Journal of Deaf Studies and Deaf Education, 2023, 1–13
https://doi.org/10.1093/deafed/enad054
Theory/Review Manuscript

Measuring speech intelligibility with deaf and hard-of-hearing children: A systematic review

Harpa Stefánsdóttir[1,*], Kathryn Crowe[1,2], Egill Magnússon[1], Mark Guiberson[3], Thora Másdóttir[1], Inga Ágústsdóttir[1], Ösp V. Baldursdóttir[1]

[1]Department of Medicine, School of Health Sciences, University of Iceland, Reykjavik, Iceland
[2]School of Education, Charles Sturt University, Bathurst, NSW, Australia
[3]Division of Communication Disorders, College of Health Sciences, University of Wyoming, United States
*Corresponding author: Charles Sturt University, School of Education, Panorama Avenue, Bathurst, NSW 2795, Australia. Email: kcrowe@csu.edu.au

Abstract

There is great variability in the ways in which the speech intelligibility of d/Deaf and hard-of-hearing (DHH) children who use spoken language as part, or all, of their communication system is measured. This systematic review examined the measures and methods that have been used when examining the speech intelligibility of children who are DHH and the characteristics of these measures and methods. A systematic database search was conducted of CENTRAL; CINAHL; Cochrane; ERIC; Joanna Briggs; Linguistics, Language and Behavior Abstracts; Medline; Scopus; and Web of Science databases, as well as supplemental searches. A total of 204 included studies reported the use of many different measures/methods which measured segmental aspects of speech, with the most common being Allen et al.'s (2001, The reliability of a rating scale for measuring speech intelligibility following pediatric cochlear implantation. *Otology and Neurotology, 22*(5), 631-633. https://doi.org/10.1097/00129492-200109000-00012) Speech Intelligibility Rating scale. Many studies included insufficient details to determine the measure that was used. Future research should utilize methods/measures with known psychometric validity, provide clear descriptions of the methods/measures used, and consider using more than one measure to account for limitations inherent in different methods of measuring the speech intelligibility of children who are DHH, and consider and discuss the rationale for the measure/method chosen.

Hearing loss is one of the most common sensory impairments and affects people in every stage of life and in every corner of the globe (Newsted et al., 2020). According to the World Health Organization (WHO), around one and a half billion people worldwide live with some degree of hearing loss, and 34 million children require rehabilitation for hearing loss (WHO, 2021). Children who are d/Deaf and hard of hearing (DHH) and acquiring a spoken language as part/all of their communication system often have reduced or different access to spoken language, as speech perception can be challenging (Spencer et al., 2011). This can have consequences for speech production skills and speech intelligibility, which refers to how clear a person's speech is, or how much of a person's speech is understood by a listener (Spencer et al., 2011). The production of intelligible speech relies on a child mastering multiple aspects of spoken language, the central aspects being speech perception, phonological knowledge, and motor control of articulation, timing, and intonation (Freeman, 2018).

The intelligibility of speech is dependent on a two-way interaction between a speaker and a listener, so factors that influence speech intelligibility are therefore related to both the speaker and the listener (McLeod, 2020). Factors related to the speaker can be the speaker's age, gender, language/dialect, and the presence of speech sound disorder or other communication disabilities or differences (Lagerberg et al., 2014; McLeod, 2020). Reduced speech intelligibility can stem from many sources such as speech disorders (e.g., speech sound disorders, childhood apraxia of speech, cleft palate, dysarthria), sensory differences (e.g., hearing loss, wearing a mask), and linguistic differences (e.g., accents, dialect differences). Factors related to the listener can be the listener's familiarity and relationship with the speaker, the speaker's language or dialect, and their familiarity with speech patterns associated with different communication disorders (Baudonck et al., 2009; McLeod, 2020; van Doornik et al., 2018). For example, Overby et al. (2007) found that second grade teachers had a negative attitude towards students with articulation disorders, and their academic, social, and behavioral expectations were significantly lower for the children with speech sound disorders than their typically developing peers. Further, speech intelligibility is also determined by the linguistic task being performed and the context/environment in which it is being performed (McLeod, 2020). Factors related to the task can be the type of speech sample examined, the tool/method used to measure intelligibility, and the response type used by a measure, and the validity and reliability of a measure/method (Ertmer, 2011; McLeod, 2020). These issues will be discussed in depth in the next section. Finally, factors related to the context/environment that impact intelligibility can, for example, be the presence of background noise which can be

Received: March 21, 2023. **Revised:** September 21, 2023. **Accepted:** October 30, 2023

distracting for both the listener and the speaker (Ishikawa et al., 2017; McLeod, 2020).

Speech Intelligibility and Children who are DHH

The central goal of intervention for children who are DHH who use spoken language as part/all of their communication system is for these children to develop intelligible connected speech (Ertmer, 2011). For children who use sign as part/all of their communication system sign intelligibility should be considered (see Crowe et al., 2019). Children who are DHH often develop intelligible speech at a slower rate than their typically hearing peers (Chin et al., 2001). However, since the 1980s, the possibilities for the development of intelligible speech have been continually improving for children who are DHH who are acquiring a spoken language as part/all of their communication system. These include universal newborn hearing screening, improvements in hearing aid technology, cochlear implantations being invented and improved, and earlier ages of identification, intervention, and implantation (Crowe & Guiberson, 2019; Flipsen Jr., 2008). Many studies have looked at factors related to better speech intelligibility outcomes for children who are DHH. These include language skills (e.g., Castellanos et al., 2014), age at implantation (e.g., Ching et al., 2010; Ching et al., 2017; Ching & Dillon, 2013; Habib et al., 2010), hearing thresholds (e.g., Sennaroğlu et al., 2016), interventions for hearing loss (e.g., Baudonck et al., 2011), family economic status (e.g., Cupples et al., 2017; Gérard et al., 2010), social integration (e.g., Freeman, 2018; Freeman & Pisoni, 2017; Most et al., 2012), and better speech perception skills (e.g., Blamey et al., 2001).

Speech intelligibility difficulties experienced by children who are DHH who use spoken language as part/all of their communication system usually stem from difficulties with speech perception, as the speech signal through which they are acquiring spoken language is degraded/restricted. Difficulties with perception can emerge in the speech of children who are DHH in many ways, such as in their use of delayed and disordered phonological processes (Ching & Dillon, 2013), vowel errors (Sosa & Bunta, 2019), and different voice characteristics (Coelho et al., 2015). Any, or all, of these differences can negatively impact the speech intelligibility of children who are DHH. The contribution of variables related to these children's individual hearing, speech, and language characteristics have been widely researched. Greater access to language through audition has consistently been found to influence speech intelligibility outcomes. Castellanos et al. (2014) found that early measures of speech intelligibility skills predict later speech and language outcomes, and better language skills were associated with better speech intelligibility. Ching et al. (2010) and Ching & Dillon (2013) reported that later age at CI switch-on was significantly associated with poorer speech and language outcomes while Ching et al. (2017) found that earlier intervention was related to better outcomes, with the benefit being greatest for the children with the poorest hearing thresholds. Similarly, Sennaroğlu et al. (2016) reported a correlation between speech intelligibility scores and hearing thresholds, with better hearing thresholds associated with better speech intelligibility. A study by Baudonck et al. (2011) found that the speech intelligibility ratings for "normal hearing children" (NH; p. 912) were the highest, followed by children with bilateral CIs (biCI), unilateral CIs (uniCI), and finally children using hearing aids (HAs). Other impacts on the speech intelligibility of children who are DHH and use spoken language as part/all of their communication system have also been considered. Gérard et al. (2010) found the communication abilities of children with hearing loss from higher socioeconomic level were consistently better compared to those from lower socioeconomic levels, and Cupples et al. (2017) reported that higher levels of maternal education were associated with better speech production outcomes.

Differences in speech intelligibility can also have broader impact on the lives and wellbeing of children who are DHH and use spoken language through listener's judgments about a person on characteristics that are unrelated to speech. Freeman (2018) found when speech samples from people with typical hearing and cochlear implant (CI) users with different levels of speech intelligibility were rated by naïve listeners on different personality traits, the typically hearing group was rated more positively on most traits, while the less intelligible CI users were rated more negatively. Research by Most et al. (2012) reported a positive correlation between better speech intelligibility, and better social competence and less feelings of loneliness in children with hearing loss. Freeman and Pisoni (2017) examined at the link between speech intelligibility and psychosocial functioning. Preschool-aged CI users' speech intelligibility was correlated with the psychosocial constructs of functional communication, attention problems, atypicality, withdrawal, and adaptability. For school-aged CI users, additional scales were significantly correlated: leadership, activities of daily living, anxiety, and depression. These findings suggest that the benefits of good speech intelligibility extend far beyond purely clear communication.

Speech Intelligibility Measurements

As children who are DHH are at risk of delayed and/or poor speech intelligibility outcomes in their development of spoken language, it is important to monitor their speech intelligibility throughout their habilitation journey (Coppens-Hofman et al., 2016; Miller, 2013). As with all assessments, it is crucial that measures/methods are reliable and valid to ensure that they are able to accurately measure children's speech intelligibility skills and change in skills (Ertmer, 2011; Miller, 2013). Two main types of speech intelligibility assessments have been used in research and clinical settings: scaling procedures and item-identification procedures (Ertmer, 2011).

Scaling Procedures

Scaling procedures are based on listeners rating speech samples along a continuum of intelligibility. The speech samples may represent different contexts, for example, general conversation, reading passages, picture descriptions, or replies to fixed questions (Ertmer, 2011; Miller, 2013). An example of a scaling procedure is when a listener rates audio-recordings of a speech sample using a Likert or descriptive numeric scale to describe how well the speaker in the recording was understood by the listener (Crowe et al., 2019; Ertmer, 2011). The speech samples can represent different linguistic contexts, e.g., single words, phrases, conversation, or reading written text (Crowe et al., 2019). Rating scales can also be used with different listeners that vary in familiarity with the speaker, for example parents, educators, clinicians, and naïve listeners (Coppens-Hofman et al., 2016; Ertmer, 2011).

The use of scaling procedures for measuring speech intelligibility has been found to have both advantages and disadvantages. The advantages of these types of scales are that they are quick and easy to complete (Ertmer, 2011). They can also represent intelligibility in real life situations. Scaling procedures have, however, often been criticized for having poor validity and reliability (Ertmer, 2011). Listener's interpretation of the level of intelligibility that each scale point represents is subjective and therefore inter-rater reliability when using scales may be poor (Ertmer, 2011). To one listener six might mean pretty good while to another it might mean not very good (Ertmer, 2011). However, this is not true for all scales, as some use descriptive words or sentences rather than numbers (e.g., Allen et al., 2001; McLeod, 2020). Scaling can also be insensitive to speech that falls in the middle range of intelligibility, as listeners might not be able to quantify the differences between speech that is 30% intelligible and speech that is 60% intelligible (Ertmer, 2011; Miller, 2013). This means that ratings on speech intelligibility with scaling procedures may be inconsistent across users and contexts (Ertmer, 2011; Miller, 2013). The use of measurement scales is commonly reported in research describing children's speech intelligibility. Two commonly used scales are the Intelligibility in Context Scale (ICS; McLeod et al., 2012) and the Speech Intelligibility Rating scale (SIR; Allen et al., 2001).

Item-Identification Procedures

Item-identification procedures are an objective way to measure speech intelligibility that is particularly well suited to measurement of segmental features of speech (e.g., consonants and vowels). This method was developed partly as a way of solving some of the reliability issues related to rating scales (Miller, 2013). These methods are based on listeners writing down the words they understand when listening to speech samples. The samples are then scored on how often the listener's written responses match the words produced and speech intelligibility is calculated from those scores (Ertmer, 2011; Miller, 2013). Again, the sample types evaluated can represent different linguistic contexts, e.g., single words, phrases, conversation, or written text. Listeners can also vary in familiarity with the speaker and the speech of children who are DHH, for example, parents, educators, experienced clinicians, and naïve listeners (Ertmer, 2011).

Item-identification procedures have an advantage over scaling procedures as they do not only rely on subjective impressions, but measure an actual number of words understood by a listener (Ertmer, 2011). They can also help clinicians with shaping therapeutic directions as they can summarize the types of words that were not understood in terms of length and complexity and other segmental features (Miller, 2013). However, item-identification procedures are not without flaws. They are more time consuming than scaling procedures and can involve difficulties related to the criteria used to construct word lists and weighting errors (Crowe et al., 2019; Ertmer, 2011). For logistic reasons, item-identification procedures usually involve a list of words that the test is limited to, so that representative sounds, sound combinations, and sound positions in a language are included. Some tests have an alternative but equivalent set of words for repeat assessment, and the sound content and syllable structure of words is not necessarily the same. This can make comparisons of the two measurements invalid (Miller, 2013). Two commonly used item-identification procedures are The Beginners' Intelligibility Test (BIT; Osberger et al., 1994), and Percentage of Consonants Correct (PCC; Shriberg & Kwiatkowski, 1982).

Objectives of this Systematic Review

Appropriate assessment of speech intelligibility is a key concern for professionals who work with children who are DHH and in the process of acquiring a spoken language for the first time. Professionals working with children who are DHH should have an appreciation of the variety of choices available for assessing speech intelligibility in terms of the measures/methods that can be used to examine speech intelligibility can the benefits and limitations of these different measures/methods. This decision is important for researchers, clinicians, and educators and evidence-based rationales should be present for decisions about the measures/methods chosen to be used. The aim of this review is to provide a systematic overview of the measures/methods that have been used in past research of the speech intelligibility of children who are DHH and the rationales that were given for the choice of assessment tools. This is not to say that such evidence-based consideration of different measures/methods of speech intelligibility is not important for professionals working with other populations of clients, for example adult rehabilitation following cochlear implantation, or people with speech difficulties that are physical (e.g., cerebral palsy, cleft palate), linguistic (e.g., speech sound disorder), or resulting from neurological changes (e.g., stroke, Parkinson's disease).

Table 1. Study questions, population, interventions, comparisons, outcomes, timing, type of study, setting (PICOTS)

Study Component	Current Study
Overall Questions	1. Which measures/methods are being used to examine speech intelligibility in children who are DHH? 2. Why are these methods/measures being used?
Population	Children who are DHH 0–12 years of age with permanent hearing loss.
Interventions	Perceptual measures of speech intelligibility based on children's speech production.
Comparisons	All
Outcomes	Speech intelligibility is to what extent a listener understands the speaker (Ertmer, 2011).
Time	All
Type of Study	All
Setting	Measurement is conducted in a clinical or education setting.

Table 1 provides Participant, Intervention, Comparison, Outcome, Timing, Type of study, Setting criteria and definitions relevant to this review. This systematic review addresses two questions:

1) Which measures/methods have been used in research studies to examine the speech intelligibility of children who are DHH?
2) What rationales do authors provide for their choice of methods/measures?

This review is a companion to that of Magnússon et al. (2022), which investigated the application of the International Classification of Functioning, Disability, and Health—Children and Youth (WHO, 2007) to research on the speech intelligibility of children who are DHH.

Method

Search Protocol

A systematic review methodology was used to identify, select, and analyze measures/methods that have been used in research studies to investigate the speech intelligibility of children who are DHH as part/all of their communication system. The method and reporting of this review follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses guidelines (PRISMA; Moher et al., 2010). Detailed information on how this review meets the PRISMA guidelines is provided in Supplementary Material S1.

Systematic Literature Search

In May 2021 an electronic search was conducted of CENTRAL via Ovid; CINAHL via EBSCO; Cochrane; ERIC via ProQuest; Joanna Briggs via Ovid; Linguistics, Language and Behavior Abstracts (LLBA) via ProQuest; Medline via PubMed; Scopus; and Web of Science databases. Development of the search protocol was guided by team members who had experience conducting systematic reviews and research and clinical experience working with children who are DHH, and a research librarian. The search blocks included terms describing (a) outcome, (b) participant type, and (c) hearing status. Search terms were developed and trialed with consultation with a research librarian. An indicative search string (which was used in Medline) was: ("Speech Intelligibility"[Mesh] OR " speech intelligibility" OR " speech production" OR oral OR phoneme OR "Communication Disorders"[Mesh]) AND (((girl OR girls OR boy OR boys) OR (Infant [MeSH] OR Child [MeSH] OR Pediatrics [MeSH] OR Infant OR infants OR baby OR babies OR child∗ OR pediatric∗ OR paediatric∗)) AND ("Hearing loss" OR "hearing impaired" OR "hearing impairment" OR "deaf" OR "hard of hearing" OR "cochlear implant∗" OR "hearing aid∗" OR "hearing disab∗" OR "hearing handicap∗" OR "hearing difficult∗" OR Hearing disorders [MeSH] OR "Persons With Hearing Impairments"[Mesh] OR "Hearing Aids"[Mesh])). Full search strategies and results for each database are presented in Supplementary Material S2 (Search Strategy). All records from the database searches were imported into Covidence (Veritas Health Innovation, 2021), a web-based platform for conducting systematic reviews. Duplicates were identified and removed by Covidence when records were uploaded. Remaining duplicates were removed manually during the screening process.

Supplementary Handsearching

Additional searches were conducted to identify records which may meet the inclusion criteria for the study but were not located in the systematic database search. These searches were conducted for the year range 2001 to 2021 (see inclusion/exclusion criteria) and were managed in an Excel worksheet. First, hand searching on the table of contents of 17 journals was performed: American Annals of the Deaf; American Journal of Speech-Language Pathology; American Speech-Language-Hearing Association Special Interest Group publications; Child Language Teaching and Therapy; Clinical Linguistics and Phonetics; Cochlear Implants International; Communication Disorders Quarterly; Deaf Education International; International Journal of Audiology; International Journal of Language and Communication Disorders; International Journal of Pediatric Otorhinolaryngology; International Journal of Speech-Language Pathology; Journal of Communication Disorders; Journal of Deaf Studies and Deaf Education; Journal of Speech, Language, and Hearing Research; Language, Speech, and Hearing Services in Schools; Seminars in Speech and Language. Second, a grey literature search was conducted on 25 websites and databases. These included research depositories (e.g., Clinical Trials (US), Chinese Clinical Trial Registry), thesis depositories (e.g., ProQuest Dissertations, Theses Global), and research and community organizations serving children who are DHH (e.g., American Society for Deaf Children, World Federation of the Deaf). Supplementary Material S2 (Search Strategy) presents the complete list of websites/databases searched and the search strings used, and the number of results in each database.

Article Screening

Inclusion and Exclusion Criteria

In the title and abstract screening stage, records were excluded if they were (a) published before 2001. This is when monitoring of early identified children with hearing loss was starting to roll out in many countries. Articles were also excluded if (b) participants did not have a permanent hearing loss, (c) more than half of the participants were 13 years or older, and (d) participants had diagnoses in addition to hearing loss which could impact communication development. Studies were included regardless of: the severity or configuration of hearing loss described, the assistive listening devices used by children in the study, the intervention/s children had participated in, and the communication mode/s children used. The majority of children who are DHH and acquiring spoken language as a mode of communication show developments in their speech production skills during their early years and early school years. It is during this period when speech intelligibility is most often examined to ascertain developmental and functional changes and skills. Even though the focus of this review was on the developmental changes in speech during childhood, a much higher cut-off age of 13 years was chosen. This was done to ensure that studies were not excluded that described children with more protracted development of speech intelligibility. The focus of the study was on speech intelligibility of children who are DHH, so studies including children with additional diagnoses that could impact communication development and speech intelligibility (e.g., cerebral palsy, autism spectrum disorder, intellectual disability, and auditory neuropathy) were excluded.

Articles were excluded if (e) no speech intelligibility data for speech production was included in the article, (f) the data were not relevant to the scope of the review (e.g., acoustic, physiological), and (g) the research was not conducted in a clinical or educational setting (e.g., it was a radiology study). Included methods were perceptual measures of the speech intelligibility of children's speech production that can be used in typical clinical or educational settings, for example rating scales and speech production accuracy measures. This also included methods/measures that focused on segmental (e.g., consonants, vowels) and suprasegmental (e.g., intonation, syllable stress, emphasis) components of speech. Acoustic and physiological measures were excluded as these typically require specialist equipment or training and are not available to educators and clinicians and not part of clinical or educational practice.

The criteria for the full text screening were the same, with one addition: (h) articles where the full text was not able to be obtained. That criterion was only used after an extensive search, authors were contacted, and interlibrary loan requests with several libraries were unable to be fulfilled. Finally, articles were excluded where (i) the authors stated that data were collected at a point much earlier than that considered in this review. Articles published in languages other than English were not excluded.

Title and Abstract Screening

Title and abstract screening was conducted in Covidence (Veritas Health Innovation, 2021). The titles and abstracts of records were screened using a double-blind method to determine whether

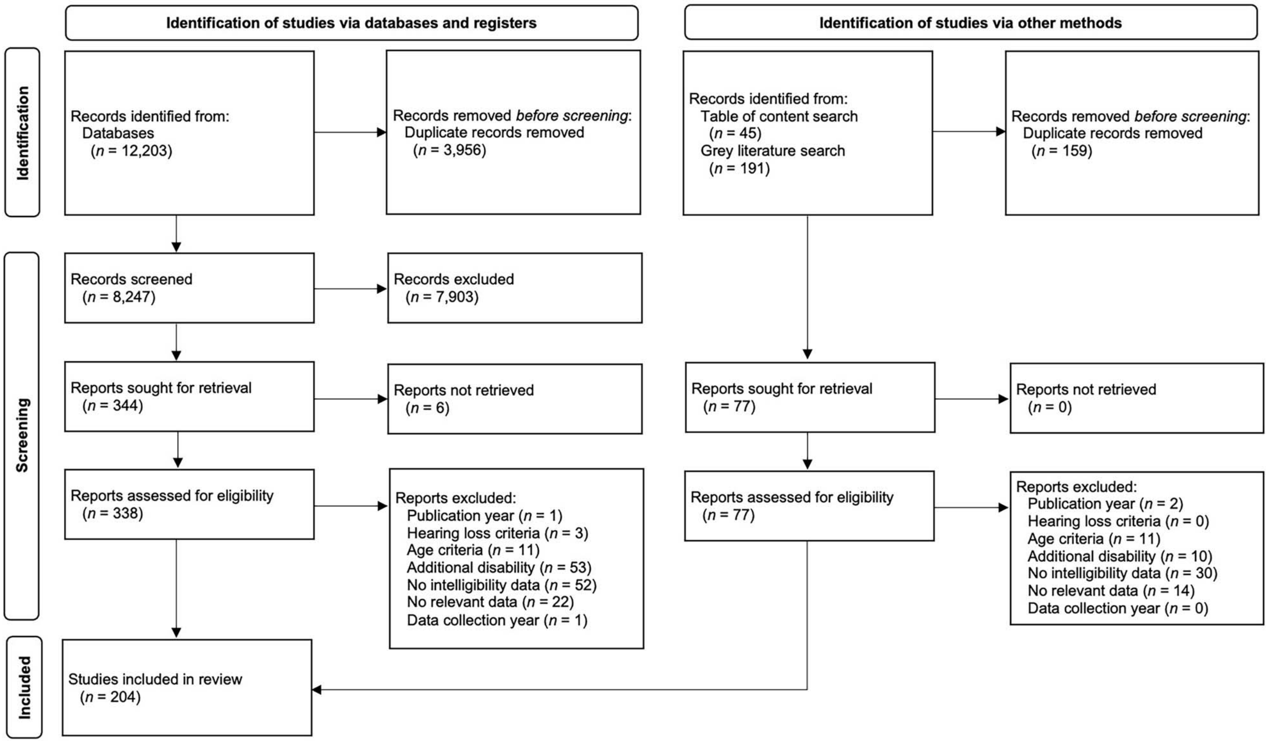

Figure 1. PRISMA flow diagram describing the search and inclusion/exclusion process.

they should move to the full text screening stage. Two reviewers independently screened each record against the exclusion criteria using Covidence, which flagged conflicts. Disagreements were discussed between the two reviewers until consensus was reached. When consensus was not reached, disagreements were discussed with a third reviewer. If articles were written in languages other than English, they were only screened during the full text screening stage.

Full Text Screening

Full texts of the records that were not excluded in the title and abstract screening stage were retrieved and assessed in detail, again using a double-blind method. Results of full-text screening are presented in the PRISMA chart in Figure 1. The two reviewers screened the articles independently against the exclusion criteria and assessed if the article should be included or excluded, and if excluded, the reason for exclusion. The exclusion criteria were applied hierarchically. Covidence flagged any conflicts between reviewers in both include/exclude decisions and in the reason chosen for exclusion. Conflicts were discussed until consensus was reached, with input from a third reviewer when needed.

In the full-text screening stage, articles in languages other than English were screened in by professionals outside the research team, with knowledge of that language. They completed a short online survey that had questions for each exclusion criteria. Possible responses were *yes*, *no*, *authors don't say*, and *I'm not sure*. Articles were excluded if they met any of the exclusion criteria or information necessary to judge the eligibility of the article was not included. If the reviewer selected *I'm not sure,* the article proceeded to the data extraction phase for careful reading. In this stage articles in the following languages were screened: Chinese (n = 11), Croatian (n = 1), Dutch (n = 2), Farsi (n = 2), Finnish (n = 1), French (n = 1), German (n = 4), Japanese (n = 1), Norwegian (n = 1), Portuguese (n = 3), Russian (n = 1), Spanish (n = 1), and Turkish (n = 1).

Data Extraction

Data were blindly extracted from articles included in the review by two reviewers and entered into a spreadsheet. Differences between the reviewers were discussed until a consensus was reached, and remaining differences were discussed with a third reviewer. Data for variables related to the articles, participants, methodology, listeners, and speech intelligibility measurements were extracted from all studies. Article variables were year of publication, language of publication, and country data. Participant variables included number, sex and age of participants, hearing status (unaided, aided, and pre-implant), language status, hearing devices used (hearing aid, cochlear implant, auditory brainstem implant), communication style (oral, total communication, sign language), and body structure (any differences in anatomy of the inner ear and/or the cochlear nerve). Methodology variables included study design (e.g., prospective, retrospective, longitudinal, comparative), what was measured (e.g., speech intelligibility, phonology, literacy, speech perception), and where the research was conducted (clinic, school, home). Listener variables included the number of listeners, listeners background, and listeners' age. Speech intelligibility variables included the name of the measurement, language tested, whether the measurement was subjective or objective, if the measurement was described in the article, and whether there was a rationale for the use of that measurement. Some studies did not explicitly state information that was needed, such as the country in which the study was conducted, but this could be inferred from other information contained in the article, such as the location of the authors and/or the language assessed. In those cases, this information was marked as inferred. Supplemental Material S3 (Data Code Book) provides additional details on this process.

When an article was not published in English, data extraction was conducted by a fluent user of that language with relevant research and/or content experience. This was the same person who conducted the full text screening in most cases. Extractors received the data extraction spreadsheet, the data extraction code book, the article to conduct, and instructions on data extraction. They also had the opportunity to meet with a member of the research team to discuss the data extraction process or to extract the data in collaboration with a team member. Six researchers chose to extract the data in collaboration with a research team member, one chose to extract the data independently after a meeting, and two chose to conduct the extraction independently.

Reliability

It was not necessary to examine reliability because at the following stages a double-blind method was used: title and abstract screening, full text screening, and data extraction.

Results

A PRISMA flowchart showing search results, numbers, and reasons for exclusion is presented in Figure 1. The final number of articles included in this review was 204. A table presenting the characteristics of each included study is available in Supplemental Material S4 (Study Information).

Publication Variables

The year of article publications ranged from 2001 to 2021 ($M = 2013$, $Mode = 2020$). The majority of the articles were published in English ($n = 189$, 92.6%). Articles were also published in: Chinese ($n = 10$), Croatian ($n = 1$), Russian ($n = 1$), Portuguese ($n = 2$), and Farsi ($n = 1$).

Participant Characteristics

The total number of participants with hearing loss in included studies was 15,520 ($Range = 1$ to 1,004, $M = 76$, $SD = 143.6$). It was difficult to compare the age of participants as age was reported in several different ways across studies. Studies reported participant age range ($n = 74$, 36.3%), mean age ($n = 71$, 34.8%), grade level ($n = 32$, 15.7%), age at implantation ($n = 65$, 31.9%), and/or did not report any age information ($n = 22$, 10.8%). Some articles reported more than one type of information. The order of inclusion above was therefore hierarchical. In the studies where age information was presented for the time of the study the age of participants ranged from 6 to 381.6 months. The participant that was 381.6 months was an outlier, and the majority of participants in that study was < 13 years old. Therefore, this study was included in the review. The minimum age ranged from 6 to 113.7 months ($M = 53.0$, $SD = 29.5$). The maximum age ranged from 20 to 381.6 months, ($M = 122.78$, $SD = 60.3$). For the studies that reported mean age, this ranged from 19.2 to 147.5 months, ($M = 79.3$, $SD = 30.6$). In total 116 (56.9%) studies did not provide precise information about the age of participants when testing occurred.

The studies included came from 28 different countries. The countries were as follows: Australia ($n = 11$), Belgium ($n = 10$), Brazil ($n = 4$), Canada ($n = 1$), China ($n = 33$), Croatia ($n = 1$), Denmark ($n = 2$), Egypt ($n = 4$), England ($n = 1$), Finland ($n = 1$), France ($n = 3$), India ($n = 23$), Iran ($n = 20$), Israel ($n = 2$), Italy ($n = 3$), Malaysia ($n = 2$), Northern Ireland ($n = 1$), Portugal ($n = 2$), Russia ($n = 1$), Saudi Arabia ($n = 4$), Slovakia ($n = 1$), South Korea ($n = 3$), Sweden ($n = 1$), Switzerland ($n = 1$), Taiwan ($n = 10$), Turkey ($n = 7$), the United Kingdom ($n = 9$), and the United States of America ($n = 36$). Five studies were conducted in two countries, the USA and Canada, and in two studies the location could not be inferred and were therefore recorded as *not mentioned*. The languages spoken were explicitly reported in 121 (59.3%) studies and could be inferred in 64 studies (31.4%), meaning this information was determined for 185 (90.7%) of the studies. The languages spoken were as follows: Arabic ($n = 8$), Croatian ($n = 1$), Danish ($n = 2$), Dutch ($n = 8$), English ($n = 62$), Farsi ($n = 20$), Finnish ($n = 1$), French ($n = 4$), Hebrew ($n = 2$), Hindi ($n = 2$), Italian ($n = 3$), Korean ($n = 3$), Malay ($n = 2$), Malayalam ($n = 1$), Mandarin ($n = 43$), Portuguese ($n = 6$), Russian ($n = 1$), Slovakian ($n = 1$), Swedish ($n = 1$), Tamil ($n = 1$), and Turkish ($n = 7$). Some studies ($n = 6$, 2.9%) included either bilingual participants or participants who spoke different languages. Only 158 studies (77.5%) explicitly reported the communication mode used by children who were DHH and in 43 studies (21.1%) the communication mode was inferred. Communication modes were as follows: cued speech ($n = 1$, .49%), oral ($n = 116$, 56.8%), total communication ($n = 3$, 1.5%), sign language (none exclusively), and simultaneous (none exclusively). Numerous studies ($n = 37$, 18.1%) included participants who used different communication modes. They were: oral and signed exact English ($n = 1$, .49%); oral and simultaneous communication ($n = 1$, .49%); oral and sign language ($n = 2$, .9%); oral and total communication ($n = 28$, 13.8%); oral, sign language, and total communication ($n = 3$, 1.5%); and sign language and total communication ($n = 2$, .9%).

Many studies reported on the hearing device used by participants. Participants used CIs in 169 (82.8%) studies, HAs in six (2.9%) studies, and ABIs in four (1.9%) studies. There was one study that did not report this information. No information on pure tone averages (PTA) was reported explicitly in 134 (65.7%) studies. However, 78 (38.2%) of those studies reported on the severity of hearing loss, which ranged from mild to profound. The PTA for unaided hearing loss on the better ear ranged from 15 dB to "not registered", or 125 dB. The PTA for aided hearing loss on the better ear ranged from 11 to 120 dB. The PTA for pre-implant on the better ear ranged from 70 dB to "not registered", or 121.7 dB.

Listener Characteristics

The number of listeners was reported in 66 (32.4%) studies, where the number ranged from 1 to 108. The listener background was reported in 131 (64.2%) studies, where the background was audiologists ($n = 4$, 1.9%), naïve listeners ($n = 33$, 16.2%), parents ($n = 17$, 8.3%), SLPs ($n = 55$, 26.9%), students ($n = 14$, 6.9%), and other professionals ($n = 36$, 17.6%). Some studies included listeners from more than one category. The listener relationship to the speaker was reported in 128 (62.17%) studies, where the relationships were clinician ($n = 59$, 28.9%), parents ($n = 17$, 8.3%), researcher ($n = 59$, 28.9%), and other professionals ($n = 8$, 3.9%). Again, some studies had listeners from more than one category. The listeners' ages were explicitly reported in only 13 (6.4%) studies, but they were inferred to be adults in 63 (30.9%) studies. In the studies that reported on the age of listeners, it ranged from 18 to 50 years.

Study Characteristics

All the studies included looked at speech intelligibility, as this was an inclusion-criteria for this review. Several studies looked at other factors as well. Other factors commonly looked at were: articulation ($n = 25$, 12.3%), phonology ($n = 21$, 10.3%), voice ($n = 10$, 4.9%), language ($n = 49$, 24.0%), speech perception ($n = 57$, 27.9%), and literacy ($n = 11$, 5.4%). Speech intelligibility measurements were a covariant in 23 (11.3%) studies. Most studies were conducted in a clinic ($n = 147$, 72.1%). Some studies were conducted in a school ($n = 23$, 11.3%), some at the child's home ($n = 19$, 9.3%),

Table 2. Evidence synthesis of speech intelligibility measurements used with children who are DHH

		Speech intelligibility measurement					
		Subjective			Objective		
		SIR (Allen et al., 2001) *n (%)*	Other/ unspecified SIR scales *n (%)*	Other subjective measures *n (%)*	BIT *n (%)*	PCC *n (%)*	Other objective measures *n (%)*
Study Characteristics	Total number of studies	69 (33.8%)	45 (22.1%)	23 (11.3%)	20 (9.8%)	32 (15.7)	26 (12.7%)
	Rationale provided	31 (44.9%)	13 (28.9%)	3 (13.0%)	1 (5%)	1 (3.1%)	2 (7.7%)
	RP: Psychometric properties	25 (80.6%)	8 (61.5%)	1 (33.3%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	RP: Easily understood	2 (6.5%)	0 (0.0%)	0 (.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	RP: Adapted for certain characteristics	1 (3.2%)	1 (7.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	2 (100%)
	RP: Time effective	3 (9.7)	2 (15.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	RP: Assessment of measure/method	1 (3.2%)	0 (0.0%)	2 (66.7%)	0 (0.0%)	1 (100%)	0 (0.0%)
	RP: Used in real-life situations	3 (9.7%)	2 (15.4%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)
	RP: Widely used	2 (6.5%)	0 (0.0%)	0 (0.0%)	1 (100%)	0 (0.0%)	0 (0.0%)
	English speaking	12 (17.4%)	5 (11.4%)	4 (17.4%)	18 (90.0%)	26 (81.3%)	16 (61.5%)

BIT = The Beginners' Intelligibility Test; PCC = Percentage consonants correct; RP = Rationale provided; SIR = Speech intelligibility rating

and some at other locations (*n* = 4, 1.9%). A few studies were conducted in more than one of these places. The methodologies used by the 204 studies were categorized by study design. The study designs identified were as follows: Descriptive case study (*n* = 65, 31.9%), Descriptive case study longitudinal (*n* = 44, 21.6%), Descriptive cohort study (*n* = 19, 9.3%), Descriptive cohort study longitudinal (*n* = 6, 2.9%), Descriptive group comparison (*n* = 55, 26.9%), Descriptive group comparison longitudinal (*n* = 10, 4.9%), Treatment study group comparison (*n* = 4, 1.9%), treatment study single-subject (*n* = 1, 0.5%).

Evidence Synthesis: Speech intelligibility Measurement Characteristics

A summary of the evidence synthesis can be found in Table 2. All studies reported some data on the speech intelligibility measurement used (as this was an inclusion criteria), but only 155 (75.9%) reported a reference for the test used. The distribution between subjective and objective measurements was 139 (68.1%) subjective and 76 (37.3%) objective. Some studies used both subjective and objective measurements (*n* = 12). Only 133 (65.5%) studies described the measurements used thoroughly.

Subjective Measurements

The most common subjective measurement tool used was The Speech Intelligibility Rating scale (SIR; Allen et al., 2001) (*n* = 69, 33.8%). The SIR, developed by Allen et al. (2001), is a subjective five-point rating scale developed for use with children who are DHH. It measures the level of speech intelligibility, the ease of understanding, with reference to everyday situations and can be used by professionals as well as naïve listeners. Each point on the scale is represented by a sentence describing how intelligible the child's speech is to the listener, (e.g., 2 = *Connected speech is unintelligible; intelligible speech is developing in single words when context and lip-reading cues are available*, and 5 = *Connected speech is intelligible to all listeners; the child is understood easily in everyday contexts*). A total of 27 (13.2%) studies reported they used SIR but did not provide a reference for it and 20 (9.8%) studies used an SIR scale other than that of Allen et al. (2001) and provided a reference, for example the Arabic speech intelligibility test (Bassiouny et al., 2013) (*n* = 2). Unnamed subjective measurements (*n* = 8, 3.9%) were described as: a Likert scale (*n* = 4), a nominal scale with four categories (*n* = 1), a rating scale (*n* = 3), and a visual analog scale (*n* = 1).

Objective Measurements

The most common objective measurements used were The Beginners' Intelligibility Test (BIT; Osberger et al., 1994) (*n* = 20, 9.8%), and Percentage of Consonants Correct (PCC; Shriberg & Kwiatkowski, 1982) (*n* = 32, 15.7%). The BIT (Osberger et al., 1994) is an objective sentence-repetition task developed for use with young children who are DHH. Each sentence on the test contains two to six words that are syntactically simple, and likely to be familiar to young children, for example: *The baby falls, Daddy runs*, and *My airplane is big*. The listener transcribes each sentence, and the BIT scores are expressed as a percentage of the number of words correctly transcribed divided by the total number of words (Osberger et al., 1994). PCC is an objective measurement, developed for estimating the severity of speech sound disorder. It expresses the percentage of correctly articulated consonants in a speech sample. It has been used to calculate scores from different types of speech samples, for example articulation tests, conversation samples, and non-conversational speech tasks (Shriberg et al., 1997; Shriberg & Kwiatkowski, 1982). The samples on which PCC calculations were based, in the studies included, came from a variety of sources: Teste de Linguagem Infantil task (ABFW) (*n* = 1), Assessment Link Between Phonology and Articulation—Revised Test (Lowe, 2000) (*n* = 1), Diagnostic Evaluation of Articulation and Phonology (Dodd et al., 2002) (*n* = 3), Goldman-Fristoe Test of Articulation (GFTA; Goldman & Fristoe, 2000) (*n* = 6), Hebrew Picture Speech Pattern Contrast (Kishon-Rabin & Henkin, 2000) (*n* = 1), the McGarr test (McGarr, 1981) (*n* = 5), unnamed picture articulation test (*n* = 2), 108 Single-Word Articulation Test (Paatsch, 1997) (*n* = 1), spontaneous speech sample (*n* = 2), and unspecified (*n* = 11). One study used both GFTA and a spontaneous speech sample.

Other named objective tests used were: Arizona Articulation Proficiency Scale (Fudala, 2000) (*n* = 1), Fanzago test (Fanzago, 1983) (*n* = 1), identifying early phonological needs in children with hearing impairment (IEPN; Paden & Brown, 1992) (*n* = 1), Intelligibility Index (Flipsen Jr. & Colvard, 2006) (n = 1), minimal pairs production test (MP2) (n = 1), percentage spraakverstaanbaarheid bij kinderen [percentage of speech intelligibility in children] (Baudonck et al., 2009) (n = 2), Persian speech intelligibility test (Darouie, 2013) (n = 2), Picture-SPINE (Monsen et al., 1988) (n = 1), The Children's Speech Intelligibility Measure (CSIM; Wilcox & Morris, 1999) (n = 1), The Monsen Indiana University sentences (MIUS; Monsen, 1981) (n = 2), and the Speech Intelligibility Test (SIT; Magner, 1972) (n = 1). Unnamed objective measurements (n = 7, 3.4%) were described as: a connected speech measure (n = 1), a magnitude estimation scale (n = 1), speech accuracy (n = 1), understandability of speech (n = 1), unnamed item-identification (n = 2), and phoneme accuracy (n = 1).

Multiple Measures

Some studies used more than one measure of intelligibility (n = 17, 8.3%). The reasons for studies using more than one measure of intelligibility, and the combinations of methods used varied. Some studies (n = 10) used both subjective and objective measurements (e.g., SIR and PCC, a visual analog scale and PCC). Typically, those articles did not provide a rationale for the choice of method, and in fact only one article did. In that case the rationale was that the study was done to compare a visual analog scale with an objective method (PCC) (Huttunen & Sorri, 2004). Two studies used both BIT and MIUS, as they had a wide age range, and needed different measures to measure the speech intelligibility for different ages (Ertmer, 2010; Svirsky et al., 2007). Four studies used BIT combined with other objective measures/methods, one used it with PCC (Chin & Kuhns, 2014), one with MP2 (Chin et al., 2001), one with a connected speech measure (Daneshmandan & Borghei, 2007), and one with CSIM (Khwaileh & Flipsen Jr., 2010). These studies did not provide a clear rationale for why they used more than one measure.

Rationales for Measurements

Rationales provided for the choice of measure/method varied in detail and quality, with only 49 (24.0%) studies explaining their choice of measurement. The rationales given were: easily understood (n = 2), good psychometric properties (n = 34), has been adapted for the participant characteristics presented in the study (e.g., a certain language, a certain age, or children following CI implantation) (n = 4), time effective (n = 5), the study was assessing a particular measure/method (n = 3), used in real-life situations (n = 4), and widely used (n = 3). A few studies (n = 6) provided a rationale that fitted into more than one group, for example good psychometric properties and time efficient. The studies that provided the rationale of the measure being easily understood both used Allen et al.'s (2001) SIR. Measures reported to be used for their psychometric properties mostly also used Allen et al.'s (2001) SIR (n = 25). Other tests used with this rationale were The Persian Speech Intelligibility Test (Heydari et al., 2011) (n = 1) and other rating scales (n = 8). The adaptability of a measure to the characteristics of the study's participant justified the choice of Allen et al.'s (2001) SIR, IEPN (Paden & Brown, 1992), CSIM (Wilcox & Morris, 1999), and BIT (Osberger et al., 1994), with each used in one study. The choice of time efficient measures were Allen et al.'s (2001) SIR (n = 3) and two different SIR scales. Studies examining a particular measure used a Likert scale, a visual analog scale and PCC from an unspecified source, and Allen et al.'s (2001) SIR. SIR scales were also used to best reflect real-life situations, including Allen et al.'s (2001) SIR and a different SIR scale. Finally, the rationale of being widely used supported the choice of Allen et al.'s (2001) SIR scale, BIT (Osberger et al., 1994), and MUIS (Monsen, 1981).

Discussion

This systematic review adds to the existing literature concerning the speech intelligibility outcomes of children who are DHH and use spoken language as part/all of their communication system by critically examining the methods and measures used in evaluating speech intelligibility. The purpose of this review was to examine which methods and measurements have been used to measure the speech intelligibility of children who are DHH and what rationales were provided for the choice of measure/method. In total, 204 articles met all the inclusion criteria and were included in the review. The variability in measures/methods used to assess speech intelligibility, the lack of rationales provides, and the lack of discussion of the psychometric properties of measures, indicates that researcher and clinicians should take caution when comparing the results across studies and when choosing intelligibility measures/methods to use. This review article shows that there is a gap in the literature concerning which measures/methods are the most reliable or useful in examining the speech of children who are DHH who communicate using speech. The lack of an agreed upon intelligibility method or measure makes it difficult to compare speech outcomes of across studies and treatment conditions.

This review article clearly found that there is variability in speech intelligibility measures and methods used to describe the speech of children who are DHH across studies. Even though some methods that are more commonly used than others (e.g., SIR [Allen et al., 2001], BIT, and PCC) there are many studies that use methods that are not thoroughly described or researched (e.g., Boonen et al., 2020; Grandon et al., 2020; Van Lierde et al., 2005; Warner-Czyz & Davis, 2008; Yoon et al., 2004). As previously mentioned, not all studies provided a reference for their measure/method (e.g., An et al., 2012; Arumugam et al., 2021; Blamey et al., 2001; Boonen et al., 2020; Daneshmandan et al., 2009; Melo et al., 2017) and not all studies described the measure/method they were using (e.g., Alenzi et al., 2020; Jalil-Abkenar et al., 2013; Venkataramani et al., 2021; Wu et al., 2008; Zhou et al., 2013). In some cases, it was therefore unclear exactly what measure/method was being used (e.g., das Neves et al., 2020; Lucchesi et al., 2015; Tomblin et al., 2008; Warner-Czyz & Davis, 2008; Zachariah & Kumaraswamy, 2015). It is an interesting result that when the subjective and objective measurements are compared, more of the studies conducted in English speaking countries used objective measures, and more of the studies conducted in different languages, used subjective scales (See Table 2 for evidence synthesis summary). Only 49 articles (24.0%) provided some type of rationale for the choice of measures and methods. Few studies, only 17%, provided high quality rationales and discussed the psychometric properties of the scales (e.g., Edwards et al., 2006; Fulcher et al., 2014; Nikolopoulos et al., 2008; Singh et al., 2004; Wilkinson & Brinton, 2003), which is concerning. Few measures of speech intelligibility have been examined for their psychometric properties and the low utilization of valid and reliable methods to examine speech intelligibility should be of concern to researchers and clinicians/educators alike.

Limitations of the Literature and Recommendations for Researchers

The literature examined in this review presented some clear limitations with the information presented, and the measures/

methods chosen. First, the information presented in the studies was often unclear or lacking. Some of the literature failed to report on numerous crucial variables, such as participant variables (e.g., age, sex, PTA, communication mode), listener characteristics (e.g., number of listeners, relationship to speaker, age of listener), and speech intelligibility measurement (e.g., description of the method, rationale for method, reference for method and language tested). This makes a comparison of children's performance between studies challenging, and in some cases impossible. Second, there is a need for a clear description of the measures/methods being used. In addition, it would be useful to provide the reader with a reference to where more information about this measure/method can be found and/or if it is a novel or custom-designed measure (usually a rating scale) for the measure to be provided with the article as an Appendix or as supplemental online material. Third, there is a need for researchers to choose and use measures that have known and strong psychometric properties, and to investigate the psychometric properties of commonly used measures that currently have not been established as being valid and/or reliable. This is also important when a measure with known strong psychometric properties is being used in a language or context in which it was not described to be used. This was the case where studies reported using a measure because of its strong psychometric properties when used in English in the United Kingdom as a rationale for the use of a translated version of the measure in a new language and context (e.g., De Raeve, 2010; Koşaner et al., 2017; Othman et al., 2020; Wu et al., 2008).

Recommendations for Clinicians and Educators

While working with children who are DHH, clear, high-quality measurements are an important source of information for measuring progress and are important in clinical decision-making and monitoring (Miller, 2013). Therefore, it is crucial for clinicians to have information about which measures and methods are valid and reliable, when used in clinical settings. Clinicians should always research the literature on the measures/methods they choose to use and see whether they have been thoroughly researched, and how they align with other types of measurements. This is necessary when using evidence-based practice. Evidence-based practice states that clinicians should integrate clinical expertise, evidence, and client's/caregiver's perspectives (ASHA, n.d.). When reviewing the literature on speech intelligibility measures/methods used with children who are DHH it is clear that the limitations of the literature poses a great concern for evidence-based practice. This stems from the problem that the evidence on speech intelligibility measurement and method is limited, and it is difficult for clinicians to compare different types of measurements across studies.

From this review it is clear that some measures/methods are better researched than others and have been more widely used. Of the subjective measures, Allen et al.'s (2001) SIR is by far the most widely used with this group, and more is known about the psychometric properties of this SIR scale, than any other SIR scale. As with many subjective measures, this SIR scale can be insensitive to speech that falls in the middle range of intelligibility, the ratings can be difficult to interpret, and although the findings of this review show that it has been used in many languages, descriptions of the psychometric properties of this scale in languages other than English are limited. Another subjective measure that has been well researched and widely used with preschool-aged children with speech sound disorder is the ICS (McLeod, 2020), although this has not been used with children who are DHH. The ICS is a five-point scale used by parents to report on how intelligible their child's speech is to different communication partners who vary both in familiarity and authority, e.g., parents, extended family strangers, teachers, and friends (McLeod et al., 2012). This scale was however, not used in any of the studies included in the review, implying that it has not been as widely used and researched with children who are DHH. The benefits of this scale are that its psychometric properties have been researched and found to be good in various languages and situations: for example: Fijian (Hopf et al., 2017), Swedish (Lagerberg et al., 2019), Korean (Lee, 2019), Cantonese (Ng et al., 2014), Vietnamese (Phạm et al., 2017), Italian (Piazzalunga et al., 2020), Jamaican Creole (Washington et al., 2017), and German (Neumann et al., 2017). It has similar disadvantages as the SIR scale being a subjective measure, and if this scale were to be used with children who are DHH, further research would be needed to assess its qualities for that specific group.

Of the objective measures, PCC was the most widely used. PCC has been well researched and its psychometric properties are well known (Shriberg et al., 1997; Shriberg & Kwiatkowski, 1982). As with many objective measures, limitations of PCC include a lack of standardization between speech samples assessed, and reflecting only on consonant production (not vowels or whole words), thus missing other aspects of speech that might impact intelligibility, (e.g., voice quality, nasality, and intonation) (Shriberg et al., 1997). Another objective measure that has been widely used and quite well researched is the BIT (Osberger et al., 1994). Its psychometric properties have also been assessed and found to be good (Ertmer, 2010; Khwaileh & Flipsen Jr., 2010). It does however have its limitations as an objective measure, being more time consuming than scaling procedures and research on its psychometric properties in languages other than English is limited.

As both subjective and objective measures have different strengths and limitations, using both types of measurements when assessing the speech intelligibility of children who are DHH may provide a stronger assessment of skills. Because of the different limitations SIR and PCC have, a combination of these two measures when assessing the speech intelligibility of children who are DHH could be a better option. As a pair, they may be able to mitigate many of the limitations that they exhibit individually. This combination of measures potentially would give a broad and strong picture of the speech intelligibility of children being assessed. When assessing the speech intelligibility of young children who are DHH, the BIT might be a better fit than PCC, as it was designed specifically for use with young children who are DHH. Using the BIT alongside an SIR scale or the ICS, which is specifically designed for use with preschool aged children, might therefore be a good fit when assessing the speech intelligibility of younger children. This being said, further research is needed that assesses different measures and combinations of measures, to state with confidence what should be the golden standard for speech intelligibility measures of children who are DHH. It should also be noted that measures of suprasegmental characteristics of speech were absent from the measure/methods identified in this review. The importance of suprasegmental features of speech, which can greatly impact on speech intelligibility, should also be considered as part of a wholistic assessment of the speech intelligibility of children who are DHH.

Limitations of this Review

There were a number of limitations to this review. First, there is a possibility some relevant studies were not identified. The risk of this was minimized by doing an extensive search of databases using broad search terms. Second, it is a possibility that due to strict exclusion criteria, articles that could have been informative were excluded. For example, we found some studies that measured speech intelligibility at a broad age range and were excluded due to the fact that > 50% of the participants were 13 years or older. The speech intelligibility measures used in studies of adolescents and adults could be applicable for children also, meaning that some relevant measures may have been missed in this review. Further research should consider speech intelligibility measures used with adolescents and adults, particularly in rehabilitation (rather than habilitation) contexts. Third, many studies failed to report important information about numerous variables (e.g., participant information, study information, speech intelligibility measurement information). This made the comparison between studies challenging, and in some cases impossible.

Future Research

There is a need to establish a gold standard for assessing the speech intelligibility of children who are DHH across different stages of speech development so that they are comparable across studies, and to reduce methodological weaknesses related to their choice of measure, as well as providing clinicians and educators with a tool to measure the speech intelligibility of their clients/students. Selecting measures/methods with good psychometric properties demonstrated across different ages, they should be researched in various languages and different contexts, and should be able to be implemented by researchers as well as clinicians and educators. It is important to note that measures and methods which considered suprasegmental characteristics of children's speech were conspicuously absent from our findings. Reporting on methods/measures that include or focus on suprasegmental characteristics of children's speech impacting intelligibility should be considered in future research on speech intelligibility in children who are DHH. When operationalizing a method for establishing speech intelligibility, it is important to establish if using a combination of measures to address the limitations that each measure presents on its own. It is also important to note that a speech intelligibility gold standard measure or method might be different for different age groups, and further research is needed on this. Speech intelligibility methods and measures should also be standardized across different languages. In order for this to be possible it is important to assess the different measures being used in every language that it will be used in. The measures need to be translated, local norms established, and their psychometric properties described in each language and context.

Conclusion

This review provided an overview of what measures/methods have been used to measure the speech intelligibility of children who are DHH and what rationale is provided for the choice of method. Speech intelligibility measures are important clinical and research measures that provide a general index of how well a person is understood by others. The results of this review highlight that a diverse range of subjective and objective measures have been used in previous research with children who are DHH. Poor and inconsistent description of measures and method and the rationale for method choice across the majority of studies hampers comparison of findings across studies. It is clear that a more standard way of measuring speech intelligibility that involves both subjective and objective elements would be beneficial. This is the first study we know of that has critically analyzed the existing research on speech intelligibility measures with children who are DHH and the results show that the literature is in need of improvements, as well as further research being needed.

Acknowledgements

The authors thank Cen Wang, Jessica Gries, Jóhanna Thelma Einarsdóttir, Katsumi Hagiwara, Lian van Berkel-van Hoof, Lina Motlagh Zadeh, Maria Claudia Franca, Olga Prokhorova, Ömer Dağlar Tanrıkulu, Pauline van der Straten Waillet, Sari Kunnari, Sisan Cuervo, Maja Kelic, and Vesna Mildner for their assistance with reviewing papers in languages other than English.

Supplementary material

Supplementary material is available online at *Journal of Deaf Studies and Deaf Education*.

Data availability

Contact authors for information on data availability.

Funding

This project was conducted in the course of the authors' usual study/employment and did not receive external funding.

Conflicts of Interest: There are no relevant conflicts of interest.

References

Alenzi, S. H., Halawani, R. T., Alshalan, A. M., Habis, S. A., & Alsanosi, A. A. (2020). Influence of family environment on the outcomes of cochlear implantation in pediatric recipients. *Saudi Medical Journal*, **41**(5), 485–490. https://doi.org/10.15537/SMJ.2020.5.25070.

Allen, C., Nikolopoulos, T. P., Dyar, D., & O'Donoghue, G. M. (2001). The reliability of a rating scale for measuring speech intelligibility following pediatric cochlear implantation. *Otology and Neurotology*, **22**(5), 631–633. https://doi.org/10.1097/00129492-200109000-00012.

American Speech-Language-Hearing Association. (n.d.). *Evidence-Based Practice (EBP)*. https://www.asha.org/research/ebp/#:~:text=Evidence%2Dbased%20practice%20(EBP)%20is%20the%20integration%20of,your%20training%20and%20professional%20experiences

An, Y. S., Kim, S. T., & Chung, J. W. (2012). Preoperative voice parameters affect the postoperative speech intelligibility in patients with cochlear implantation. *Clinical and Experimental Otorhinolaryngology*, **5**(Suppl 1), S69–S72. https://doi.org/10.3342/ceo.2012.5.S1.S69.

Arumugam, S. V., Mathews, S., Paramasivan, V. K., & Kameswaran, M. (2021). Quality of life assessment in paediatric cochlear implant recipients in South India: Parental assessment and correlation with rehabilitation scores using a bilingual quality of life assessment questionnaire. *Cochlear Implants International*, **22**(3), 128–135. https://doi.org/10.1080/14670100.2020.1841436.

Bassiouny, S. E., Mona, A. H., Jilan, F. N., Amal, S. S., & Abdelhamid, A. (2013). Development of an Arabic speech intelligibility test for children. The Egyptian test for children. *Egyptian Journal of Otolaryngology*, **29**, 202–206. https://doi.org/10.7123/01.EJO.0000429577.63818.de.

Baudonck, N. L. H., Buekers, R., Gillebert, S., & Van Lierde, K. M. (2009). Speech intelligibility of Flemish children as judged by their parents. *Folia Phoniatrica et Logopaedica*, **61**, 288–295. https://doi.org/10.1159/000235994.

Baudonck, N., Lierde, K., D'haeseleer, E., & Dhooge, I. (2011). A comparison of the perceptual evaluation of speech production between bilaterally implanted children, unilaterally implanted children, children using hearing aids, and normal-hearing children. *International Journal of Audiology*, **50**(12), 912–919. https://doi.org/10.3109/14992027.2011.605803.

Blamey, P. J., Sarant, J. Z., Paatsch, L. E., Barry, J. G., Bow, C. P., Wales, R. J., Wright, M., Psarros, C., Rattigan, K., & Tooher, R. (2001). Relationships among speech perception, production, language, hearing loss, and age in children with impaired hearing. *Journal of Speech, Language, and Hearing Research*, **44**(2), 264–285. https://doi.org/10.1044/1092-4388(2001/022).

Boonen, N., Kloots, H., & Gillis, S. (2020). Is the spontaneous speech of 7-year-old cochlear implanted children as intelligible as that of their normally hearing peers? *International Journal of Pediatric Otorhinolaryngology*, **133**, 109956. https://doi.org/10.1016/j.ijporl.2020.109956.

Castellanos, I., Kronenberger, W. G., Beer, J., Henning, S. C., Colson, B. G., & Pisoni, D. B. (2014). Preschool speech intelligibility and vocabulary skills predict long-term speech and language outcomes following cochlear implantation in early childhood. *Cochlear Implants International*, **15**(4), 200–210. https://doi.org/10.1179/1754762813Y.0000000043.

Chin, S. B., Finnegan, K. R., & Chung, B. A. (2001). Relationships among types of speech intelligibility in pediatric users of cochlear implants. *Journal of Communication Disorders*, **34**(3), 187–205. https://doi.org/10.1016/S0021-9924(00)00048-4.

Chin, S. B., & Kuhns, M. J. (2014). Proximate factors associated with speech intelligibility in children with cochlear implants: A preliminary study. *Clinical Linguistics and Phonetics*, **28**, 532–542. https://doi.org/10.3109/02699206.2014.926997.

Ching, T. Y. C., Crowe, K., Martin, V., Day, J., Mahler, N., Youn, S., Street, L., Cook, C., & Orsini, J. (2010). Language development and everyday functioning of children with hearing loss assessed at 3 years of age. *International Journal of Speech-Language Pathology*, **12**(2), 124–131. https://doi.org/10.3109/17549500903577022.

Ching, T. Y. C., & Dillon, H. (2013). Major findings of the LOCHI study on children at 3 years of age and implications for audiological management. *International Journal of Audiology*, **52**(2), S65–S68. https://doi.org/10.3109/14992027.2013.866339.

Ching, T. Y. C., Dillon, H., Marnane, V., Hou, S., Day, J., Seeto, M., Crowe, K., Street, L., Thomson, J., Van Buynder, P., Zhang, V., Wong, A., Burns, L., Flynn, C., Cupples, L., Cowan, R. S. C., Leigh, G., Sjahalam-King, J., & Yeh, A. (2017). Outcomes of early- and late-identified children at 3 years of age: Findings from a prospective population-based study. *Ear and Hearing*, **34**(5), 535–552. https://doi.org/10.1097/AUD.0b013e3182857718.

Coelho, A. C., Medved, D. M., & Brasolotto, A. G. (2015). Hearing loss and the voice. In F. Bahmad (Ed.), *Update on hearing loss* (pp. 103–128). InTechOpen.

Coppens-Hofman, M. C., Terband, H., Snik, A. F., & Maassen, B. A. (2016). Speech characteristics and intelligibility in adults with mild and moderate intellectual disabilities. *Folia Phoniatrica et Logopaedica*, **68**(4), 175–182. https://doi.org/10.1159/000450548.

Crowe, K., & Guiberson, M. (2019). Evidence-based interventions for learners who are deaf and/or multilingual: A systematic quality review. *American Journal of Speech-Language Pathology*, **28**, 964–983. https://doi.org/10.1044/2019_AJSLP-IDLL-19-0003.

Crowe, K., Marschark, M., & McLeod, S. (2019). Measuring intelligibility in signed languages. *Clinical Linguistics and Phonetics*, **33**(10–11), 991–1008. https://doi.org/10.1080/02699206.2019.1600169.

Cupples, L., Ching, T. Y. C., Button, L., Seeto, M., Zhang, V., Whitfield, J., Gunnourie, M., Martin, L., & Marnane, V. (2017). Spoken language and everyday functioning in 5-year-old children using hearing aids or cochlear implants. *International Journal of Audiology*, **57**(2), S55–S69. https://doi.org/10.1080/14992027.2017.1370140.

Daneshmandan, N., & Borghei, P. (2007). Speech intelligibility development in severe to profound hearing-impaired children and establishment of a data collection for early intervention in hearing-impaired children. *Acta Medica Iranica*, **45**(1), 35–42.

Daneshmandan, N., Borghei, P., Yazdany, N., Soleimani, F., & Vameghi, R. (2009). Oral communication development in severe to profound hearing impaired children after receiving aural habilitation. *Acta Medica Iranica*, **47**(5), 363–367.

Darouie, A. (2013). *Comparison of speech intelligibility evaluation procedures in Persian hearing impaired children*. [Doctoral dissertation Tehran: University of Social Welfare and Rehabilitation]. Faculty of Speech Therapy.

De Raeve, L. (2010). A longitudinal study on auditory perception and speech intelligibility in deaf children implanted younger than 18 months in comparison to those implanted at later ages. *Otology and Neurotology*, **31**(8), 1261–1267. https://doi.org/10.1097/MAO.0b013e3181f1cde3.

Dodd, B., Zhu, H., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic Evaluation of Articulation and Phonology (DEAP)*. Psychology Corporation.

van Doornik, A., Gerrits, E., McLeod, S., & Terband, H. (2018). Impact of communication partner familiarity and speech accuracy on parents' ratings of their child for the Intelligibility in Context Scale: Dutch. *International Journal of Speech-Language Pathology*, **20**(3), 350–360. https://doi.org/10.1080/17549507.2018.1472808.

Edwards, L. C., Frost, R., & Witham, F. (2006). Developmental delay and outcomes in paediatric cochlear implantation: Implications for candidacy. *International Journal of Pediatric Otorhinolaryngology*, **70**(9), 1593–1600. https://doi.org/10.1016/j.ijporl.2006.04.008.

Ertmer, D. J. (2010). Relationships between speech intelligibility and word articulation scores in children with hearing loss. *Journal of Speech, Language, and Hearing Research*, **53**(5), 1075–1086. https://doi.org/10.1044/1092-4388(2010/09-0250).

Ertmer, D. J. (2011). Assessing speech intelligibility in children with hearing loss: Toward revitalizing a valuable clinical tool. *Language, Speech and Hearing Services in Schools*, **42**(1), 52–58. https://doi.org/10.1044/0161-1461(2010/09-0081).

Fanzago, F. (1983). Test di valutazione dell'articolazione from 'Trattamento logopedico delle dislalie e delle insufficienze velofaringee' [Articulation test for 'Speech-language pathology treatment of articulation difficulties and velopharyngeal insufficiency]. *Quaderni di Acta Phoniatrica Latina*, **2**, 80–85.

Flipsen, P. Jr. (2008). Intelligibility of spontaneous conversational speech produced by children with cochlear implants: A review. *International Journal of Pediatric Otorhinolaryngology*, **72**(5), 559–564. https://doi.org/10.1016/j.ijporl.2008.01.026.

Flipsen, P. Jr., & Colvard, L. G. (2006). Intelligibility of conversational speech produced by children with cochlear implants. *Journal of Communication Disorders*, **39**(2), 93–108. https://doi.org/10.1016/j.jcomdis.2005.11.001.

Freeman, V. (2018). Speech intelligibility and personality peer-ratings of young adults with cochlear implants. *Journal of Deaf Studies and Deaf Education*, **23**(1), 41–49. https://doi.org/10.1093/deafed/enx033.

Freeman, V., & Pisoni, D. B. (2017). Speech rate, rate-matching, and intelligibility in early-implanted cochlear implant users. *Journal of the Acoustical Society of America*, **142**(2), 1043–1054. https://doi.org/10.1121/1.4998590.

Fudala, J. B. (2000). *Arizona Articulation Proficiency Scale, Third Revision*, *Western Psychological Services*.

Fulcher, A., Baker, E., Purcell, A., & Munro, N. (2014). Typical consonant cluster acquisition in auditory-verbal children with early-identified severe/profound hearing loss. *International Journal of Speech-Language Pathology*, **16**(1), 69–81. https://doi.org/10.3109/17549507.2013.808698.

Gérard, J.-M., Deggouj, N., Hupin, C., Buisson, A.-L., Monteyne, V., Lavis, C., Dahan, K., & Gersdorff, M. (2010). Evolution of communication abilities after cochlear implantation in prelingually deaf children. *International Journal of Pediatric Otorhinolaryngology*, **74**(6), 642–648. https://doi.org/10.1016/j.ijporl.2010.03.010.

Goldman, R., & Fristoe, M. (2000). *Goldman-Fristoe Test of Articulation-2*. American Guidance Service.

Grandon, B., Martinez, M. J., Samson, A., & Vilain, A. (2020). Long-term effects of cochlear implantation on the intelligibility of speech in French-speaking children. *Journal of Child Language*, **47**(4), 881–892. https://doi.org/10.1017/S0305000919000837.

Habib, M. G., Waltzman, S. B., Tajudeen, B., & Svirsky, M. A. (2010). Speech production intelligibility of early implanted pediatric cochlear implant users. *International Journal of Pediatric Otorhinolaryngology*, **74**(8), 855–859. https://doi.org/10.1016/j.ijporl.2010.04.009.

Heydari, S., Torabi, N. F., Agha, R. Z., & Hoseyni, F. (2011). ساخت آزمون اندازه‌گیری وضوح گفتار برای کودکان بهنجار 3 تا 5 ساله [Development of speech intelligibility measurement test for 3 to 5years old normal children]. *Auditory and Vestibular Research*, **20**(1), 47–53.

Hopf, S. C., McLeod, S., & McDonagh, S. H. (2017). Validation of the Intelligibility in Context Scale for school students in Fiji. *Clinical Linguistics and Phonetics*, **31**(7–9), 487–502. https://doi.org/10.1080/02699206.2016.1268208.

Huttunen, K., & Sorri, M. (2004). Methodological aspects of assessing speech intelligibility among children with impaired hearing. *Acta Oto-Laryngologica*, **124**(4), 490–494. https://doi.org/10.1080/00016480310000557.

Ishikawa, K., Boyce, S., Kelchner, L., Powell, M. G., Schieve, H., De Alarcon, A., & Khosla, S. (2017). The effect of background noise on intelligibility of dysphonic speech. *Journal of Speech, Language, and Hearing Research*, **60**(3), 1919–1929. https://doi.org/10.1044/2017_JSLHR-S-16-0012.

Jalil-Abkenar, S. S., Ashori, M., Pourmohamadreza-Tajrishi, M., & Hasanzadeh, S. (2013). Auditory perception and verbal intelligibility in children with cochlear implant, hearing aids and normal hearing. *Practice in Clinical Psychology*, **1**(3), 141–147.

Khwaileh, F. A., & Flipsen, P. Jr. (2010). Single word and sentence intelligibility in children with cochlear implants. *Clinical Linguistics and Phonetics*, **24**(9), 722–733. https://doi.org/10.3109/02699206.2010.490003.

Kishon-Rabin, L., & Henkin, Y. (2000). Age-related changes in the visual perception of phonologically significant contrasts. *British Journal of Audiology*, **34**(6), 363–374. https://doi.org/10.3109/03005364000000152.

Koşaner, J., Deniz, H., Uruk, D., Deniz, M., Kara, E., & Amann, E. (2017). Assessment of early language development in Turkish children with a cochlear implant using the TEDIL test. *Cochlear Implants International*, **18**(3), 153–161. https://doi.org/10.1080/14670100.2017.1299392.

Lagerberg, T. B., Åsberg, J., Hartelius, L., & Persson, C. (2014). Assessment of intelligibility using children's spontaneous speech: Methodological aspects. *International Journal of Language and Communication Disorders*, **49**(2), 228–239. https://doi.org/10.1111/1460-6984.12067.

Lagerberg, T. B., Hellström, A., Lundberg, E., & Hartelius, L. (2019). An investigation of the clinical use of a single word procedure to assess intelligibility (Swedish Test of Intelligibility for Children) and an evaluation of the validity and reliability of the Intelligibility in Context Scale. *Journal of Speech, Language, and Hearing Research*, **62**(3), 668–681. https://doi.org/10.1044/2018_JSLHR-S-18-0018.

Lee, Y. (2019). Validation of the Intelligibility in Context Scale for Korean-speaking pre-school children. *International Journal of Speech-Language Pathology*, **21**(4), 395–403. https://doi.org/10.1080/17549507.2018.1485740.

Lowe, R. (2000). *Assessment Link Between Phonology and Articulation—Revised (ALPHA–R)*. www.speechlanguage-therapy.com.

Lucchesi, F. D. M., Almeida-Verdu, A. C. M., Buffa, M. J. M. B., & Bevilacqua, M. C. (2015). Effects of a reading teaching program on the speech intelligibility of children using cochlear implant. *Psicologia: Reflexao eCritica*, **28**(3), 500–510. https://doi.org/10.1590/1678-7153.201528309.

Magner, M. E. (1972). *A Speech Intelligibility Test for Deaf Children*. Clarke School for the Deaf.

Magnússon, E., Crowe, K., Stefánsdóttir, H., Guiberson, M., Másdóttir, T., Ágústsdóttir, I., & Baldursdóttir, Ö. V. (2022). Representation of the ICF-CY in research of speech intelligibility: A systematic review of literature describing deaf and hard-of-hearing children.

McGarr, N. S. (1981). The effect of context on the intelligibility of hearing and deaf children's speech. *Language and Speech*, **24**(3), 255–264. https://doi.org/10.1177/002383098102400305.

McLeod, S. (2020). Intelligibility in Context Scale: Cross-linguistic use, validity and reliability. *Speech, Language and Hearing*, **23**(1), 9–16. https://doi.org/10.1080/2050571X.2020.1718837.

McLeod, S., Harrison, L. J., & McCormack, J. (2012). *Intelligibility in Context Scale*. Charles Sturt University https://www.csu.edu.au/research/multilingual-speech/ics.

Melo, A. S., Martins, J., Silva, J., Quadros, J., & Paiva, A. (2017). Cochlear implantation in children with anomalous cochleovestibular anatomy. *Auris Nasus Larynx*, **44**(5), 509–516. https://doi.org/10.1016/j.anl.2017.02.003.

Miller, N. (2013). Measuring up to speech intelligibility. *International Journal of Language and Communication Disorders*, **48**(6), 601–612. https://doi.org/10.1111/1460-6984.12061.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group (2010). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *International Journal of Surgery*, **8**(5), 336–341. https://doi.org/10.1016/j.ijsu.2010.02.007.

Monsen, R. B. (1981). A usable test for the speech intelligibility of deaf talkers. *American Annals of the Deaf*, **126**(7), 845–852. https://doi.org/10.1353/aad.2012.1333.

Monsen, R. B., Moog, J. S., & Geers, A. E. (1988). Picture Speech Intelligbility Evaluation. *Central Institute for the Deaf*.

Most, T., Ingber, S., & Heled-ariam, E. (2012). Social competence, sense of loneliness, and speech intelligibility of young children with hearing loss in individual inclusion and group inclusion. *Journal of Deaf Studies and Deaf Education*, **17**(2), 259–272. https://doi.org/10.1093/deafed/enr049.

Neumann, S., Rietz, C., & Stenneken, P. (2017). The German Intelligibility in Context Scale (ICS-G): Reliability and validity evidence. *International Journal of Language and Communication Disorders*, **52**(5), 585–594. https://doi.org/10.1111/1460-6984.12303.

das Neves, A. J., Almeida-Verdu, A. C. M., do Nascimento Silva, L. T., & Moret, A. L. M. (2020). Acquisition of speech accuracy of sentences in children with cochlear implant. *Revista de Psicologia (Peru)*, **38**(2), 387–421. https://doi.org/10.18800/PSICO.202002.002.

Newsted, D., Rosen, E., Cooke, B., Beyea, M. M., Simpson, M. T. W., & Beyea, J. A. (2020). Approach to hearing loss. *Canadian Family Physician*, **66**(11), 808–809.

Ng, K. Y. M., To, C. K. S., & McLeod, S. (2014). Validation of the Intelligibility in Context Scale as a screening tool for preschoolers in Hong Kong. *Clinical Linguistics and Phonetics*, **28**(5), 316–328. https://doi.org/10.3109/02699206.2013.865789.

Nikolopoulos, T. P., Archbold, S. M., Wever, C. C., & Lloyd, H. (2008). Speech production in deaf implanted children with additional disabilities and comparison with age-equivalent implanted children without such disorders. *International Journal of Pediatric Otorhinolaryngology*, **72**(12), 1823–1828. https://doi.org/10.1016/j.ijporl.2008.09.003.

Osberger, M., Robbins, A., Todd, S., & Riley, A. (1994). Speech intelligibility of children with cochlear implants. *Volta Review*, **96**(5), 169–180.

Othman, I. A., Abdullah, A., See, G. B., Umat, C., & Tyler, R. S. (2020). Auditory performance in early implanted children with cochleovestibular malformation and cochlear nerve deficiency. *Journal of International Advanced Otology*, **16**(3), 297–302. https://doi.org/10.5152/iao.2020.8563.

Overby, M., Carrell, T., & Bernthal, J. (2007). Teachers' perceptions of students with speech sound disorders: A quantitative and qualitative analysis. *Language, Speech and Hearing Services in Schools*, **38**(4), 327–341. https://doi.org/10.1044/0161-1461(2007/035).

Paatsch, L. E. (1997). *The Effectiveness of the Auditory Skills Program in Developing Auditory Skills in Severe to Profound Hearing Impaired Children*. (Unpublished). [Master's thesis, University of Melbourne]. Melbourne, Australia.

Paden, E., & Brown, C. (1992). *Identifying early phonological needs in children with hearing loss*. MED-EL Group.

Phạm, B., McLeod, S., & Harrison, L. J. (2017). Validation and norming of the Intelligibility in Context Scale in Northern Viet Nam. *Clinical Linguistics and Phonetics*, **31**(7–9), 665–681. https://doi.org/10.1080/02699206.2017.1306110.

Piazzalunga, S., Salerni, N., Ambrogi, F., Limarzi, S., Visconti, G., & Schindler, A. (2020). Normative data and construct validity of a cross-linguistic functional speech outcome, the Intelligibility in Context Scale: Italian (ICS–I). *International Journal of Pediatric Otorhinolaryngology*, **132**, 109924. https://doi.org/10.1016/j.ijporl.2020.109924.

Sennaroğlu, L., Sennaroğlu, G., Yücel, E., Bilginer, B., Atay, G., Bajin, M. D., Mocan, B. Ö., Yaral, M., Aslan, F., Çnar, B. Ç., iÖzkan, B., Batuk, M. Ö., Kirazl, Ç. E., Karakaya, J., Ataş, A., Saraç, S., & Ziyal, İ. (2016). Long-term results of ABI in children with severe inner ear malformation. *Otology and Neurotology*, **37**(7), 865–872. https://doi.org/10.1097/MAO.0000000000001050.

Shriberg, L. D., Austin, D., Lewis, B. A., McSweeny, J. L., & Wilson, D. L. (1997). The percentage of consonants correct (PCC) metric: Extensions and reliability data. *Journal of Speech Language and Hearing Research*, **40**(4), 708–722. https://doi.org/10.1044/jslhr.4004.708.

Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: a procedure for assessing severity of involvement. *Journal of Speech and Hearing Disorders*, **47**(3), 256–270. https://doi.org/10.1044/jshd.4703.256.

Singh, S., Liasis, A., Rajput, K., Towell, A., & Luxon, L. (2004). Event-related potentials in pediatric cochlear implant patients. *Ear and Hearing*, **25**(6), 598–610. https://doi.org/10.1097/00003446-200412000-00008.

Sosa, A. V., & Bunta, F. (2019). Speech production accuracy and variability in monolingual and bilingual children with cochlear implants: A comparison to their peers with normal hearing. *Journal of Speech, Language, and Hearing Research*, **62**(8), 2601–2616. https://doi.org/10.1044/2019_JSLHR-S-18-0263.

Spencer, P. E., Marschark, M., & Spencer, L. J. (2011). Cochlear implant: Advances, issues, and implications. In M. Marschark & P. E. Spencer (Eds.), *The Oxford Handbook of Deaf Studies, Language, and Education Vol. 1* (2nd Ed., pp. 452–457). Oxford University Press.

Svirsky, M. A., Chin, S. B., & Jester, A. (2007). The effects of age at implantation on speech intelligibility in pediatric cochlear implant users: Clinical outcomes and sensitive periods. *Audiological Medicine*, **5**(4), 293–306. https://doi.org/10.1080/16513860701727847.

Tomblin, J. B., Peng, S.-C., Spencer, L. J., & Lu, N. (2008). Long-term trajectories of the development of speech sound production in pediatric cochlear implant recipients. *Journal of Speech, Language, and Hearing Research*, **51**(5), 1353–1368. https://doi.org/10.1044/1092-4388(2008/07-0083).

Van Lierde, K. M., Vinck, B. M., Baudonck, N., De Vel, E., & Dhooge, I. (2005). Comparison of the overall intelligibility, articulation, resonance, and voice characteristics between children using cochlear implants and those using bilateral hearing aids: A pilot study. *International Journal of Audiology*, **44**(8), 452–465. https://doi.org/10.1080/14992020500189146.

Venkataramani, N., Anbuchezhian, R., Maheswari, S. S., Arumugam, S. V., Raghu Nandhan, S., & Kameswaran, M. (2021). Comparison of clinician versus parental perspectives of outcomes in cochlear implantees: A south Indian experience. *Indian Journal of Otolaryngology and Head and Neck Surgery*, **73**(1), 41–44. https://doi.org/10.1007/s12070-020-01959-x.

Veritas Health Innovation. (2021). *Covidence Systematic Review Software*. Author. www.covidence.org.

Warner-Czyz, A. D., & Davis, B. L. (2008). The emergence of segmental accuracy in young cochlear implant recipients. *Cochlear Implants International*, **9**(3), 143–166. https://doi.org/10.1179/cim.2008.9.3.143.

Washington, K. N., McDonald, M. M., McLeod, S., Crowe, K., & Devonish, H. (2017). Validation of the intelligibility in context scale for Jamaican Creole-speaking preschoolers. *American Journal of Speech-Language Pathology*, **26**, 750–761. https://doi.org/10.1044/2016_AJSLP-15-0103.

Wilcox, K. M., & Morris, S. (1999). *Children Speech Intelligibility Measure*. The Psychological Corporation.

Wilkinson, A. S., & Brinton, J. C. (2003). Speech intelligibility rating of cochlear implanted children: Inter-rater reliability. *Cochlear Implants International*, **4**(1), 22–30. https://doi.org/10.1179/cim.2003.4.1.22.

World Health Organization (2007). *International Classification of Functioning, Disability and Health: Children & Youth Version*.

World Health Organization. (2021). *Deafness and hearing loss*. https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss.

Wu, C.-M., Sun, Y.-S., & Liu, T.-C. (2008). Long-term categorical auditory performance and speech intelligibility in Mandarin-speaking prelingually deaf children with early cochlear implantation in Taiwan. *Clinical Otolaryngology*, **33**(1), 35–38. https://doi.org/10.1111/j.1749-4486.2007.01585.x.

Yoon, M. S., Kim, C. S., Chang, S. O., & Sim, H. S. (2004). The predicting variables of speech-production abilities in prelingually deafened children with cochlear implantation. *Cochlear*, **5**, 130–132. https://doi.org/10.1179/cim.2004.5.Supplement-1.130.

Zachariah, S. A., & Kumaraswamy, S. (2015). Speech intelligibility of Malayalam speaking cochlear implant children. *Language in India*, **15**(12), 211–237.

Zhou, H., Chen, Z., Shi, H., Wu, Y., & Yin, S. (2013). Categories of auditory performance and speech intelligibility ratings of early-implanted children without speech training. *PLoS One*, **8**(1), e53852. https://doi.org/10.1371/journal.pone.0053852.