Sign Language and Spoken Language for Children With Hearing Loss: A Systematic Review

Elizabeth M. Fitzpatrick, PhD,a,b Candyce Hamel, MSc,c Adrienne Stevens, MSc,c,d Misty Pratt, MES,c David Moher, PhD,c,e Suzanne P. Doucet, MEd,1 Deirdre Neuss, PhD,5,6 Anita Bernstein, MSc,7 Eunjung Na, MSc,a,b

abstract

CONTEXT: Permanent hearing loss affects 1 to 3 per 1000 children and interferes with typical communication development. Early detection through newborn hearing screening and hearing technology provide most children with the option of spoken language acquisition. However, no consensus exists on optimal interventions for spoken language development.

OBJECTIVE: To conduct a systematic review of the effectiveness of early sign and oral language intervention compared with oral language intervention only for children with permanent hearing loss.

DATA SOURCES: An a priori protocol was developed. Electronic databases (eg, Medline, Embase, CINAHL) from 1995 to June 2013 and gray literature sources were searched. Studies in English and French were included.

STUDY SELECTION: Two reviewers screened potentially relevant articles.

DATA EXTRACTION: Outcomes of interest were measures of auditory, vocabulary, language, and speech production skills. All data collection and risk of bias assessments were completed and then verified by a second person. Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) was used to judge the strength of evidence.

RESULTS: Eleven cohort studies met inclusion criteria, of which 8 included only children with severe to profound hearing loss with cochlear implants. Language development was the most frequently reported outcome. Other reported outcomes included speech and speech perception.

LIMITATIONS: Several measures and metrics were reported across studies, and descriptions of interventions were sometimes unclear.

CONCLUSIONS: Very limited, and hence insufficient, high-quality evidence exists to determine whether sign language in combination with oral language is more effective than oral language therapy alone. More research is needed to supplement the evidence base.

Early detection of permanent childhood hearing loss through population-based newborn screening has become standard care in much of the world. Expectations are that early intervention through hearing technology will improve spoken language outcomes for children.\textsuperscript{1,2} Childhood hearing loss is a relatively frequent disorder, affecting 1 to 3 per 1000 live births;\textsuperscript{3–5} that disrupts typical language acquisition, placing children at risk for delays in language, literacy, and social development.\textsuperscript{6–8} There is strong consensus that specialized intervention must be combined with early identification to develop communication skills in these children.\textsuperscript{1,9}

Historically, considerable debate about optimal outcomes for children with hearing loss have resulted in the evolution of a plethora of intervention methods that constitute two broad but distinct philosophies. The oral approach aims to facilitate spoken language and inclusion with normal-hearing peers, and the manual approach focuses on visual communication systems (sign language) and a Deaf culture identity.\textsuperscript{10,11} Although there is wide recognition that various options should be available to families, 2 events in the past two decades (newborn hearing screening and cochlear implant technology) have made it possible for even children with profound deafness to develop spoken language.\textsuperscript{7,12–14} Epidemiologic data confirm that >90% of children with hearing impairment are born to parents with normal hearing.\textsuperscript{15}

Although there is substantial evidence that children with hearing loss can develop oral language skills,\textsuperscript{7,14,16} there is no consensus about optimal interventions.\textsuperscript{17} There is a common expectation that children receiving oral language intervention should develop better language skills than those who are also exposed to sign language, which might disrupt or delay spoken language acquisition. However, another body of research suggests that visual languages such as American Sign Language (ASL) are processed in the brain in the same manner as spoken languages and are complementary to auditory stimulation, and therefore provide a strong foundation for learning oral language.\textsuperscript{18,19} Therefore, adding sign language may develop bilingual skills and facilitate transition to spoken language acquisition. Anecdotal evidence supports various intervention options, but there is little scientifically based consensus. This information is essential to (1) guide families in making decisions about care for their children in infancy, (2) inform clinicians so that they can tailor treatment plans to achieve the desired outcomes, and (3) inform policy makers so that they can make optimal investments in early intervention services.

This review is timely given the worldwide proliferation of screening programs in the past 10 years. In 2008, the US Preventive Services Task Force recommendations on newborn hearing screening highlighted the need for further research to demonstrate the effectiveness of the entire screening-to-intervention process.\textsuperscript{2} New possibilities due to screening and technology have reignited the discussion on best practices for children with hearing loss. Accordingly, the primary purpose of this research was to examine the evidence for the effects of various intervention options for early-identified children with hearing loss when the desired outcome is spoken communication. Our interest was in whether adding sign language facilitates spoken language, because of the increased focus on this outcome since the advent of cochlear implantation. Specifically, the review addressed the following question: Do children with hearing loss have better spoken language outcomes when exposed to early intervention that uses signs to support language compared with oral language intervention without sign language?

\textbf{METHODS}

The protocol and report for this review were prepared according to Preferred Reporting Items for Systematic Reviews and Meta-Analyses for Protocols,\textsuperscript{20} and the protocol was published\textsuperscript{21} and registered at PROSPERO (Registration #CRD42013005426),\textsuperscript{22} an international register of systematic review protocols. The review was undertaken using an integrated knowledge translation approach involving knowledge-user clinicians from health and education in the early stages and as required, to ensure relevance of the project for parents, clinicians, and decision-makers.

\textbf{Literature Search}

Ovid Medline In-Process & Other Non-Indexed Citations and Ovid Medline (1946 to June week 2, 2013), Embase (1974 to June 25, 2013), PsycINFO (1806 to June week 3 2013), and Cochrane CENTRAL (through May 2013 issue) databases were searched using the Ovid interface. CINAHL was searched on June 26, 2013, using the EbscoHost interface, and SpeechBITE was searched on June 26, 2013. The search strategy (Appendix 1) was developed in Medline by an experienced information specialist, peer-reviewed using the Peer Review of Electronic Strategies (PRESS) standard,\textsuperscript{23} and adapted for the other databases. The search was restricted to the pediatric age group,\textsuperscript{24} and the date was limited to material published in 1995 or later. A broad methodological filter was applied, incorporating published filters for controlled trials and other study designs.\textsuperscript{25,26} In addition,
gray literature websites (A.G. Bell Association for the Deaf, National Acoustics Laboratory, National Health Services UK Newborn Hearing Screening Program, CADTH Gray Matters Checklist), conference proceedings (Newborn Hearing Screening Conference Abstracts 2010 and 2012), two journals (Ear and Hearing and Journal of Deaf Studies and Deaf Education), and six books were searched.

Eligibility Criteria

Study Designs

The following study designs were included: randomized controlled trials; controlled clinical trials and other quasi-experimental designs that include comparator groups; and prospective and retrospective cohort studies.

Population

Studies were eligible for inclusion if they included children (1) with early-onset (before age 3 years) hearing loss of any severity using hearing aids and/or cochlear implants; (2) receiving early intervention (age ≤3 years); and (3) <18 years old at assessment.

Based on several studies that categorized participants with cutoffs at the 3-year age level, the early intervention criterion was adjusted from age 2 in our original protocol to age 3. We excluded studies that reported only outcomes for children with hearing loss who had developmental disabilities that interfered with spoken language.

Intervention

We included studies addressing early intervention aimed at spoken language development, which involved (1) oral language intervention and (2) any form of sign language (ASL) or sign support (e.g.,Signing Exact English, baby sign), commonly described under various names (e.g., total communication, simultaneous communication).

Relevancy of papers was assessed based on the components of the intervention described and not on the program label.

The comparator of interest was oral language intervention without sign language.

Outcomes

Primary outcomes included all measures of spoken language including auditory, receptive, and expressive language skills (e.g., vocabulary), speech production, and intelligibility. These outcomes were selected as clinically relevant based on a large body of literature. Secondary outcomes included electrophysiologic outcomes (e.g., cortical responses). In addition, any adverse outcomes (e.g., parent stress) were noted.

Time Frame

We included studies from 1995 onward to capture studies after wide implementation of cochlear implantation and newborn hearing screening. Earlier studies were excluded, as previous generations of children were unlikely to receive the same standards of early intervention and technology.

Language

Articles written in English and French were included.

Study Selection

All records were compiled in a Reference Manager database, checked for duplication, and exported to Distiller SR software for study selection, which involved 2 distinct stages. Screening forms were developed from inclusion and exclusion criteria and calibrated among reviewers with a subset of records before each screening stage. One reviewer assessed titles and abstracts for potential relevance; a second reviewer verified as not relevant all records coded as such. Two independent reviewers screened all potentially relevant full-text articles. Disagreements were resolved by consultation and consensus with a third researcher and knowledge-user clinicians as needed.

Data Extraction

A study-specific data extraction form, finalized with input from knowledge-user clinicians, was developed in Distiller SR to extract predetermined data variables. Items extracted included (1) study characteristics (citation, year, setting, country, language, publication status, and source of funding or other potential conflict of interest), (2) study design, (3) population characteristics (e.g., sample size, gender, ethnicity, etiology, age, severity of hearing loss, hearing technology, and time with hearing technology), (4) details (type) of intervention, (5) details of comparison groups, and (7) outcome data. One researcher extracted data and a second researcher independently verified data. Discrepant and unclear data were resolved through consensus and included a third researcher if required. Outcomes reported only in graph format were estimated using Engauge software.

We contacted study authors for clarifications related to (1) intervention age, (2) intervention method, and (3) metric (e.g., standard score) reported for a test. Responses were received from all authors. No data were imputed for any outcomes.

Study Quality Assessment

Studies were assessed using the Qualitative Assessment Tool for Quantitative Studies, developed by the Effective Public Health Practice Project at McMaster University to assess the quality of studies in a systematic review. The tool, accompanied by a reviewers’ dictionary, provides a methodological rating of studies based on an appraisal of 8 areas: selection
bias, study design, confounders, blinding, data collection methods, withdrawals and dropouts, integrity of intervention, and study analysis. One researcher conducted the quality assessment after 2 studies were first assessed by 2 researchers, and the second verified the remaining assessments. A third researcher with content expertise rechecked all ratings and discussed any differences to reach consensus.

Grading Strength of Evidence

We applied methodology developed by the Grades of Recommendation, Assessment, Development, and Evaluation (GRADE) working group\textsuperscript{38} to rate the quality of the overall body of evidence for each outcome. Primary outcomes were assessed across the domains of risk of bias, consistency, directness, precision, and publication bias. GRADE results in a rating of the quality of the body of evidence as high (very confident that true effect is close to the effect estimate), moderate (moderately confident), low (limited confidence), or very low (little confidence).

Data Synthesis

We summarized the characteristics of study populations and interventions in tabular and narrative form. A meta-analysis was not possible due to heterogeneity in designs, methods, and outcome metrics. A narrative synthesis and table summary of data were therefore completed. We present continuous outcomes as mean differences (MDs) with 95% confidence intervals, where possible. Planned subgroup analyses to examine variables related to hearing loss severity, age of identification, and hearing technologies could not be conducted owing to the lack of studies.

RESULTS

We identified 432 records through database searching and 113 records through other sources, of which 470 were retained for screening after duplicates were removed. Figure 1 outlines the flow of records through the screening process. We assessed 352 full-text documents. As shown, the primary reasons for exclusion were (1) not a study design (58%) and (2) not an intervention of interest (32%). Eleven unique studies were included.

Study Characteristics

Table 1 summarizes the study characteristics. All were prospective cohort studies published between 1999 and 2013, 4 since 2004. The majority (n = 8) were conducted in the United States\textsuperscript{32,39–45} and one each was conducted in Spain\textsuperscript{46} the UK\textsuperscript{47} and Denmark\textsuperscript{48} Three studies\textsuperscript{44,46,47} were conducted specifically to examine whether signs added to spoken language intervention (sign + oral) improved outcomes. The remaining studies evaluated hearing technology or provided data about language intervention as 1 of several predictor variables in the analysis.

Participants and Interventions

Study size for reported outcomes ranged from 13 to 90 participants. It is not possible to rule out overlap.
<table>
<thead>
<tr>
<th>Study, Year, Country</th>
<th>Design</th>
<th>Main Study Objective and Setting</th>
<th>Participants, Age, Hearing Loss, Other Characteristics</th>
<th>Intervention</th>
<th>Comparator</th>
<th>Study Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td>Miyamoto et al, 1999, US</td>
<td>Cohort</td>
<td>To examine speech perception, speech production, and language abilities in children with CI as a function of age at implantation; each age group included some children who used TC and some who used OC; Indiana University CI program</td>
<td>n = 14 for review; NR by group; CI; only children implanted &lt;3 y of age = 14 included in this review; age: TG 4.5 ± 0.3 y vs OC 4.6 ± 0.3 y; HL: profound, 118 vs 109 dBHL; age onset: 0.0 vs 0.2 y; length CI use: 1.8 vs 2.3 y; age at testing: 4.5 vs 4.6 y</td>
<td>TC: combined use of sign and spoken English</td>
<td>OC</td>
<td>Weak</td>
</tr>
<tr>
<td>Robbins et al, 1999, US</td>
<td>Cohort</td>
<td>To examine language development in children who received a CLARION multistrategy CI; to examine language changes postimplant as a function of preoperative communication mode</td>
<td>n = 25, 8 TC, 15 OC; mean age Cl 3 y 2 mo (range 24–65 mo); HL: profound, 116 vs 110 dBHL; implant duration &lt;3 y (extracted for this review)</td>
<td>TC: simultaneous spoken English and Signing Exact English</td>
<td>OC: access to lip-reading in addition to audition in everyday learning environments</td>
<td>Weak</td>
</tr>
<tr>
<td>Osberger et al, 1999, US</td>
<td>Cohort</td>
<td>To examine performance in speech perception in children with 18-mo CI experience who were enrolled in CLARION multistrategy CI trial; data were analyzed as a function of preoperative communication mode</td>
<td>n = 36 for review; 12 TC, 24 OC; CI; only younger group included in this review; followed pre-CI up to 18 mo post-CI; age CI: 3.1 vs 3.6 y; HL: profound, 116 vs 108 dBHL; age onset: 0.5 vs 0.4 y</td>
<td>TC: English-based sign system as well as audition, speech-reading, and speech</td>
<td>OC: audition, speechreading, and speech</td>
<td>Moderate</td>
</tr>
<tr>
<td>Kirk et al, 2000, US</td>
<td>Cohort</td>
<td>To examine the effects of age at CI on communication skills in early-implanted children; to examine the effects of communication mode on speech perception and language skills in early-implanted children; Indiana University CI program</td>
<td>n = 14 for review; 6 TC, 8 OC; CI; only children implanted &lt;2 y of age included in this review; age: 3.99 vs 4.11 y (based on age + duration CI); HL: severe to profound, 103 vs 113 dBHL; age onset: 0.16 vs 0.0 y; age CI: 1.72 vs 1.61 y; duration CI: 2.22 vs 2.30 y</td>
<td>TC: combined use of Signed Exact English and spoken English</td>
<td>OC: does not involve use of sign language</td>
<td>Weak</td>
</tr>
<tr>
<td>Kirk et al, 2002, US</td>
<td>Cohort</td>
<td>To examine effects of age at CI on communication abilities; effect of age at CI was examined separately for children with TC and OC; Indiana University CI program</td>
<td>n = 41 for review; 16 TC, 25 OC (younger group only included in this review); age: NR; tested preimplant and at 6-mo post-CI intervals for ≥3 y; HL: prelingual severe to profound; age CI &lt;5 y; mean age onset: 0.18 y; age CI fitting: 2.13 vs 2.24 y; unaided hearing: 109 vs 111 dBHL</td>
<td>TC: combined use of signed and spoken English</td>
<td>OC</td>
<td>Weak</td>
</tr>
<tr>
<td>Janjua et al, 2002, UK</td>
<td>Cohort</td>
<td>To determine if there are differences in parent–child interaction between children following an aural/ oral versus a bilingual intervention approach</td>
<td>n = 13 were evaluated (16 started study); 9 TC, 4 OC; 48.15% female; age: 9 to 36 mo; HL: 4 severely deaf (70–90 dBHL); 9 profoundly deaf (90+ dBHL); NR by group; families varied in financial status, years of parental education, number of children, marital status, SES (NR by group)</td>
<td>Bilingual: same services as oral/ aural group + BSL optional weekly classes for parents, weekly support of deaf adult in nursery program, weekly sign enrichment group for children in School for the Deaf, optional weekly sign classes</td>
<td>Oral/aural: support from home/nursery visiting teacher; speech training for some; placement in mainstream nursery encouraged; weekly parents’ group with support for spoken language</td>
<td>Weak</td>
</tr>
<tr>
<td>Study, Year, Country</td>
<td>Design</td>
<td>Main Study Objective and Setting</td>
<td>Participants, Age, Hearing Loss, Other Characteristics</td>
<td>Intervention</td>
<td>Comparator</td>
<td>Study Quality</td>
</tr>
<tr>
<td>---------------------</td>
<td>--------</td>
<td>---------------------------------</td>
<td>------------------------------------------------------</td>
<td>--------------</td>
<td>------------</td>
<td>--------------</td>
</tr>
<tr>
<td>Nicholas and Geers, 2003, US&lt;sup&gt;14&lt;/sup&gt;</td>
<td>Cohort</td>
<td>To examine whether deaf children who have the addition of signs show an early linguistic advantage and greater communicative function maturity than those with spoken language input only; to examine whether deaf children who receive speech and sign lag behind children in OC in their use of speech in preschool</td>
<td>n = 78, 38 SC, 38 OC (OC from previously published study); age: 18–54 mo; age tested: 18–54 mo; group scores reported at 18–30 mo (n = 16 per intervention group) and 36–54 mo (n = 22 per intervention group); HL: better-ear PTA threshold ≥80 dBHL, 103.21 vs 104.41 dBHL; CI: 27 vs 30; CI: 7 vs 5; tactile aid: 0 vs 3; no device: 4 vs 0; age onset: birth (n = 34 vs 33) and 1–18 mo (n = 4 vs 5); age diagnosis: 10.42 vs 12.03 mo; maternal education: 14.14 vs 14.71 y; duration CI: 7.5 (n = 7) vs 3.8 mo (n = 5)</td>
<td>SC: emphasizes the combination of spoken language with corresponding manual signs</td>
<td>OC: emphasizes the auditory signal with or without visual cues from speaker’s lips</td>
<td>Moderate</td>
</tr>
<tr>
<td>Connor and Zwolan, 2004, US&lt;sup&gt;15&lt;/sup&gt;</td>
<td>Cohort</td>
<td>To examine the effects of multiple variables (SDT, age, SES, communication method, pre- and postimplant vocabulary) on reading comprehension&lt;sup&gt;3&lt;/sup&gt; (pre- and postimplant vocabulary scores were reported by group)</td>
<td>n = 91, 43 TC, 48 OC; 49.45% male; CI: age: 11.16 ± 2.83 vs. 10.81 ± 2.53 y (for reading evaluation); age pre–post vocabulary tests: NR; HL: &gt;80 dBHL; age onset: 0.12 vs 0.19 y; age CI: 7.40 vs 6.23 y; duration CI: 3.82 ± 1.56 vs. 4.63 ± 2.51 y; HL:SDT: 55.22 vs 51.72</td>
<td>TC: use of sign language in combination with spoken language, sign systems based on English, including Signed English and Signing Exact English</td>
<td>OC: use of spoken language only</td>
<td>Weak</td>
</tr>
<tr>
<td>Jiménez et al, 2009, Spain&lt;sup&gt;16&lt;/sup&gt;</td>
<td>Cohort</td>
<td>To compare speech/language development after CI in children educated using spoken language versus spoken and sign language</td>
<td>n = 18, NR by group; 61.11% male; CI: age: mean 6.25 y, range 4 y 3 mo to 8 y; HL: profound; groups similar for gender; age (avg 6.25 y), age diagnosis (avg 10 mo); CI (avg 3.2 y), duration CI (avg 3.1 y) (P &gt; .05 for all)</td>
<td>Bilingual: spoken + sign language</td>
<td>Monolingual: spoken language</td>
<td>Moderate</td>
</tr>
<tr>
<td>Nittrouer, 2010, US&lt;sup&gt;17&lt;/sup&gt;</td>
<td>Cohort</td>
<td>To explain the contributions of various independent measures to development outcomes for children with hearing loss, 1 of which was sign support</td>
<td>n = 118, 44 sign support; 74 spoken language; 55% male, early-identified (&lt;6 mo) n = 29 vs 48; late-identified (&gt;6 mo) n = 15 vs 26; age: 12–48 mo; tested at 6-mo intervals, n varied per interval: 16 (12 mo) to 90 (48 mo); HL: moderate to profound, 38 HA, 80 CI by end of study; groups similar for gender, SES, HL severity, age diagnosis, age first hearing technology/intervention</td>
<td>Spoken + sign support; 28 of 44 parents used ASL and 16 a manually coded English system; parents reported using signs &lt;50% of time</td>
<td>Spoken language</td>
<td>Moderate</td>
</tr>
<tr>
<td>Percy-Smith et al, 2013, Denmark&lt;sup&gt;18&lt;/sup&gt;</td>
<td>Cohort</td>
<td>To examine language understanding and receptive and active vocabulary in Danish children with CI and to identify factors associated with these outcomes; examined the effects of 18 factors including communication mode on outcomes; all children born in Denmark January 2005 to January 2011 who had 6-mo experience with CI</td>
<td>n = 83, 12 spoken + sign, 70 spoken (1 missing info); different n for different outcomes; n by group for different outcomes; NR with additional disabilities but unclear if included in analysis; age: 46.3 mo (range 17–74 mo); PPVT (n = 68), and Reynell (n = 71) at ≥24 mo; Viborgmaterialet test (n = 49) at ≥35 mo; HL: CI or CI + HA (88% with bilateral CI); mean age HA 8 mo (range 3–36 mo), mean age CI 19.6 mo (range 5–55 mo); mean duration CI 25.8 mo (range 6–59 mo), NR by intervention group</td>
<td>Spoken + sign language</td>
<td>Spoken language</td>
<td>Weak</td>
</tr>
</tbody>
</table>

<sup>3</sup> BSIL, British Sign Language; CI, cochlear implant; dBHL, decibel hearing level; HA, hearing aid; HL, hearing loss; NR, not reported; OC, oral communication; PTA, pure-tone average; SC, simultaneous communication; SDT, speech detection threshold; SES, socioeconomic status; TC, total communication.
<sup>4</sup> Vocabulary and language tests administered in child’s preferred modality, TC or OC.
<sup>5</sup> Authors stated that test administration procedures were followed but children were permitted to answer using sign and/or spoken language.
of participants in some studies; for example, 339,42,43 from 1999 to 2002 included a subset of children from 1 US cochlear implant program. Eight studies included only children with severe to profound deafness who had received cochlear implants. The 3 remaining studies included a mix of children with respect to technology use.32,44,47 As shown in Table 1, authors used various descriptions and terms to describe both the intervention (eg, sign, bilingual) and comparator (oral, spoken language) groups.

Quality Assessment

Study quality and scores are shown in Table 1. Studies were rated as moderate (n = 4) or weak (n = 7) quality. Applying the McMaster Quality Assessment ratings,37 an overall weak rating is assigned if ≥2 areas are rated as weak. It is important to note that in these studies it was not possible to blind the examiners to the intervention, as children’s communication mode is generally reflective of intervention method; therefore, lack of blinding resulted in a weak rating as per the tool’s criteria. Other primary reasons contributing to an overall weak quality rating included limited information on confounders or data collection methods, which involved sign language adaptation of standardized test administration such that responses may not entirely reflect spoken language abilities. As noted, most included studies were not specifically designed to answer the review question. Consequently, the different focus for these studies may have influenced their quality ratings; that is, studies may have included fewer details about intervention methods, as it was not their primary question.

Outcomes

For reporting purposes, outcomes have been categorized into key domains (eg, language, speech) commonly used in auditory-language therapy (Table 2). Most studies (n = 10) reported multiple outcomes of interest, whereas 1 study41 reported speech perception only (Table 3). As shown in Table 2, different metrics (eg, standard score, raw score, age equivalency, language quotient) were reported in different studies. Studies also reported variations in test administration. In 1 study, Spanish test versions (which authors reported as validated) were administered.46 Danish test adaptations48 were used with American norms in another study except for the expressive vocabulary test, which was developed and standardized in Danish. As noted in Table 1, tests were administered in the child’s preferred modality (sign or spoken language) in 4 studies39,40,42,43 or with adaptations (for 1 test) in 1 study.32 and responses were accepted in sign and/or spoken language in another.45 These adjustments in test procedures can potentially affect the psychometric properties of the tests, which were standardized on children with normal hearing, and hence interpretation of the results. No study reported any adverse effects of the intervention. No electrophysiologic (secondary) outcomes were reported.

Language Results: Vocabulary

Receptive vocabulary was reported in 4 studies using the Peabody Picture Vocabulary Test (PPVT), a norm-referenced test for ages 2.6 to 90+ years.49 All 4 studies included only children with cochlear implants.39,42,46,48 The 2 US studies39,42 were conducted in 1999 to 2000 and therefore included children implanted in the early years of cochlear implant availability. These studies reported no significant effect of intervention (communication mode). The third study46 reported standard scores from a Spanish version of the test, with no statistical difference between groups. One additional Danish study48 used a translated version of the PPVT with American norms. In that study, authors did not report scores separately by group but concluded that not using total communication was 1 factor associated with greater odds of age-equivalent receptive vocabulary.

Three studies also reported results for expressive vocabulary. The Danish study48 examined expressive vocabulary using a Danish norm-referenced test but found no effect of intervention mode. The US preschool study42 found that expressive vocabulary results favored oral intervention only for late-identified children (>12 months) and showed no significant effects of intervention for early-identified children (<6 months). The late-identified group consisted of 12 children exposed to sign and 26 with no sign exposure (oral). One US cochlear implant study,45 whose primary goal was to investigate reading comprehension, also reported pre- and postimplant expressive vocabulary scores. The authors reported that sign + oral intervention was significantly associated with better vocabulary outcomes preimplant. However, as noted in Table 1, children in the study responded using sign and/or spoken words; therefore, this finding may not represent a measurement of spoken vocabulary. Postimplant vocabulary standard scores (74.60 vs 67.12) for both groups were well below the average for normal-hearing peers but slightly favored the sign + oral group. However, the authors did not state whether the difference was statistically significant.

Language Results: Receptive and Expressive Language

The most commonly reported outcomes for receptive and expressive language (5 studies) were from the Reynell Developmental Language Scales, a norm-referenced test for children ages 1 to 6 years. Four studies39,40,42,43 reported results for US children who
<table>
<thead>
<tr>
<th>Outcome Measure, Test, and Length of Follow-Up or Test Age&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sign + Oral vs Oral Only&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Studies (Participants)</th>
<th>Effect Estimate&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Statistical Information Provided by Author&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Quality of Evidence&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language</strong>&lt;sup&gt;i&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PPVT&lt;sup&gt;39,42,46&lt;/sup&gt;</td>
<td>In 2 studies, mean LQ 0.30–0.72 vs 0.32–0.62</td>
<td>2 (28)</td>
<td>MD (no CI): −0.02 to 0.10</td>
<td>NS for both studies</td>
<td>Very low</td>
</tr>
<tr>
<td>Mean 5.4 y&lt;sup&gt;39&lt;/sup&gt;</td>
<td>Mean standard score 66.9 vs 73.6</td>
<td>1 (18)</td>
<td>MD (no CI): −6.7</td>
<td>NS (P &gt; .05)</td>
<td>Very low</td>
</tr>
<tr>
<td>Mean 3.2 y&lt;sup&gt;39&lt;/sup&gt; (range 6 to &gt;36 mo)</td>
<td>Scores NR by group; reported as odds of having age-equivalent vocabulary</td>
<td>1 (68)</td>
<td>OR estimated to be infinity by authors</td>
<td>P = .012 (favors oral)</td>
<td>Very low</td>
</tr>
<tr>
<td>Viborgmaterialet&lt;sup&gt;39&lt;/sup&gt; (Danish vocabulary)</td>
<td>Scores NR by group; reported as odds of having age-equivalent vocabulary</td>
<td>1 (49)</td>
<td>OR not reported</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>Mean 3.2 y&lt;sup&gt;39&lt;/sup&gt; (range 6 to &gt;36 mo)</td>
<td>Raw score: 32.89 ± 8.76 vs 32.18 ± 10.81 (early id); 21.50 ± 13.40 vs 32.12 ± 7.86 (late id)</td>
<td>1 (early: 52; late: 38)</td>
<td>MD-early (95% confidence interval): 0.71 (−4.72 to 6.14); MD-late: −10.62 (−18.78 to −2.45)</td>
<td>Early id: NS; late id: P = .002 (favors oral)</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>Picture Vocabulary Test/EOWPVT&lt;sup&gt;32,45&lt;/sup&gt;</strong></td>
<td>Pre- and post implant scores 65.54–74.60 vs 53.08–67.12</td>
<td>1 (91)</td>
<td>MD (no CI): 7.48 to 13.46</td>
<td>NS preimplant (P &gt; .05); postimplant NR</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>EOWPVT&lt;sup&gt;32&lt;/sup&gt;</strong></td>
<td>Raw score: 32.89 ± 8.76 vs 32.18 ± 10.81 (early id); 21.50 ± 13.40 vs 32.12 ± 7.86 (late id)</td>
<td>1 (early: 52; late: 38)</td>
<td>MD-early (95% confidence interval): 0.71 (−4.72 to 6.14); MD-late: −10.62 (−18.78 to −2.45)</td>
<td>Early id: NS; late id: P = .002 (favors oral)</td>
<td>Very low</td>
</tr>
<tr>
<td>Reynell-R&lt;sup&gt;39,40,42,43,48&lt;/sup&gt;</td>
<td>Mean LQ 0.61–0.89 vs 0.64–0.91</td>
<td>3 (69)</td>
<td>MD (no CI): −0.02 to −0.03</td>
<td>NS for all studies</td>
<td>Very low</td>
</tr>
<tr>
<td>Reynell-E&lt;sup&gt;39,40,42,43&lt;/sup&gt;</td>
<td>Mean 8- vs 9-mo increase in language age (no CIs)</td>
<td>1 (23)</td>
<td>MD (no CI): −1-mo language age</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>Mean 3.2 y, unknown by group&lt;sup&gt;40&lt;/sup&gt;</td>
<td>Scores NR by group; odds of age-equivalent language</td>
<td>1 (72)</td>
<td>OR estimated to be infinity by authors</td>
<td>P = .013 (favors oral)</td>
<td>Very low</td>
</tr>
<tr>
<td>Reynell-L&lt;sup&gt;39,40,42,43,46&lt;/sup&gt;</td>
<td>Mean LQ 0.51–0.68 vs 0.69–0.93</td>
<td>3 (68)</td>
<td>MD (no CI): −0.02 to −0.26</td>
<td>1 NS&lt;sup&gt;41&lt;/sup&gt;; 1 NR&lt;sup&gt;42&lt;/sup&gt;; 1 significant&lt;sup&gt;38&lt;/sup&gt;</td>
<td>Very low</td>
</tr>
<tr>
<td>Unknown&lt;sup&gt;39,42,43&lt;/sup&gt;</td>
<td>Mean 7- vs 10-mo increase in language age</td>
<td>1 (25)</td>
<td>MD (no CI): −3-mo language age</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>ITPA (Spanish) Auditory Reception&lt;sup&gt;46&lt;/sup&gt;</strong></td>
<td>Mean standard score 34.8 vs 56.0</td>
<td>1 (18)</td>
<td>MD (no CI): −21.4</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>ITPA (Spanish) Auditory Association&lt;sup&gt;46&lt;/sup&gt;</strong></td>
<td>Mean standard score 37.8 vs 71.6</td>
<td>1 (18)</td>
<td>MD (no CI): −33.8</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>ITPA (Spanish) Auditory sequential memory&lt;sup&gt;46&lt;/sup&gt;</strong></td>
<td>Mean standard score 44.8 vs 59.0</td>
<td>1 (18)</td>
<td>MD (no CI): −14.2</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>ITPA (Spanish) Verbal Expression&lt;sup&gt;46&lt;/sup&gt;</strong></td>
<td>Mean standard score 52.5 vs 39.2</td>
<td>1 (18)</td>
<td>MD (no CI): 13.3</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>ITPA (Spanish) Grammatical closure&lt;sup&gt;46&lt;/sup&gt;</strong></td>
<td>Mean standard score 39.2 vs 71.3</td>
<td>1 (18)</td>
<td>MD (no CI): −32.1</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td><strong>ICAP (Spanish) Social communication&lt;sup&gt;46&lt;/sup&gt;</strong></td>
<td>Mean standard score 38.75 vs 53.0</td>
<td>1 (18)</td>
<td>MD (no CI): −14.25</td>
<td>NS</td>
<td>Very low</td>
</tr>
</tbody>
</table>

<sup>a</sup>Outcome Measure, Test, and Length of Follow-Up or Test Age.<br><sup>b</sup>Sign + Oral vs Oral Only.<br><sup>c</sup>Effect Estimate.<br><sup>d</sup>Statistical Information Provided by Author.<br><sup>e</sup>Quality of Evidence.
<table>
<thead>
<tr>
<th>Outcome Measure, Test, and Length of Follow-Up or Test Age&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Sign + Oral vs Oral Only&lt;sup&gt;b&lt;/sup&gt;</th>
<th>Studies (Participants)</th>
<th>Effect Estimate&lt;sup&gt;c&lt;/sup&gt;</th>
<th>Statistical Information Provided by Author&lt;sup&gt;d&lt;/sup&gt;</th>
<th>Quality of Evidence&lt;sup&gt;e&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Language Comprehension–SIB-R&lt;sup&gt;32j&lt;/sup&gt;  &lt;br&gt;3.5 y (early id), 2–3 y (late id)&lt;sup&gt;k&lt;/sup&gt;; test age: 48 mo  &lt;br&gt;Mean raw score: 45.39 ± 8.82 vs 42.59 ± 10.02 (early id); 35.92 ± 12.80 vs 43.73 ± 8.59 (late id)</td>
<td></td>
<td>1 (early: 52; late: 38)</td>
<td>MD early id (95% confidence interval): −2.80 (−5.48 to 5.08); MD late id: −7.81 (−15.76 to 0.14)</td>
<td>Early id: NS; P = .011 (favors oral)</td>
<td>Very low</td>
</tr>
<tr>
<td>Language Expression - SIB-R&lt;sup&gt;32j&lt;/sup&gt;  &lt;br&gt;3.5 y (early id), 2–3 y (late id)&lt;sup&gt;k&lt;/sup&gt;; test age: 48 mo  &lt;br&gt;Mean raw score: 28.56 ± 1.89 vs 27.03 ± 3.59 (early id); 23.42 ± 6.80 vs 28.86 ± 3.19 (late id)</td>
<td></td>
<td>1 (early: 52; late: 38)</td>
<td>MD early id (95% confidence interval): 1.53 (0.04 to 3.01); MD late id: −3.48 (−7.49 to 0.57)</td>
<td>Early id: similar number of words; late id: sign had fewer words (author report)</td>
<td>Very low</td>
</tr>
<tr>
<td>Number of spoken words-LDS&lt;sup&gt;32j&lt;/sup&gt;  &lt;br&gt;3.5 y - early-id; 2-3 y - late-id group&lt;sup&gt;k&lt;/sup&gt;; test age: 30 mo  &lt;br&gt;Mean: 185.79 ± 103.19 vs 157.08 ± 104.56 (early id); 38.75 ± 29.64 vs 122.25 ± 69.34 (late id)</td>
<td></td>
<td>1 (early: 51 late: 16)</td>
<td>MD early id (95% confidence interval): 28.71 (0.04 to 3.01); MD late id: −83.5 (−132.31 to −34.68)</td>
<td>Early id: NS; late id: P = .012 (favors oral)</td>
<td>Very low</td>
</tr>
<tr>
<td>BLADES&lt;sup&gt;47&lt;/sup&gt;  &lt;br&gt;Unknown  &lt;br&gt;Delay 6–24 vs 12–24 mo</td>
<td></td>
<td>1 (13)</td>
<td>See next column</td>
<td>No difference between groups (author report)</td>
<td>Very low</td>
</tr>
<tr>
<td>Number of different words (sign or spoken)&lt;sup&gt;64&lt;/sup&gt;  &lt;br&gt;Range&lt;sup&gt;1&lt;/sup&gt; 7.6–19.6 mo (sign + oral), 6–18 mo (oral); test age 19–30 mo  &lt;br&gt;Mean 13.88 ± 17.55 vs 2.65 ± 5.87 words</td>
<td></td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 11.23 (5.35 to 17.11)</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>&lt;br&gt;Range&lt;sup&gt;1&lt;/sup&gt; 25.6–43.6 (sign + oral), 24–42 mo (oral); test age: 36–54 mo  &lt;br&gt;Mean 45.50 ± 29.08 vs 38.90 ± 3.96 words</td>
<td></td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 6.60 (−2.73 to 15.93)</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>Words per utterance (sign or spoken)&lt;sup&gt;64&lt;/sup&gt;  &lt;br&gt;Test age 18–30 mo  &lt;br&gt;Mean 0.99 ± 0.44 vs 0.31 ± 0.5 words</td>
<td></td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 0.68 (0.47 to 0.89)</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td>Test age 36–54 mo  &lt;br&gt;Mean 1.25 ± 0.18 vs 1.33 ± 0.36 words</td>
<td></td>
<td>1 (76)</td>
<td>MD (95% confidence interval): −0.08 (−0.21 to 0.05)</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>Total number of spoken words&lt;sup&gt;64&lt;/sup&gt;  &lt;br&gt;Test age 18–30 mo  &lt;br&gt;Mean 3.56 ± 5.19 vs 5.47 ± 14.1 words</td>
<td></td>
<td>1 (76)</td>
<td>MD (95% confidence interval): −1.91 (−6.68 to 2.87)</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>Test age 36–54 mo  &lt;br&gt;Mean 23.09 ± 33.03 vs 117.88 ± 207.75 words</td>
<td></td>
<td>1 (76)</td>
<td>MD (95% confidence interval): −94.79 (−161.67 to −27.91)</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>Outcome Measure, Test, and Length of Follow-Up or Test Age&lt;sup&gt;a&lt;/sup&gt;</td>
<td>Sign + Oral vs Oral Only&lt;sup&gt;b&lt;/sup&gt;</td>
<td>Studies (Participants)</td>
<td>Effect Estimate&lt;sup&gt;c&lt;/sup&gt;</td>
<td>Statistical Information Provided by Author&lt;sup&gt;d&lt;/sup&gt;</td>
<td>Quality of Evidence&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>---------------------------------------------------------------</td>
<td>---------------------------------</td>
<td>------------------------</td>
<td>-----------------------------</td>
<td>-----------------------------------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td>Number of different words spoken&lt;sup&gt;f&lt;/sup&gt;</td>
<td>Mean 2.62 ± 3.59 vs 1.94 ± 4.29</td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 0.68 (−1.10 to 2.46)</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>Test age 18–30 mo</td>
<td>Mean 12.91 ± 17.38 vs 38.67 ± 44 words</td>
<td>1 (76)</td>
<td>MD (95% confidence interval): −25.76 (−40.80 to −10.72)</td>
<td>Significant</td>
<td></td>
</tr>
<tr>
<td>Test age 36–54 mo</td>
<td>Mean 0.23 ± 0.2 vs 0.04 ± 0.12 words</td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 0.19 (0.12 to 0.21)</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td>Test age 36–54 mo</td>
<td>Mean 0.46 ± 0.17 vs mean 0.34 ± 0.21 words</td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 0.12 (0.03 to 0.21)</td>
<td>Significant</td>
<td></td>
</tr>
<tr>
<td>PIHF&lt;sup&gt;g&lt;/sup&gt;</td>
<td>Mean 41.56 ± 35.71 vs 36.76 ± 35.68 words</td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 4.80 (−11.27 to 20.87)</td>
<td>NS</td>
<td>Very low</td>
</tr>
<tr>
<td>Test age 36–54 mo</td>
<td>Mean 121 ± 75.25 vs 131.67 ± 70.79 words</td>
<td>1 (76)</td>
<td>MD (95% confidence interval): −10.67 (−43.52 to 22.18)</td>
<td>NS</td>
<td></td>
</tr>
<tr>
<td>Test age 18–30 mo</td>
<td>Mean 0.23 ± 0.2 vs 0.04 ± 0.12 words</td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 0.19 (0.12 to 0.21)</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td>Test age 36–54 mo</td>
<td>Mean 0.46 ± 0.17 vs mean 0.34 ± 0.21 words</td>
<td>1 (76)</td>
<td>MD (95% confidence interval): 0.12 (0.03 to 0.21)</td>
<td>Significant</td>
<td></td>
</tr>
<tr>
<td>Speech Perception Mr. Potato Head - Words&lt;sup&gt;h,i&lt;/sup&gt;</td>
<td>Mean % correct 24%–58% vs 33%–95% (no CI)</td>
<td>4 (105)</td>
<td>MD (no CI): −9 to −39%</td>
<td>1 NS&lt;sup&gt;i&lt;/sup&gt;, 1 NR&lt;sup&gt;j&lt;/sup&gt;, 2 significant&lt;sup&gt;k&lt;/sup&gt;</td>
<td>Very low</td>
</tr>
<tr>
<td>Speech Perception Mr. Potato Head – Sentences&lt;sup&gt;i&lt;/sup&gt;</td>
<td>Percent correct 39.76% vs 68.88%</td>
<td>1 (14)</td>
<td>MD (no CI): −30.12</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td>Speech Perception ESP-Low Verbal&lt;sup&gt;i&lt;/sup&gt;</td>
<td>Percent correct 82% vs 90%</td>
<td>1 (36)</td>
<td>MD (no CI): −8%</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td>Speech Perception GAEL-P Words recognized&lt;sup&gt;i&lt;/sup&gt;</td>
<td>Percent correct 46%–87.38% vs 53%–96.68%</td>
<td>2 (28)</td>
<td>MD (no CI): −7 to −9.3%</td>
<td>Significant&lt;sup&gt;k&lt;/sup&gt;</td>
<td>Very low</td>
</tr>
</tbody>
</table>

<sup>a</sup> Outcome Measure, Test, and Length of Follow-Up or Test Age<sup>a</sup>  
<sup>b</sup> Sign + Oral vs Oral Only<sup>b</sup>  
<sup>c</sup> Effect Estimate<sup>c</sup>  
<sup>d</sup> Statistical Information Provided by Author<sup>d</sup>  
<sup>e</sup> Quality of Evidence<sup>e</sup>
received cochlear implants before 2002. According to the authors, there were no significant effects of intervention compared with controls (3 studies reported language quotients, and 1 study, language age scores). Only the Danish study using a Danish adaptation with American norms of the Reynell receptive subtest reported spoken language intervention to be statistically significant, for language comprehension only. The same 4 US cochlear implant studies also reported similar findings for expressive language measured by the Reynell test. The Danish study did not report expressive language.

Three additional studies did not report expressive language, for expressive language measured for a variety of other receptive and expressive language measures, all yielding different results. As shown in Table 1, a US study did not report expressive language, for expressive language measured with the Reynell language intervention. Only 1 study that favored spoken language intervention reached statistical significance. The study produced mixed findings with respect to expressive language.

Various subtest scores were also reported in the study of 18 children. The authors reported significantly higher standard scores in auditory reception, auditory association, and grammatical closure for the spoken language-only group. Mean scores were significantly higher for the sign language-only group on the verbal expression subtest. The authors also reported a significant difference in interventions on the Reynell receptive subtest.

**TABLE 2**

<table>
<thead>
<tr>
<th>Outcome Measure, Test, and Length of Follow-Up</th>
<th>Sign + Oral vs Oral Only</th>
<th>Studies (Participants)</th>
<th>Effect Estimate</th>
<th>Statistical Information Provided by Author</th>
<th>Quality of Evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unknown</td>
<td>Percent correct 50% vs 83%</td>
<td>1 (38)</td>
<td>MD (no CI): −13%</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td>PBK-Words</td>
<td>Percent correct 20% vs 24%</td>
<td>1 (38)</td>
<td>MD (no CI): −4%</td>
<td>Significant</td>
<td>Very low</td>
</tr>
<tr>
<td>PBK-Phonemes</td>
<td>Percent correct 28% vs 50%</td>
<td>1 (38)</td>
<td>MD (no CI): −21%</td>
<td>Significant</td>
<td>Very low</td>
</tr>
</tbody>
</table>

BIT: Beginner's Intelligibility Test; BLADES, Bristol Language Developmental Scales; CI, confidence interval; CSIM, Children's Speech Intelligibility Measure; EOWPVT, Expressive One Word Picture Vocabulary Test; ESP, Early Speech Perception; GAIL, Grammatical Analysis of Elicited Language (used as a closed-set speech perception measure); GASP, Glendale Auditory Screening Procedure; ICA, intentionally communicative act; ICAP, Inventory for Client and Agency Planning; ITPA, Illinois Test of Psycho-linguistic Abilities; LDS, Language Development Survey; i/d, identified; LQ, language quotient; MD, mean difference; NR, not reported; NS, not statistically significant; OR, odds ratio; PBK, Phonetically Balanced Kindergarten Test; PIHF, Proportion of Informative or Heuristic Functions; PLS, Preschool Language Scale; PPVT, Peabody Picture Vocabulary Test; SIB-R, Scales of Independent Behavior-Revised

Follow-up measured from time of intervention.

Outcomes data: higher score = better performance.

Minus sign (−) indicates higher score in oral group.

No meta-analysis due to insufficient data and different metrics reported; statistics reported are those from authors.

Based on GRADE ratings.

Length of follow-up (from diagnosis of hearing loss) with language intervention of interest is unknown. Studies report duration of cochlear implant use and age at assessment (details in Table 1).

Authors stated that language measures were adapted and administered in child's preferred communication modality. Signed and spoken responses accepted; or instructions given nonverbally and certain vocabulary considered less likely to be known by children using sign was removed.

Follow-up measured from time of intervention.

Outcomes data: higher score = better performance.

Minus sign (−) indicates higher score in oral group.

No meta-analysis due to insufficient data and different metrics reported; statistics reported are those from authors.

Based on GRADE ratings.

Length of follow-up (from diagnosis of hearing loss) with language intervention of interest is unknown. Studies report duration of cochlear implant use and age at assessment (details in Table 1).

Authors stated that language measures were adapted and administered in child's preferred communication modality. Signed and spoken responses accepted; or instructions given nonverbally and certain vocabulary considered less likely to be known by children using sign was removed.

Follow-up measured from time of intervention.

Outcomes data: higher score = better performance.

Minus sign (−) indicates higher score in oral group.

No meta-analysis due to insufficient data and different metrics reported; statistics reported are those from authors.

Based on GRADE ratings.
developmental scales to measure expressive language in the areas of pragmatics, semantics, and syntax. The authors reported delays ranging from 6 to 24 months in both groups and concluded that intervention did not influence outcomes.

**Language Results: Natural Language Sample Analysis**

Finally, 2 studies reported scores for various aspects of communication development. A study by Nicholas and Geers reported findings for 2 age groups, 18 to 30 months and 36 to 54 months, for 38 children who received sign + oral intervention and compared results to a historical cohort of 38 children in oral language intervention. As shown in Table 2, of 12 different results, statistically significant differences were reported only for words per utterance (sign or spoken words), which favored the sign + oral group at 18 to 30 months; number of different words spoken, which favored the oral group at 36 to 54 months; and the Proportion of Informative or Heuristic Functions, which favored sign + oral in both age groups. Of the 19 natural language results reported in the Nittrouer study, 18 showed no statistically significant difference for early-identified children (18 of 51 had sign + oral), whereas 1 measure (number

### TABLE 3 Overview of Outcome Measures for All Included Studies

<table>
<thead>
<tr>
<th>Measure/tool</th>
<th>Study, reference no.</th>
<th>Sign + Oral</th>
<th>Oral</th>
</tr>
</thead>
<tbody>
<tr>
<td>PPVT(^1)</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Viborg materialet(^4)</td>
<td></td>
<td></td>
<td>\ *\</td>
</tr>
<tr>
<td>PVT/EOWPVT</td>
<td>1</td>
<td></td>
<td>1(^b)</td>
</tr>
<tr>
<td>Reynell Receptive</td>
<td>18–30 mo</td>
<td></td>
<td>\ *\</td>
</tr>
<tr>
<td>Reynell Expressive</td>
<td>36–54 mo</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ITPA(^2) Auditory Reception</td>
<td>18–30 mo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITPA Auditory Association</td>
<td>36–54 mo</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ITPA Auditory Sequential Memory</td>
<td>18–30 mo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ITPA Verbal Expression</td>
<td>36–54 mo</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ITPA Grammatical Closure</td>
<td>18–30 mo</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ICAP</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PLS-4 Auditory Comprehension</td>
<td>18–30 mo</td>
<td></td>
<td>1(^b)</td>
</tr>
<tr>
<td>Language Comprehension SIB-R</td>
<td>36–54 mo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Language Comprehension SIB-R</td>
<td>18–30 mo</td>
<td></td>
<td>1(^b)</td>
</tr>
<tr>
<td>BLADES Expressive</td>
<td>18–30 mo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of different words(^d)</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Words per utterance(^d)</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Total number of spoken words(^d)</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Number of different words spoken(^d)</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>PHP(^3)</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Intentionally communicative acts(^d)</td>
<td></td>
<td></td>
<td>7(^f)</td>
</tr>
<tr>
<td>Language-11 communication acts(^d)</td>
<td></td>
<td></td>
<td>5(^b)</td>
</tr>
<tr>
<td>Language-5 grammar/syntax(^d)</td>
<td></td>
<td></td>
<td>2(^d)</td>
</tr>
<tr>
<td>Language-3 vocalizations(^d)</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>Beginner’s Intelligibility Test</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>CSIM</td>
<td></td>
<td></td>
<td>1(^b)</td>
</tr>
<tr>
<td>Mr. Potato Head—Words</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>Mr. Potato Head—Sentences</td>
<td></td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>ESP-Low Verbal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GAEL-P Words(^d)</td>
<td></td>
<td></td>
<td>2</td>
</tr>
<tr>
<td>GASP—Words</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBK—Words</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PBK—Phonemes</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>4</td>
</tr>
</tbody>
</table>

Numbers in boxes refer to number of studies reporting statistical significance for sign + oral or oral group. BLADES, Bristol Language Developmental Scales; CSIM, Children’s Speech Intelligibility Measure; ESP, Early Speech Perception; EOWPVT, Expressive One Word Picture Vocabulary Test; GAEL-P, Grammatical Analysis of Expressive Language: Presentense Level; GASP, Glendonald Auditory Screening Procedure; ICAP, Inventory for Client and Agency Planning; Lang, language; PBK, Phonetically Balanced Kindergarten Test; PLS-4, Preschool Language Scale; PVT, Picture Vocabulary Test of the Woodcock Johnson Test of Cognitive Ability; SIB-R, Scales of Independent Behavior-Revised.

\(^a\) Danish version of test.
\(^b\) Nittrouer reported significant difference for late-identified sign group only; results for early-identified group were not significant.
\(^c\) Spanish version of ITPA for all subtests.
\(^d\) Results for 2 age groups (18–30 mo and 36–54 mo for each measure in this study).
\(^e\) Various aspects of communication/language reported from analyses of language samples (e.g., number of responses, total words, number of pronouns, jargon).
\(^f\) Both early- and late-identified groups showed significant difference for 1 communication act (number of inquiries); 6 others showed significance for late-identified group only.
\(^g\) GAEL-P administered as closed-set speech perception test.
of inquiries made by child) favored oral intervention. However, results were markedly different for the late-identified children (12 of 37 had sign + oral), in whom 14 of 19 results favored oral intervention.

We retrieved only 1 study that addressed other aspects of language functioning considered relevant to this review. Using the Inventory for Client and Agency Planning, Jiménez et al46 reported no significant differences between groups in social communication skills.

Speech Results
Three studies addressed aspects of speech production, 2 related to speech intelligibility32,39 and 1 to phonological production.46 In 2 studies, scores were statistically better for the oral intervention groups based on author reports, whereas in the US preschool study,32 scores were significantly better for oral intervention in the late-identified but not early-identified children.

Speech Perception
Four cochlear implant studies39,41–43 (1999 to 2002) specifically investigated speech perception and included 3 closed-set measures (response selected from defined word set) and 2 open-set measures (no word set available). All scores were significantly better for the oral group, with the exception of 1 closed-set test (Mr. Potato Head, words) in 141 (n = 36) of the 4 studies, which was not statistically significant. However, for open-set monosyllabic words, the differences were clinically small despite statistical significance (eg, 20% vs 24%). No speech recognition measures were reported for children with hearing aids.

Table 3 provides an overview of measures extracted in the included studies and the interventions that showed statistically significant differences. Of 61 total test scores reported, only 4 showed statistically significant results for sign + oral intervention and 36 for oral language only (9 related to speech perception, 17 for late-identified group in US preschool study32).

DISCUSSION
There continue to be questions about the advantages or disadvantages of adding sign language to interventions focused on spoken language development in children with hearing loss. This review found that few studies have systematically addressed the issue. Eight of the 11 studies in this review included only children with severe to profound deafness who were using cochlear implants. This is likely because many children received preimplant intervention with a signing component in addition to spoken language, owing to the profound nature of hearing loss, particularly in the 1990s when cochlear implantation was a new option. Evidence to address the issue for children with less than profound degrees of hearing loss is severely lacking, as this review retrieved only 3 studies32,44,47 that included children with hearing loss using hearing aids.

Measures of language development, including receptive vocabulary and receptive and expressive language skills, contributed the most information to this review. When comparing sign and oral language intervention with oral language intervention alone, individual studies showed no differences between groups for receptive and expressive vocabulary outcomes, with the exception of 1 study’s32 results that favored oral intervention only for late-identified children (n = 12 in sign group). However, a variety of measures and metrics were used across studies that limit the generalizability of the evidence. Overall, mixed results of statistical significance and no differences were observed in studies for receptive and expressive language outcomes. Three studies addressing speech significantly favored oral language therapy only (except 1 early-identified group32). Closed-set speech perception studies mostly favored spoken language only, whereas open-set speech perception studies statistically favored spoken language but with small clinical significance. The majority of studies were methodologically weak, with a few scoring a moderate rating. However, when rated according to GRADE criteria, the overall quality of the evidence per outcome is very low.

According to GRADE, cohort studies are initially considered as low quality before applying the assessment criteria, as they are at a greater risk of bias, generally. The risk of bias for the body of evidence for outcomes was deemed to have serious limitations. The overall quality of evidence for each outcome we assessed was limited because of the small number of included studies, which resulted in low ratings for the precision of results and an inability to assess for consistency. There was also considerable variation in the measurements of outcomes, which made comparison across studies difficult. Furthermore, the full process of care from time of diagnosis is unknown, as is whether crossover interventions occurred. Therefore, additional studies in contemporary cohorts of children are needed to supplement the evidence base. Most studies were conducted before 2002 with children with severe to profound deafness and therefore may not reflect current practice, in which children with congenital deafness are likely to be identified and receive early intervention well before 1 year of age. Accordingly, new information about the effects of intervention in children across the spectrum of hearing loss severity has the potential to change the conclusions. Since completing our review, we are aware of 1 recent retrospective
cohort study conducted in Australia with young children who used cochlear implants that compared a sign and oral language intervention group (n = 10) with 2 oral intervention groups (auditory-oral, n = 14, and auditory-verbal, n = 18).50 No significant differences in receptive vocabulary, auditory comprehension, or expressive language were found across the 3 groups after controlling for confounders such as age of hearing loss management and family involvement in the intervention program. These results are aligned with the overall findings of our review and would not change the conclusions.

Limitations

Despite a comprehensive search by an information specialist, it is possible that some studies were not included owing to the range of descriptions of intervention methods. During study selection, program descriptions were sometimes difficult to discern. For example, it was difficult to know whether interventions included the same level of emphasis on spoken language when sign language was added. Although we contacted authors for clarity of interventions and outcomes, we did not contact them for additional sample characteristics or statistical data. A further important limitation is that 8 of the 11 studies, although meeting inclusion criteria, were not designed to directly examine the central question of this review.

Implications for Practice

Although the question of intervention effectiveness and specifically the contribution of sign language combined with oral language is longstanding, there is strikingly limited new research to guide clinical or policy decision-making. The question of the contribution of sign language to the development of spoken language remains elusive, and there has been little growth in studies that include comparator groups since previous reviews.51,52 Despite strong conclusions based on expert opinion and intervention results without comparator groups, this review indicates that very limited, and hence insufficient, evidence exists to determine whether adding sign language to spoken language is more effective than spoken language intervention alone to foster oral language acquisition.

CONCLUSIONS

This review showed that there are important gaps in knowledge concerning the effectiveness of sign and oral language intervention, compared with oral language intervention only, for children with hearing loss when spoken language is the intended outcome. To date, there is no evidence that adding sign language facilitates spoken language acquisition. However, this review also found no conclusive evidence that adding sign language interferes with spoken language development. Overall, the literature related to intervention methods for children with hearing loss lacks properly designed cohort studies of today’s generation of children. It will be important to conduct more research to supplement this evidence base and to update this review as those studies become available. Given the current context of widespread neonatal screening, data can be collected prospectively from diagnosis, allowing for more accurate data regarding onset, severity, and changes in hearing loss, as well as specific characteristics and changes related to intervention, than was possible for study participants in this review. Interventions should be reported with specific detail, and if definitions can be developed and adopted on an international level, comparisons of similar programs would be more feasible. In addition, agreement on a common set of core outcomes and reporting metrics would be very useful in comparing future study results.

There is insufficient information to guide parents about contemporary cohorts of children who have benefited from early detection and early access to hearing technology. The field would gain from carefully controlled prospective studies with today’s children.

APPENDIX 1: MEDLINE SEARCH STRATEGY

MEDLINE

1. exp Hearing Loss/
2. (hearing adj (loss or impair$ or disorder*)).tw.
3. deaf$.tw.
4. (prelingual$ or pre-lingual$).tw.
5. (sensorineural$ or sensori-neural$).tw.
6. congenital.mp.
7. or/1-6
8. auditory verbal.tw.
9. ((speech or auditory) adj2 feedback).tw.
10. cued speech.tw.
11. (listen* and (spoken or speak*)).tw.
12. oral approach$.tw.
13. aural.tw.
14. Lipreading/
15. (lipread$ or lip read$ or speechread$ or speech read$).tw.
16. or/8-15
17. Manual Communication/ or Sign Language/
18. (sign$ language or sign$ english).tw.
19. ASL.tw.
20. visual language.tw.
21. (baby sign or infant sign).tw.
22. or/17-21
23. Communication Methods, Total/
24. total communication.tw.
25. simultaneous communication.tw.
26. (multilingual or multi-lingual).tw.
27. (bicultural or bi-cultural).tw.
28. (bilingual or bi-lingual).tw.
29. bi bi.tw.
30. or/23-29
31. exp clinical trial/
32. clinical trial.pt. or randomized.ti,ab. or placebo.ti,ab. or randomly.ti,ab. or trial.ti,ab. or group.ti,ab.
33. ((control$ or clinical or comparative$) adj2 (trial$ or study$ or program or design)).mp.
34. exp Epidemiologic studies/ or Case-control studies/ or Retrospective studies/ or Cohort studies/ or Longitudinal studies/ or Cross-sectional studies/
35. between group design$.mp.
36. control group$.mp.
37. (cohort stud$ or longitudinal).mp.
38. (case adj2 (series or control$)).mp.
39. ((consecutive or clinical) adj2 case$).tw.
40. ((control$ or intervention or evaluation or comparative or effectiveness or evaluation or feasibility) adj3 (trial or studies or study or program or design)).tw.
41. or/31-40
42. 7 and ((16 and 22) or 30) and 41
43. 42 and ((Infan* or newborn* or new-born* or perinat* or neonat* or baby or baby* or babies or toddler* or minors or minors* or boy or boys or boyfriend or boyhood or girl* or kid or kids or child or child* or children* or schoolchild* or schoolchild).mp. or school.child.ti,ab. or school child*.ti,ab. or (adolescen* or juvenil* or youth* or teen* or under*age* or pubescen*).mp. or exp pediatrics/ or (pediatric* or paediatric* or peadiatric*).mp. or school.ti,ab. or school$.ti,ab. or (prematur* or preterm*).mp.
44. limit 43 to (“in data review” or in process or “pubmed not medline”)
45. 42 and (child* or adolescent or infant*).mp.
46. 44 or 45
47. limit 46 to yr=“1985 -Current”

ACKNOWLEDGMENTS
We thank Margaret Sampson, MLIS, PhD, AHIP (Children’s Hospital of Eastern Ontario), for developing the electronic search strategies and Janet Joyce, MLIS, for peer review of the MEDLINE search strategy. We also thank Pauline Quach for assistance with the quality assessment.

ABBREVIATIONS
ASL: American Sign Language
GRADE: Grades of Recommendation, Assessment, Development, and Evaluation
ITPA: Illinois Test of Psycholinguistic Abilities
MD: mean difference
PPVT: Peabody Picture Vocabulary Test

was involved in extracting data and quality assessment, conducted the GRADE ratings, assisted with data interpretation; Dr Moher assisted with the development of the study protocol and consulted on the quality assessment and data interpretation as required; Ms Doucet (a knowledge-user clinician), Dr Neuss (a knowledge-user clinician), and Ms Bernstein (a knowledge-user clinician) provided input into the inclusion criteria and data abstraction form, helped clarify data interpretation when needed, and commented on the manuscript; Ms Na verified and provided input into the synthesis tables and made comments on the manuscript; and all authors approved the final manuscript.

This review is registered at PROSPERO: #CRD42013005426, http://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42013005426#.VWeaU-vqc6I.

DOI: 10.1542/peds.2015-1974

Accepted for publication Oct 7, 2015

Address correspondence to Elizabeth M. Fitzpatrick, Audiology and Speech-Language Pathology Program, Faculty of Health Sciences, University of Ottawa, 451 Smyth Road, Ottawa, Ontario, Canada K1H 8L1. E-mail: elizabeth.fitzpatrick@uottawa.ca

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).

Copyright © 2016 by the American Academy of Pediatrics

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.

FUNDING: This systematic review was funded by a Canadian Institutes of Health Research (CIHR) Knowledge Synthesis Grant (FRN-124600). The funding agency was not involved in developing the protocol and was not involved in any aspect of the review (data collection, analysis, or interpretation) or publications. Dr Fitzpatrick’s work is also supported by a CIHR New Investigator and a Canadian Child Health Clinician Scientist Program award. Dr Moher holds a University of Ottawa Research Chair. All researchers are independent from the funding agencies.

POTENTIAL CONFLICT OF INTEREST: The authors have indicated they have no potential conflicts of interest to disclose.
REFERENCES


28. Brisk ME, Harrington MM. Literacy and bilingualism: a handbook for all


Sign Language and Spoken Language for Children With Hearing Loss: A Systematic Review
Elizabeth M. Fitzpatrick, Candyce Hamel, Adrienne Stevens, Misty Pratt, David Moher, Suzanne P. Doucet, Deirdre Neuss, Anita Bernstein and Eunjung Na
Pediatrics 2016;137; DOI: 10.1542/peds.2015-1974 originally published online December 18, 2015;
Sign Language and Spoken Language for Children With Hearing Loss: A Systematic Review
Elizabeth M. Fitzpatrick, Candyce Hamel, Adrienne Stevens, Misty Pratt, David Moher, Suzanne P. Doucet, Deirdre Neuss, Anita Bernstein and Eunjung Na
Pediatrics 2016;137; DOI: 10.1542/peds.2015-1974 originally published online December 18, 2015;

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/137/1/e20151974