and the MacArthur Communicative Development Inventories.[80, 93] The most widely used outcome measure was mean length utterances, used by 6 studies.[73, 75, 77, 80, 85]

Studies rated good or fair quality are described below by age categories according to the youngest ages included, although many studies included children in overlapping categories

Ages 0 to 2 years. No studies examined this age group exclusively, although 1 good-quality study enrolled children 18 to 42 months old.[72] The clinician-directed, 12-month intervention consisted of 10-minute weekly sessions focusing on multiple language domains, expressive and

receptive language, and phonology. Treatment for receptive auditory comprehension led to significant improvement for the intervention group compared with control group, however, results did not differ between groups for several expressive and phonology outcomes.[72]

Ages 2 to 3 years. One good-[72] and 6 fair-quality studies[77-80, 84, 85] evaluated speech and language interventions for children who were 2 to 3 years old. Studies reported improvement on a variety of communication domains including clinician-directed treatment for expressive and receptive language,[80] parent-directed therapy for expressive delay,[77, 78] and clinician-directed receptive auditory comprehension.[72] Lexical acquisition was improved with both clinician-directed therapy[84, 91] and group therapy approaches.[84] In 3 studies, there were no between group differences for clinician-directed expressive[72, 85] or receptive language therapy,[72, 85] parent-directed expressive or receptive therapy,[85] or parent-directed phonology treatment.[79]

Ages 3 to 5 years. Five fair-quality studies reported significant improvements for children 3 to 5 years old undergoing interventions compared with controls,[73, 74, 76, 81, 82] while 2 studies reported no differences.[75, 83] Both group-based interventions[81] and clinician-directed interventions[74] were successful in improving expressive and receptive competencies.

Systematic review. A Cochrane systematic review included a meta-analysis utilizing data from 25 RCTs of interventions for speech and language delay for children up to adolesence.[35] Twenty-three of these studies also met criteria for this review and were included in Table 7,[72-92, 95] and 2 trials were unpublished. The review reported results in terms of standard mean differences (SMD) in scores for a number of domains (phonology, syntax, and vocabulary). Effectiveness was considered significant for both the phonological (SMD=0.44; 95% CI, 0.01-

0.86) and vocabulary (SMD=0.89; 95% CI, 0.21-1.56) interventions. Less effective was the receptive intervention (SMD=-0.04; 95% CI, 0.64-0.56), and results were mixed for the expressive syntax intervention (SMD=1.02; 95% CI, 0.04-2.01). In the analysis, when interventions were comparable in duration and intensity, there were no differences between interventions when administered by trained parents or clinicians for expressive delays. Use of normal-language peers as part of the intervention strategy also proved beneficial.[81]

Key Question 6. Do Interventions for Speech and Language Delay Improve Other Non-Speech and Language Outcomes?

Four good-[72] or fair-quality[80, 81, 85] intervention studies included functional outcomes other than speech and language. Increased toddler socialization skills,[80] improved child self-esteem,[85] and improved play themes[81] were reported for children in intervention groups in 3 studies. Improved parent-related functional outcomes included decreased stress[80] and increased positive feelings toward their children.[85] Functional outcomes that were studied but did not show significant treatment effects included well being, levels of play and attention, and socialization skills in 1 study.[72]

Key Question 7. Does Improvement in Speech and Language Outcomes Lead to Improved Additional Outcomes?

No studies addressed this question.

Key Question 8. What Are the Adverse Effects of Interventions?

No studies addressed this question. Potential adverse effects of treatment programs include the impact of time and cost of interventions on clinicians, parents, children, and siblings.

Loss of time for play and family activities, stigmatization, and labeling may also be potential adverse effects.

CONCLUSIONS

Studies are not available addressing the overarching key question about the effectiveness of screening (key question 1), adverse effects of screening (key question 3), the role of enhanced surveillance in primary care (key question 4), long-term effectiveness of interventions on non-speech and language outcomes for children identified with delay (key questions 7), and adverse effects of interventions (key question 8). No studies determine the optimal ages and frequency for screening (key question 2d). Relevant studies are available regarding the use of risk factors for screening (key question 2a), techniques for screening (key question 2b and 2c), and effectiveness of interventions on short-term speech and language and non-speech and language outcomes for children identified with delay (key questions 5 and 6).

The use of risk factors for selective screening has not been evaluated and a list of specific risk factors to guide primary care physicians has not been developed or tested. Sixteen studies about potential risk factors for speech and language delay in children enrolled heterogeneous populations, had dissimilar inclusion and exclusion criteria, and measured different risk factors and outcomes. The most consistently reported risk factors included a family history of speech and language delay, male gender, and perinatal factors. Other risk factors that were reported less consistently included educational levels of the mother and father, childhood illnesses, birth order, and family size.

Although brief evaluations are available and have been used in a number of settings with administration by professional and nonprofessional individuals, including parents, the optimal method of screening for speech and language delay has not been established. The performance

characteristics of evaluation techniques taking 10 minutes or less to administer were described in 24 studies relevant to screening. Studies rated good-to-fair quality reported wide ranges of sensitivity and specificity when compared with reference standards (sensitivity 31% to 100%; specificity 45% to 100%). In these studies, the instruments providing the highest sensitivity and specificity included the Early Language Milestone Scale, Clinical Linguistic and Auditory Milestone Scale, Language Development Survey, Screening Kit of Language Development, and the Levett-Muir Language Screening Test. Most of the evaluations, however, were not designed for screening purposes, the instruments measured different domains, and the study populations and settings were often outside primary care. No gold standard has been developed and tested for screening, reference standards varied across studies, few studies compared the performance of 2 or more screening techniques in 1 population, and comparisons of a single screening technique across different populations are lacking.

RCTs of multiple types of interventions reported significantly improved speech and language outcomes compared with control groups. Improvement was demonstrated in several domains including articulation, phonology, expressive language, receptive language, lexical acquisition, and syntax among children in all age groups studied and across multiple therapeutic settings. However, studies were small, heterogeneous, may be subject to plateau effects, and reported short-term outcomes based on various instruments and measures. As a result, long-term outcomes are not known, interventions could not be directly compared to determine optimal approaches, and generalizability is questionable.

There are many limitations of the literature relevant to screening for speech and language delay in preschool-aged children including lack of studies specific to screening as well as difficulties inherent in this area of research. This evidence review is limited by use of only published studies of instruments and interventions. Data about performance characteristics of

instruments, in particular, are not generally accessible and are often only available in manuals that must be purchased. Interventions vary widely and may not be generalizable. In addition, studies from countries with different health care systems, such as the U.K., may not translate well to U.S. practice.

Although speech and language development is multi-dimensional, the individual constructs that comprise it are often assessed separately. Numerous evaluation instruments and interventions that accommodate children across a wide range of developmental stages have been developed to identify and treat specific abnormalities of these functions. As a result, studies include many different instruments and interventions that are most often designed for purposes other than screening. Also, studies of interventions typically focus on 1 or a few interventions. In clinical practice, children are provided with individualized therapies consisting of multiple interventions. The effectiveness of these complex interventions may be difficult to evaluate. Adapting results of this heterogeneous literature to determine benefits and adverse effects of screening is problematic. Also, behavioral interventions are difficult to conduct in long-term randomized trials, and it is not possible to blind parents or clinicians. Randomizing children to therapy or control groups when clinical practice standards support therapy raises ethical dilemmas.

Speech and language delay is defined by measurements on diagnostic instruments in terms of a position on a normal distribution. Measures and terminology are inconsistently used and there is no recognized gold standard. This is challenging when defining cases and determining performance characteristics of screening instruments in studies.

Identification of speech and language delay may be associated with benefits and adverse effects that would not be captured by studies of clinical or health outcomes. The process of screening alerts physicians and caretakers to developmental milestones and focuses attention on

the child's development, potentially leading to increased surveillance, feelings of caregiver support, and improved child self esteem. Alternatively, caretakers and children may experience increased anxiety and stress during the screening and evaluation process. Detection of other conditions during the course of speech and language evaluation, such as hearing loss, is an unmeasured benefit if appropriate interventions can improve the child's status.

Future research should focus on determining optimal approaches of identifying preschool children with speech and language delay in primary care settings who would be appropriate candidates for further evaluations and possibly speech and language interventions. These approaches should be integrated into routine developmental surveillance practices of clinicians caring for children.[97] Studies that evaluate the effectiveness of validated brief screening instruments that include child and caretaker components could lead to a more standardized approach. Studies of specific speech and language components of currently available broad developmental screening instruments, such as Ages and Stages, would be useful. Incorporation of risk factors and parent report in studies of screening approaches could provide information about their added value. Additional studies that compare screening instruments and methods in large primary care populations could lead to defining gold standards and acceptable referral criteria. Evaluating these criteria in different populations of children would minimize cultural and language biases.

Additional work about the effectiveness of interventions, including speech and language domain-specific results, may provide new insights. School-based efforts could be designed to complement strategies developed for young children improving long-term outcomes. Results of these studies may help determine optimal ages and intervals for screening. Functional long-term outcomes such as school performance, high school graduation rates, in-grade retention, special education placement/duration, and social adjustment need to be addressed more thoroughly.

Cost-effectiveness evaluations of effective approaches that consider cost of treatment, the time that caregivers spend in transit to treatment locations, the time they spend participating in the program on site or in the home, and long-term outcomes, among other factors, would be useful.

ACKNOWLEDGEMENTS

The investigators thank Andrew Hamilton, MLS, MS for conducting the literature searches, expert reviewers for commenting on draft versions of this report, and members of the USPSTF who served as leads for this project including Alfred Berg, MD, MPH, Paul Frame, MD, Leon Gordis, MD, DrPH, Jonathan Klein, MD, MPH, Virginia Moyer, MD, MPH, and Barbara Yawn, MD, MSc.

Corresponding author: Heidi D. Nelson MD, MPH, Oregon Health & Science University, Mail code BICC 504, 3181 SW Sam Jackson Park Road, Portland, Oregon 97239; Telephone: (503) 494-1566; Fax: (503) 494-4551; Email: nelsonh@ohsu.edu

Other authors:

Peggy Nygren, MA

Oregon Health & Science University

Mail code BICC 504

3181 SW Sam Jackson Park Road

Portland, Oregon 97239

Telephone: (503) 494-8358; Fax: (503) 494-4551

Miranda Walker, BA

Oregon Health & Science University

Mail code BICC 504

3181 SW Sam Jackson Park Road

Portland, Oregon 97239

Telephone: (503) 418-3740; Fax: (503) 494-4551

Rita Panoscha, MD

Child Development and Rehabilitation Center

Oregon Health & Science University

707 SW Gaines

Portland, Oregon 97239

Telephone: (503) 494-0581; Fax: (503) 494-6868

REFERENCES

1. Schuster MA. Developmental screening. In: McGlynn EA, editor. Quality of care for children and adolescents: A review of selected clinical conditions and quality indicators. Santa Monica, CA: RAND; 2000. p. 157-168.
2. Catts HW, Fey ME, Tomblin JB, Zhang X. A longitudinal investigation of reading outcomes in children with language impairments. *J Speech Lang Hear Res 2002*;45(6):1142-1157.
3. Scarborough HS, Dobrich W. Development of children with early language delay. *J Speech Hear Res 1990*;33(1):70-83.
4. Richman N, Stevenson J, Graham PJ. Pre-school to school: A behavioural study. In: Schaffer R, ed. *Behavioural Development: A Series of Monographs* London, United Kingdom: Academic; *1982*:228.
5. Silva PA, Williams SM, McGee R. A longitudinal study of children with developmental language delay at age three: later intelligence, reading and behaviour problems. *Dev Med Child Neurol 1987*;29:630-640.
6. Bishop D, Clarkson B. Written Language as a Window into Residual Language Deficits: A Study of Children With Persistent and Residual Speech and Language Impairments. *Cortex 2003*;39:215-237.
7. Stern LM, Connell TM, Lee M, Greenwood G. The Adelaide preschool language unit: results of follow-up. *Journal of Paediatrics & Child Health. 1995*;31(3):207-212.
8. Welcome to ASHA. In: American Speech-Language-Hearing Association; 2004. Available at www.asha.org/default.htm.

9. Law J, Boyle J, Harris F, Harkness A, Nye C. Screening for primary speech and language delay: a systematic review of the literature. *Int J Lang Commun Disord 1998*;33(Suppl):21-23.

10. Randall D, Reynell J, Curwen M. A study of language development in a sample of 3 year old children. *Br J Disord Commun 1974*;9(1):3-16.

11. Burden V, Stott CM, Forge J, Goodyer I. The Cambridge Language and Speech Project (CLASP). I .Detection of language difficulties at 36 to 39 months. *Dev Med Child Neurol 1996*;38(7):613-631.

12. Rescorla L, Hadicke-Wiley M, Escarce E. Epidemiological investigation of expressive language delay at age two. *First Language 1993*;13:5-22.

13. Silva PA, McGee R, Williams SM. Developmental language delay from three to seven years and its significance for low intelligence and reading difficulties at age seven. *Dev Med Child Neurol 1983*;25(6):783-793.

14. Stevenson J, Richman N. The Prevelence of Language delay in a Population of Three-year-old Children and its association with General Retardation. *Dev Med Child Neurol 1976*;18:431-441.

15. Wong V, Lee PWH, Mak-Lieh F, Yeung CY, Leung PWL, Luk SL, et al. Language screening in preschool Chinese children. *Eur J Disord Commun 1992*;27(3):247-264.

16. Roulstone S, Peters TJ, Glogowska M, Enderby P. A 12-month follow-up of preschool children investigating the natural history of speech and language delay. *Child: Care, Health & Development. 2003*;29(4):245-255.

17. Bashir AS, Scavuzzo A. Children with language disorders: natural history and academic success. *Journal of Learning Disabilities. 1992*;25(1):53-65; discussion 66-70.

18. Young AR, Beitchman JH, Johnson C, Douglas L, Atkinson L, Escobar M, et al. Young adult academic outcomes in a longitudinal sample of early identified language impaired and control children. *J Child Psychol Psychiatry 2002*;43(5):635-645.

19. Felsenfeld S, Broen PA, McGue M. A 28-Year Follow-up of Adults With a History of Moderate Phonological Disorder: Educational and Occupational Results. *J Speech Hear Res 1994*;37:1341-1353.

20. Cohen NJ, Barwick MA, Horodezky N, Vallance DD, Im N. Language, Achievement, and Cognitive Processing in Psychiatrically Disturbed Children with Previously Identified and Unsuspected Language Impairments. *J Child Psychol Psychiatry 1998*;39(865-877).

21. Cohen NJ, Menna R, Vallance DD, Barwick MA, Im N, Horodezky N. Language, Social Cognitive Processing, and Behavioral Characteristics of Psychiatrically Disturbed Children with Previously Identified and Unsuspected Language Impairments. *J Child Psychol Psychiatry 1998*;39:853-864.

22. American Academy of Pediatrics. Guidelines for Health Supervision III. Elk Grove Village, IL; 1997.

23. Ireton H, Glascoe FP. Assessing children's development using parents' reports: The child development inventory. *Clin Pediatr 1995*;34:248-255.

24. Bricker D, Squires J. Ages & Stages Questionnaires: A parent-completed, child-monitoring system. 2nd ed: Paul H. Brookes Publishing Company; 1999.

25. Capute AJ, Palmer FB, Shapiro BK, Wachtel RC, Schmidt S, Ross A. Clinical linguistic and auditory milestone scale: prediction of cognition in infancy. *Dev Med Child Neurol 1986*;28(6):762-771.

26. Frankenburg WK, Dodds J, Archer P, Shapiro H, Bresnick B. The Denver II: a major revision and restandardization of the Denver Developmental Screening Test.[see comment]. *Pediatrics. 1992*;89(1):91-97.

27. Fenson L, Pethick SJ, Renda C, Cox JL, Dale PS, Reznick JS. Short-form versions of the MacArthur Communicative Development Inventories. *Applied Psycholinguistics 2000*;21:95-96.

28. Coulter L, Gallagher C. Piloting new ways of working: evaluation of the WILSTAAR Programme. *Int J Lang Commun Disord 2001*;36(Suppl):270-275.

29. Fluharty. The design and standardization of a speech and language screening test for use with preschool children. *J Speech Hear Disord 1973*;39:75-88.

30. Coplan J, Gleason JR, Ryan R, Burke MG, Williams ML. Validation of an early language milestone scale in a high-risk population. *Pediatrics 1982*;70(5):677-683.

31. Halfon N, et al. Summary statiscs from the National Survey of Early Childhood Health, 2000. National Center for Health Statistics. *Vital and Health Statistics 2002*;15(4).

32. Harris RP, Helfand M, Woolf SH, Lohr KN, Mulrow CD, Teutsch SM, et al. Current methods of the US Preventive Services Task Force. A reivew of the process. *Am J Prev Med 2001*;30(3S):21-35.

33. Nelson HD, Nygren P, Walker M, Panoscha R. Screening for Speech and Language Delay in Preschool Children: Systematic Evidence Synthesis. Rockville, MD: Agency for Healthcare Research and Quality; 2005.

34. Law J, Boyle J, Harris F, Harkness A, Nye C. Screening for primary speech and language delay: A systematic review of the literature. *Health Technol Assess 1998*;2(9):1-200.

35. Law J, Garrett Z, Nye C. Speech and language therapy interventions for children with primary speech and language delay or disorder. *Cochrane Database Syst Rev 2003*(3):CD004110.

36. Brookhouser PE, Hixson PK, Matkin ND. Early childhood language delay: the otolaryngologist's perspective. *Laryngoscope 1979*;89(12):1898-1913.

37. Campbell TF, Dollaghan CA, Rockette HE, Paradise JL, Feldman HM, Shriberg LD, et al. Risk factors for speech delay of unknown origin in 3-year-old children. *Child Dev 2003*;74(2):346-357.

38. Cantwell DP, Baker L. Psychiatric and learning disorders in children with speech and language disorders: A descriptive analysis. *Advances in Learning & Behavioral Disabilities 1985*;4.

39. Choudhury N, Benasich AA. A family aggregation study: The influence of family history and other risk factors on language development. *J Speech Lang Hear Res 2003*;46(2):261-272.

40. Singer LT, Siegel AC, Lewis B, Hawkins S, Yamashita T, Baley J. Preschool language outcomes of children with history of bronchopulmonary dysplasia and very low birth weight. *J Dev Behav Pediatr 2001*;22(1):19-26.

41. Tallal P, Ross R, Curtiss S. Familial aggregation in specific language impairment. *J Speech Hear Disord 1989*;54(2):167-173.

42. Tomblin JB, Hardy JC, Hein HA. Predicting poor-communication status in preschool children using risk factors present at birth. *J Speech Hear Res 1991*;34(5):1096-1105.

43. Tomblin J, Smith E, Zhang X. Epidemiology of specific language impairment: Prenatal and perinatal risk factors. *J Commun Disord 1997*;30(4):325-344.

44. Whitehurst GJ, Arnold DS, Smith M, Fischel JE, Lonigan CJ, Valdez-Menchaca MC. Family history in developmental expressive language delay. *J Speech Hear Res 1991*;34(5):1150-1157.

45. Fox A, Dodd B, Howard D. Risk factors for speech disorders in children. *Int J Lang Commun Disord 2002*;37(2):117-131.

46. Klein PS, Tzuriel D. Preschoolers type of temperament as predictor of potential difficulties in cognitive functioning. *Isr J Psychiatry Relat Sci 1986*;23(1):49-61.

47. Kloth S, Janssen P, Kraaimaat F, Brutten G. Communicative behavior of mothers of stuttering and nonstuttering high-risk children prior to the onset of stuttering. *J Fluency Disord 1995*;20(4):365-377.

48. Lyytinen H, Ahonen T, Eklund K, Guttorm TK, Laakso ML, Leinonen S, et al. Developmental pathways of children with and without familial risk for dyslexia during the first years of life. *Dev Neuropsychol 2001*;20(2):535-554.

49. Peters SA, Grievink EH, van Bon WH, van den Bercken JH, Schilder AG. The contribution of risk factors to the effect of early otitis media with effusion on later language, reading and spelling. *Dev Med Child Neurol 1997*;39(1):31-39.

50. Weindrich D, Jennen-Steinmetz C, Laucht M, Esser G, Schmidt MH. Epidemiology and prognosis of specific disorders of language and scholastic skills. *Eur Child Adolesc Psychiatry 2000*;9(3):186-194.

51. Yliherva A, Olsen P, Maki-Torkko E, Koiranen M, Jarvelin MR. Linguistic and motor abilities of low-birthweight children as assessed by parents and teachers at 8 years of age.[comment]. *Acta Paediatr 2001*;90(12):1440-1449.

52. Sherman T, Shulman BB, Trimm RF, Hoff C. Plaster: predicting communication impairments in a NICU follow-up population... Pediatric Language Acquistion Screening

Tool for Early Referral. *Infant-Toddler Intervention: the Transdisciplinary Journal 1996*;6(3):183-195.

53. Sturner RA, Heller JH, Funk SG, Layton TL. The Fluharty Preschool Speech and Language Screening Test: a population-based validation study using sample-independent decision rules. *J Speech Hear Res 1993*;36(4):738-745.

54. Dixon J, Kot A, Law J. Early language screening in City and hackney: work in progress. *Child: Care, Health & Development 1988*;14:213-229.

55. Black MM, Gerson LF, Freeland CA, Nair P, Rubin JS, Hutcheson JJ. Language screening for infants prone to otitis media. *J Pediatr Psychol 1988*;13(3):423-433.

56. Glascoe FP. Can clinical judgment detect children with speech-language problems? *Pediatrics. 1991*;87(3):317-322.

57. Glascoe FP, Byrne KE. The Accuracy of Three Developmental Screening Tests. *Journal of Early Intervention 1993*;17(4):368-379.

58. Clark JG, Jorgensen SK, Blondeau R. Investigating the validity of the clinical linguistic auditory milestone scale. *Int J Pediatr Otorhinolaryngol 1995*;31:63-75.

59. Klee T, Carson DK, Gavin WJ, Hall L, Kent A, Reece S. Concurrent and predictive validity of an early language screening program. *J Speech Lang Hear Res 1998*;41(3):627-641.

60. Klee T, Pearce K, Carson DK. Improving the positive predictive value of screening for developmental language disorder. *J Speech Lang Hear Res 2000*;43(4):821-833.

61. Rescorla L, Alley A. Validation of the language development survey (LDS): a parent report tool for identifying language delay in toddlers. *J Speech Lang Hear Res 2001*;44(2):434-445.

62. Macias MM, Saylor CF, Greer MK, Charles JM, Bell N, Katikaneni LD. Infant screening: the usefulness of the Bayley Infant Neurodevelopmental Screener and the Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale. *J Dev Behav Pediatr 1998*;19(3):155-161.

63. Laing GJ, Law J, Levin A, Logan S. Evaluation of a structured test and a parent led method for screening for speech and language problems: prospective population based study. *Bmj 2002*;325(7373):1152-1154.

64. Levett L, Muir J. Which three year olds need speech therapy? Uses of the Levett-Muir language screening test. *Health Visitor 1983*;56(12):454-456.

65. Blaxley L, Clinker M, Warr-Leeper GA. Two Language Screening Tests Compared with Developmental Sentence Scoring. *Language, Speech, and Hearing Services in the Schools 1983*;14:38-46.

66. Bliss LS, Allen DV. Screening Kit of Language Development: a preschool language screening instrument. *J Commun Disord 1984*;17(2):133-141.

67. Law J. Early language screening in City and Hackney: The concurrent validity of a measure designed for use with 2!12-year-olds. *Child: Care, Health & Development 1994*;20(5):295-308.

68. Walker D, Gugenheim S, Downs MP, Northern JL. Early Language Milestone Scale and language screening of young children.[comment]. *Pediatrics 1989*;83(2):284-288.

69. Allen DV, Bliss LS. Concurrent validity of two language screening tests. *J Commun Disord 1987*;20(4):305-317.

70. Merrell AW, Plante E. Norm-referenced test interpretation in the diagnostic process. *Language, Speech & Hearing Services in the Schools 1997*;28(1):50-58.

71. Sturner RA, Funk SG, Green JA. Preschool speech and language screening: further validation of the sentence repetition screening test. *J Dev Behav Pediatr 1996*;17(6):405-413.

72. Glogowska M, Roulstone S, Enderby P, Peters TJ. Randomised controlled trial of community based speech and language therapy in preschool children.[see comment]. *BMJ 2000*;321(7266):923-926.

73. Almost D, Rosenbaum P. Effectiveness of speech intervention for phonological disorders: a randomized controlled trial.[erratum appears in Dev Med Child Neurol 1998 Oct;40(10):719]. *Dev Med Child Neurol 1998*;40(5):319-325.

74. Barratt J, Littlejohns P, Thompson J. Trial of intensive compared to weekly speech therapy in preschool children. *Archives of Disease in Childhood 1992*;671:106-108.

75. Cole KN, Dale PS. Direct language instruction and interactive language instruction with language delayed preschool children: a comparison study. *J Speech Hear Res 1986*;29(2):206-217.

76. Courtright JA, Courtright IC. Imitative modeling as a language intervention strategy: the effects of two mediating variables. *J Speech Hear Res 1979*;22(2):389-402.

77. Gibbard D. Parental-based intervention with pre-school language-delayed children. *Eur J Disord Commun 1994*;29(2):131-150.

78. Girolametto L, Pearce PS, Weitzman E. Interactive focused stimulation for toddlers with expressive vocabulary delays. *J Speech Hear Res 1996*;39(6):1274-1283.

79. Girolametto L, Pearce PS, Weitzman E. Effects of lexical intervention on the phonology of late talkers. *J Speech Lang Hear Res 1997*;40(2):338-348.

80. Robertson SB, Weismer SE. Effects of treatment on linguistic and social skills in toddlers with delayed language development. *J Speech Lang Hear Res 1999*;42(5):1234-1248.

81. Robertson SB, Weismer SE. The influence of peer models on the play scripts of children with specific language impairment. *J Speech Lang Hear Res 1997*;40(1):49-61.

82. Rvachew S, Nowak M. The effect of target-selection strategy on phonological learning. *J Speech Lang Hear Res 2001*;44:610-623.

83. Shelton RL, Johnson AF, Ruscello DM, Arndt WB. Assessment of parent-administered listening training for preschool children with articulation deficits. *J Speech Hear Disord 1978*;43(2):242-254.

84. Wilcox MJ, Kouri TA, Caswell SB. Early language intervention: a comparison of classroom and individual treatment. *American Journal of Speech-Language Pathology 1991*;1(1):49-61.

85. Law J, Kot A, Barnett G. A comparison of two methods for providing intervention to three year old children with expressive/receptive language impairment. London: Department of Language and Communication Science, City University; 1999. Report No.: 002.

86. Fey ME, Cleave PL, Ravida AI, S.H. L, Dejmal AE, Easton DL. Effects of grammar facilitation on phonological performance of children with speech and language impairments. *J Speech Hear Res 1994*;37:594-607.

87. Fey ME, Cleave PL, Long SH. Two models of grammar facilitation in children with language impairments: phase 2. *J Speech Lang Hear Res 1997*;40(1):5-19.

88. Mulac A, Tomlinson CN. Generalization of an operant remediation program for syntax with language delayed children. *J Commun Disord 1977*;10(3):231-243.

89. Ruscello DM, Cartwright LR, Haines KB, Shuster LI. The use of different service delivery models for children with phonological disorders. *J Commun Disord 1993*;26(3):193-203.

90. Rvachew S. Speech perception training can facilitate sound production learning. *J Speech Hear Res 1994*;37(2):347-357.

91. Schwartz RG, Chapman K, Terrell BY, Prelock P, Rowan L. Facilitating word combination in language-impaired children through discourse structure. *J Speech Hear Disord 1985*;50(1):31-39.

92. Fey ME, Cleave PL, Long SH, Hughes DL. Two approaches to the facilitation of grammar in children with language impairment: an experimental evaluation. *J Speech Hear Res 1993*;36(1):141-157.

93. Girolametto L, Pearce PS, Weitzman E. The effects of focuses stimulation for promoting vocabulary in young children with delays: a pilot study. *Journal of Childhood Communication Development 1996*;17(2):39-49.

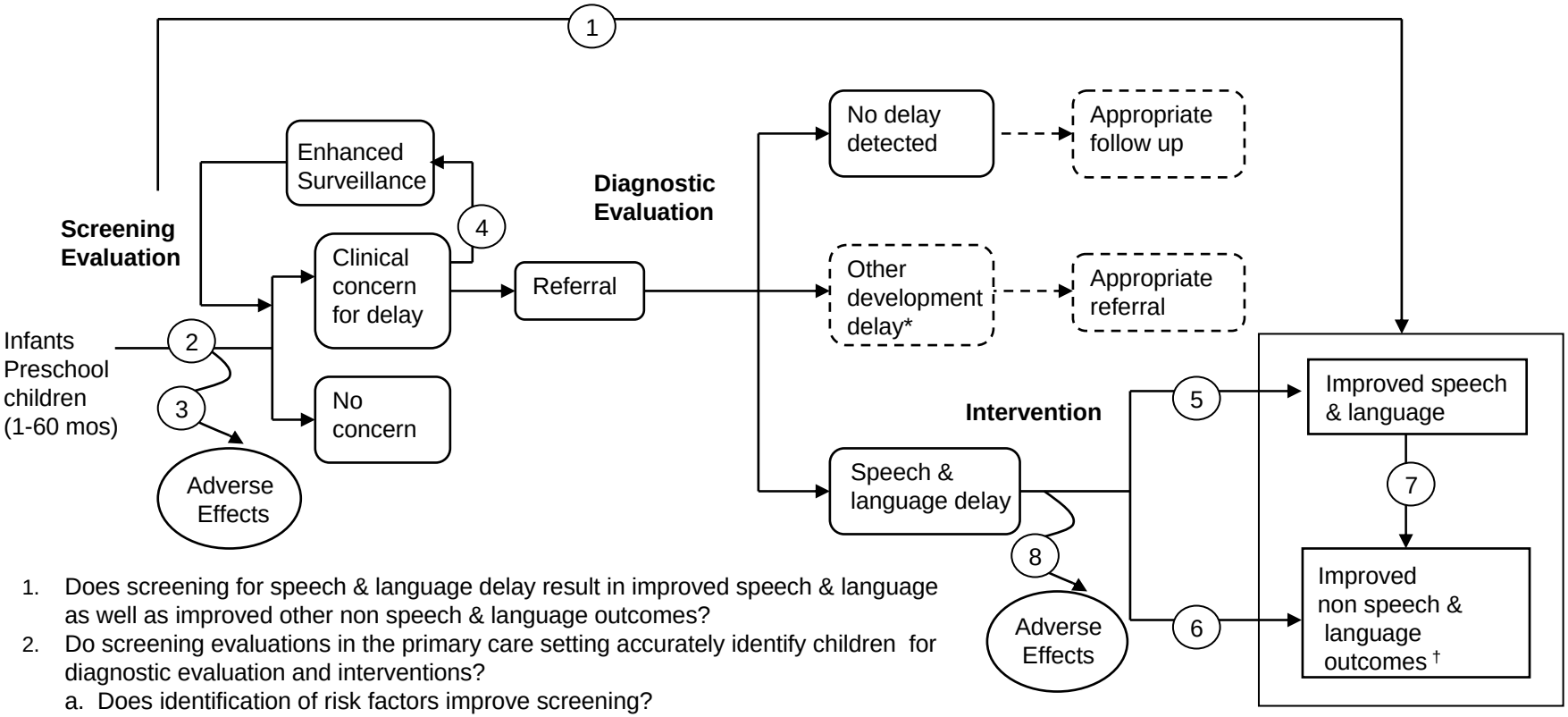
94. Glogowska M, Campbell R, Peters TJ, Roulstone S, Enderby P. A multimethod approach to the evaluation of community preschool speech and language therapy provision. *Child: Care, Health & Development. 2002*;28(6):513-521.

95. Reid J, Donaldson ML. The effectiveness of therapy for child phonolgical disorder: the Metaphon approach. In: Aldridge M, editor. Child Language. Clevedong, Avon: Multiligual Matters; 1996.

96. Sutton L, Tapper L. Investigating WILSTAAR. *Bulletin of the Royal College of Speech and Language Therapists 1999*;August.

97. King TM, Glascoe FP. Developmental surveillance of infants and young children in pediatric primary care. *Current Opinion in Pediatrics 2003*;15(6):624-629.

Figure 1. Analytic Framework and Key Questions. The analytic framework represents an outline of the evidence review and includes the patient population, interventions, outcomes, and adverse effects of the screening process. The key questions examine a chain of evidence about the effectiveness, accuracy, and feasibility of screening children age 5 years and younger for speech and language delay in primary care settings (key questions 1and 2), adverse effects of screening (key question 3), the role of enhanced surveillance in primary care (key question 4), effectiveness of interventions for children identified with delay (key questions 5, 6, and 7), and adverse effects of interventions (key question 8).



1. Does screening for speech & language delay result in improved speech & language as well as improved other non speech & language outcomes?
2. Do screening evaluations in the primary care setting accurately identify children for diagnostic evaluation and interventions?
   a. Does identification of risk factors improve screening?
   b. What are screening techniques and how do they differ by age?
   c. What is the accuracy of screening techniques and how does it vary by age?
   d. What are the optimal ages and frequency for screening?
3. What are the adverse effects of screening?
4. What is the role of enhanced surveillance by primary care clinicians?
5. Do interventions for speech & language delay improve speech & language outcomes?
6. Do interventions for speech & language delay improve other non speech & language outcomes?
7. Does improvement in speech & language outcomes lead to improved additional outcomes?
8. What are the adverse effects of interventions?

* Autism, mental retardation, Fragile X, hearing loss, degenerative and other neurologic disorders.

† School performance, family function, social behavior, and others.

Table 1. Definitions of terms

Term	Definition
Articulation	The production of speech sounds.
Dysfluency	Interrupted flow of speech sounds, such as stuttering.
Expressive language	The use of language to share thoughts, protest, or comment.
Language	The conceptual processing of communication which may be receptive and or expressive.
Morphology	The rules governing meanings of word units.
Phonology	The set of rules for sound production.
Pragmatics	Adaption of language to the social context.
Prosody	Appropriate intonation, rate, rhythm and loudness of speech utterances.
Receptive language	Understanding of language.
Semantics	A set of words known to a person that are a part of a specific language (vocabulary).
Speech	Verbal production of language.
Syntax	The way linguistic elements are put together to form phrases or clauses (grammar).
Voice disorders	Difficulty with speech sound production, at the level of the larynx, may be related to motor or anatomical issues, e.g., hypernasal or hoarse speech.

Table 2. Summary of studies of risk factors

Author/year	Population	Age range (months)	Speech & language domains	Family history	Male sex	SES	Birth order	Perinatal factors	Parental education	Medical conditions	Other associations
English Language											
Brookhouser, 1979[36]	24 referred from Boys Town Institute.	28-62	Language	0	NR	NR	NR	NR	NR	X	NR
Campbell, 2003[37]	398 cases and 241 controls from a large, prospective study in Pittsburgh, PA.	36	Speech	X	X	X	NR	NR	X Mother	NR	NR
Cantwell, 1985[38]	600 children referred from a speech and hearing clinic in Los Angeles, CA.	20-191	Multiple types	NR	NR	NR	NR	NR	NR	NR	X Psychiatric, behavioral, or developmental disorder
Choudhury, 2003[39]	42 cases with positive family histories and 94 controls from New York City, NY area.	36	Language	X	X	0	NR	NR	0 Mother or father	0 Asthma	X Older parents, more children in family
Singer, 2001[40]	98 cases (VLBW/BPD), 70 VLBW/non BPD controls, and 95 term controls from Cleveland, OH region hospitals.	36	Language	NR	NR	X	NR	NR	NR	X BPD, PDA	X Neurologic risk, minority race
Tallal, 1989[41]	76 cases and 54 controls from the San Diego, CA Longitudinal Study.	48-59	Language	X	NR	NR	NR	NR	Mother X Father	NR	NR

Table 2. Summary of studies of risk factors (continued)

Author/year	Population	Age range (months)	Speech & language domains	Family history	Male sex	SES	Birth order	Perinatal factors	Parental education	Medical conditions	Other associations
English Language (continued)											
Tomblin, 1991[42]	662 from a longitudinal cohort.	30-60	Speech and language	X	X	NR	X Born later	NR	O Mother X Father	NR	NR
Tomblin, 1997[43]	177 cases and 925 controls from metro regions of Iowa or Illinois.	Kindergarten age	Speech and language	X	NR	NR	NR	0 Low birth weight	X Mother or father	NR	X Younger mother, less breastfeeding
Whitehurst, 1991[44]	62 cases and 55 controls from Long Island, NY.	24-38	Expressive language	0	NR	NR	NR	NR	NR	NR	NR
Non-English Language											
Fox, 2002[45] (Germany)	65 cases and 48 controls.	32-86	Speech	X	NR	NR	NR	X Birth difficulties, sucking habits	NR	NR	NR
Klein, 1986[46] (Israel)	72 kindergarten children from a middle-class urban area.	48-108	Vocabulary	NR	NR	NR	NR	NR	NR	NR	0 Child's behavior
Kloth, 1995[47] (Netherlands)	93 referred because one or both parents were stutterers or had a history of stuttering.	23-58	Stuttering	NR	NR	NR	NR	NR	NR	NR	0 Mother stutters, mother's speaking style or rate
Lyytinen, 2001[48] (Finland)	107 with familial risk of dyslexia and 93 without.	0-54	Speech and language	X	NR	NR	NR	NR	NR	NR	NR

Table 2. Summary of studies of risk factors (continued)

Author/year	Population	Age range (months)	Speech & language domains	Family history	Male sex	SES	Birth order	Perinatal factors	Parental education	Medical conditions	Other associations
Non-English Language (continued)											
Peters, 1997[49] (Netherlands)	946 from a Dutch birth cohort in Nijmegen.	84-96	Language and educational attainment	NR	0	NR	NR	0 Preterm or low birth weight	X	NR	X Dutch as a second language
Weindrich, 2000[50] (Germany)	320 recruited at birth at a German hospital.	Tested at 54 and 96 months	Receptive and expressive language and articulation (54 months); reading and spelling (96 months)	NR	NR	NR	NR	X Preterm, toxemia, low birth weight	X Mother or father	NR	X Parental psychiatric disorder, overcrowding, parental broken home or delinquency, one-parent family, unwanted pregnancy
Yliherva, 2001[51] (Finland)	8,370 recruited at birth from 2 northern provinces of Finland (99% of pregnant women in 1985-1986).	96	Speech, language, learning, motor abilities	NR	X	NR	NR	X Preterm, low birth weight	0 Mother	X Impaired hearing	0 Mother's age X >4 children in family, reconstructed family status

Key/Abbreviations
X=statistically significant association
0=variable examined and not associated with delay
NR=not reported
SES=socioeconomic status
VLBW=very low birth weight
BPD=bronchopulmonary dysplasia
PDA=patent ductus arteriosis

Table 3. Instruments used in studies

Instrument	Abbreviation	Components	References
Bayley Infant Neurodevelopmental Screener*	BINS	Assesses 4 areas: 1) Neurological function/intactness 2) Receptive function 3) Expressive function 4) Cognitive processes	Macias, 1998
Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale	CAT/ CLAMS	Includes psychometrics and speech and language milestones. CAT: 19 age sets with 12 instruments and 57 items for visual motor skills. CLAMS: 19 age sets with 3 instruments up to 24 months and 4 instruments after 24 months, includes 43 items for language skills.	Clark, 1995
Denver Developmental Screening Test - II*	DDST II	Domains include: 1) Language 2) Fine motor-adaptive 3) Personal-social 4) Gross motor	Glascoe, 1993
Developmental Profile-II*	DP-II	5 subsets: 1) Physical 2) Self-help 3) Social 4) Academic 5) Communication	Glascoe, 1993
Early Language Milestone Scale		41 items covering 4 areas: 1) Auditory expressive 2) Auditory receptive 3) Visual expressive 4) Visual receptive	Coplan, 1982; Black, 1988; Walker, 1989
Fluharty Preschool Speech and Language Screening Test		35 items separated into 3 sections (A, B, C) including identification of 15 common objects (phoneme), nonverbal responses to 10 sentences (syntax), and imitation of 10 one-sentence picture descriptions. Assess identification, articulation, comprehension, and repetition.	Blaxley, 1983; Sturner, 1993; Allen, 1987
Hackney Early Language Screening Test		20-item test in 7 sections: 1) Comprehension - following instructions to manipulate toys. 2) Expression - tester manipulates toys and asks child questions about this. 3) Comprehension - following instructions for placing toys. 4) Comprehension - child chooses picture from 3 options. 5) Expression - child answers question about pictures. 6) Expression - child names objects. 7) Comprehension - child chooses picture from 4 options.	Dixon, 1988; Law, 1994

Table 3. Instruments used in studies (continued)

Instrument	Abbreviation	Components	References
Language Development Survey	LDS	310 words arranged in 14 semantic categories. Parents indicate which words their child has spoken and describe word combinations of 2 or more words that their child has used.	Klee, 1998; Klee, 2000; Rescorla, 2001
Levett-Muir Language Screening Test		Test is divided into 6 sections: 1) Comprehension - child is asked to pick toys from group. 2) Vocabulary - child's ability to name the toys. 3) Comprehension - using pictures child is required to respond to questions. 4) Vocabulary - child's ability to name what's in the pictures. 5) Comprehension & representation - child's ability to answer "what" and "who" questions. 6) Overall - child is asked to explain the detailed composite picture.	Levett, 1983
Parent Evaluation of Developmental Status	PEDS	2 questions for parents to elicit concerns in general and in specific areas. Other items determine reasons for parents' concerns.	Glascoe, 1991
Parent Language Checklist*	PLC	12 questions for parents about their child's receptive and expressive language including one question assessing hearing problems.	Burden, 1996
Pediatric Language Acquisition Screening Tool for Early Referral	PLASTER	Communication development milestones by age with 7 individual areas. Each area contains 10 questions (5 relate to receptive language and 5 expressive language).	Sherman, 1996
Screening Kit of Language Development	SKOLD	Vocabulary comprehension, story completion, sentence completion, paired sentence repetition with pictures, individual sentence repetition with pictures, individual sentence repetition without pictures, auditory comprehension of commands.	Bliss, 1984
Sentence Repetition Screening Test	SRST	15 sentences repeated one at a time by the child after demostration by the tester.	Sturner, 1996
Structured Screening Test		20 questions covering both expressive and receptive language skills.	Laing, 2002
Test for Examining Expressive Morphology	TEEM	54 items targeting a variety of morphosyntactic structures using a sentence completion task.	Merrell, 1997

Key
*Speech and language are part of a broader screening instrument.

Table 4. Studies of screening instruments for children up to 2 years old

Author/ year	N	Instrument	Reference standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Study quality rating
Under 5 minutes to administer										
Glascoe, 1991[56]	157	Parent Evaluation of Developmental Status	Clinical assessment	Expressive language, articulation	From outpatient clinic or private practice; 78% Caucasian; 54% male; 6-77 months.	Clinic	Doctoral students in psychology or special education	72%	83%	Good
Coplan, 1982[30]	191	Early Language Milestone Scale	Clinical assessment	Expressive and receptive language	From private practices and pediatric outpatients of hospital; 80% Caucasian; 50% male; 0-36 months.	Physician's office	Medical students	97%	93%	Fair
Black, 1988[55]	48	Early Language Milestone Scale	Receptive-Expressive Emergent Language Scale, Bayley Scales of Infant Development	Expressive and receptive language	From low socioeconomic groups; 8-22 months.	Large pediatric clinic, university teaching hospital	Not reported	83%	100%	Poor
5-10 minutes to administer										
Glascoe, 1993[57] *Study 1*	89	denver Developmental Screening Test II (communica-tion components)	Battery of measures	Fine motor adaptive, personal social, gross motor, and language	From 5 day care centers; 52% male; 7-70 months.	Day care centers	Psychologist	73%	76%	Fair

Table 4. Studies of screening instruments for children up to 2 years old (continued)

Author/ year	N	Instrument	Reference standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Study quality rating
5-10 minutes to administer (continued)										
Sherman, 1996[52]	173	Pediatric Language Acquisition Screening Tool for Early Referral (PLASTER)	Early Language Milestone Scale	Expressive and receptive language	123 high risk infants; 50 normal controls; 3-36 months.	*High risk:* neonatal developmental follow-up clinic *Control:* speech and hearing clinic	Speech and language pathologist and graduate students	53%	86%	Fair
10 minutes to administer										
Macias, 1998[62]	78	Bayley Infant Neurodevelopmental Screener	Bayley Scales of Infant Development II	Expressive and receptive language	Randomly selected from those presenting for routine neonatal high-risk follow-up; 54% male; 62% African American; 6-23 months.	Physician's office	Developmental pediatrician	*Using middle-cut scores:* 73%	*Using middle-cut scores:* 66%	Fair
Klee, 1998[59]	306	Language Development Survey	Infant Mullen Scales of Early Learning	Expressive vocabulary	Toddlers turning 2-years old during the study in Wyoming; 52% male; 24-26 months.	Home	Parent	91%	87%	Good-Fair
Klee, 2000[60]	64	Language Development Survey	Infant Mullen Scales of Early Learning	Expressive vocabulary	Children turning 2 years in a specific month in an area of Wyoming.	Home	Parent	*At age 2:* 83% *At age 3:* 67%	*At age 2:* 97% *At age 3:* 93%	Fair

Table 4. Studies of screening instruments for children up to 2 years old (continued)

Author/ year	N	Instrument	Reference standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Study quality rating
10 minutes to administer (continued)										
Rescorla, 2001[61]	422	Language Development Survey	Bayley Scales of Infant Development, Stanford-Binet, Reynell Developmental Language Scales	Expressive vocabulary Delay 1: <30 words and no word combinations Delay 2: <30 words or no word combinations Delay 3: <50 words or no word combinations	Toddlers in four townships of Delaware County, PA turning 2-years old during the study.	Home	Parent and research assistant	*Delay 1:* Bayley - 70%; Binet - 52%; Reynell - 67% *Delay 2:* Bayley - 75%; Binet - 56%; Reynell - 89% *Delay 3:* Bayley - 80%; Binet - 64%; Reynell - 94%	*Delay 1:* Bayley - 99%; Binet - 98%; Reynell - 94% *Delay 2:* Bayley - 96%; Binet - 95%; Reynell - 77% *Delay 3:* Bayley - 94%; Binet - 94%; Reynell - 67%	Fair
Clark, 1995[58]	99	Clinical Linguistic and Auditory Milestone Scale	Sequenced Inventory of Communication Development	Syntax, pragmatics	Infants turning 1 or 2 years old during study; 55% male; 0-36 months.	Home or school for the deaf	Speech and language pathologist	*Receptive:* 14-24 months: 83% 25-36 months: 68% *Expressive:* 14-24 months: 50% 25-36 months: 88%	*Receptive:* 14-24 months: 93% 25-36 months: 89% *Expressive:* 14-24 months: 91% 25-36 months: 98%	Fair

Table 4. Studies of screening instruments for children up to 2 years old (continued)

Author/ year	N	Instrument	Reference standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Study quality rating
10 minutes to administer (continued)										
Glascoe, 1993[57] *Study 2*	89	Denver Developmental Screening Test II (communication components)	Battery of measures	Physical, self-help, social, academic, and communication	Children from five day care centers; 52% male: 7-70 months.	Day care centers	Psychologist	22%	86%	Fair

Table 5. Studies of screening instruments for children 2 to 3 years old

Author/ year	N	Instrument	Reference standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Study quality rating
5 minutes to administer										
Burden, 1996[11]	2,590	Parent Language Checklist	Clinical judgement	Expressive and receptive language	All children turning 36 months; 52% male; 41% urban.	Home (mailed)	Parent	87%	45%	Good
Laing, 2002[63]	376	Structured Screening Test	Reynell Developmental Language Scales	Expressive and receptive language	Children from 2 low SES counties in London; mean age 30 months.	Physician's office	Health visitor	*Severe:* 66% *Needs therapy:* 54%	*Severe:* 89% *Needs therapy:* 90%	Fair
Levett, 1983[64]	140	Levett-Muir Language Screening Test	Reynell Developmental Language Scales, Goldman-Fristoe Test of Articulation, Language Assessment and Remediation Procedure	Receptive language, phonology, syntax	Private practice population; 34-40 months.	Physician's office	Medical practitioners	100%	100%	Fair
Sturner, 1993[53] *Study 1*	279	Fluharty Preschool Speech and Language Screening Test	Arizona Articulation Proficiency Scale Revised, Test of Language Development Primary	Expressive and receptive language, articulation	46% male; 74% Caucasian; 86% rural; 24-72 months.	Preschool	Teacher	*Speech & Language:* 43% *Speech:* 74% *Language:* 38%	*Speech & Language:* 82% *Speech:* 96% *Language:* 85%	Fair

Table 5. Studies of screening instruments for children 2 to 3 years old (continued)

Author/ year	N	Instrument	Reference standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Study quality rating
5 minutes to administer (continued)										
Sturner, 1993[53] *Study 2*	421	Fluharty Preschool Speech and Language Screening Test	Test for Auditory Comprehension of Language Revised, Templin-Darley Test of Articulation	Expressive and receptive language, articulation	52% male; 75% Caucasian; 24-72 months.	Preschool	Teacher	*Speech & Language:* 31% *Speech:* 43% *Language:* 17%	*Speech & Language:* 93% *Speech:* 93% *Language:* 97%	Fair
10 minutes to administer										
Law, 1994[67]	1,205	Hackney Early Language Screening Test	Reynell Developmental Language Scales	Expressive language	Children attending routine developmental checkups; mean age 30 months.	Home	Health visitor	98%	69%	Good-Fair
Blaxley, 1983[65]	90	Fluharty Preschool Speech and Language Screening Test	Developmental Sentence Scoring	Expressive and receptive language, articulation	Children referred for speech &/or language assessment and intervention and controls; 24-72 months.	Speech and hearing clinic in western Ontario	Clinician	*10th percentile:* 36% *25th percentile:* 30%	*10th percentile:* 95% *25th percentile:* 100%	Fair
Bliss, 1984[66]	602	Screening Kit of Language Development	Sequenced Inventory of Communication Development	Expressive and receptive language	From day care centers in Detroit; 30-48 months.	Speech and language hearing clinic, day-care center, physician's office, educational and health facilities	Paraprofessionals and speech and language pathologists	*30-36 months:* 100% *37-42 months:* 100% *43-48 months:* 100%	*30-36 months:* 98% *37-42 months:* 91% *43-48 months:* 93%	Fair

Table 5. Studies of screening instruments for children 2 to 3 years old (continued)

Author/ year	N	Instrument	Reference standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Study quality rating
10 minutes to administer (continued)										
Dixon, 1988[54]	25	Hackney Early Language Screening Test	Clinical judgement	Expressive language	Pilot study at one clinic setting in Hackney; mean age 30 months.	Physician's office	Health visitor	95%	94%	Poor
Walker, 1989[68]	77	Early Language Milestone Scale	Sequenced Inventory of Communication Development	Expressive and receptive language	All children attending a study clinic; mean age 36 months.	Clinic	Speech and language pathologist	*0-12 months:* 0% *13-24 months:* 100% *25-36 months:* 100%	*0-12 months:* 86% *13-24 months:* 60% *25-36 months:* 75%	Poor

Table 6. Studies of screening instruments for children 3 to 5 years old

Author/ year	N	Instrument	Reference standard	Speech & language domains	Subjects	Setting	Screener	Sensitivity	Specificity	Study quality rating
10 minutes or less to administer										
Allen, 1987[69]	182	Fluharty Preschool Speech and Language Screening Test	Sequenced Inventory of Communication Development	Expressive and receptive language, articulation	From day care programs; 36-47 months.	Clinic	Speech and language pathologists	60%	80%	Fair
Sturner, 1996[71]	76	Sentence Repetition Screening Test	Speech and Language Screening Questionnaire	Receptive and expressive language, articulation	Children registering for kindergarten; 48% male; 65% Caucasian; 54-66 months.	School	Nonspecialists or school speech and language pathologists	*Receptive and expressive:* 62% *Articulation:* 57%	*Receptive and expressive:* 91% *Articulation:* 95%	Fair
Merrell, 1997[70]	40	Test for Examining Expressive Morphology	Kaufman Assessment Battery for Children, Structured Photographic Expression Lanuage Test II	Expressive vocabulary, syntax	20 impaired and 20 unimpaired; 52% male 73% Caucasian; 48-67 months.	School or clinic	Speech and language pathologists	90%	95%	Poor

Table 7. Randomized controlled trials of interventions

Author/ year	Speech & language domains	N	Age (months)	Interventions	Speech and language outcomes	Function and health outcomes	Study quality rating
Up to 2 years*							
Glogowska, 2000[72]	Expressive and receptive language and phonology	159 in 2 groups	18-42	Clinician-directed individual intervention routinely offered by the therapist for 12 months vs. none	Improved auditory comprehension in intervention vs. control group; no differences for expressive language, phonology error rate, language development, or improvement on entry criterion	No differences in well being, attention level, play level, or socialization skills	Good
2 to 3 years							
Gibbard, 1994[77] *Study 1*	Expressive language	36 in 2 groups	27-39	Parent-directed individual therapy 60-75 minutes every other week for 6 months vs. none	Improved scores on several measures for intervention vs. control group	Not reported	Fair
Girolametto, 1996[78]	Expressive language	25 in 2 groups	23-33	Parent-directed individual focused stimulation intervention 150 minutes per week for 11 weeks vs. none	Larger vocabularies, use of more different words, more structurally complete utterances and multiword utterances in intervention group vs. control; no differences in several other measures	Not reported	Fair
Law, 1999[85]	Expressive and receptive language	38 in 3 groups	33-39	Clinician-directed 450 minutes per week for 6 weeks vs. parent-directed 150 minutes per week for 10 weeks vs. none	No differences between groups	Improved parent perception of child's behavior and positivity towards child, improved child self-esteem	Fair
Robertson, 1999[80]	Expressive and receptive language	21 in 2 groups	21-30	Clinician-directed individual therapy 150 minutes per week for 12 weeks vs. none	Improved mean length of utterances, total number of words, lexical diversity, vocabulary size, and percentage of intelligible utterances in intervention group vs. control	Improved socialization skills, decreased parental stress for intervention group	Fair

Table 7. Randomized controlled trials of interventions (continued)

Author/ year	Speech & language domains	N	Age (months)	Interventions	Speech and language outcomes	Function and health outcomes	Study quality rating
2 to 3 years (continued)							
Gibbard, 1994[77] *Study 2*	Expressive language	25 in 3 groups	27-39	Clinician-directed individual therapy 60-75 minutes every other week for 6 months vs. parent-directed 60-75 minutes every other week for 6 months vs. none	Improved scores on all 5 measures for parent-directed group vs. control; improvement on 2 measures for clinician-directed group vs. control; improvement on 1 measure for parent vs. clinician group	Not reported	Poor
Girolametto, 1996[93]	Expressive and receptive language	16 in 2 groups	22-38	Parent-directed individual therapy 150 minutes per week for 10 weeks vs. none	More target words used in intervention group vs. control; no differences in vocabulary development	Increased symbolic play gestures, decreased aggressive behavior in intervention group	Poor
Schwartz, 1985[91]	Expressive language and lexical acquisition	10 in 2 groups	32-39	Clinician-directed individual therapy for 3 weeks vs. none	Improved multiword utterances from baseline in intervention group; no between group differences reported	Not reported	Poor
Wilcox, 1991[84]	Lexical acquisition	20 in 2 groups	20-47	Clinician-directed individual intervention 90 minutes per week for 3 months vs. classroom intervention 360 minutes per week for 3 months	No differences between groups in use of target words; more use of words at home in classroom group vs. individual group	Not reported	Fair
Girolametto, 1997[79]	Lexical acquisition and phonology	25 in 2 groups	23-33	Parent-directed individual therapy eight 150-minute sessions and 3 home sessions for 11 weeks vs. none	Improved level of vocalizations and inventory of consonants for intervention group vs. control; no differences in the number of vocalizations	Not reported	Fair

Table 7. Randomized controlled trials of interventions (continued)

Author/ year	Speech & language domains	N	Age (months)	Interventions	Speech and language outcomes	Function and health outcomes	Study quality rating
3 to 5 years							
Barratt, 1992[74]	Expressive and receptive language	39 in 2 groups	37-43	Clinician-directed interactive language therapy for 40 minutes weekly for 6 months (traditional group) vs. 40 minutes for 4 days per week for 3 weeks in two 3 month blocks (intensive group)	Improved expression score on Reynell scale for intensive group vs. weekly (or traditional) therapy group; no difference in comprehension scores, both improved	Not reported	Fair
Courtright, 1979[76]	Expressive language	36 in 3 groups	47-83	3 clinician-directed approaches are compared for 5 months: mimicry, clinician modeling, 3rd person modeling for 5 months	Increased number of correct responses in modeling groups vs. mimicry group	Not reported	Fair
Robertson, 1997[81]	Expressive and receptive language	30 in 3 groups	44-61	2 clinician-directed play groups with language impairments (treatment vs control) with normal peers for 20 minutes per week for 3 weeks	More words used, greater verbal productivity, more lexical diversity, and more use of linguistic markers by normal peer play group (not normal group, treatment group with language impairment) vs. control	Play-theme-related acts increased for the normal peer play group (not normal group, treatment group with language impairment)	Fair
Glogowska, 2002[72]	Expressive and receptive language and phonology	159 in 2 groups	≤42	Clinician-directed for 12 months vs. none	Improved receptive language in intervention group vs. control; no differences between groups for 4 other measures	Improved family response to child in intervention group	Poor
Almost, 1998[73]	Phonology	26 in 2 groups	33-61	Clinician-directed individual therapy two 30-minute sessions per week for 4 months vs. none	Higher scores on 3 of 4 measures for intervention vs. control group	Not reported	Fair

Table 7. Randomized controlled trials of interventions (continued)

Author/ year	Speech & language domains	N	Age (months)	Interventions	Speech and language outcomes	Function and health outcomes	Study quality rating
3 to 5 years (continued)							
Rvachew, 2001[82]	Phonology	48 in 2 groups	50 (mean)	Clinician-directed individual therapy 30-40 minutes per week for 12 weeks; compares interventions for phonemes that differ (most knowledge/early developing group vs. least knowledge/latest developing group)	Improved scores on measures from baseline for both intervention groups; greater improvement for most knowledge/early developing phonemes group vs. comparison (least knowledge/latest developing) group	Not reported	Fair
Shelton, 1978[813]	Phonology and articulation	45 in 3 groups	27-55	Parent-directed individual therapy 5 minutes per day (listening group) vs. 15 minutes per day (reading and talking group) for 57 days vs. none	No improvements for intervention groups vs. control	Not reported	Fair
Fey, 1994[86]	Phonology and syntax	26 in 3 groups	44-70	Clinician-directed sessions (individual and group) for 3 hours per week for 20 weeks vs. parent-directed sessions for 8 hours per week for weeks 1-12 (includes intensive parent training) then 4 hours per week for weeks 13-20 vs. none	Improved grammatical output (developmental sentence scores) for both intervention groups vs. control; no significant difference between groups for phonological output (percentage consonants correct)	Not reported	Poor
Reid, 1996[95]	Phonology	30 in 2 groups	42-66	Clinician-directed individual therapy 30 minutes per week for 6-10 weeks vs. none	Improved scores on some measures from baseline for intervention and control groups; no between group comparisons reported	Not reported	Poor
Ruscello, 1993[89]	Phonology	12 in 2 groups	49-68	Clinician-directed vs. clinician & parent-directed individual therapy 120 minutes per week for 8 weeks	Improved scores on measures from baseline for both intervention groups; no between group comparisons reported	Not reported	Poor

Table 7. Randomized controlled trials of interventions (continued)

Author/ year	Speech & language domains	N	Age (months)	Interventions	Speech and language outcomes	Function and health outcomes	Study quality rating
3 to 5 years (continued)							
Rvachew, 1994[90]	Phonology	27 in 3 groups	42-66	Clinician-directed individual therapy 45 minutes per week for 6 weeks; compares 3 groups listening to different sets of words	Improved scores on measures for 2 intervention groups vs. third group	Not reported	Poor
Cole, 1986[75]	Syntax	44 in 2 groups	38-69	Clinician-directed individual directive approach vs. interactive approach for 600 minutes per week for 8 months	Improved scores on 6 of 7 measures from baseline for both intervention groups; no significant differences between groups	Not reported	Fair
Fey, 1993[92] *First phase*	Syntax	29 in 3 groups	44-70	Clinician-directed sessions (individual and group) for 3 hours per week for 20 weeks vs. parent-directed sessions for 8 hours per week for weeks 1-12 (includes intensive parent training) then 4 hours per week for weeks 13-20 vs. none	Improved scores on 3 of 4 measures for both intervention groups vs. control; no differences between intervention groups	Not reported	Poor
Fey, 1997[87] *Second phase*	Syntax	28 in 3 groups	44-70	Clinician-directed vs. parent-directed vs. none for 5 months continuing from prior study	Improved some developmental sentence scores from baseline in both intervention groups vs. control; no between group comparisons reported, except that clinician-directed treatment groups had larger and more consistent gains than parent-directed treatment groups or control	Not reported	Poor

Table 7. Randomized controlled trials of interventions (continued)

Author/ year	Speech & language domains	N	Age (months)	Interventions	Speech and language outcomes	Function and health outcomes	Study quality rating
3 to 5 years (continued)							
Mulac, 1977[88]	Syntax	9 in 3 groups	52-75	Clinician-directed individual Monteray language program vs. Monteray language program with extended transfer training for 67 minutes per week for 4 weeks vs. none	Improved scores for both intervention groups vs. control; no significant differences between intervention groups	Not reported	Poor

Key
*Studies with a range of ages are not repeated across categories in the table.

Appendix. USPSTF Quality Rating Criteria

Diagnostic Accuracy Studies

Criteria

- Screening test relevant, available for primary care, adequately described
- Study uses a credible reference standard, performed regardless of test results
- Reference standard interpreted independently of screening test
- Handles indeterminate results in a reasonable manner
- Spectrum of patients included in study
- Sample size
- Administration of reliable screening test

Definition of Ratings Based on Above Criteria

Good: Evaluates relevant available screening test; uses a credible reference standard; interprets reference standard independently of screening test; reliability of test assessed; has few or handles indeterminate results in a reasonable manner; includes large number (more than 100) broad-spectrum patients with and without disease.

Fair: Evaluates relevant available screening test; uses reasonable although not best standard; interprets reference standard independent of screening test; moderate sample size (50 to 100 subjects) and a "medium" spectrum of patients.

Poor: Has important limitations such as: uses inappropriate reference standard; screening test improperly administered; biased ascertainment of reference standard; very small sample size of very narrow selected spectrum of patients.

Randomized Controlled Trials (RCTs) and Cohort Studies

Criteria

- Initial assembly of comparable groups**:** RCTs—adequate randomization, including concealment and whether potential confounders were distributed equally among groups; cohort studies—consideration of potential confounders with either restriction or measurement for adjustment in the analysis; consideration of inception cohorts
- Maintenance of comparable groups (includes attrition, cross-overs, adherence, contamination)
- Important differential loss to follow-up or overall high loss to follow-up
- Measurements: equal, reliable, and valid (includes masking of outcome assessment)
- Clear definition of interventions
- Important outcomes considered
- Analysis: adjustment for potential confounders for cohort studies, or intention-to-treat analysis for RCTs (i.e. analysis in which all participants in a trial are analyzed according to the intervention to which they were allocated, regardless of whether or not they completed the intervention)

Appendix 1. USPSTF Quality Rating Criteria (continued)

Definition of Ratings Based on Above Criteria

Good: Meets all criteria: Comparable groups are assembled initially and maintained throughout the study (follow-up at least 80 percent); reliable and valid measurement instruments are used and applied equally to the groups; interventions are spelled out clearly; important outcomes are considered; and appropriate attention to confounders in analysis.

Fair: Studies will be graded "fair" if any or all of the following problems occur, without the important limitations noted in the "poor" category below: Generally comparable groups are assembled initially but some question remains whether some (although not major) differences occurred in follow-up; measurement instruments are acceptable (although not the best) and generally applied equally; some but not all important outcomes are considered; and some but not all potential confounders are accounted for.

Poor: Studies will be graded "poor" if any of the following major limitations exists: Groups assembled initially are not close to being comparable or maintained throughout the study; unreliable or invalid measurement instruments are used or not applied at all equally among groups (including not masking outcome assessment); and key confounders are given little or no attention.

Excluded Studies of Instruments

Author/Year	Instruments	Reason for Excluding
Satz, 1978	Abbreviated Screening Battery	Diagnostic, administration time not reported (8 tests in battery)
German, 1982	Alpern-Boll Developmental Profile II	Diagnostic, <40 minutes to administer, unable to calculate sensitivity/specificity from data
Craig, 2000	Assessment Battery for African American English	Diagnostic, 40 minutes to administer
Ward, 1984	Author created screening tool	Most of the important information is not reported
Blaxley, 1983	Bankson Language Screening Test (BLST)	25 minutes to administer
Macias, 1998	Bayley Infant Neurodevelopmental Screener (BINS)	Developmental pediatrician must administer
Conti-Ramsden, 2003	Children's Test of Nonword Repetition (CNRep)	15 minutes to administer
Gray, 2003	Children's Test of Nonword Repetition (CNRep)	School setting, 15 minutes to administer
Macias, 1998	Clinical Adaptive Test/Clinical Linguistic Auditory Milestone Scale (CAT/CLAMS)	Developmental pediatrician must administer, 15-20 minutes to administer
Plante, 1995	Clinical Evaluation of Language Fundamentals - Preschool (CELF-P)	Diagnostic, Speech-Language Pathologist required, 30-45 minutes to administer Cost = $348
Capute, 1986	Clinical Linguistic and Auditory Milestone Scale (CLAMS)	Linguistic part only, unable to calculate sensitivity/specificity from data
Leppert, 1998	Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS)	Developmental pediatrician must administer, 15-20 minutes to administer
Pittock, 2002	Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS)	Looked at ease of administration, not validity study, no data found useful
Voigt, 2003	Cognitive Adaptive Test/Clinical Linguistic and Auditory Milestone Scale (CAT/CLAMS)	15-20 minutes to administer, unable to calculate sensitivity/specificity from data
McCathren, 2000	Communication and Symbolic Behavior Scales (Communication Composite Only)	Training required, 20 minutes to administer, correlation data, unable to calculate sensitivity/specificity from data
Wetherby, 2002	Communication and Symbolic Behavior Scales Developmental Profile (CSBS DP)	Most of the important information is not reported

Excluded Studies of Instruments

Author/Year	Instruments	Reason for Excluding
Alberts, 1995	Davis Observation Checklist for Texas (DOCT)	>15 minutes to administer, 1 day DOCT administration training class to be able to administer.
Drumwright, 1973	Denver Articulation Screening Exam (DASE)	Training required, 10-15 minutes to administer Cost = $265
Bryant, 1974	Denver Developmental Screening Test	Broad screener, speech & language components not reported separately
Frankenburg, 1992	Denver Developmental Screening Test - II (DDST - II)	Diagnostic, training required, no sensitivity/specificity
Stokes, 1996	Developmental Nurse Screen	Not standardized or validated
Wright, 1980	Developmental Test of Visual Motor Integration	Correlation data only, unable to calculate sensitivity/specificity
Drillien, 1988	Dundee Developmental Screening Program (DDSP)	Diagnostic, not validation study, ages for follow-up were higher than preschool
Dale, 1989	Early Language Inventory	20-30 minutes to administer, correlation data
Black, 1988	Early Language Milestone Scale	1/2 of subjects were at risk for otitis media Cost = $825
Fell, 1999	Early Vocalization Analyzer (EVA)	Diagnostic, Speech-Language Pathologist must administer
Vance, 1987	Expressive One-Word Picture Vocabulary Test	Diagnostic, Speech-Language Pathologist must administer
Gray, 1999	Expressive one-word Picture Vocabulary Test-Revised (EOWPVT-R)	Diagnostic, Speech-Language Pathologist must administer, 1-hour session for all 4 tests
Gray, 1999	Expressive Vocabulary Test (EVT)	Diagnostic, Speech-Language Pathologist must administer, 1-hour session for all 4 tests
Fluharty, 1973	Fluharty Preschool Speech and Language Screening Test	Correlation data only, unable to calculate sensitivity/specificity.
Stott, 2002	General Language Screen (GLS)	Not standardized or validated.
Bountress, 1981	Goldman-Fristoe-Woodcock Test of Auditory Discrimination	Diagnostic, Speech-Language Pathologist must administer, 20-30 minutes to administer, comparisions of tests - no gold standard, unable to calculate sensitivity/specificity from data
Bloom, 1988	Kaufman Assessment Battery for Children (K-ABC)	Diagnostic, 35-85 minutes to administer, unable to calculate sensitivity/specificity from data

Excluded Studies of Instruments

Author/Year	Instruments	Reason for Excluding
Paul, 1991	Language Development Survey (LDS)	30 minutes to administer, study poorly defined, unable to extract important information
Rescorla, 2002	Language Development Survey (LDS)	30 minutes to administer, compared to other screener not a gold standard
Klee, 2000	Language Development Survey (LDS)	30 minutes to administer
Eno, 1995	Lollipop Test Cognitive Language Profile of the Early Screening Profiles Fluharty Preschool Speech and Language Screening Test	15-20 minutes to administer, correlation data only, unable to calculate sensitivity/specificity from data Cost = $35
Dale, 1991	MacArthur Communicative Development Inventory: Toddlers	Diagnostic, 20 minutes to administer, correlation data, unable to calculate sensitivity/specificity from data Cost = $155
Dale, 2003	MacArthur Communicative Development Inventory: UK Short Form (MCDI:UKSF)	20 min to administer, speech & language components not reported separately (unable to determine sensitivity/specificity for specific area)
Thal, 1999	MacArthur Communicative Development Inventory: Words and Gestures (CDI:WG)	Most of the important information is not reported
Thal, 1999	MacArthur Communicative Development Inventory: Words and Sentences (CDI:WS)	Most of the important information is not reported
Miller, 1988	Miller Assessment for Preschoolers (MAP)	Diagnostic, school psychologist must administer, 30-40 minutes to administer, unable to calculate sensitivity/specificity from data Cost = $695
Widerstrom, 1986	Miller Assessment for Preschoolers (MAP)	Diagnostic, looked at discrimination factor of test
Egan, 1986	Miniature Toys Test	Developmental delay study, verbal just small part, did not separate results for speech and language
Tomblin, 1999	Minnesota Child Development Inventory	Correlation data only, unable to calculate sensitivity/specificity
Kopparthi, 1991	Minnesota Child Development Inventory (MCDI)	Diagnostic, neonatal intensive care unit infants
Gorrell, 1981	North Carolina Psychoeducational Screening Test (PET)	Diagnostic, unable to determine administration time

Excluded Studies of Instruments

Author/Year	Instruments	Reason for Excluding
Ratusnik, 1975	Northwestern Syntax Screening Test	20 minutes to administer, children have comorbidities
Ratusnik, 1980	Northwestern Syntax Screening Test	Speech-Language Pathologist required, 20 minutes to administer, unable to calculate sensitivity/specificity from data
Allen, 1987	Northwestern Syntax Screening Test (NSST)	20 minutes to administer
Conti-Ramsden, 2003	Noun Plural Task	Diagnostic, unable to determine administration time
Laing, 2002	Parent Led Method	10-30 minutes to administer, not standardized method
Stokes, 1996	Parent Questionnaire	Not standardized or validated
Whitworth, 1993	Parent Questionnaire	Diagnostic, unable to determine administration time, unable to calculate sensitivity/specificity
Dale, 2003	Parental concern	Not standardized or validated
Conti-Ramsden, 2003	Past Tense Task (PTT)	Diagnostic, unable to determine administration time
Merrell, 1997	Patterned Elicitation Syntax Test with Morphophonemic Analysis (PEST)	Diagnostic, Speech-Language Pathologist required, unable to determine administration time
Vance, 1987	Peabody Picture Vocabulary Test - Revised	Most of the important information is not reported
Gray, 1999	Peabody Picture Vocabulary Test (PPVT-III)	Diagnostic, Speech-Language Pathologist must administer, 1-hour session for all 4 tests
Choong, 1983	Peabody Picture Vocabulary Test -Revised (PPVT-R)	10-20 minutes to administer, unable to calculate sensitivity/specificity from data
Blackman, 1990	Pediatric Examination of Educational Readiness (PEER)	Diagnostic, 30 minutes to administer
Blackman, 1992	Pediatric Examination of Educational Readiness (PEER)	Diagnostic, 30 minutes to administer
Palmer, 1990	Pediatric Examination of Educational Readiness (PEER)	Diagnostic, 30 minutes to administer, developmental pediatricians required
Shriberg, 1986	Photo Articulation Test (PAT)	13-30 minutes to administer, not much information provided
Fagundes, 1998	Preschool Language Assessment Instrument (PLAI)	30 minutes to administer, looking at racial/cultural differences, unable to calculate sensitivity/specificity from data.

Excluded Studies of Instruments

Author/Year	Instruments	Reason for Excluding
Dodge, 1980	Preschool Language Scale (PLS)	Diagnostic, 25-30 minutes to administer
Stevenson, 1976	Pre-School Version of the English Picture Vocabulary Test (EPVT)	Unable to calculate sensitivity/specificity from data, much of the important information is not reported
Jordan, 1973	Queensland University Aphasia and Language Test (QUALT)	Diagnostic, 45-60 minutes to administer, four subtests.
Conti-Ramsden, 2003	Recall of Digits	Diagnostic, Subtests from British Abilities Scale (BAS), unable to determine administration time
Rome-Flanders, 1998	Receptive Expressive Emergent Language Scales (REEL)	20 minutes to administer, correlation study comparing tests, unable to calculate sensitivity/specificity from data
Gray, 1999	Receptive one-word Picture Vocabulary Test (ROWPVT)	Diagnostic, Speech-Language Pathologist must administer, 1-hour session for all 4 tests
German, 1982	Revised Denver Developmental Screening Test (R-DDST) (Language Subtest)	Diagnostic, < 40 minutes to administer, unable to calculate sensitivity/specificity from data
Allen, 1981	Sequenced Inventory of Communication Development (SICD)	Diagnostic, 30-75 minutes to administer Cost = $395-$1050
Sturner, 1996	Speech and Language Screening Questionnaire	Diagnostic, much of the important information is not reported
Fiedler, 1971	Speech Evaluation Form	Not standardized or validated
Culatta, 1983	Story Retelling	Not standardized or validated
Plante, 1995	Structured Photographic Expressive Language Test - Preschool (SPELT-P)	Speech-Language Pathologist required, 10-15 minutes to administer $320
Laing, 2002	Structured Screening Test	Not standardized or validated
Brown, 1986	Test for Auditory Comprehension of Language (TACL)	Diagnostic, 35 minutes to administer, test re-test coefficient
Wright, 1980	Test for Auditory Comprehension of Language (TACL)	Diagnostic, 35 minutes to administer, correlation data only, unable to calculate sensitivity/specificity from data
Di Simoni, 1982	Token Test for Children	Diagnostic, teacher administered, language comprehension score, not vocabulary, syntax or grammar, correlation data, unable to calculate sensitivity/specificity from data

Excluded Studies of Instruments

Author/Year	Instruments	Reason for Excluding
Colligan, 1974	Vane Kindergarten Test (VKT) & Mecham Verbal Language Development Scale (revised)	Diagnostic, technician/Specialist, 10-20 minutes to administer
Bountress, 1981	Wepman Auditory Discrimination Test	Classroom, Speech-Language Pathologist must administer, comparisons of tests - no gold standard, unable to calculate sensitivity/specificity from data Cost = $125
Coulter, 2001	WILSTAAR Screen	Broad screener, speech & language components not reported separately, unable to determin administration time, unable to calculate sensitivity/specificity from data
Oakenfull, 2001	WILSTAAR Screen	Speech-Language Pathologist must administer, program not just screener, unable to determine administration time