Analysis of Amount and Style of Oral Interaction Related to Language Outcomes in Children with Hearing Loss: A Systematic Review (2006-2016)

Nuzhat Sultana[1], Lena L. N. Wong [1], Suzanne C. Purdy[2]

[1]Division of Speech and Hearing Sciences, Faculty of Education, the University of Hong Kong, Hong Kong.

[2]School of Psychology, the University of Auckland, Auckland, New Zealand.

Corresponding Author

Address for correspondence: Nuzhat Sultana, Faculty of Education (Division of Speech and Hearing Sciences), PGS room no. 730, 7th floor, Meng Wah Complex, the University of Hong Kong.

Telephone: +852-5349-7001

E-mail: u3003554@connect.hku.hk

The authors declared no relevant conflicts of interest.

Financial Disclosures

This study did not receive any funding for the current work.

Abstract

Purpose: This systematic review summarizes the evidence for differences in the amount of language input between children with and without hearing loss (HL). Of interest to this review is evaluating the associations between language input and language outcomes (receptive, and expressive) in children with HL in order to enhance insight regarding what oral language input is associated with good communication outcomes.

Method: A systematic review was conducted using keywords in three electronic databases: Scopus, PubMed, and Google Scholar. Keywords were related to language input, language outcomes, and HL. Titles and abstracts were screened independently, and full-text manuscript meeting inclusion criteria were extracted. An appraisal checklist was used to evaluate the methodological quality of studies as poor, good, or excellent.

Results: After removing duplicates, 1545 study results were extracted, with 27 eligible for full-text review. After the appraisal, eight studies were included in this systematic review. Differences in the amount of language input between children with and without HL were noted. Conversational exchanges, open-ended questions, expansions, recast, and parallel talk were positively associated with stronger receptive and expressive language scores. The quality of evidence was not assessed as excellent for any of the included studies.

Conclusions: This systematic review reveals low-level evidence from eight studies that specific language inputs (amount, style) are optimal for oral language outcomes in children with HL. Limitations were identified as: sample selection bias, lack of information on control of confounders and assessment protocols, and limited duration of observation/recordings. Future research should address these limitations.

Keywords: hearing loss, children, language input, language outcome, systematic review

Introduction

Language input is important to stimulate the brain for functional activation of language comprehension and expression (Feyten, 1991; Hoff, 2006) and plays a crucial role in social (Bloom, 1998; Hoff, 2006; Ratner, 1994) and academic achievements (Gilkerson & Richards, 2009; Hart & Risley, 1995). Hearing loss is a major cause of delay in oral language development (Gravel & O'Gara, 2003; Nelson, Nygren, Walker, & Panscha, 2006). Early identification of hearing impairment via newborn hearing screening, diagnostic evaluations to confirm the type and degree of hearing loss, the fitting of appropriate hearing devices (e.g., hearing aids and cochlear implants), and enrolling children with hearing loss (CwHL) in early intervention programs during infancy (Bachmann & Arvedson, 1998; Duchesne, Sutton, & Bergeron, 2009; Moeller, 2000) may reduce the risk of poor communication outcomes in CwHL. However, there is considerable variability in oral language development in young CwHL even if they have undergone newborn hearing screening and been fitted early with hearing devices (Wake at al., 2005; Stika et al., 2015; Tomblin et al., 2014) and parental input is a factor that importantly determines oral language development. Aural rehabilitation programs for CwHL provide advice and training to encourage families and caregivers to engage their children in oral interaction through language input (Dalzell et al., 2000; Fairgray, Purdy, & Smart, 2010; Harrison, Roush, & Wallace, 2003; Spivak, Sokol, Auerback, & Gershkovich, 2009).

Adult-child oral communication is crucial for language development in young children (Hoff, 2006) because receptive and expressive language skills are developed in day-to-day oral interactions with primary caregivers and parents (Quittner et al., 2010; Sandall, 2005). Parental interaction provides a basis for building language abilities in young children with typical development (Bell, 1979), and the same is expected for CwHL. However, sometimes, parents

of CwHL use fewer communication exchanges with the belief that their children may not have the language skills to participate in oral communicational activities (DesJardin et al., 2014).

Parents vary in their quantity of talk based on counts of the number of words and oral communication exchanges during parent-child oral interactions; these have significant associations with receptive and expressive language outcomes in typically hearing children (Gilkerson & Richards, 2009; Hart & Risley, 1995). The children of more talkative parents showed stronger lexical abilities in comparison to less talkative parents (Hoff & Naigles, 2002; Hurtado, Marchman, & Fernald, 2008; Huttenlocher & et al., 1991).

Parents also work as communication partners during oral interactions with their children (Masur, 1982; Masur, Flynn, & Eichorst, 2005; McNally et al., 1991; Olson & Masur, 2011, 2013; Vohr, Topol, Watson, St Pierre, & Tucker, 2014) adjusting their language behaviors for the child—for example, by portraying excitement in verbal responsiveness, and being involved in oral interactional activities to drag the child's attention towards active participation in the oral communication (DesJardin, 2005). Thus, language exposure does not only depend on the quantity of adult word exposure but also on the degree to which turn-taking takes place.

Children's natural language exposure varies greatly from parent to parent not only in terms of quantity (Gilkerson & Richards, 2008) but also styles of oral interaction (Suter, 2006) that are positively linked with stronger oral language development in typically developing children (Karras & Braungart-Rieker, 2005). While certain styles of oral interaction seem to be prevalent in parent-child interaction (e.g., labeling, pointing, directive, & close-ended question), these styles of oral interaction are observed more frequently in CwHL. DesJardin et al. (2014) reported that mothers of CwHL used lower-level strategies like pointing and labeling more frequently without trying to use high-level (parallel talk, "Wh" question, expansion, and

recast), in contrast with the mothers of CwNH. Notably, hearing mothers of CwHL naturally adjust their language input, compensating for their children's HL by using more lower-level strategies like close-ended questions (Lederberg & Everhart, 2000).

While good language is the desired outcome, the measures used to define this outcome must be commensurate with the child's development. In this review, where children aged 0-8 years are the target population, language outcomes have been quantified in several ways, including auditory comprehension and verbal expression (e.g., Zimmerman, 2002), total language scores (e.g., Gilkerson & Richards, 2009), and the number of child vocalizations (e.g., McDaniel & Purdy, 2011), measured using age-appropriate tools.

Amount of Language Input

Research conducted in general pediatric populations has demonstrated the contributions of language input to communication outcomes in children with normal hearing (CwNH) (Greenwood, Thiemann-Bourque, Walker, Buzhardt, & Gilkerson, 2011; Weisman & Snow, 2001), For example, Hart and Risley's highly cited 1995 study showed that the amount of parental language input in terms of the number of words is associated with child vocabulary size. Forty-two families and their young children, with ages ranging from 7 to 36 months, were audiotaped for three years. Oral interactions were coded to quantify the amount of language input and output. The data were derived from three different social classes (referred to by the authors as welfare class, working class, and professional class). At the age of 7-9 months, children were exposed to welfare class, 616 words/hour; working class, 1251 words/ hour; professional class, 2152 words/hour and showed vocabulary growth rates at 3 years of age of: welfare class, 550 words; working class, 750 words; and professional class, 1100 words. Thus, there were differences in parental language input and child vocabulary growth in this study across families differing in socioeconomic status and parental education – more input was associated with stronger vocabulary growth. This has been replicated in later

studies such as Huttenlocher et al. (1991), who focused on the frequency of maternal words and the relationship to vocabulary development in children with typical language development. There was a linear relationship between the frequency of maternal words and vocabulary development at several points in time, from 14 to 26 months ($r = 0.65$, $p < 0.01$). Similarly, Hoff and Naigles (2002) found that the number of language input utterances was significantly correlated with the number of different words types in children's speech.

Using the LENA® (Language ENvironment Analysis) technology, Zimmerman et al. (2009) determined the number of adult words (AWs) and conversational turns (CTs), and their association with PLS-4 score outcomes (Preschool Language Scale, receptive and expressive language). Cross-sectional regression analysis showed that an increase of 1,000 AWs per day was associated with a 0.44 increase in PLS-4 total scores, while an increase of 100 CTs per day was associated with a 1.92 increase in PLS-4 total scores. In a longitudinal analysis that Zimmerman et al. (2009) conducted with 71 children aged 2 to 48 months, followed over 18 months, CTs were also found to be significantly associated with stronger receptive and expressive language skills in CwNH.

It is vital to understand whether hearing parents communicate with their children differently because of their child's hearing loss. As mentioned above, there is evidence that parents adjust their communication styles with the belief that their modified interactional styles may increase their children's potential responses during the interaction. The literature on CwNH suggests that language input greatly influences language outcomes; this information may help clinicians working with CwHL and their parents/families in early intervention.

However, the consideration of several demographic factors (e.g., number of siblings, child's birth order, and parental involvement) is also essential during the analysis of language input. For example, Bridges and Hoff (2014) have suggested that siblings interacted with each other more frequently in daily routine than the other family members. It is possible that there

is an increased amount of language input in those households that have more siblings and other family members in addition to parents. Child's birth order is another crucial factor in terms of the variation in the amount of parental language input; Phillips (1973) found increased maternal mean length of utterance (MLU) with older children (28-month-olds) compared to younger CwNH (18-month-olds). Another important factor that may enhance language input is parental involvement in developing oral language skills in CwHL. According to DesJardin (2005), highly involved parents were more concerned about the development of oral languasge skills in their CwHL, and more parental involvement may enhance language input during adult-child oral interactions. These concerns regarding the potential influence of demographic factors on language input indicate that it is essential to account for these confounders when examining the results of research in this area.

Quality of Language Input

The quality of language input refers to a wide range of linguistic characteristics, including the use of different word types, utterance complexity, and the use of different styles and techniques to engage children in oral conversation (Hurtado, Marchman, & Fernald, 2008). These oral language strategies are crucial for enhancing children's participation in oral interaction (Girolametto, Weitzman, Wiigs, & Pearce, 1999). A longitudinal study by Rowe (2012) conducted with 50 typically developing children aged 18, 30, and 42 months to examine the influence of quality of parental language input on children's vocabulary development used videotapes of parent verbal input during parent-child interaction at 18, 30, and 42 months to analyze the type of interaction. Word types (i.e., different word roots), rare word types (i.e., list of less common dictionary words/rare vocabulary), explanations, and narrations were examined to determine associations with children's vocabulary outcomes. Vocabulary outcomes were assessed one year later at 30, 42, and 54 months of age using the Peabody Picture Vocabulary Test (PPVT). There were significant associations between parental use of

different types of words (30 months of age) and vocabulary outcomes measured 12 months later (r = .43, p < .01), and the use of rare word types at 30 months of age and vocabulary scores obtained 12 months later (r = .35, p < .05). Further, parental use of explanations at age 42 months was significantly correlated with vocabulary scores 12 months later (r = .29, p < .05) and parental use of narrations at age 42 months was also significantly correlated with vocabulary scores assessed 12 months later (r = .34, p < .05) (Rowe, 2012). Thus, it seems that the type of words, expansions, and narrations used by parents importantly related to vocabulary development. Overall, the available evidence suggests that quantity and quality of language input are positively associated with expressive language outcomes in young children.

Methods to Measure Language during Adult-Child Oral Interactions

Many studies have collected data from short semi-structured observations of mother-child interaction via video or audio recordings to calculate the amount of maternal language input manually (DesJardin & Eisenberg, 2007; Hoff, 2003; Huttenlocher & et al., 1991). However, the instructions and settings in these studies (e.g., videotaped book reading) can create an artificial environment for children and adults. Highly structured situations do not reflect how parents interact with children on a daily basis (Gardner, 2000). Due to time and resource limitations, analyses are often restricted to short recordings that may not be the most representative of natural language input.

Alternatively, a standardized objective measure of language amount such as the LENA® technology could be used to ensure less disturbance of the natural language context and requires less time commitment for data analysis when automated coding is used. The LENA® digital language processor records verbal interactions for up to a 16-hour period, and the software can estimate the number of AWs, CTs, and the child's vocalizations for 5-minute, 1-hour, and day-long durations. Traditionally the number of AWs has been used to quantify language input, but the LENA® also measures CTs, which is a dyadic variable reflecting the

participation of both the adult and the child. Although CTs include both adult input and child output, this variable is referred to as a language input variable in the current review, consistent with earlier LENA® studies. The inter-rater reliability and consistency of the LENA® tool have been reported for English (Xu & Gray, 2009; Zimmerman et al., 2009), Spanish (Weisleder & Fernald, 2013), and French (Canault et al., 2016) speaking children. These studies achieved 82% inter-rater reliability for coding of adult speech (AWs) and 73-76% for child speech (CVs) when manual coding by a human transcriber and LENA® coding were compared.

LENA® reliability has been questioned, however, by Busch, Sangen, Vanpoucke, and Wieringen (2018) who conducted a study of six Dutch-speaking 2-5-years old normal hearing children in which they obtained eight full-day LENA® audio recordings to determine the reliability of the LENA® automatic calculations. These researchers found mean differences between manual counts and LENA® automatic estimates of number of AWs (LENA® counts, $M = 228.5$, $SD = 231.7$; manual counts, $M = 284.4$, $SD = 253.7$), CTs (LENA® counts, $M = 8.4$, $SD = 7$; manual counts, $M = 22.9$, $SD = 21.9$). In general, LENA® calculations were fewer than manual counts, with the greatest under-estimation for CTs.

LENA® is more efficient than collecting interactions in the traditional way which used short recordings in non-naturalistic settings to evaluate language input in the naturalistic environment (e.g., at home), but accuracy is best for AWs and poorest for CTs. Also, LENA® can only record and analyze oral linguistic input and does not capture gestures and body language that play an integral role in everyday communication activities. Unlike human raters, LENA® software cannot undertake intensive analyses regarding the quality and style of interaction during conversational exchanges.

The Current Study

A previous systematic review has examined parent-infant interactions and children's language development from birth to 3 years without focusing on CwHL (Topping et al., 2013).

Sixty studies with good methodological quality evidence were included in this review from different disciplines based on a search of nine databases. Results from seven strong and three moderate methodological quality studies suggested that parental interaction during playing, picture description, and other educational activities enhance language skills during childhood. None of the previous studies in Topping et al.'s review summarized the evidence related to the group differences in the amount of naturalistic language input, and the association between the styles of oral interaction and language outcomes in CwHL. The current systematic review has three foci, examining evidence in regards to (1) group differences between CwHL and CwNH based on specific variables used in the literature to quantify language input (number of adult words and conversational turns), including studies using both manual and LENA® calculations; (2) statistically significant association between the amount of language input (e.g., number of adult words, number of conversational turns, and the different styles of oral interaction), and language outcomes (e.g., number of child's vocalisations, receptive and expressive language scores) in CwHL; and (3) statistically significant association between the child's oral language environment and language outcomes in CwHL. The current review not only summarizes findings from previous literature but also assesses the methodological quality (e.g., study design, data collection process, measurement tools), to determine whether the evidence is trustworthy and also to help researchers to consider how future research could be done to improve the level of evidence.

This review considered all studies conducted over a ten-year period examining the automatic calculation of a number of adult words, conversational turns, and child vocalizations for CwHL. Studies for the amount and style of oral interaction and their association with language outcomes relative to aged-matched CwNH published during the past ten years (2006 to July 2016) were reviewed. The implementation of universal hearing screening in many developed countries and hearing device technological advancements have dramatically

changed the language outcomes of CwHL and thus research findings from the past ten years provide more relevant information compared to older studies.

The review addresses three research questions: (1) Are there significant differences in the amount of language input (AWs, CTs) between CwHL and CwNH groups? (2) Is there a statistically significant association between the amount of adult oral language input and language outcomes in CwHL? (3) Is there a statistically significant association between the styles of adult oral interaction and language outcomes in CwHL?

Method

The current systematic review considered peer-reviewed journal articles published in English from 2006 to July 2016. The search procedure followed the five phases of the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) guidelines: identification, screening, eligibility, inclusion, and analysis (Moher, Liberati, Tetzlaff, & Altman, 2009). The current systematic review was performed by three researchers (authors of the review). Inclusion criteria were established with the 100% mutual consensus of three reviewers. The title and abstract were reviewed by two reviewers (authors 1 and 2) independently during phase 1 and 2. Before and during the eligibility review period (phase 3), face-to-face meetings were held to reach consensus about study selection criteria and to determine which studies should be included and excluded. Author 1 and author 3 analyzed studies independently for phase 4 (inclusion) and 5 (analysis). The decision regarding the inclusion of a full-text article was made after discussion during an organized face-to-face meeting between authors 1 and 3. To reduce the chances of subjective biases affecting judgments, both authors (1 & 3) independently reviewed the included studies. The final selected studies were discussed between three reviewers and discrepancies were resolved by 100% mutual consensus.

Phase 1: Identification

Three online databases (e.g., Scopus. PubMed, and Google Scholar) were used to retrieve relevant articles for this systematic review. The following keywords were used to search for relevant studies for the review: [(language input) OR (linguistic input) OR (Adult-child interaction) OR (Adult-child conversation) OR (Caregiver-child interaction) OR (Quantity of language input) OR (Quantity of natural language input) OR (amount of language input) OR (amount of natural language input) OR (Quality of language input) OR (Quality of natural language input) OR (Facilitative language techniques) OR (Language strategies) OR (Linguistic style)] AND [(Language development) OR (Language skills) OR (Language acquisition)] AND [(Hearing impairment) OR (Children with hearing loss) OR (Hard of hearing) OR (Children with hard of hearing) OR (Deaf children)]. The Boolean operators "OR" and "AND" were used along with keywords when searching databases. The terms amount/quantity and style/quality are not necessarily equivalent, however, they were included in the search in order to capture a range of studies examining language input.

Phase 2: Screening

Three sequential steps were followed for the screening phase: a) duplication removal, b) title screening and c) abstract screening. Title and abstract screening were based on three factors: language input, language outcomes, and CwHL. To be considered for inclusion, studies were required to have at least two of these factors in the title or the abstract.

Phase 3: Eligibility

To complete the eligibility phase, papers were required to measure language input and language outcomes in CwHL. The eligibility assessment procedure was similar to the screening procedure but more in-depth. All studies that included children with additional disabilities and did not distinguish those CwHL also were excluded. Studies examining language input with older children (above 8 years of age) were excluded. Studies that measured language input only by fathers, therapists, teachers, and peers were excluded. Studies published in languages other

than English were also excluded at this phase. These inclusion and exclusion criteria were used to narrow the search for studies to minimize participant and study design variability that would reduce the reliability of findings. The focus on children up to 8 years of age recognizes the importance of parental input in early language development at preschool and early school age.

Phase 4: Inclusion

In phase 4, the final inclusion criteria specified only those studies that focused on language input variables in terms of the number of AWs, the number of CTs, and the style of oral interactions specified as: questions, expansions, recasts, imitations, directives, labels, linguistic mapping, and/or parallel talk, etc. Additionally, included papers were also required to report the associations between language input and language outcomes. Finally, the references included in identified articles were hand searched and did not identify additional relevant articles. Commentaries, opinion-based articles, non-peer-reviewed articles, articles published in peer-reviewed magazines, poster presentations, unpublished dissertations, and short letters were excluded.

Phase 5: Analysis

For the final phase of analysis, eight full-length published articles were selected and sub-categorized into two groups. Category (1) focused on the amount of oral language input, while category (2) studied the style of oral interaction.

Appraisal of Methodological Quality

The eight included studies were evaluated for their methodological quality and the level of evidence based on the research design. The critical appraisal tool for cross-sectional studies (CAT-CSS) prepared by Mohamed, Ahmed, and Soliman (2017) was used to evaluate the methodological quality of the reviewed studies. Mohamed et al. (2017) reported the content validity of the scale as 93.3% and internal consistency as 0.76-0.97. Only section 2 (study validity) of this checklist was used for detailed critical analysis. Section 2 contains eight sub-

sections, with a total of 50 items used to determine whether the study met the criteria. A rubric evaluation included in Mohamed et al.'s (2017) study (see Table, Supplemental Digital Content 1) was used to assess the overall quality of each study. For each criterion met, a score of "1" was assigned when the information was clearly stated in the reviewed studies, according to the questions mentioned in each item of the checklist. The total score for each sub-section was calculated to determine its quality as "*poor*," "*good*," or "*excellent*," using the pre-set scoring criteria, which was based on the total percentage of evaluated items meeting the criteria. Biases in the rating of the quality of reviewed studies during the appraisal process are possible. To minimize the potential effect of biases, the CAT-CSS was used by two independent appraisers to rate the quality of each study. The few discrepancies were noted in quality ratings of introduction and discussion sections; these were resolved with 100% agreement between two appraisers.

After the three quality parameters were assigned to the individual studies, levels of evidence were determined (American Speech-Language-Hearing Association, 2004). The American Speech-Language-Hearing Association (ASHA) evidence levels have been used in earlier systematic reviews to examine the efficacy of using auditory-verbal therapy to facilitate the language development of CwHL (Kaipa & Danser, 2016) and treatment intensity for speech disorders (Kaipa & Peterson, 2016). According to ASHA (2004), a well-designed meta-analysis of >1 randomized-controlled trial (RCT) is rated at an evidence level of 1a; a well-designed RCT is assigned evidence level 1b; level 2a refers to well-established, controlled studies without randomization; well-designed quasi-experimental studies are referred to as level 2b; well-formed, non-experimental studies, including correlation and case studies, are considered to have an evidence level of 3; and finally level 4 is assigned to expert committee reports, consensus conferences, and clinical experiences of respected authorities (Kaipa & Danser, 2016; Kaipa & Peterson, 2016).

Results

Figure 1 displays a hierarchical representation of the articles selected according to the inclusion and exclusion criteria based on the PRISMA guidelines (Moher et al., 2009). The database search yielded a total of 1,665 articles, from Scopus (*n* = 468), PubMed (*n* = 664), and Google Scholar (*n* = 533) (phase 1, identification). From the initial screening, 122 papers were eliminated due to replication across databases. After the review of titles and abstracts, other 1,455 articles were excluded as non-relevant. The remaining 88 papers were retrieved for further evaluation (phase 2, screening). Sixty-one articles were excluded because they did not meet the eligibility criteria. Eventually, 27 studies were selected for evaluation in phase 3 (eligibility). Based on the inclusion criteria of the three research questions, ten more studies were excluded. A final nine studies were then excluded because they did not meet the peer review requirements (phase 4, inclusion).

In total, eight full-text articles were included in this systematic review and categorized into the two subcategories. Five articles were grouped under category (1), "amount of oral language input" (Ambrose, VanDam, & Moeller, 2014; Aragon & Yoshinaga-Itano, 2012; McDaniel & Purdy, 2011; VanDam, Ambrose, Moeller, 2012; Vohr et al., 2014) two articles were grouped under category (2), "style of oral interaction" (DesJardin & Eisenberg, 2007; Cruz, Quittner, Marker, & DesJardin, 2013), and one article was categorized as both (1) and (2) "amount and style of oral interaction" (Ambrose, Walker, Unflat-Berry, Oleson, & Moeller, 2015). The analysis yielded five cross-sectional, two longitudinal, and one cross-sectional and longitudinal mix study designs (phase 5, analysis). The search procedure is illustrated in Figure 1.

Classification of Methodological Quality and Level of Evidence

The quality of studies was assessed using a CAT-CSS rubric (Mohamed et al., 2017). Based on total methodological quality scores, five studies (Ambrose et al., 2015,2014; Cruz et

al., 2013; DesJardin & Eisenberg, 2007; VanDam et al., 2012) were classified as good quality, achieving scores of 50% or higher. All included studies were found to be at evidence level 3 (ASHA, 2004) (see Table, Supplement Digital Content 1).

Critique of Method and Procedure of the Studies

The search yielded eight articles examining the amount and styles of oral interaction (see Table 2). Synthesis of these studies' findings will be used to explore factors associated with the development of age-appropriate language in CwHL. The body of included evidence had limitations in the following specific areas: sample size, sample selection, duration of observations, measurement tools, effect sizes, effects of confounding variables, and lack of analysis regarding the influence of audiological variables.

Sample Size

One study involved only eight CwHL (McDaniel & Purdy 2011), and one study involved 24 English CwHL and 10 Spanish CwHL (Aragon & Yoshinaga-Itano 2012), comparing results to the LENA® normative sample. Two studies involved 28-32 CwHL (Ambrose et al. 2014; DesJardin & Eisenberg 2007), and one study (Cruz et al., 2013) had a larger sample size 93 CwHL without a comparison group of CwNH. Two studies involved 22-23 CwHL with a comparison group of 8-41 CwNH (VanDam et al., 2012; Vohr et al. 2014). One longitudinal study had 71-85 CwHL, and 18-41 CwNH tested at two different time points (18 months and 3 years) (Ambrose et al., 2015). No study included a justification of sample size or comments about the representativeness of the sample. Due to the small sample sizes of CwHL, it is difficult to determine whether the results can be generalized to the target population of CwHL. Consequently, there may not be sufficient statistical power, and the results may not be reliable for accurate interpretation and application in clinical situations.

Sample Selection

None of the included studies described the sampling technique for the recruitment of potential participants. It is unclear whether the researchers were directly involved or blinded during participant recruitment or data collection. There was no indication of how the researcher or participant bias was prevented during the study. Three studies (Aragon & Yoshinaga-Itano 2012; Ambrose et al. 2014; VanDam et al. 2012) did not provide enough information about their population. Only one study (Vohr et al., 2014) described the method of sample selection clearly, with inclusion and exclusion criteria for the matched groups. Many studies did not report the child's birth order.

Duration of Observation

Amongst studies analyzing qualitative language input, two studies (Ambrose et al., 2015; Cruz et al., 2013) used audio and video recordings for direct observation of mother-child interactions for periods of 5-10 minutes during structured activities, on only one or two occasions. Analyses of data based on recordings over short time periods may not be representative and could be an unreliable measure of everyday interactions due to the adult participant's awareness of the recording, which can influence oral interactions. Only one study used the videotape approach over multiple videotaped interactions (DesJardin & Eisenberg, 2007); the use of multiple recording samples should improve result validity but be rarely done.

Five studies used LENA® recordings in natural settings, allowing automatic analysis of language input based on all adults interacting with the child throughout the day (Ambrose et al., 2014; Aragon & Yoshinaga-Itano, 2012; McDaniel & Purdy, 2011; VanDam et al., 2012; Vohr et al., 2014).

There are variations in the number of days for recordings across LENA® studies. Three studies (Aragon & Yoshinaga-Itano, 2012; VanDam et al., 2012; Vohr et al., 2014) considered only one full day of LENA® recordings in natural situations. The other two LENA® studies included three to six days of recordings (Ambrose et al., 2014; McDaniel & Purdy, 2011). Only

one study (McDaniel & Purdy, 2011) justified the selection of recording over several days as ensuring that the data reflected the variety of language input that children were exposed to naturally. Consideration of information regarding the observed time is important to guide future empirical research in this domain. This will assist in determining the stability of measures such as language input in everyday interaction.

Language Assessment Tools and Reliability

Significant associations between adult input and language outcomes were examined in six studies (Ambrose et al., 2014; Cruz et al. 2013; DesJardin & Eisenberg 2007; McDaniel & Purdy 2011; VanDam et al. 2012; Vohr et al. 2014). Three studies (Cruz et al., 2013; DesJardin & Eisenberg, 2007; Vohr et al., 2014) used the Reynell Developmental Language Scales, 3rd Edition (RDLS-3), which has reported test-retest reliability (comprehension 0.96; expressive language 0.97) but content validity is not reported (Edwards et al. 1990). This test should be administered by a trained speech and language therapist (SLT). Out of three studies, only one (Cruz et al., 2013) reported that assessments were done by a qualified SLT, as per the test administration protocol of RDLS-3. One study (McDaniel & Purdy, 2011) used the Pre-School Language Scale, 4th edition (PLS-4), administered by an SLT, which also has good reported test-retest reliability (0.90-0.97) and content validity (0.82) (Zimmerman et al., 2002). The PLS-4 can be used by SLTs, special educators, and researchers. Two studies (Ambrose et al., 2014; VanDam et al., 2012) used the Mullen Scales of Early Learning, which has strong test-retest reliability (0.85-0.96) and low-moderate content validity (no statistics) (Mullen, 1995). Ambrose et al. (2014) also used the Comprehensive Assessment of Spoken Language, with a reported moderate-good test-retest reliability (0.65-0.90), but content validity is not reported (Carrol-Woolfolk, 1999). The Mullen Scales of Early Learning combines clinician-elicited and parent-reported items (Mullen, 1995); the Comprehensive Assessment of Spoken Language should only be administered and interpreted by qualified SLTs. Ambrose et al. (2014) and

VanDam et al. (2012) did not provide any information regarding the test administrator/s. The use of different outcome measures across studies makes comparison less straightforward.

Effect Sizes

Information regarding effect sizes is widely acknowledged to broaden insight into the clinical implications of results. However, only one study (Ambrose et al., 2015) reported effect sizes. This study found significant differences in caregiver's input between CwHL and CwNH, with mostly medium effect sizes. Caregivers of CwHL spoke less to their children at age 18 months (i.e., fewer total words) when compared to CwNH of the same chronological age.

The calculation of effect sizes for the difference in the amount of language input (AWs, and CTs) between CwHL and CwNH using an online calculator for Cohen's d was possible for only two included studies (VanDam et al., 2012; Vohr et al., 2014) due to incomplete information on descriptive statistics in other two studies (Aragon & Yoshinaga, 2012; McDaniel & Purdy, 2011). One study (VanDam et al., 2012) showed a medium effect size (d = 0.41, p = .52) for the differences in AWs/hr between CwHL and CwNH, and a small effect size (d = 0.29, p = .51) for CTs/hr. One study by Vohr et al. (2014) showed small effect size (d = 0.10, p = .526) for the differences in AWs/hr, and small effect size (d = 0.29, p = .511) for CTs/hr between CwHL and CwNH.

Control of Confounding Variables

Participant recruitment was not blinded in any of the reviewed studies. Blinding is possible by having the institution's or clinic's head or another independent person approach potential participant and obtain consents without direct interaction with researchers at this stage. The lack of blinding could lead to a strong self-selection influence, as more engaged families and children with better outcomes may be more likely to be enrolled. Effects of confounding variables such as family and child characteristics were generally not controlled for. Several demographic and family environment confounders associated with language

outcomes have been identified in other research, such as a number of siblings, birth order, and parental involvement (Berglund et al., 2005; DesJardin, 2005; Gilkerson & Richards, 2009; Hoff, 1998). Four studies (Ambrose et al., 2015; Aragon & Yoshinaga-Itano, 2012; VanDam et al., 2012; Vohr et al., 2014) compared language input between CwHL and CwNH without a description of these factors. However, one included study did not provide information on the severity of hearing loss (Aragon & Yoshinaga-Itano, 2012). Two studies did not mention the mode of communication used with the children during interaction (Ambrose et al., 2015; Aragon & Yoshinaga-Itano, 2012). Four studies did not indicate the age of identification of hearing loss (Ambrose et al., 2015, 2014; Aragon & Yoshinaga-Itano, 2012; Vohr et al., 2014). One study did not consider the type of hearing device (Aragon & Yoshinaga-Itano, 2012). Only two studies provided information regarding the type and intensity of therapeutic interventions (DesJardin & Eisenberg, 2007; McDaniel & Purdy 2011). Due to incomplete demographic information, it was difficult to infer how these studies control the possible effect of these variables on obtained results.

Consideration of Audiological Factors

Audiological factors (i.e., severity of children's hearing loss, age of hearing loss identification), preferred communication mode (oral or/and sign), use of hearing devices, type of hearing devices used (cochlear implant or/and hearing aid), the age at fitting of devices, length of device use importantly influence the amount of language input and thus language outcomes (Moeller & Tomblin, 2015). Only a few included studies provided additional information on the analysis of possible audiological factors that affect language input and outcomes in CwHL. McDenial & Purdy (2011) failed to find a significant relationship between CTs/hr as language input, and a number of weeks engaged in auditory verbal therapy. However, they reported a higher number of CTs for those children who spent more time in auditory verbal therapy. Vohr et al. (2014) also supported McDaniel's report indicating a trend for CTs/hr to

be higher for those CwHL who was involved in early intervention for more than three months than those who were involved less. VanDam et al. (2012) concluded that audibility (speech intelligibility index) might influence the amount of language input (AWs, CTs) and in turn, more exposure to language input that enhanced language output was likely to be associated with better speech intelligibility in CwHL.

Differences in the Amount of Language Input between CwHL and CwNH

Table 3 compares the amount of language input (i.e., AWs, CTs) between CwHL and CwNH. No significant difference in the number of AWs as a form of language input was found between CwHL and CwNH. However, studies comparing CTs between CwHL and CwNH show variable results.

Adult Words

Five studies (Ambrose et al., 2015; Aragon & Yoshinaga-Itano, 2012; McDaniel & Purdy, 2011; VanDam et al., 2012; Vohr et al., 2014) compared total number of adult words (AWs) during the observation period between CwHL and CwNH (age-matched control group or LENA® normative data). Only one of these studies used audio-recordings of structured interactions with a caregiver rather than the LENA® technology to examine AWs (Ambrose et al., 2015). There is considerable variation in the calculated number of AWs within and across studies. Two studies (McDaniel & Purdy, 2011; Vohr et al., 2014) described an increased number of AWs in CwHL relative to CwNH. All eight participants in McDaniel and Purdy's (2011) study had daily average AW rates above the LENA® norm. Vohr et al. (2014) included an age-matched control group, finding that AWs tended to be higher (range 659-2460, M =1416, SD = 486, n = 23) in the CwHL than in the controls (range 396-3114, M = 1358, SD = 625, n = 41) but this difference was not significant (p = .071). Similarly, another study (Aragon & Yoshinaga-Itano, 2012) compared CwHL to the LENA® norms for 2-36-month olds, finding increased AWs (range 5,292-42,536; Mdn = 12,297; M = 17,605) in native English-speaking

CwHL when compared with the normative range of 6,003-29,428 AWs (*Mdn* = 12,297). This study also reported the number of AWs for Spanish-speaking CwHL (range 4,081-23,382; *M* = 13,914; *Mdn* = 14,062) and compared these results to a control group of typically developing Spanish-speaking children (range 4,812-14,790; *M* = 8,796; *Mdn* = 7,422). Aragon and Yoshinaga-Itano (2012) did not compare groups statistically but did express concern over the narrow range of AWs for typically developing children from Spanish-speaking homes, highlighting the need for future studies to account for families' linguistic and cultural backgrounds.

One study (Ambrose et al., 2015) compared the total number of caregivers' (79% mothers, 13% fathers, 1% grandmothers) words across CwHL and CwNH (age-matched control group) at two time points, 18 months and 3 years. Results showed lower numbers of AWs for CwHL (*M* = 320.9) than CwNH (*M* = 364.9) at 18 months and also at 3 years (M_{CwHL} = 383.1; M_{CwNH} = 433.8). This difference was not significant at 18 months (p = .018) but was significant at 3 years (p = .010). VanDam et al. (2012) reported equal number of AWs (p = .526) for CwHL and an age-matched control group of CwNH, for children aged 24-36 months.

Conversational Turns

Four studies (Aragon & Yoshinaga-Itano, 2012; McDaniel & Purdy, 2011; VanDam et al., 2012; Vohr et al., 2014) compared the total number of adult-child CTs between CwHL and CwNH (age-matched control group or LENA® normative data). Overall, CTs appeared to be higher for CwHL compared with CwNH; however, there was also some individual variability. In McDaniel and Purdy's (2011) study, all eight children had AWs above the LENA® norms, but CTs varied. For two out of three participants with receptive language scores below the norm, CTs were also below the norm. The other five participants with typical or above average receptive language scores all had CTs at or above the norm.

Aragon and Yoshinaga-Itano (2012) found higher CTs in CwHL (M = 644) compared to LENA® normative data (M = 462) for 2-36-month-old children. While LENA® software provides normative descriptive statistics (mean, standard deviation, and percentile rank) for age matched 2-48-month-old children's AWs and CTs. Similar results were found by Vohr et al. (2014), comparing CTs (range 22-126, M = 55, SD = 24) for CwHL to CwNH (range 3-97, M = 48, SD = 23) for six- to eight-year-olds (seven years on average), but this difference was not significant (p = .26). In both studies, CT measures are based on one day of LENA® recording, for 12 (Vohr et al., 2014) or 12-16 hours (Aragon & Yoshinaga-Itano, 2012). Variability in the results could reflect age differences across participant groups. Also using one-day LENA® recordings, VanDam et al. (2012) reported fewer CTs in CwHL when compared to age-matched controls for children age with age range 24-36 months (M_{CwHL} = 61, SD_{CwHL} =17; $M_{\text{control group}}$ = 66, $SD_{\text{control group}}$ = 17; p = .51).

Association between Amount of Language Input and Language Outcomes

Table 4 illustrates the association between language input (i.e., AWs and CTs) and outcomes in terms of a number of child's vocalisations, and receptive and expressive language scores. In total, four studies (Ambrose et al., 2014, 2015; McDaniel & Purdy, 2011; VanDam et al., 2012) were identified as examining associations between the quantity of input (number of AWs) and language outcomes (receptive language, expressive language). None of the included studies examined the association between a number of adult words and child vocalisations. Ambrose et al. (2014) reported non-significant associations between a high rate of AWs and total receptive language scores using the Mullen Scales of Early Learning (r = 0.339, p > .05), and total expressive language scores (r = 0.138, p > .05). Ambrose et al. (2015) examined the association between quantity of language input (i.e., number of total words, number of total utterances) at 18 months and language outcomes at three years of age. The regression model for quantity of caregiver talk was not significant (R^2 = 0.09, p = .24). VanDam

et al. (2012) found no association between AWs and total receptive language scores using the Mullen Scales of Early Learning ($p > .05$). Similar results were obtained by McDaniel and Purdy (2011), who were also unable to detect associations between higher rates of AWs and stronger PLS-4 receptive ($Rs = 0.71$, $p > .05$) or expressive ($Rs = 0.27$, $p > .05$) total scores.

Four studies included in the review (Ambrose et al., 2014; McDaniel & Purdy, 2011; VanDam et al., 2012; Vohr et al., 2014) analysed the number of CTs and other language outcomes (number of children's vocalisations, receptive and expressive language scores). McDaniel and Purdy (2011) found a positive association between number of CTs and children's vocalisation ($Rs = 1.0$, $p < .05$), which is expected, but found no expected correlations between CTs and PLS-4 receptive ($Rs = 0.35$, $p > .05$) or expressive ($Rs = 0.00$, $p > .05$) scores. Three studies (Ambrose et al., 2014; Vohr et al., 2014; VanDam et al., 2012) found significant correlations between CTs and language abilities, measured using standardized tests. Ambrose et al. (2014) used the Mullen Scales of Early Learning to test two-year-olds and the Comprehensive Assessment of Spoken Language to test three-year-olds, finding significant correlations between CTs and language outcomes based on the receptive scores using the Mullen Scales of Early Learning ($r = 0.61$, $p < .01$), total expressive scores ($r = 0.45$, $p < .05$), and the Comprehensive Assessment of Spoken Language total scores ($r = 0.45$, $p < .05$). Vohr et al. (2014) found that higher rates of CTs were associated with stronger RDL-3 receptive ($r = 0.35$, $p = .03$), expressive ($r = 0.89$, $p = .02$), and total language scores. VanDam et al. (2012) also reported significant correlations between CTs and receptive language outcomes for two-year-olds tested with the Mullen Scales of Early Learning ($r = 0.62$, $p < .01$). Generally, higher rates of CTs, indicating that children who are engaged by their parents in conversational exchange, have better language outcomes (see Table 4). Although an increased number of AWs and CTs may not capture well the construct of language input, they are nonetheless an important consideration for enhancing language outcomes. Equally, one could argue that CTs

are more important to consider for oral language development compared to AWs as previous research has shown that CTs are based on joint attention (Tamasello, 1999; Vanniarajan, 2000) and social interactions (Chapman, 2000), which play a crucial role in promoting oral language skills in typically hearing children. Hence, the first step towards increasing the amount of language input is to increase CTs, not just AWs as increased CTs showed stronger a correlation with oral language development in CwHL.

Association between Style of Oral Interaction and Language Outcomes

Table 5 illustrates the association between the style of oral interaction and language outcomes in CwHL.

Three studies (DesJardin & Eisenberg, 2007; Cruz et al., 2013; Ambrose et al., 2015) measured associations between different linguistic styles and language outcomes. DesJardin and Eisenberg (2007) calculated proportional scores for "facilitative language techniques" based on transcriptions of structured videotaped interactions; they found negative associations between three different linguistic styles (linguistic mapping, labeling, directives) and RDLS-3 receptive and expressive language scores. A higher proportion of linguistic mapping was associated with poorer receptive (r = -0.50, p < .01) and expressive (r = -0.42, p < .05) language scores. Labeling were also negatively associated with receptive (r = -0.44, p < .05) and expressive (r = -0.45, p < .05) language. The use of directives was likewise negatively associated with receptive (r = -0.58, p < .01) and expressive (r = -0.49, p < .05) language. Positive results were found for recasts and open-ended questions. The proportion of recasts was associated with stronger language comprehension (r = 0.47, p < .01), while proportion of open-ended questions was positively associated with expressive language (r = 0.51, p < .01).

A longitudinal study by Cruz et al. (2013), which also examined videotaped interactions, divided verbal input style into "high-level" and "low-level" facilitative language techniques and used latent growth curve modeling to examine associations between facilitative

language techniques and language outcomes at three-time points while controlling for socioeconomic status. This study has the largest sample size (n = 93) and is the only study included in this review that controlled for socioeconomic status. At 12, 24, and 36 months, low-level facilitative language techniques (linguistic mapping, comments, imitation, labeling, directives, and close-ended question) were not associated with receptive or expressive language skills. The use of high-level facilitative language techniques (parallel talk, open-ended questions, expansion, and recast) was positively associated with both receptive and expressive language scores. Another study by Ambrose et al. (2015) with a relatively large data set (n = 156 CwHL, n = 59 CwNH), reported a longitudinal analysis of caregivers' language input (quality variables: numbers of different words, mean length of utterance in morphemes ($MLU_{m)}$, proportion of directing, and high-level eliciting). This study found a significant negative contribution of 'directing', with unique variance, to the Comprehensive Assessment of Spoken Language scores at age 3 years. A high proportion of directing (e.g., "Look right here," "No, don't touch that," "Count the bugs," and "Say elephant") at 18 months was negatively associated with CwHL's language scores using the Comprehensive Assessment of Spoken Language at age 3 years (r= -0.41, p = .03). Thus, CwHL exposed to more directing utterances at 18 months had weaker language skills at three years of age than those who were exposed to a lower proportion of directing at 18 months of age. The other three examined variables did not show a significant contribution (i.e., number of different words: $\beta = -0.20$, p = .32; MLUm: β = 0.39, p = .06; High-Level: β = −0.24, p = .18). Overall, CwHL were significantly delayed in their auditory comprehension and use of verbal language compared to CwNH.

Discussion

The current study was designed to review the evidence of findings concerning three research questions, in addition to the assessment of the methodological quality of the included

studies. The three research questions were as follows: (1) Are there significant differences in the amount of language input between CwHL and CwNH groups? (2) Is there a statistically significant association between the amount of adult oral language input and language outcomes in CwHL? (3) Is there a statistically significant association between the styles of adult oral interaction and language outcomes in CwHL?

General Methodological Considerations

All eight studies included in this systematic review were found to have low-level evidence (level 3). The methodological quality of five studies was determined to be good, according to preset quality analysis scoring criteria in CAT-CSS (i.e., 50% or higher scores marked to have good quality), whereas the other three studies were of poor quality (less than 50% scores were marked to have poor quality) due to methodological limitations. For example, none of these studies justified their sample sizes, sample selection criteria, and data collection method (i.e., the number of days and which day). However, according to Sihoe (2015), sampling procedures and size should be justified to ensure reliable results. Another important consideration was related to the lack of information on the type of included days (i.e., weekend and weekdays) for data collection. Variations in activities across different days of the week, for instance, weekdays with more structured school activities versus weekend days with varied family activities, may cause differences in the quantity and quality of adult-child interactions (Booth et al., 2002). In most of the included studies, it was not evident whether those who administered language outcome evaluations were certified to administer these tools, thereby resulting in possible variations in language assessment results. A number of important considerations including audiological variables, such as the age of identification of hearing loss and severity of hearing loss (Moeller & Tomblin, 2015), type of hearing device used and age of fitting of device (Moeller, Tomblin, Yoshinaga-Itano, Connor, & Jerger, 2007), preferred mode of communication (Fairgray et al., 2010), intensity of therapeutic interventions

(Chapman, 2000), birth order (Zambrana et al., 2012), and parental involvement (Calderon, 2000), which have a significant influence on language input and language outcomes, were not mentioned. Future research should include a comprehensive examination of these factors to yield a high-level of evidence for the association between language input and outcomes, with these potential confounders controlled for across studies.

The difference in the Amount of Language Input between CwHL and CwNH

For the current review, the amount of language input was defined as the total number of AWs and CTs for the first research question. AWs are defined as the number of single words produced during the recording by adults in the child's environment (father, mother, caregivers, etc.), regardless of the child's communication during the interaction.

Overall, the included studies showed inconsistency across studies in the amount of language input between CwHL and CwNH. Generally, it is believed that CwHL requires a higher amount of linguistic input than CwNH to achieve the same level of verbal language abilities (Lederberg et al., 2000; Lederberg & Spencer, 2009; Pittman et al., 2005; Stelmachowicz et al., 2004). Thus, it is expected that being exposed to more AWs will be required to facilitate language development in CwHL. However, the current review found great variations in CT rates across studies and between CwHL and CwNH. This variation in CT rates is not unusual due to the involvement of many environmental and familial factors. For example, primary caregivers' socioeconomic status, level of education, age, mental health, and cultural differences play a crucial role in increasing or decreasing the number of AWs and CTs during oral interactions (Ganek et al., 2018; Hoff, 2003; Pan et al., 2005; Rowe, 2005, 2012), in addition to the day to day variations in natural environment. Most of these factors were not considered in the included studies. While it would be difficult to control the stability of the recordings across samples, researchers should provide clear instructions to caregivers in the selection of the periods to be recorded and frame their research findings within the context of

these communication settings. These investigations can be conducted using generalizability and decision studies based on generalizability theory, according to which one needs to consider different characteristics of exposure, effect modifiers, confounders, and outcome (Webb & Shavelson, 2005). Sandbank and Yoder (2014) measured two variables: 1) rate of intentional communication acts, and 2) rate of different words, across three assessment contexts at four communication sampling periods. Results verified that measurement stability increased with time and development for both variables, regardless of the type of assessment procedure used. This type of study design with multiple measures, contexts, and time points could address some of the limitations of the studies examined in this systematic review.

Association between the Amount of Language Input and Language Outcomes

The present systematic review studied the association between the amount of language input and language outcomes to address the second research question. Verbal language outcomes were defined as the total receptive and expressive language scores and number of CVs. Rates of CVs were obtained through LENA® calculations, while receptive and expressive language skills were measured using standardized tests. Out of the six included studies, only McDaniel and Purdy (2011) considered the number of CVs to be a language outcome and found a positive association between higher rates of CVs and CTs as an indication that children were engaged in oral communication exchanges.

Overall, findings from the review are consistent in terms of the association between the number of AWs, CTs, and language outcomes. Three studies (Ambrose et al., 2014; Vohr et al., 2014; VanDam et al., 2012) found a statistically significant positive association between higher rates of CTs and stronger receptive and expressive language scores in CwHL. Therefore, it can be concluded that the highest rates of CTs in adult-child oral interactions promote children's oral language skills (Most et al., 2010). High CTs indicate joint attention and shared focus during targeted conversations (Scofield & Behrend, 2011), and active participation in

oral conversation (Weisleder & Fernald, 2013) and highlight the importance of the child's engagement.

However, three studies (Ambrose et al., 2014; McDaniel & Purdy, 2011; VanDam et al., 2012) found no association between the number of AWs and the total receptive and expressive language scores. One possible explanation regarding this lack of association between the number of AWs and language outcomes is that the LENA® technology estimates AWs based on automatic adult speech calculations in the language environment. The software may underestimate the number of AWs (Busch et al., 2018) and does not indicate specific types of adult words (Gilkerson & Richards, 2008), which may be more important for language outcomes (Ambrose et al., 2015).

Association between Style of Oral Interaction and Language Outcomes

The third research question of the current systematic review relates to the statistically significant associations between the style of verbal input and language outcomes in CwHL. Generally, the findings of all the included studies showed a statistically significant positive association between greater use of open-ended questions, recasts, expansions, and stronger receptive and expressive language outcomes in CwHL. Studies on typically developing children support these results. For example, Salomo et al. (2013) and Girolametto et al. (2002) reported that specific styles (i.e., asking different types of questions, formulating children's words and sentences into question forms, repetition, and expansions of children's oral responses into correct grammatical structures) of parental interactions with children play a crucial role in acquiring adequate oral language skills. Information on how specific styles of oral interactions between an adult/primary caregiver and a child influence the language outcomes may help readers focus on an optimal style of interaction for oral language development.

Limitations of the Current Review and Future Implications

To our knowledge, this is the first systematic review to evaluate the body of literature and summarize the findings regarding the associations between the amount and style of linguistic inputs and language outcomes in CwHL. The limitations are mentioned in two levels.

1. *Limitations at the primary study level*

The evidence is mostly low-level, focusing on observational cross-sectional analysis. The methodological quality was rated as 64% for only one study (Ambrose et al., 2015) on the rating scale (CAT-CSS) due to considerable limitations such as small samples, incomplete demographic and audiological information, and incomplete descriptive and statistical analyses including effect sizes, as discussed above in the critical appraisal section.

2. *Limitations at the systematic review level*

The current systematic review has some limitations. First, the review focused on studies examining the amount and style of communication interaction in children who reportedly did not have additional disabilities. It is expected that children with additional disabilities may show inconsistencies in their language outcomes due to variation in capabilities. Second, the review examined studies involving participants aged 0 to 8 years and did not measure the long-term effects of language input on outcomes. Third, language outcomes were limited to the number of CVs and overall receptive and expressive language scores. These broad measures do not consider the specific structure of language such as semantics, syntax, and pragmatics. Fourth, only articles written in English were considered, and it is possible that other studies may contribute greatly in term of the structure of studies and finding on language input. Fifth, the review examined studies involving primary caregivers, solely mothers, and studies that measured language input from teachers, peers, or therapists were excluded. Input from teachers and peers may also contribute in an important way to enhancing conversational exchanges with the child.

Finally, although efforts were made to reduce bias, the current systematic review cannot be truly bias-free because the subjective judgment was used in the analysis to classify the methodological quality of the studies. A meta-analysis would be preferred for validation of findings but was not possible due to the wide range in the severity of hearing loss (mild to profound), inclusion of unilateral or/and bilateral, conductive or sensorineural hearing loss, variations in the age of hearing loss identification, differences in types of hearing device, influence of different modes of communication (inclusion of oral and sign), the wide range of age of fitting of hearing devices either one ear or both , influence of confounders, and diversity in the age of children in published studies. We acknowledge the effects of bias in the available evidence due to the small sample sizes and lack of control of confounding factors in most studies included in this analysis. Variations in chronological age, exposure to the intervention, socioeconomic status, and maternal involvement across studies lead to difficulties in concluding how much and what type of input is needed for optimizing language development. It is important to note the limitations of the available evidence when designing future studies.

Conclusions

This systematic review analyzed the best available evidence. The results show that early oral language development relates to the specific amount and optimal styles of oral interaction (i.e., "high-level") engaging children in oral communication exchanges. More structured and detailed analysis of language input is needed to see its impact on stronger oral language development in CwHL. The problem in comparing findings from included studies regarding amount and styles of interaction is due to incomplete information on demographics, control of confounding variables, and methodology (e.g., assessment protocols). The recording standards and testing parameters should be reported in detail in the methodology section, e.g., instructions given to parents for the recordings, number of recordings done, and criteria for

794 selecting parts of recordings for analysis that perhaps people can follow when they report
795 studies.

796 **Acknowledgments**

797 We would like to thank all reviewers who have greatly helped to improve this
798 manuscript. We gratefully acknowledge the efforts of a Ph.D. colleague who has great
799 knowledge in research for assistance with the analysis of the quality measures of all included
800 studies using the rating scale. A poster reporting finding of this study was presented at the
801 World Congress Audiology (WCA), held at Cape Town, South Africa, in October 2018. An
802 oral presentation reporting finding of this study in detail was presented at the European
803 Federation of Audiology Societies (EFSA), held at Lisbon, Portugal, in May 2019.

804

805

806

References

[References marked with an asterisk (*) indicate the studies included in the systematic review]

***Ambrose, S. E., VanDam, P. M., & Moeller, P. M.** (2014). Linguistic input, electronic media, and communication outcomes of toddlers with hearing loss. *Ear and Hearing, 35(*2), 139-147.

*** Ambrose, S. E., Walker, E. A., Unflat-Berry, L. M., Oleson, J. J., & Moeller, M. P.** (2015). Quantity and quality of caregivers' linguistic input to 18-month and 3-year-old children who are hard of hearing. *Ear and Hearing, 36*(1), 48S-59S.

American Speech-Language-Hearing Association. (2004). Evidence-based practice in communication disorders: An introduction [Technical report]. doi: 10.1044/policy. TR2004-0000. Retrieved from https://www.asha.org/policy/TR2004-00001/

***Aragon, M., & Yoshinaga-Itano, C.** (2012). Using Language ENvironment Analysis to improve outcomes for children who are deaf or hard of hearing. *Seminars in Speech and Language, 33*(4), 340-353.

Bachmann, K. R., & Arvedson, J. C. (1998). Early identification and intervention for children who are hearing impaired. *Pediatrics in review / American Academy of Pediatrics, 19*(5), 155-165.

Bell, R. Q. (1979). Parent, child, and reciprocal influences. *American Psychologist, 34*(10), 821-826.

Berglund, E., Eriksson, M., & Westerlund, M. (2005). Development and aging: Communicative skills in relation to gender, birth order, childcare and socioeconomic status in 18-month-old children. *Scandinavian Journal of Psychology, 46*(6), 485-491.

Bloom, L. (1998). Language acquisition in its developmental context. In W. Damon (Ed.), *Handbook of child psychology: Vol. 2. Cognition, perception, and language* (pp. 309-370). Hoboken, NJ, US: John Wiley & Sons Inc.

Booth, C. L., Clarke-Stewart, K. A., Vandell, D. L., McCartney, K., & Owen, M. T. (2002). Child-care usage and mother-infant "quality time". *Journal of Marriage and Family, 64*(1), 16-26.

Bridges, K., & Hoff, E. (2014). Older sibling influences on the language environment and language development of toddlers in bilingual homes. *Applied Psycholinguistics, 35*(2), 225-241.

Busch, T., Sangen, A., Vanpoucke, F., & Wieringen, A. (2018). Correlation and agreement between Language ENvironment Analysis (LENA™) and manual transcription for Dutch natural language recordings. *Behavior Research Methods, 50*(5), 1921-1932.

Calderon, R. (2000). Parental involvement in deaf children's education programs as a predictor of child's language, early reading, and social-emotional development. *Journal of Deaf Studies and Deaf Education, 5*(2), 140-155.

Canault, M., Normand, M.-T., Foudil, S., Loundon, N., & Thai-Van, H. (2016). Reliability of the Language ENvironment Analysis system (LENA^{TM}) in European French. *Behaviour Research Methods, 48*(3), 1109-1124.

Carrol-Woolfolk, E. (1999). *Comprehensive Assessment of Spoken Language (CASL)*. Circle Piness, MN: American Guidance Service.

Chapman, R. S. (2000). Children's language learning: an interactionist perspective. *Journal of Child Psychology & Psychiatry & Allied Discipline, 41*(1), 33-54.

***Cruz, I., Quittner, A. L., Marker, C., DesJardin, J. L.** (2013). Identification of effective strategies to promote language in deaf children with cochlear implants. *Child Development, 84*(2), 543-559.

Dalzell, L., Orlando, M., MacDonald, M., Berg, A., Bradley, M., Cacace, A., ... & Gross, S. (2000). The New York State universal newborn hearing screening demonstration

project: Ages of hearing loss identification, hearing aid fitting, and enrollment in early intervention. *Ear and Hearing, 21*(2), 118-130.

DesJardin, J. L. (2005). Maternal perceptions of self-efficacy and involvement in the auditory development of young children with prelingual deafness. *Journal of Early Intervention, 27*(3), 193-209.

***DesJardin, J. L., & Eisenberg, L. S.** (2007). Maternal contributions: Supporting language development in young children with cochlear implants. *Ear and Hearing, 28*(4), 456-469.

DesJardin, J. L., Doll, E. R., Stika, C. J., Eisenberg, L. S., Johnson, K. J., Ganguly, D. H., . . . Henning, S. C. (2014). Parental support for language development during joint book reading for young children with hearing loss. *Communication Disorders Quarterly, 35*(3), 167-181.

Duchesne, L., Sutton, A., & Bergeron, F. (2009). Language achievement in children who received cochlear implants between 1 and 2 years of age: Group trend and individual patterns. *Journal of Deaf Studies and Deaf Education, 14*(4), 465-485.

Dunn, L. M., & Dunn, L. M. (2007). *Peabody Picture Vocabulary Test (PPVT-IV)* (4th ed.). United States: Pearson Assessments.

Edwards, S., Garman, M., Hughes, A., Letts, C., & Sinka, I. (1990). Assessing the comprehension and production of language in young children: An account of the Reynell Developmental Language Scales: RDLS-3 (3rd ed.). *International Journal of Language & Communication Disorders, 34*(2), 151-191.

Fairgray, E., Purdy, S. C., & Smart, J. L. (2010). Effects of auditory-verbal therapy for school-aged children with hearing loss: An exploratory study. *Volta Review, 110*(3), 407-433.

Feyten, C. M. (1991). The power of listening ability: An overlooked dimension in language acquisition. *The Modern Language Journal, 75*(2), 173-180.

Gardner, F. (2000). Methodological issues in the direct observation of parent-child interaction: Do observational findings reflect the natural behavior of participants? *Clinical Child and Family Psychology Review, 3*(3), 185-198.

Gilkerson, J., & Richards, J. (2008). The LENA national language study [Technical report LTR-02-2]. Boulder, CO: LENA Research Foundation.

Gilkerson, J., & Richards, J. A. (2009). Impact of adult talk, conversational turns, and TV during the critical 0-4 years of child development [Technical report ITR-01-2]. *The Power of Talk (2nd ed.)*. Boulder, CO: LENA Research Foundation.

Girolametto, L., Bonifacio, S., Visini, C., Weitzman, E., Zocconi, E., & Pearce, P.S. (2002). Mother-child interactions in Canada and Italy: Linguistic responsiveness to late-talking toddlers. *International Journal of Language & Communication Disorders, 37*(2), 153-171.

Girolametto, L., Weitzman, E., Wiigs, M., & Pearce, P. S. (1999). The relationship between maternal language measures and language development in toddlers with expressive vocabulary delays. *American Journal Speech-Language Pathology, 8*(4), 364-374.

Gravel, J., & O'Gara, J. (2003). Communication options for children with hearing loss. *Developmental Disabilities Research Review, 9*(4), 243-251.

Greenwood, C. R., Thiemann-Bourque, K., Walker, D., Buzhardt, J., Gilkerson, J. (2011). Assessing children's home language environments using automatic speech recognition technology. *Communication Disorders Quarterly, 32*(2), 83-92.

Harrison, M., Roush, J., & Wallace, J. (2003). Trends in age of identification and intervention in infants with hearing loss. *Ear and Hearing, 24*(1), 89-95.

Hart, B., & Risley, T. R. (1995). *Meaningful Differences in the Everyday Experience of Young American Children* (1st ed.). Baltimore, MD, US: Paul H Brookes Publishing.

Hoff, E. (2006). How social contexts support and shape language development. *Developmental Review, 26*(1), 55-88.

Hoff, E. (2003). The specificity of environmental influence: Socioeconomic status affects early vocabulary development via maternal speech. *Child Development, 74*(5), 1368-1378.

Hoff, E., & Naigles, L. (2002). How children use input to acquire a lexicon. *Child Development, 73*(2), 418-433.

Hoff, G., E. (1998). The relation of birth order and socioeconomic status to children's language experience and language development. *Applied Psycholinguistics, 19(04)* 603-629.

Hurtado, N., Marchman, V. A., & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science, 11*(6), f31-f39.

Huttenlocher, J., Haight, W., Bryk, A., Seltzer, M., & Lyons, T. (1991). Early vocabulary growth: Relation to language input and gender. *Developmental Psychology, 27*(2), 236-248.

Kaipa, R., & Danser, M. L. (2016). Efficacy of auditory-verbal therapy in children with hearing impairment: A systematic review from 1993 to 2015. *International Journal of Pediatric Otorhinolaryngology, 86*, 124-134.

Kaipa, R., & Peterson, A. M. (2016). A systematic review of treatment intensity in speech disorders. *International journal of speech-language pathology, 18*(6), 507-520.

Karrass, J., & Braungart-Rieker, J. M. (2005). Effects of shared parent-infant book reading on early language acquisition. *Journal of Applied Developmental Psychology: An International Lifespan Journal, 26*(2), 133-148.

928 **Lederberg, A. R., & Spencer, P. E.** (2009). Word- learning abilities in deaf and hard-of-
929 hearing preschoolers: Effect of lexicon size and language modality. *Journal of Deaf*
930 *Studies and Deaf Education, 14*(1), 44-62.
931 **Lederberg, A. R., & Everhart, V. S.** (2000). Conversations between deaf children and their
932 hearing mothers: Pragmatic and dialogic characteristics. *Journal of Deaf Studies and*
933 *Deaf Education, 5*(4), 303-322.
934 **Lederberg, A. R., Prezbindowski, A. K., & Spencer, P. E.** (2000). Word-learning skills of
935 deaf preschoolers: The development of novel mapping and rapid word-learning
936 strategies. *Child Development, 71*(6), 1571-1585.
937 **Masur, E. F.** (1982). Object-related gestures: Influences on lexical development. *Journal of*
938 *Child Language, 9*(1), 23-30.
939 **Masur, E. F., Flynn, V., & Eichorst, D. L.** (2005). Maternal responsive and directive
940 behaviors and utterances as predictors of children's lexical development. *Journal of*
941 *Child Language, 32*(1), 63-91.
942 ***McDaniel, R., Purdy, S. C.** (2011). Using the Language ENvironment Analysis (LENA)
943 system to investigate the language environment and outcomes of deaf children. *New*
944 *Zealand Journal of Speech-Language Therapy, 1*, 1-27.
945 **McNally, S., Eisenberg, N., & Harris, J. D.** (1991). Consistency and change in maternal
946 child‐rearing practices and values. *Child Development, 62*(1), 190-198.
947 **Moeller, P. M., & Tomblin, B. J.** (2015). Epilogue: Conclusions and implications for research
948 and practice. *Ear and Hearing, 36*(1), 92S-98S.
949 **Moeller, P. M., Tomblin, B. J., Yoshinaga-Itano, M. C., Connor, M. C., & Jerger, M. S.**
950 (2007). Current state of knowledge: language and literacy of children with hearing
951 impairment. *Ear and Hearing, 28*(6), 740-753.

Moeller, M. P. (2000). Early intervention and language development in children who are deaf and hard of hearing. *Pediatrics, 106*(3), e43-e43.

Mohamed, R. A., Ahmed, A. I., & Soliman, S. M. (2017). Development of Critical Appraisal Tool for Cross-Sectional Studies (CAT-CSS). *International Journal of Nursing Didactics, 7*(04), 09-20.

Moher, D., Liberati, A., Tetzlaff, J., & Altman, D. G. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *Annals of Internal Medicine, 151*(4), 264-269.

Most, T., Shina-August, E., & Meilijson, S. (2010). Pragmatic abilities of children with hearing loss using cochlear implants or hearing aids compared to hearing children. *Journal of Deaf Studies and Deaf Education,* 15(4), 422-437.

Mullen, E. M. (1995). *Mullen Scales of Early Learning (MSEL)* (pp 58-64). Circle Pines, MN: American Guidance Service.

Nelson, H. D., Nygren, P., Walker, M., & Panoscha, R. (2006). Screening for speech and language delay in preschool children: systematic evidence review for the US preventive services task force. *Pediatrics, 117,* E298-E319.

Olson, J., & Masur, E. F. (2013). Mothers respond differently to infants' gestural versus nongestural communicative bids. *First Language,* 33(4), 372-387.

Pan, B. A., Rowe, M. L., Singer, J. D., & Snow, C. E. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development, 76*(4), 763-782.

Pittman, L. A., Lewis, E. D., Hoover, M. B., & Stelmachowicz, G. P. (2005). Rapid word-learning in normal-hearing and hearing-impaired children: Effects of age, receptive vocabulary, and high-frequency amplification. *Ear and Hearing, 26*(6), 619-629.

Phillips, J. R. (1973). Syntax and vocabulary of mothers' speech to young children: Age and sex comparisons. *Child Development, 44*(1), 182-185.

Quittner, A. L., Barker, D. H., Cruz, I., Snell, C., Grimley, M. E., …& Cdaci Investigative Team, M. (2010). Parenting stress among parents of deaf and hearing children: Associations with language delays and behavior problems. *Parenting,* 10(2), 136-155.

Rowe, M. L. (2012). A longitudinal investigation of the role of quantity and quality of child-directed speech in vocabulary development. *Child Development, 83*(5), 1762-1774.

Rowe, M. L. (2008). Child-directed speech: Relation to socioeconomic status, knowledge of child development, and child vocabulary skill. *Journal of Child Language, 35*(1), 185-205.

Salomo, D., Lieven, E., & Tomasello, M. (2013). Children's ability to answer different types of questions. *Journal of Child Language, 40*(02), 469-491.

Sandbank, M., & Yoder, P. (2014). Measuring representative communication in young children with developmental delay. *Topics in Early Childhood Special Education, 34*(3), 133-141.

Sandall, S. R., & Council for Exceptional Children Division for Early, C. (2005). *DEC recommended practices: A comprehensive guide for practical application in early intervention/early childhood special education* (3rd ed.). Longmont, CO: Sopris West, DEC.

Scofield, J., & Behrend, D. A. (2011). Clarifying the role of joint attention in early word learning. *First Language, 31*(3), 326-341.

Sihoe, A. D. L. (2015). Rationales for an accurate sample size evaluation. *Journal of Thoracic Disease, 7*(11), E531.

Spivak, L., Sokol, H., Auerbach, C., & Gershkovich, S. (2009). Newborn hearing screening follow-up: Factors affecting hearing aid fitting by 6 months of age. *American Journal of Audiology, 18*(1), 24-33.

Stelmachowicz, G. P., Pittman, L. A., Hoover, M. B., & Lewis, E. D. (2004). Novel-word learning in children with normal hearing and hearing loss. *Ear and Hearing, 25*(1), 47-56.

Stika, C. J., Eisenberg, L. S., Johnson, K. C., Henning, S. C., Colson, B. G., Ganguly, D. H., & Desjardin, J. L. (2015). Developmental outcomes of early-identified children who are hard of hearing at 12 to 18 months of age. *Early Human Development, 91*(1), 47-55.

Suter, S. (2006). Meaningful differences in the everyday experience of young American children. *European Journal of Pediatrics, 165*, 282-282.

Tomasello, M. (1999). *The cultural origins of human cognition.* Cambridge, MA: Harvard University Press.

Tomblin, J. B., Oleson, J. J., Ambrose, S. E., Walker, E., & Moeller, M. P. (2014). The influence of hearing aids on the speech and language development of children with hearing loss. *JAMA Otolaryngology-Head & Neck Surgery, 140*(5), 403-409.

Topping, K., Dekhinet, R., & Zeedyk, S. (2013). Parent-infant interaction and children's language development. *Educational Psychology, 33*(4), 391-426.

Vanniarajan, S. (2000). The Cultural origins of human cognition by Michael Tomasello. Cambridge, MA: Harvard University Press, 1999, 248 pp. *Applied Linguistics, 11*(1).

***VanDam, M., Ambrose, S. E., Moeller, M. P.** (2012). Quantity of parental language in the home environments of hard-of-hearing 2-year-olds. *Journal of Deaf Studies and Deaf Education, 17*(4), 402-420.

*** Vohr, B. R., Topol, D., Watson, V., St Pierre, L., & Tucker, R.** (2014). The importance of language in the home for school-age children with permanent hearing loss. *Acta Paediatrica, 103*(1), 62-69.

Wake, M., Poulakis, Z., Hughes, E. K., Carey-Sargeant, C., & Rickards, F. W. (2005). Hearing impairment: A population study of age at diagnosis, severity, and language outcomes at 7-8 years. *Archives of Disease in Childhood, 90*(3), 238-244.

Webb, N. M., & Shavelson, R. J. (2005). Generalizability theory: Overview. *Encyclopedia of statistics in behavioral science*, 2, 717-719.

Weisleder, A., & Fernald, A. (2013). Talking to children matters. *Psychological Science, 24*(11), 2143-2152.

Xu, D., Yapanel, U., & Gray, S. (2009). Reliability of the LENA Language Environment Analysis system in young children's natural home environment [Technical report LTR-05-2]. Boulder, CO: LENA Research Foundation.

Zambrana, M. l., Ystrom, M. E., & Pons, M. F. (2012). Impact of gender, maternal education, and birth order on the development of language comprehension: A longitudinal study from 18 to 36 months of age. *Journal of Development & Behavioral Pediatrics, 33*(2), 146-155.

Zimmerman, F. J., Gilkerson, J., Richards, J. A., Christakis, D. A., Xu, D., Gray, S., & Yapanel, U. (2009). Teaching by listening: the importance of adult-child conversations to language development. *Pediatrics, 124* (1), 342-349.

Zimmerman, I. L., Steiner, V. G., & Pound, R. E. (2002). *Preschool Language Scales (PLS-4)* (4th ed.). Camberwell, Australia: Pearson Australia Group.

Fig.1 Hierarchical representation of selected articles according to inclusion and exclusion criteria, based on the PRISMA Checklist (Moher et al., 2009)

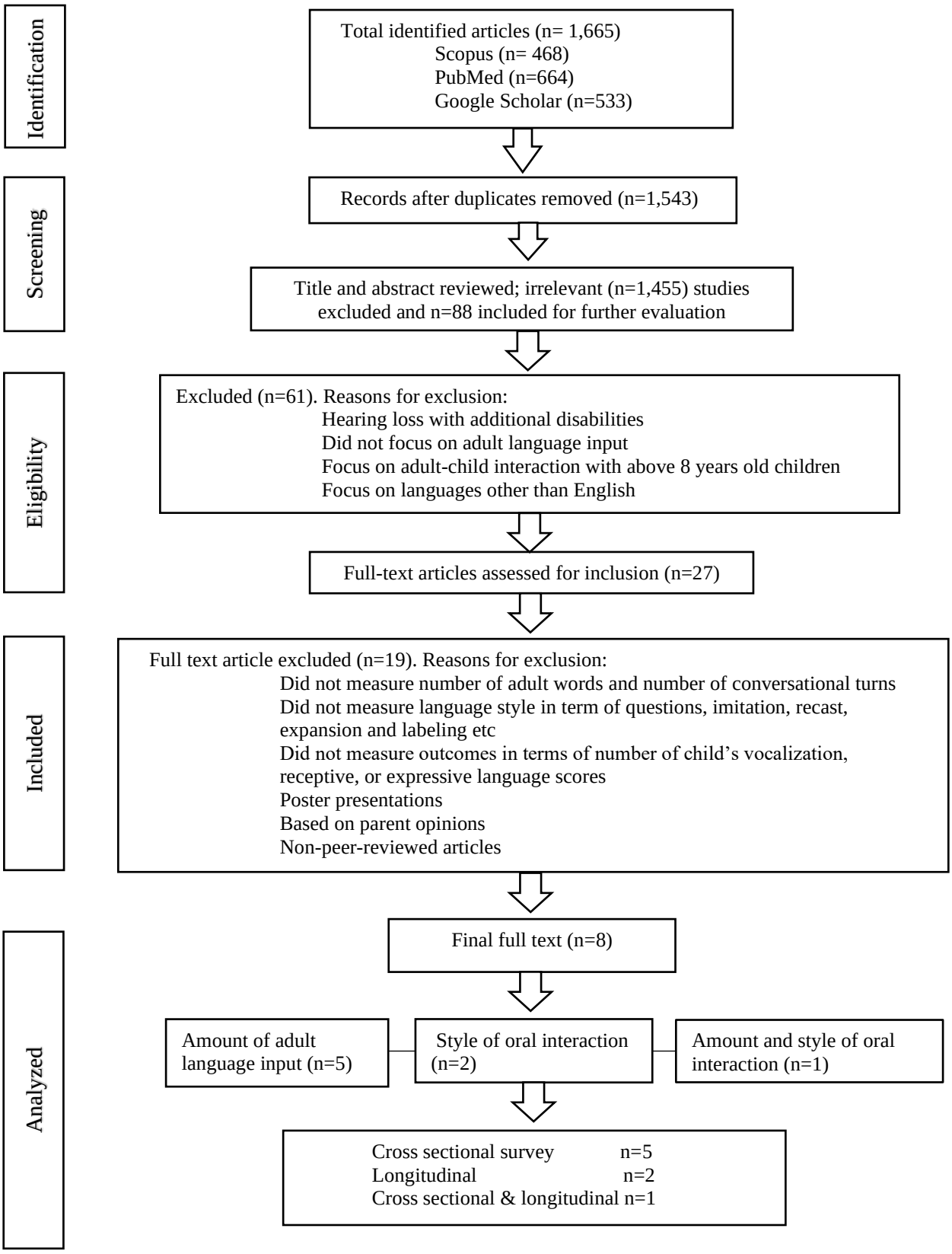

TABLE 1. Quality of studies based on the critical appraisal tool for cross-sectional studies (CAT-CSS) and level of evidence for all studies those were reviewed for the amount and style of oral interaction in relation to language outcomes in CwHL

Section 2, Study Validity

For each criterion, document the appropriate response, according to how you think it is addressed:

Studies	DesJardin & Eisenberg (2007)	McDaniel & Purdy (2011)	Aragon & Yoshinaga-Itano (2012)	VanDam et al. (2012)	Cruz et al. (2013)	Ambrose et al. (2014)	Vohr et al. (2014)	Ambrose et al. (2015)
ABSTRACT AND INTRODUCTION OF THE STUDY								
1. Abstract is presented in an informative and balanced summary **of what was done and what was found.**	✓	✓	×	✓	×	✓	✓	✓
2. Sufficient scientific background information on the topic.	×	×	×	×	×	✓	×	✓
3. Introduction is **focused, relevant, in logical fashion and justifiable to the research question.**	×	×	×	×	×	×	×	✓
4. Burden of disease/**condition is quantified** to magnify the magnitude of the problem in a particular population	×	×	×	×	×	×	×	×
5. Introduction is zoomed into **regional or national** perspective if applicable.	×	×	×	×	×	×	×	×
6. Introduction is ended with the **aim** of the study.	✓	✓	×	✓	✓	✓	✓	✓
Number of covered criteria out of 6	**2**	**2**	**0**	**2**	**1**	**3**	**2**	**4**
Grading judgment (index score=6)								
Poor **< 3 criteria**	✓	✓	✓	✓	✓		✓	
Good **3-4 criteria**						✓		✓
Excellent **> 4 criteria**								
AIM AND QUESTION/S OF THE STUDY								
1. **Aim** is descriptive and clearly stated**.**	✓	✓	×	✓	✓	✓	✓	✓
2. Aim is SMART**: Specific, Measurable, Achievable**, resourced (within the project budget) and Time Bound.	✓	✓	×	✓	✓	✓	✓	✓
3. **Question**/s of study is adequately described.	✓	✓	✓	×	×	×	×	✓

4. Type of research question/s is corresponded to the study design**.**	✓	✓	✓	×	×	×	×	✓
Number of covered criteria out of 4	**4**	**4**	**2**	**2**	**2**	**2**	**2**	**4**
Grading judgment (index score=4)								
Poor **<2 criteria**								
Good **2-3 criteria**			✓	✓	✓	✓	✓	
Excellent **> 3 criteria**	✓	✓						✓
METHODS								
STUDY DESIGN/SETTING AND TIMEFRAME								
1. Study **design** is **clearly presented**.	×	×	×	✓	✓	×	✓	✓
2. Study **design** is **justified**.	×	×	×	✓	✓	×	✓	✓
3. Study **setting, or a location** is described.	✓	✓	×	×	✓	×	✓	✓
4. Study **timeframe** is clearly illustrated.	✓	✓	×	✓	✓	✓	×	✓
5. Study timeframe seems **appropriate.**	×	×	×	×	✓	✓	×	✓
Number of covered criteria out of 5	**2**	**2**	**0**	**3**	**5**	**2**	**3**	**5**
Grading judgment (index score=5)								
Poor **< 3 criteria**	✓	✓	✓			✓		
Good **3-4 criteria**				✓			✓	
Excellent **> 4 criteria**					✓			✓
SAMPLING								
1. Sample is selected and **representative of reference** population.	×	×	×	×	×	×	×	×
2. The **methods of sample selection** are clearly described.	×	×	×	×	×	×	×	×
3. Appropriate **sample technique** is used with ensured randomization.	×	×	×	×	×	×	×	×

4. Specific description of **inclusion criteria.**	✓	✓	×	✓	✓	✓	✓	✓
5. Specific description of **exclusion criteria.**	×	×	×	×	×	×	✓	×
6. **Sample size estimates** have been performed.	×	×	×	×	×	×	×	×
7. Sample size seems feasible (considering resources/ prevalence of disease/ study population, etc.).	×	×	×	×	×	×	×	×
8. The chosen level of precision, confidence limit, and variability) estimated proportion of an attribute that is present in the population) are adequate for the study question	×	×	×	×	×	×	×	×
9. A high participation level.	×	×	×	×	×	×	×	×
10. The subjects covered in the study could be sufficiently similar from your population to cause concern.	✓	×	✓	✓	✓	✓	✓	✓
Number of covered criteria out of 10	**2**	**1**	**1**	**2**	**2**	**2**	**3**	**2**

Grading judgment (index score=10)

Poor **<5 criteria**	✓	✓	✓	✓	✓	✓	✓	✓
Good **5-6 criteria**								
Excellent **> 6 criteria**								

DATA COLLECTION AND ETHICAL ISSUES

1. The methods for data collection are described for each of the variables collected (**where, by who and when**)	✓	×	×	×	✓	×	×	✓
2**. Content and face validity of the all tools are** well described	×	×	×	✓	×	✓	×	×
3. Data collection tools are tested for its **reliability.**	✓	✓	×	✓	✓	✓	✓	✓
4. The study specifies **who are the data collectors** and their background.	✓	×	×	×	✓	×	×	×
5. Exposure factor/s is/are identified	✓	✓	✓	✓	✓	✓	✓	✓
6. Outcome/s is/are ascertained:	✓	✓	✓	✓	✓	✓	✓	✓
7. Exposure and outcomes are measured at one **specific point in time.**	✓	✓	✓	✓	✓	✓	✓	✓
8. **Potential confounding factors** are measured accurately.	✓	✓	×	✓	✓	✓	✓	✓

9. Measures were made to contact **non-responders**	×	×	×	×	×	×	×	×
10. **Ethical issues** are mentioned clearly (if appropriate).	×	✓	×	×	×	✓	✓	×
Number of covered criteria out of 10	**7**	**6**	**3**	**6**	**7**	**7**	**6**	**6**

Grading judgment (index score=10)

Poor **< 5 criteria**		✓	✓				✓	
Good **5-6 criteria**	✓			✓	✓	✓		✓
Excellent **> 6 criteria**								

RESULTS

1. The results are **adequately, objectively, and explicitly described**	✓	✓	×	✓	✓	✓	✓	✓
2. Characteristics of study participants (e.g. **demographic, clinical, and social**) are presented.	×	✓	×	✓	✓	×	✓	✓
3. Exposure variables **are associated with outcome variables.**	✓	✓	×	✓	✓	✓	✓	✓
4. Tables and figures are adequate, clear, and appropriately titled.	✓	✓	✓	✓	✓	✓	✓	✓
5. Is appropriate statistical analyses be used?	✓	×	×	✓	✓	✓	✓	✓
6. The study mentions if **negative results** or results of **no effect**/difference are considered for publication.	×	×	×	×	×	×	×	×
Number of covered criteria out of 6	**4**	**4**	**1**	**5**	**5**	**4**	**5**	**5**

Grading judgment (index score=6)

Poor **< 3 criteria**			✓					
Good **3-4 criteria**	✓	✓				✓		
Excellent **> 4 criteria**				✓	✓		✓	✓

DISCUSSION/CONCLUSION AND RECOMMENDATIONS

1. The results are **summarized** and discussed in relation to the original research questions	✓	✓	✓	✓	×	×	×	✓
2. The researcher has **discussed the credibility** of their results.	×	×	×	×	×	×	×	×

3. There is adequate discussion of the evidence for the researchers' **arguments**	✓	×	×	✓	✓	×	×	✓
4**. Limitations** of the study are discussed, taking into account sources of potential bias or imprecision.	×	✓	×	✓	✓	✓	✓	✓
5. Discussion shows the contribution of the study to the body of knowledge and existing evidence base.	✓	✓	×	✓	✓	✓	×	✓
6. The results suggest a **more rigorous** study is needed.	×	✓	×	✓	×	✓	×	✓
7. The authors mention how the study results will be used, i.e. potential implications for actions.	✓	×	✓	✓	✓	✓	✓	×
Number of covered criteria out of 7	**4**	**4**	**2**	**6**	**4**	**4**	**2**	**5**

Grading judgment (index score=7)

Poor **< 4 criteria**			✓					
Good **4-5 criteria**	✓	✓			✓	✓	✓	✓
Excellent **> 5 criteria**				✓				

REFERENCES

1.References are adequate and relevant to the study topic	✓	✓	×	✓	✓	✓	✓	✓
2. References are up-to-date.	×	×	×	✓	✓	✓	×	×
Number of covered criteria out of 2	**1**	**1**	**0**	**2**	**2**	**2**	**1**	**1**

Grading judgment (index score=2)

Poor **< 1 criteria**			✓					
Good **1 criteria**	✓						✓	✓
Excellent **2 criteria**		✓		✓	✓	✓		

Overall Quality Scoring of the Study

Number of covered criteria out of 50

% of covered criteria: (obtaining score/ total score× 100= covered %)	26/50×100=52%	24/50×100=48%	9/50×100=18%	28/50×100=56%	28/50×100=56%	26/50×100=52%	24/50×100=48%	32/50×100=64%
Poor: if less than 50%		✓	✓				✓	
Good: if 50% to 65%	✓			✓	✓	✓		✓
Excellent: if more than 65%								
Level of evidence according to ASHA guidelines	Level 3	Level 3	Level 3	Level 3	Level 3	Level 3	Level 3	Level 3

*Studies are arranged in ascending ordered from left to right according to the year of publication

TABLE 2. Summary of critique of method and procedure of included studies

Study	Sample Size	Age Range; Mean Age (M); Standard Deviation (SD)	Sample Selection Procedure	Study Design	Measurement Tools	Observation Duration	Effect Size	Confounders/Gaps
DesJardin & Eisenberg (2007)	CwHL = 32	30-86 mos.; *M* = 57.3; *SD* = 15.68	PR	Cross- sectional study	Video recordings SPISE RDLS-3	10 video tape sessions for 5-7 mis	NA	• Not considered natural language environments • Short time observations/ recordings were considered • No comparison with control group • Information of number of siblings and birth order was not reported
McDaniel & Purdy (2011)	CwHL = 8	2;6-4;2 yrs.; *M* = 3;1; *SD* = 0.65	PR	Cross- sectional study	LENA® recordings PLS-4	3 full day LENA® recordings for 8-12 hrs/day	Effect sizes could not be estimated due to lack of information on mean and SD	• Small sample size • No comparison with control group • Information of number of siblings and birth order was not reported
Aragon & Yoshinaga-Itano (2012)	English CwHL = 24 LENA® normative Data = 329 Spanish CwHL = 10 Spanish CwNH = 10	2-36 mos.; NR; NR	PR	Cross- sectional study	LENA® recordings	1 full day LENA® recording for 10-16 hrs	Effect sizes could not be estimated due to lack of information on SD	• No comparison with control group with English CwHL • Information of severity of hearing loss, age of hearing loss identification, use of type of hearing devices, age at fitting of hearing devices and, number of siblings and birth order were not reported
VanDam et al. (2012)	CwHL = 22 CwNH = 8	24-36 mos.; M_{HL} = 29.4; SD_{HL} = 2.9, M_{NH} = 30.1; SD_{NH} = 3.8	PR	Cross- sectional study	LENA® recordings MSEL SII PTA	1full-day LENA®, no time description	Effect sizes calculated based on reported data for AWs/hr in CwHL vs CwNH (Cohen's d) = 0.41, CTs in CwHL vs CwNH (Cohen's d) = 0.29	• Information of type and duration of therapeutic intervention, and number of siblings and birth order was not reported

Cruz et al. (2013)	CwHL = 93	12-36 mos.; *M* = 14.67; *SD* =5.76	PR	Longitudinal study	Video recordings RDLS-3 Questionnaire	2 video recordings, no time description	NA	• Not considered natural language environments • Information of number of siblings was not reported
Ambrose et al. (2014)	CwHL = 28	20-30 mos.; *M* = 25.8; *SD* = 3.1	PR	Cross-sectional study	LENA® recordings MSEL CASL	6 full day LENA® recordings, 8- 16 hrs/day	NA	• No comparison with control group • Information of age of hearing loss identification, type and duration of therapeutic intervention, and number of siblings and birth order was not reported
Vohr et al. (2014)	CwHL = 23 CwNH = 41	6-8 yrs.; M_{HL} = 82.7 mos.; SD_{HL} = 6, M_{NH} = 81.9 mos.; SD_{NH} = 7	PR	Longitudinal study	LENA® recordings RDLS-3	1 full day LENA® recording for 12 hrs	Effect sizes calculated based on reported data for AWs in CwHL vs CwNH (Cohen's d) = 0.10, CTs in CwHL vs CwNH (Cohen's d) = 0.30	• Information of age of hearing loss identification, and number of siblings and birth order was not reported
Ambrose et al. (2015)	CwHL = 156 CwNH = 59	<u>At 18 mos. visit:</u> *CwHL*: 16-21; *M* = 18.6; *SD* = 1.2, *CwNH*: 17-21; *M* = 19.1; *SD*=1.2, <u>At 3 yrs. visit:</u> *CwHL*: 34-44; *M* = 37.7; *SD*=2.8, *CwNH*: 33-44; *M* = 37.5; *SD* = 2.8	CR	Cross-sectional & Longitudinal study	Video recordings CASL	1 video recording for 5 mis	Reported effect sizes for CwHL vs CwNH group differences of NTW (Cohen's d) = 0.37 at18-mos. & NTW (Cohen's d) = 0.50 at 3 yrs. visit	• Not considered natural language environments • Information of type and duration of therapeutic intervention, and number of siblings was not reported

*Studies are arranged in ascending ordered according to the year of publication

CASL = Comprehensive Assessment of Spoken Language; CR = Completely Reported; CTs = Conversational Turns; CwHL = Children with Hearing Loss; CwNH = Children with Normal Hearing; Hrs. = Hours; LENA = Language ENvironment Analysis; Mis = Minutes; Mos. = Months; MSEL = Mullen Scales of Early Learning; NTW = Number of Total Words (adult); NA = Not applicable; NR = Not reported; PLS-4 = Pre-School Language Scale 4th edition; PR = Partially Reported; PTA = Pure Tone Audiometry; RDLS-3 = The Reynell Developmental Language Scales 3rd edition; SII = Speech Intelligibility Index; Yrs. = Years

TABLE 3. Summary of studies for the difference in the amount of language input between CwHL and CwNH

Study	Comparison Groups		Input: Mean; Standard Deviation		Significant Difference between Groups (CwHL vs CwNH)		General Conclusion
	CwHL	CwNH	AWs	CTs	AWs	CTs	
McDaniel & Purdy (2011)	8	Missing number (LENA® normative data)	M_{HL}= 0.52, $M_{LENA®}$ = NR; SD_{HL}= 0.31, $SD_{LENA®}$ = NR	M_{HL}= 0.31, $M_{LENA®}$ = NR; SD_{HL} = 0.59, $SD_{LENA®}$ = NR	NR	NR	The number of AWs and CTs were higher in CwHL than the data published by LENA®
Aragon & Yoshinaga-Itano (2012)	24	329 (LENA® normative data)	M_{HL}=17,605, $M_{LENA®}$ = NR; SD_{HL}= NR, $SD_{LENA®}$ = NR	M_{HL}= 644, $M_{LENA®}$ = 462; SD_{HL}= NR, $SD_{LENA®}$ = NR	NR	NR	The comparison between the number of AWs in CwHL and CwNH is not clear due insufficient information on mean and SD of CwNH
VanDam et al. (2012)	22	8	M_{HL} = 1249, M_{NH} = 1397; SD_{HL}= 326, SD_{NH} = 386	M_{HL} = 61, M_{NH} = 66; SD_{HL}= 17, SD_{NH} = 17	P = .526	P = .511	The number of AWs and CTs were equal between CwHL and CwNH
Vohr et al. (2014)	23	41	M_{HL}= 1415.9, M_{NH} = 1358.2; SD_{HL}= 486, SD_{NH} = 625	M_{HL} = 55.1, M_{NH} = 48.1; SD_{HL} = 24, SD_{NH} = 23	P = .070	P = .260	The number of AWs and CTs were higher in CwHL than CwNH
Ambrose et al. (2015)	71	18	At 18 mos.; M_{HL} = 320.9, M_{NH} = 364.9; SD_{HL} = 125.4, SD_{NH} = 115.2	NA	P = .18	NA	The number of AWs were lower in CwHL than CwNH at 18 mos
	85	41	At 3 yrs.; M_{HL} = 383.1, M_{NH} = 433.8; SD_{HL} = 91.3, SD_{NH} = 109.5	NA	P = .01	NA	The number of AWs were higher in CwHL than CwNH at 3 yrs

*Studies are arranged in ascending ordered according to the year of publication

AWs = Adult Words; CTs = Conversational Turns; CwHL = Children with Hearing Loss; CwNH = Children with Normal Hearing; Mos. = Months; NA = Not Applicable; NR = Not Reported; Yrs. = Year

TABLE 4. Summary of studies for the association between the amount of language input and language outcomes in CwHL

Study	Input	Association between input and outcomes			General Conclusion
		CVs	Receptive	Expressive	
McDaniel & Purdy (2011)	AWs	NR	PLS-4: *Rs* = 0.71, *p* > .05	PLS-4: *Rs* = 0.27, *p* > .05	The number of AWs were negatively associated with receptive and expressive language scores
	CTs	*Rs* = 1.0, *p* < .05	PLS-4: *Rs* = 0.35, *p* > .05	PLS-4: *Rs* = 0.00, *p* > .05	The number of CTs were positively associated with number of CVs but not receptive and expressive language scores
VanDam et al. (2012)	AWs	NA	MSEL: *p* > .05	NA	The number of AWs were not significantly associated with receptive language scores
	CTs	NA	MSEL: *r* = .62, *p* < .01	NA	The number of CTs were positively associated with receptive language scores
Ambrose et al. (2014)	AWs	NA	Age = 2 yrs, MSEL: *r* = 0.339, *p* > .05	MSEL: r = 0.138; *p* > .05	The number of AWs were not significantly associated with receptive and expressive language outcomes
			Age = 3yrs, CASL (composite) receptive and expressive language score: *r* = 0.02; p > .05		
	CTs	NA	Age = 2 yrs MSEL: *r* = 0.61, *p* < .01	MSEL: *r* = 0.45; *p* < .05	The number of CTs were significantly associated with receptive and expressive language outcomes
			Age = 3 yrs, CASL (composite) receptive and expressive language score: *r* = 0.45, p < .05		
Vohr et al. (2014)	NA	NA	NA	NA	
	CTs	NA	RDLS-3: *r* = 0.35, *p* = .03	RDLS-3: *r* = 0.89, *p* = .02	The number of CTs were significantly associated with receptive and expressive language scores
Ambrose et al. (2015)	Age = 18 mos, Quality variables (NTW, NTU)	N/A	Age = 3yrs, CASL score: R^2 = 0.09, *p* = .24		The quantity variables (NTW, NTU) was not significantly associated with language outcomes

*Studies are arranged in ascending ordered according to the year of publication

AWs = Adult Words; CTs = Conversational Turns; MSEL = Mullen Scales of Early Learning; Mos. = months; NA = Not Applicable; NTW= Number of total words; NTU = Number of total utterances; NR = Not Reported; PLS-4 = Pre-School Language Scale - 4th edition; RDLS-3 = The Reynell Developmental Language Scales 3rd edition; Yrs. = Years

TABLE 5. Summary of studies for the association between the style of language input and language outcomes in CwHL

Study	Input (style)	Association between input (style) and language outcomes		General Conclusion
		Receptive	Expressive	
DesJardin & Eisenberg (2007)	Parallel Talk	RDLS-3; $r = 0.27$, $p > .05$	RDLS; $r = 0.13$, $p > .05$	Concurrent correlation showed the frequency of recast was significantly associated with receptive language scores, use of open-ended questions were significantly associated with expressive language scores
	Expansion	RDLS-3; $r = 0.33$, $p > .05$	RDLS; $r = 0.21$, $p > .05$	
	Recast	RDLS-3; $r = 0.47$, $p < .01$	RDLS; $r = 0.27$, $p > .05$	
	Open-ended question	RDLS-3; $r = 0.34$, $p > .05$	RDLS; $r = 0.51$, $p < .01$	
	Linguistic mapping	RDLS-3; $r = -0.50$, $p < .01$	RDLS-3; $r = -0.42$, $p < .05$	
	Close-ended question	RDLS-3; $r = 0.17$, $p > .05$	RDLS-3; $r = 0.11$, $p > .05$	
	Imitation	RDLS-3; $r = -0.27$, $p > .05$	RDLS-3; $r = -0.23$, $p > .05$	
	Label	RDLS-3; $r = -0.44$, $p < .05$	RDLS-3; $r = -0.45$, $p < .05$	
	Directive	RDLS-3; $r = -0.58$, $p < .01$	RDLS-3; $r = -0.49$, $p < .05$	
	Comment	RDLS-3; $r = 0.15$, $p > .05$	RDLS-3; r = 0.07, $p > .05$	
Cruz et al. (2013)	Lower-lever FLTs (linguistic mapping, comments, imitation, labeling, directive, and close-ended question)	RDLS-3 (12 mos.); $t = 1.03$, $p < .05$ RDLS-3 (24 mos.); $t = 1.03$, $p < .05$ RDLS-3 (36 mos.); $t = 1.04$, $p < .05$	RDLS-3 (12 mos.); $t = 1.88$, $p > .05$ RDLS-3 (24 mos.); $t = 1.88$, $p > .05$ RDLS-3 (36 mos.); $t = 1.98$, $p > .05$	Repeated measures indicated lower level FLTs did not predict change in expressive language or receptive language scores over 3 yrs
	High-level FLTs (parallel talk, open-ended question, expansion, recast)	RDLS-3 (12 mos.); $t = 1.74$, p < .08 RDLS-3 (24 mos.); $t = 1.82$, p < .08 RDLS-3 (36 mos.); $t = 1.85$, p < .08	RDLS-3 (12 mos.); $t = 3.00$, $p < .01$ RDLS-3 (24 mos.); $t = 2.86$, $p < .01$ RDLS-3 (36 mos.); $t = 2.79$, $p < .01$	Repeated measures indicated high level FLTs were a strong predictor for the development of receptive and expressive language skills over 3 yrs
Ambrose et al. (2015)	Directives	CASL (cumulative receptive and expressive language score); $r = -0.14$, $p = .03$		Longitudinal relationship of proportion of use of direct utterances by parents at 18 mos was negatively correlated with language scores at 3 yrs. High-level at 18 mos did not contribute a significant variance for language outcome at 3 yrs
	High-level "conversational eliciting"	CASL (cumulative receptive and expressive language score); $\beta = -0.24$, $p = .18$		

*Studies are arranged in ascending ordered according to the year of publication

FLTs = Facilitative Language Techniques; CASL = Comprehensive Assessment of Spoken Language; Mos. = Months; NR = Not Reported; RDLS-3 = The Reynell Developmental Language Scales 3rd edition; Yrs. = Years