Effects of Noise and a Speaker's Impaired Voice Quality on Spoken Language Processing in School-Aged Children: A Systematic Review and Meta-Analysis

Isabel S. Schiller[1], Angélique Remacle[1,2], Nancy Durieux[3], and Dominique Morsomme[1]

[1]Faculty of Psychology, Speech and Language Therapy, and Educational Sciences, University of Liège, Liège, Belgium

[2]Faculty of Psychological Sciences and Education, Université Libre de Bruxelles, Brussels, Belgium

[3]ULiege Library, University of Liège, Liège, Belgium

Author Note

Isabel S. Schiller https://orcid.org/0000-0003-2387-7625

Angélique Remacle https://orcid.org/0000-0001-9338-977X

Nancy Durieux https://orcid.org/0000-0002-4688-293X

Dominique Morsomme https://orcid.org/0000-0002-7697-0498

Disclosure: We have no conflict of interest to disclose. We presented a poster with preliminary data of this systematic review at the 12th Speech in Noise Workshop in Toulouse, France, 9-10 January 2020.

Funding: Isabel S. Schiller was supported by a PhD grant from the University of Liège (no. RD/DIR-vdu/2016.7166)

Correspondence concerning this article should be addressed to Isabel S. Schiller, Rue de l'Aunaie, 30 (B38), 4000 Liège, Belgium, e-mail: isabel.schiller@uliege.be

Abstract

Purpose: Background noise and voice problems among teachers can degrade listening conditions in classrooms. The aim of this literature review is to understand how these acoustic degradations affect spoken language processing in 6-18-year-old children. *Method:* In a narrative report and meta-analysis, we review 31 studies that examined the effects of noise and/or impaired voice on children's answer accuracy and response time (RT) in listening tasks. We propose the SPADE framework to classify relevant findings according to three processing dimensions – speech perception, listening comprehension, and auditory working memory – and highlights potential moderators. *Results:* Our meta-analysis shows that noise can impede children's accuracy in listening tasks across all processing dimensions (Cohen's *d* between –0.67 and –2.65, depending on signal-to-noise ratio), and that impaired voice lowers children's accuracy in listening comprehension tasks (*d* = –0.35). A handful of studies assessed RT, but results are inconclusive. The impact of noise and impaired voice can be moderated by listener-, task-, environmental-, and exposure factors. The interaction between noise and impaired voice remains under-investigated. *Conclusions:* Overall, this review suggests that children have more trouble perceiving speech, processing verbal messages, and recalling verbal information when listening to speech in noise or to a dysphonic speaker. Impoverished speech input could impede pupils' motivation and academic performance at school.

Keywords: spoken language processing, acoustically degraded speech, meta-analysis

Effects of Noise and a Speaker's Impaired Voice Quality on Spoken Language Processing in School-Aged Children: A Systematic Review and Meta-Analysis

Children's learning and academic attainment at school may depend on a range of factors, such as student-teacher interactions (Vandenbroucke, Spilt, Verschueren, Piccinin, & Baeyens, 2018), teachers' classroom management skills (Korpershoek, Harms, de Boer, van Kuijk, & Doolaard, 2016), and teachers' ability to transfer knowledge to their students (Kraft, Blazar, & Hogan, 2018). From an acoustical perspective, effective classroom listening may be hindered by classroom noise and reverberation (Gheller, Lovo, Arsie, & Bovo, 2020; Klatte, Bergström, & Lachmann 2013; Shield & Dockrell, 2003) as well as by a teacher's poor voice quality (Martins, Pereira, Hidalgo, & Tavares, 2014; Rodrigues, Medeiros, & Teixeira, 2017). In this review, we investigate the effects of background noise and impaired (dysphonic) voice on spoken language processing in regular school-aged children.

Classrooms should provide children with acoustic conditions allowing them to understand their teacher, focus on the lesson, and learn new subject matter. Therefore, the American National Standards Institute (ANSI, 2010) recommends that noise levels in unoccupied classrooms do not exceed 35 dB(A) and reverberation times do not surpass 0.6 s (0.7 s for larger classrooms). Reverberation time is defined as the time it takes for a sound signal to decay by 60 dB in an enclosed room (International Organization for Standardization [ISO], 2008). Considering that classroom noise increases in the presence of children, signal-to-noise ratios (SNRs) should be at least +15 dB, meaning that the speech-sound level should surpass the noise level by 15 dB (Bradley & Sato, 2008; Mealings, 2016). Quite alarmingly, a recent review indicates that, in many classrooms around the world, unoccupied noise levels vary between 40-50 dB(A) – far beyond the threshold of 35 dB(A) – and SNRs roughly range between +10 dB to –10 dB (Mealings, 2016). Long reverberation times may aggravate the problem of noise (Neuman, Wroblewski, Hajicek, & Rubinstein, 2010; Valente, Plevinsky, Franco, Heinrichs-Graham, & Lewis, 2012; Wróblewski, Lewis, Valente, & Stelmachowicz, 2012). Such conditions may hinder speech communication in classrooms.

When speaking in noise, teachers may be forced to raise their voice and adapt their speaking style to make themselves understood. This phenomenon, known as Lombard speech (Garnier & Henrich, 2014), may involve over-articulation, thereby aiding audio-visual speech perception (Garnier, Ménard, & Alexandre, 2018). Although Lombard speech may promote speech communication, speaking in noisy classrooms for several hours a day can be vocally exhausting. Unsurprisingly, classroom noise is thought to elevate teachers' risk for voice disorders (Martins et al., 2014). Typical symptoms of voice disorders include hoarseness, roughness, breathiness, asthenia, and strain (De Bodt et al., 2015). The prevalence of voice disorders among teachers has been reported to range around 20-50%, sometimes even higher (Martins et al., 2014). At the same time, relatively few of the concerned teachers seek medical treatment (Van Houtte, Claeys, Wuyts, & Van Lierde, 2011). As they continue to teach with an impaired voice quality, students might find it difficult to follow the lesson – particularly in the presence of noise.

Both noise and a speaker's impaired voice are claimed to reduce speech intelligibility (Ishikawa, Nudelman, Park, & Ketring, in press; Shield & Dockrell, 2003), although they do so in different ways. Noise interferes with the transmission of the speech signal by generating masking effects (Mattys et al., 2012). Energetic masking reduces speech intelligibility due to the physical overlap between the acoustic characteristics of the speech and noise signal (Mattys, Brooks, & Cooke, 2009). Informational masking refers to the impeding effects of noise on a higher level of listening, after energetic masking effects have been accounted for (Mattys et al., 2009). In contrast to noise, impaired voice directly impedes the speech signal during its production, on a laryngeal level, and can therefore be considered a source degradation (Mattys et al., 2012). For example, irregular vocal fold vibrations create air turbulences that may be perceived as “phonation noise” (Schiller, Morsomme, Kob, & Remacle, 2020). Whether or not background noise and a speaker’s impaired voice have similar impacts on children’s spoken language processing is unclear.

Spoken language processing is the process in which the acoustic speech signal is translated into linguistic representations, allowing listeners to interpret speech and memorize speech-encoded information (Medwetsky, 2011). It involves a complex interplay of auditory, linguistic, and cognitive mechanisms (Fitzpatrick & Wheeldon, 2000). On a low level, auditory information is perceived and decoded, which is, for example, necessary for phoneme identification and discrimination (Holt & Lotto, 2010). On a higher level, understanding the content of speech demands the listener’s capacity to integrate semantic, syntactic, and pragmatic components (O’Malley, Uhl Chamot, & Küpper, 1989). This is critical to comprehend longer utterances, such as sentences or passages. Spoken language processing is also linked to working memory, because speech-information must be temporarily stored, manipulated, and recalled (Rönnberg et al., 2013). Narrative reviews indicate that children experience a higher listening effort (as can be indirectly assessed by means of RT measures; see e.g., Schiller et al., 2020), and make more processing errors when listening to speech that is degraded by noise (Gheller et al., 2020; Klatte et al., 2013; Shield & Dockrell, 2003) or a speaker’s impaired voice (Rodrigues et al., 2017). However, these effects have never been synthesized and quantified in a systematic review and meta-analysis.

Beyond that, little is known about how noise and impaired voice affect different subtasks or dimensions of children’s spoken language processing. Traditionally, children’s processing of acoustically degraded speech has been assessed by means of intelligibility tasks, in which children are asked to repeat back speech segments of different lengths. However, as highlighted above, spoken language processing entails more than just auditory-perceptual mapping. In fact, it can be disturbed even under circumstances of high intelligibility (Gheller et al., 2020; Klatte, Lachmann, & Meis, 2010). Researchers have become aware that evaluating the consequences of acoustic adversity within naturalistic listening contexts calls for a greater variety of listening tasks. Tasks that have been used include sentence-comprehension tasks (Lyberg-Åhlander, Haake, Brännström, Schötz, & Sahlén, 2015), passage-comprehension tasks (Morsomme et al., 2011), oral-instruction tasks (Klatte et al., 2007), veracity-judgement tasks (Osman & Sullivan, 2014), and recall tasks (Sullivan, Osman, & Schafer, 2015). However, unless we evaluate the findings from such tasks in a broader context, it is difficult to determine their implications for everyday listening situations.

There is also a need for analysing moderators (i.e., independent variables that influence the relationship between a predictor and the dependent variable) of the effects of noise and impaired voice on children's spoken language processing. Understanding under which circumstances children might be most vulnerable to acoustically degraded speech is critical to developing purposeful strategies for improving classroom listening.

Purpose of the Present Study

The aim of this study is to systematically review the effects of acoustically degraded listening conditions on spoken language processing in school-aged children. Specifically, we set out to qualitatively and quantitatively analyze how noise and a speaker's impaired voice influence children's answer accuracy and RT in listening tasks along different processing dimensions. We propose the Speech Processing under Acoustic DEgradations (SPADE) framework to classify and evaluate the findings from the relevant literature into broader categories and identify potential moderators. Four research questions are investigated:

- Research Question 1: To what extent does noise affect children’s spoken language processing?
- Research Question 2: To what extent does a speaker’s impaired voice affect children’s spoken language processing?
- Research Question 3: Is there an interaction between the effects of noise and a speaker’s impaired voice?
- Research Question 4: Which factors moderate the effects of noise and a speaker’s impaired voice?

The Speech Processing under Acoustic DEgradations (SPADE) framework

Our SPADE framework was developed in a bottom-up as well as a top-down manner. It is an outcome of prior knowledge deduced from psycholinguistic theory (Rueschemeyer & Gaskell, 2018) and the reviewed literature on the effects of noise and impaired voice on children’s spoken language processing. The SPADE framework classifies listening tasks into three dimensions of spoken language processing, which we will call SPADE dimensions: speech perception, listening comprehension, and auditory working memory (see S1 for a description of specific listening tasks and their categorization). The purpose of the SPADE framework is twofold. First, we seek to gain a better understanding of the effects of noise and/or a speaker’s impaired voice on children’s ability to *hear* what is being said, *understand* the content of a verbal message, and *memorize* what they have been told. To achieve this, we will meta-analytically determine these effects along each SPADE dimension. Second, we aim to identify moderators of the effects of noise and/or impaired voice on children’s spoken language processing. This will be accomplished by means of a qualitative analysis of interactions reported across the included studies. Below, we describe each SPADE dimension.

Speech perception. The dimension of speech perception includes findings from listening tasks that primarily assessed auditory-perceptual processing. An example of auditory-perceptual processing is the mapping of speech input onto linguistic representations, such as phonemes, syllables, or words (Holt & Lotto, 2010). Listening tasks assigned to this dimension are speech-intelligibility tasks presented under acoustic adversity. Semantic and syntactic skills as well as recall skills are of secondary importance. Our synthesis of research findings within this dimension will help evaluate how well children *hear* under acoustically degraded conditions.

Listening comprehension. The dimension of listening comprehension contains findings from listening tasks that primarily assess children's ability to grasp the meaning of acoustically degraded spoken utterances (see e.g., Klatte et al., 2007). Listening comprehension builds on speech perception and requires semantic and syntactic integration. It is strongly linked to working memory, because information must be temporarily retained. The tasks within this dimension test children's comprehension of longer speech segments presented under conditions of acoustic adversity. Results within this dimension will give an insight into how well children *understand* speech under acoustically degraded conditions.

Auditory working memory. The dimension of auditory working memory focuses on research finding from listening tasks that test children's ability to store, manipulate, and recall speech-encoded information under acoustically adverse conditions. Despite these skills being critical for listening comprehension, we decided to present auditory working memory as a separate dimension to highlight the aspect of time-delayed manipulation and recall. Results within the dimension of auditory working memory enable us to evaluate how well children *memorize* speech-encoded information under acoustically degraded conditions.

We will return to the SPADE framework in the results section of this review. That is, we will determine the effects of noise and a speaker's impaired voice along the three SPADE dimensions, and we will present a collection of moderating factors identified in our qualitative analysis. A scheme of the SPADE framework that visualizes the results of this systematic review in simplified form is presented later in the discussion.

Method

We conducted a systematic review and meta-analysis according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) statement (Moher, Liberati, Tetzlaff, Altman, 2009). Our protocol is available on PROSPERO (registration number: CRD42019137275). The meta-analysis was added post hoc in response to the reviewers' request.

Inclusion and Exclusion Criteria

Participants. We included studies that investigated regular school-aged children (6-18 years old), while studies that contained adult participants were excluded (e.g., Leibold & Buss, 2013). We also excluded studies that examined children with hearing impairments or developmental disorders, such as specific language impairment (e.g., Leibold, Hillock-Dunn, Duncan, Roush, and Buss, 2013).

Exposure. We included studies that assessed the impact of noise (in the following referred to as noise studies), a speaker's impaired voice (in the following referred to as voice studies), or the combined impact of noise and impaired voice in listening tasks. Noise was defined as any interfering sound that masks speech (e.g., classroom noise or speech noise). The meta-analysis only includes studies that implemented exposure conditions of ≤ 10 dB signal-to-noise ratio (the lower the SNR the higher the noise exposure). This threshold reflects the upper range of a classroom-typical noise exposure (Mealings, 2016). Impaired voice was defined as any perceptual deviance from a speaker's normal voice quality. Speech stimuli of impaired voice could result from (a) a healthy speaker imitating dysphonia, (b) a healthy speaker who underwent a vocally demanding task that temporarily provoked an impaired voice, or (c) a dysphonic patient (Schiller, Remacle, & Morsomme, 2019).

Comparison. Studies needed to implement a control or comparison condition with a lower degree of exposure. For noise studies, this implied a listening condition with no added noise or with a higher SNR than in the control. The meta-analysis only includes studies that implemented a comparison condition with no added noise or ≥ 15 dB SNR. The threshold of ≥ 15 dB SNR was applied, because it is the minimum SNR considered to provide a good classroom listening condition for children (Bradley & Sato, 2008; Mealings et al., 2016). For voice studies, the comparison condition needed to be a condition in which children listened to a speaker with a normal voice quality.

Outcome measures. We included studies that measured answer accuracy as a measure of task performance (primary outcome) and RT as a measure of listening effort (secondary outcome). We excluded studies that measured the outcome SNR in dB to assess the threshold at which listeners would perform at a certain level (e.g., Leibold et al., 2013).

Study design and publication aspects. We included interventional studies with repeated-measures or between-subjects designs. Eligibility was restricted to English-language studies published in scientific journals.

Literature Search

To identify eligible studies published up to November 2019, we searched the databases PsycINFO/Ovid, Medline/Ovid, Eric/Ovid, and Scopus (search dates: 02/2018, 08/2018, 11/2019). Our search strategies combined text words and – when relevant – controlled vocabulary (see Appendix A for our Medline/Ovid search strategy). Boolean operators connected the search terms of interest as follows: *child* AND *spoken language processing* AND (*noise* OR *impaired voice*). Additional studies were sought by hand searching the reference lists of all eligible studies.

Study Selection and Data Extraction

As a first step, two investigators (IS [first author] and DM [last author]) independently screened the title and the abstract of each study according to predefined eligibility criteria to exclude irrelevant studies. As a second step, the same two investigators independently read the full texts of all the studies selected in the first step. Disagreements about the inclusion of a study were discussed and, when necessary, conflicts were solved together with a third investigator (AR [second author]).

Coding procedure. We used an a-priori coding scheme to code the eligible studies for (1) year of publication (2) participant characteristics (i.e., number of participants included in the statistical analysis, age, and gender ratio), (3) experimental setting, including room specifications (i.e., whether the experiment was conducted in a classroom, sound-proof booth, quiet room at school), form of testing (i.e., group vs. individual testing), and auditory presentation mode (i.e., earphones vs. loudspeakers), (4) task type (e.g., word-picture matching; see supplemental Table S1 for definitions) and assigned SPADE dimension (i.e., speech perception, listening comprehension, or auditory working memory), (5) type of exposure (i.e., source of noise or impaired voice), (6) listening conditions, including the SNR levels (noise studies) and the degree of voice impairment (voice studies), (7) outcome (i.e., answer accuracy and/or RT), (8) effect-size data (i.e., means and standard deviations [*SD*s]; when only median and interquartile ranges [IQR] were available, we followed Hozo, Djulbegovic, and Hozo’s [2005] approach and converted them into means and *SD*s.), (9) main finding regarding the effect in question.

The data extraction and coding were independently performed by two investigators (IS and DM). Conflicts were discussed and solved in consensus with a third investigator

(AR). Authors of studies were contacted to obtain any information that could not be retrieved from the text.

Quality Assessment

Two reviewers (IS and DM) independently appraised the quality of included studies using a shortened version of the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies of the National Heart, Lung, and Blood Institute of the National Institutes of Health (NHLBI, 2019). Despite this assessment tool being developed for observational and cross-sectional studies, we chose it because it included most aspects we considered critical to appraise the quality of the included studies (e.g., whether the study population was clearly described and recruited based on pre-defined criteria, or whether key confounding variables were considered). The original tool contains 14 questions (see Appendix B). We removed questions 6, 7, 10, and 13 because they were irrelevant to our studies of interest (i.e., studies providing a snapshot of exposures on children's spoken language processing at a fixed point in time). The remaining questions were answered with yes, no, cannot determine, not reported, and not applicable. Overall study quality was rated on a 3-point scale (i.e., good, fair, poor), based on personal appraisal. Note that with the NHLBI quality assessment tool, the overall quality is not rated on fixed criteria. Raters are supposed to evaluate the internal validity of each study based on the risk of bias resulting from their responses to the sub-question of the tool.

Data Synthesis and Statistical Analysis

The qualitative analysis includes a description of the included studies in the form of a table, which is organized into studies investigating the effect of noise, impaired voice, and their combined effect. In addition, with respect to the SPADE framework, the qualitative analysis entails a narrative report on moderators of the effects of noise and impaired voice that have been identified across the included studies.

The quantitative analysis was performed on a subset of the studies included in this systematic review (i.e., studies from which we could determine effect sizes). Using RStudio software, version 1.2.5019 (RStudio Team, 2019), we statistically quantified the effects of noise and impaired voice on children's answer accuracy and RT by means of random-effects models. Random-effects models were used because heterogeneity was expected between the studies. The Standardized Mean Difference (SMD, Cohen's *d*) with a 95% CI and *p*-value was assumed as an effect size, considering the variety of listening tasks applied across the studies. Heterogeneity was assessed using Cochrane's *Q* statistic and I^2, a percentage estimate of inconsistency across studies (an I^2 of 0% indicates no observed heterogeneity, while an I^2 of 100% indicates maximal observed heterogeneity; Higgins et al., 2003). We ran several meta-analyses and subgroup-analyses to identify differences in the effects of noise and impaired voice with regard to the SPADE dimensions and, in the case of noise-studies, SNR.

To assess the impact of noise on children's answer accuracy, we stratified data into five SNR-bins: (1) +6 to +10 dB, (2) +1 to +5 dB, (3) 0 dB, (4) –1 to –5 dB, (5) –6 to –12 dB, taking into account that children's susceptibility to noise varies with exposure level. Separate meta-analyses were performed for each SNR bin. Whenever possible, we carried out subgroup-analyses to test whether effects would vary with respect to SPADE dimension. Differences between groups were assessed using χ^2 –tests. Some studies assessed the effects of (1) children listening through an L2 (i.e., a second language) instead of their native language, (2) different noise sources, (3) different processing dimensions, or (4) different SNRs falling within the range of the same SNR bin. In these cases, data considered for the meta-analysis was restricted to (1) data from children listening through their native language, (2) classroom noise, (3) the dimension of listening comprehension, or (4) the lower SNR condition. These restrictions were necessary to avoid data from the same participants feeding into the meta-analysis more than once. The effect of noise on RT was assessed in a single meta-analysis; no SNR-stratification was applied due to the low number of studies, but a subgroup analysis according to SPADE dimension.

To assess the impact of a speaker's impaired voice on children's answer accuracy in listening tasks, we ran separate meta-analyses for listening comprehension and auditory working memory (speech perception was only assessed in one study; Morsomme et al., 2011). In contrast to how we proceeded with the noise studies, we took into account variation regarding SPADE dimensions by performing separate meta-analyses. The reason was that several of the impaired-voice studies assessed the effects of impaired voice within more than one processing dimension and we needed to ensure that data from the same participants would be considered only once in each meta-analysis. Some studies assessed the effect of (1) degree of dysphonia or (2) children listening through an L2 instead of their native language. In these cases, data considered for the meta-analysis was restricted to (1) the more severe dysphonic-voice exposure, and (2) data from children listening through their native language. No meta-analysis was carried out on the impact of impaired voice on children's RT, as there was only one relevant study available (Sahlén et al, 2017).

Results

Our literature search resulted in a total of 5853 records identified through database searching, and another three studies identified through reference-list searching (Figure 1). After removing duplicates and performing the study-screening steps, we included 31 studies in our systematic review, twenty of which fed into the meta-analysis. In the following, we first describe and qualitatively synthesize the studies included in the systematic review. We also present the results from the study-quality assessment. Second, we quantitatively synthesize the studies included in the meta-analysis. Third, we narratively report on moderators of the effects of noise and impaired voice.

Qualitative Synthesis

Thirty-one studies were included in the qualitative synthesis (see Table 1). The effect of noise was assessed in 21 studies, the effect of impaired voice in eight studies, and the combined effect of noise and impaired voice in two studies. The qualitative synthesis comprises data from over 3,000 listeners between the age of six and 17 years. A variety of listening tasks (e.g., word-picture matching, passage comprehension, etc.) and interventional settings (e.g., group testing in a classroom, individual testing in a sound-proof booth, etc.) were applied. The most frequent noise source was multi-talker babble (*n* = 13), followed by speech-shaped noise (*n* = 5), and classroom noise (*n* = 4). SNRs ranged from +30 dB to –10 dB. Impaired-voice was either imitated (*n* = 4), provoked (*n* = 5), or obtained from dysphonic patients (*n* = 1). Dysphonia severity ranged from mild to severe. Answer accuracy was assessed in 97% of the studies. RT was assessed in 23% of the studies.

Most noise studies reported negative effects of noise on children’s answer accuracy and RT (see last column of Table 1). Likewise, most voice studies found that impaired voice lowered children’s answer accuracy, and – in the one relevant study (Sahlén et al., 2017) – prolonged RT. Neither of the two studies that assessed the combined effect of noise and impaired voice revealed a statistically significant interaction in children’s answer accuracy (Brännström, von Lochow et al., 2018; von Lochow et al., 2018).

Quality assessment. Figure 2 shows the results from the quality assessment of the included studies. Overall study quality was rated good in 71% of the studies and fair in the remaining 29%. All studies clearly stated the study aim (Q1). Fifty-two percent of the studies used suitable outcome measures (Q11; i.e., the tools or methods applied for assessing outcomes were accurate and reliable). Different levels of exposure (i.e., different SNRs) were assessed in 45% of the studies and not applicable in 48% (Q8). Key confounding factors, such as gender or potential speech-language difficulties, were considered and analyzed in 32% of studies. However, this aspect could not be determined in another 39% (Q14).

The quality assessment also revealed some methodological weaknesses. Fifty-two percent of the studies did not provide a power estimation or effect size measure (Q5). In 48% of the studies, the study population was not clearly specified (Q2), mainly due to an insufficient assessment of language skills. Eligibility criteria were not clearly reported in 48% (Q4).

Results were inconclusive regarding participation rate (Q3; i.e., at least 50% of eligible children actually participated in the study), suitability of exposures (Q9), and blinding of outcome assessors (Q12). Participation rate of eligible persons was not reported in 74% of studies, although the > 50% participation criterion was likely met by most of them. In 39% of studies, we could not determine whether the exposure measure was suitable and reliable. None of the studies reported whether outcome assessors were blinded.

Meta-Analysis

The quantitative analysis is based on 20/31 studies (from the 11 remaining studies, we could either not obtain effect-size data or there was no control condition of ≥ 15 dB SNR). Results for each exposure are reported separately. First, we present the effects of noise on children’s spoken language processing, taking into account differences regarding SNR and SPADE dimension. Second, we present the effects of impaired voice.

Effects of noise on children’s spoken language processing. Noise significantly impeded children’s answer accuracy in each SNR bin (i.e., +6 to +10 dB, +1 to +5 dB, 0 dB; –1 to –5 dB, and –6 to –12 dB), with medium to large effect sizes of between –0.67 and –2.65 (Cohen’s *d*) and *p*-values below 0.01 (see Table 1 and, for additional forest plots, the supplemental Figures S1-S5). Taken together, results from these meta-analyses indicated that effect sizes increased as SNRs decreased (i.e., became more adverse). Results from subgroup analyses were only meaningful for two of the five SNR bins (i.e., +6 to +10 dB and 0 dB) and no significant subgroup differences was found (i.e., +6 to +10 dB SNR: $\chi^2 = 0.25$, $p = 0.62$; 0 dB SNR: $\chi^2 = 3.01$, $p = 0.08$). Regarding the remaining SNR bins, subgroups were either made up of only one study (this applies to the +6 to +10 dB bin and the –1 to –5 dB bin) or all studies assessed the same dimension (this applies to the –6 dB to –12 dB bin). With one exception, there was considerable heterogeneity across the studies (I^2 values of 89% and above, with p-values below 0.01). No heterogeneity was found in the +6 to +10 dB SNR bin. However, a look at the forest plot (Figure S1) indicates that this is due to the high variance in the study of Nelson et al. (2005).

The effect of noise on children's RT in listening tasks was only small ($d = 0.2$, CI [–0.11, 0.5]) and statistically not significant ($z = 1.28$, $p = 0.20$). Figure S6 presents the forest plot, grouped by speech perception (3 studies), listening comprehension (1 study), and auditory working memory (1 study). Substantial heterogeneity was found across the studies ($I^2 = 67\%$, $p = 0.02$).

Effects of impaired voice on children's spoken language processing. Two meta-analyses were carried out to investigate the effect of impaired voice on children's performance in listening comprehension tasks and auditory working memory tasks. As shown in Table 3 and the corresponding forest plots (Figures S7 and S8), children's answer accuracy in the dimension of listening comprehension significantly dropped when the speaker's voice was dysphonic ($d = –0.35$, CI [–0.59; –0.11], $z = –2.28$, $p < 0.01$), which was not the case for the dimension of auditory working memory ($d = –0.13$, CI [–0.72; –0.46], $z = –0.42$, $p = 0.67$). Substantial heterogeneity was found across the studies (listening comprehension: $I^2 = 73\%$, $p < 0.01$; auditory working memory: $I^2 = 67\%$, $p = 0.08$). Although we could not run a meta-analysis on the dimension of speech perception, it is worth mentioning that the study that assessed this dimension (Morsomme et al., 2011) revealed a significant drop in children's answer accuracy when the speaker's voice quality changed from normal to impaired (M = 7.83, SD = 2.7 vs. M = 4.54, SD = 1.71).

Narrative Report

Moderators of the effects of noise and impaired voice. To better evaluate the challenges of classroom listening, it is important to consider other factors that may influence children's processing of speech in noise and a speaker's impaired voice. In the qualitative analysis of the 31 studies included in this systematic review, we identified several moderators. These relate to the listener (i.e., age, gender, language proficiency), the task (i.e., cognitive demands induced by the listening task), the environment (i.e., reverberation), and the exposure (i.e., noise source).

Age. Several studies demonstrated that the effect of noise on children's spoken language processing is age-dependent (Bradley and Sato, 2008; Howard et al., 2010; Nelson et al., 2005; Peng et al., 2016; Peng & Jiang, 2016; Peng & Wu, 2018; Yacullo & Hawkins, 1987). The younger the child, the higher their susceptibility to noise. For example, Bradley and Sato (2008) found that children's answer accuracy from an optimal to a 0-dB-SNR listening condition dropped by 10%, 16%, and 24% in 11-, 8-, and 6-year-olds, respectively. There was no evidence suggesting that the effect of impaired voice might be age-dependent.

Gender. The processing of spoken language under degraded listening conditions may be affected by children's gender, although results were inconclusive. Sahlén et al. (2017) showed that listening to an impaired voice significantly increased RT in girls. However, Morsomme et al. (2011) did not observe a significant interaction between voice quality and children's gender on answer accuracy. Likewise, Prodi, Visentin, Borella, et al. (2019) found no significant interaction between noise and gender on answer accuracy or RT.

Language proficiency. The degree to which noise affects children's spoken language processing may vary with the child's familiarity with the speaker's language (Crandell & Smaldino, 1996; Nelson et al., 2005). Crandell and Smaldino (1996) found that speech perception among L2 learners was significantly more susceptible to decreasing SNRs than among children who listened through their native language. Likewise, Nelson et al. (2005) observed that children's speech perception in noise (but not in quiet) deteriorated when speech was presented in their L2. On the other hand, no such interaction was found regarding a speaker's impaired voice quality (Chui & Ma, 2018).

Cognitive demands. High task demands appear to be met less readily when listening to speech in noise (Howard et al., 2016; Prodi, Visentin, Peretti, et al., 2019; Sullivan et al., 2015) or to a dysphonic speaker (Lyberg-Åhlander, Haake, et al., 2015; Sahlén et al., 2017). Howard et al. (2016) found that the effect of noise on children's speech perception was significantly more impeding when children were asked to simultaneously retain digits. Regarding voice quality, Lyberg-Åhlander, Haake, et al. (2015) and Sahlén et al. (2017) showed that impaired voice may impede children's sentence comprehension, but only in difficult sentences. The interplay between task demand and listening conditions depends on children's cognitive ability to respond to these demands (Brännström, Kastberg et al., 2018).

Reverberation time. Although poor room acoustics – particularly long reverberation times – may generally impede children's spoken language processing (Bradley & Sato; Peng et al., 2016; Peng & Jiang, 2016; Peng & Wu, 2018), results regarding interactions with noise were inconclusive. In Peng et al. (2016), the effect of noise on children's answer accuracy in a speech perception task significantly increased as the reverberation time changed from 0.83 s to 1.30 s, while Picou et al. (2019) found no such interaction for reverberation times of 0.1 s vs. 0.83 s. Yet another study suggested that the effect of reverberation time on children's speech-in-noise perception might depend on the SNR (Hurtig et al., 2016). There was no study on the interaction between impaired voice and reverberation time.

Noise source. The effect of noise may depend on the noise source (Klatte et al., 2007; Peng et al., 2016; Prodi, Visentin, Borella, et al., 2019; Zhang et al., 2019). In Peng et al. (2016), speech-shaped noise and babble noise had greater impacts on children's answer accuracy in a speech perception task than impact, fan, or traffic noise. Pointing in a similar direction, several other studies indicate that classroom- and babble noise may be more detrimental for children's answer accuracy and RT than traffic noise (Klatte et al., 2007; Prodi, Visentin, Borella, et al., 2019; Zhang et al., 2019).

Summary of the Results

Taken together, the qualitative and quantitative syntheses demonstrated that both noise and a speaker's impaired voice have serious consequences for children's spoken language processing. Acoustically degraded speech can lower children's performance in all three SPADE dimensions. Regarding Research Question 1, noise effects on children's answer accuracy were medium to large (Cohen's *d* varied between –0.67 and –2.65), and deteriorated with decreasing SNR. Regarding Research Question 2, impaired-voice effects were small and only statistically significant in the dimension of listening comprehension (*d* = –0.35, CI [–0.59, –0.11]). Regarding Research Question 3, there was no evidence for an interaction between noise and a speaker's impaired voice, but this was based on only two studies. Regarding Research Question 4, the effect of noise may be moderated by children's age and language proficiency, the cognitive demands induced by the task, reverberation, and the noise source. The effect of impaired voice may be moderated by children's gender and the cognitive demands induced by the task.

Discussion

Children’s ability to process acoustically impoverished speech has received increasing attention in the context of education, due to the high levels of classroom noise (Mealings, 2016) and the problem of dysphonia among teachers (Martins et al., 2014). In this systematic review and meta-analysis, we synthesized and quantified the effects of noise and a speaker’s impaired voice on spoken language processing in regular school-aged children, based on accuracy and RT measures. We found evidence that noise and impaired voice exert their influence along different areas of spoken language processing, spanning from the basic identification of phonemes to the complex comprehension of text.

Effects of Noise on Children’s Spoken Language Processing

Regarding Research Question 1 (i.e., to what extent does noise affect children’s spoken language processing), our meta-analysis revealed that noise-induced impediments on answer accuracy decreased with increasing SNR, but even in the most favorable SNR bin (i.e., +6 to +10 dB SNR), effect sizes were still medium to large. Viewed from another angle, small SNR gains of about 5 dB may already improve children’s spoken language processing in noise. Our meta-analysis revealed a small RT increase in noise, which was however not statistically significant. The substantial heterogeneity in study outcomes probably contributed to the fact that the significance level was not reached. Interestingly, however, one of the included studies showed that noise slowed down children’s processing of spoken language even when performance was unaltered (Prodi, Visentin, and Borella, et al., 2019), which was confirmed in a later study by Schiller et al. (2020). More RT studies should be carried out to better understand subtle noise effects.

In summary, these findings are worrisome, especially since most classrooms exhibit SNRs roughly varying between +10 dB to –10 dB (Mealings, 2016). Educational stakeholders are in demand to find solutions to tackle the problem of classroom noise. Our review underlines the important of adhering to official guidelines for classroom acoustics, such the ANSI guideline (2010), which states that maximum unoccupied noise levels should not exceed 35 dB(A) and reverberation time should not exceed 0.6-0.7 s. In light of our review, we agree with authors of previous studies who stressed that SNRs in classrooms should be at least +15 dB, preferably higher (Bradley & Sato, 2008; Mealings, 2016).

Our subgroup analyses showed that noise may affect all SPADE dimensions, although we could not confirm that one SPADE dimension might be particularly susceptible to noise. Whenever applicable, out meta-analysis showed no significant subgroup difference between speech perception and listening comprehension (auditory working memory was not assessed in a sufficient number of studies). Even minor noise disruptions generated large effects sizes in both dimensions. This highlights that, beyond bottom-up processing, listening to speech in noise impedes top-down processing, probably due to the increased mental effort. Children become less efficient in processing verbal information. The concept of listening effort is thoroughly discussed in Pichora-Fuller et al. (2016).

In real-life classroom listening, various factors may enhance or attenuate the effect of noise. Regarding Research Question 4 (i.e., which factors moderate the effects of noise and a speaker’s impaired voice?), our qualitative analysis showed that the effect of noise may interact with factors relating to the listener (i.e., language proficiency and age), the environment (i.e., reverberation), the exposure (i.e., noise source), and the task (i.e., the cognitive demands induced by the task).

Age was the most frequently investigated moderator of children’s speech-in-noise perception. Literature suggests that up to the age of 12, pupils require at least +10 dB SNR for optimal speech perception (Howard et al., 2010; Nelson et al., 2005; Peng et al., 2016; Peng

& Jiang, 2016; Peng & Wu, 2018; Yacullo & Hawkins, 1987). Younger children, aged between 6 and 10 years, may already show performance deficits at relatively favorable SNRs of +19 to +15 dB (Bradley and Sato, 2008; Peng & Wu, 2018). These findings generally align with narrative reviews (Gheller et al., 2020; Shield & Dockrell, 2003; Klatte et al., 2013), according to which children up to about 13 years are particularly susceptible to noise. The age effect relates to the trajectory of spoken language development: younger children's auditory and language systems are less developed and their ability to integrate sensory auditory information is less mature than in adults (Talarico et al., 2007).

Noise source was also identified as a moderator of noise effects on children's spoken language processing. The finding that different noise sources induce different masking effects was made in several of the included studies (Klatte et al., 2007; Peng et al., 2016; Prodi, Visentin, Borella, et al., 2019; Zhang et al., 2019). This variation is probably associated with the degree of energetic masking and informational masking of each individual noise source in relation to the speech signal (Mattys et al., 2009, 2012). We wish to stress that noise sources children typically encounter at school (i.e., classroom noise, competing speech, and babble noise) were more impeding than those less relevant for classroom listening (e.g., train noise; Klatte et al., 2007; Prodi, Visentin, Borella, et al., 2019; Zhang et al., 2019).

Another factor that moderated speech-in-noise perception was children's language proficiency. Two studies concurringly indicated that children are more disturbed by noise when listening through a second language instead of their native language (Crandell & Smaldino, 1996; Nelson et al., 2005). It is assumed that the earlier a child acquires an L2, the better their phonologic coding abilities and the more robust their speech-in-noise perception (Gheller et al., 2020). Beyond that, we assume that L2 learners experience a lower benefit from top-down effects (e.g., semantic and syntactic priming). The increasing number of L2 learners in classrooms (Geay, McNally, & Telhaj, 2013) stresses the need for noise control.

The cognitive demands associated with the listening task may moderate the effect of noise on children's spoken language processing (Howard et al., 2016; Prodi, Visentin, Peretti, et al., 2019; Sullivan et al., 2015). However, compared to the moderators discussed thus far, findings from the included studies do not warrant a firm conclusion. The critical question seems to be, whether a child has the cognitive ability to meet the demands induced by a given task. Future studies should focus on this aspect.

Finally, reverberation may moderate the effect of noise on children's spoken language processing. However, while reverberation is generally a well-recognized predictor of classroom listening (Gheller et al., 2020), evidence from the included studies on its interaction with noise was weak. In one study, the effect of noise increased with longer reverberation time (Peng et al., 2016), while in another study it did not (Picou et al., 2019). A third study indicated that interplay between reverberation and noise might depend on SNR (Hurtig et al., 2016). While these findings provide little clarity, several studies not included in this review (mostly because the study population contained adults) have already demonstrated that reverberation might boost the effect of noise on spoken language processing (Neuman et al., 2010; Valente et al., 2012; Wróblewski et al., 2012).

Effects of Impaired Voice on Children's Spoken Language Processing

In line with the impeding effect of noise, a speaker's impaired voice can also lower children's answer accuracy in listening tasks. Regarding Research Question 2 (i.e., to what extent does a speaker's impaired voice affect children's spoken language processing), our meta-analysis revealed a small to medium, yet significant effect of impaired voice on children's listening comprehension. The effect of impaired voice on auditory working memory was small and not statistically significant. Unfortunately, we could not quantify the effect of impaired voice on speech perception, because only one study assessed this dimension (Morsomme et al., 2011). This was also true for RT, which was only assessed by

Sahlén et al. (2017). Still, the results from these studies provide a first indication that listening to a dysphonic voice lowers children's accuracy in speech perception tasks (Morsomme et al., 2011) and prolongs their RT in listening comprehension tasks (Sahlén et al., 2017. More recent works support these findings (Schiller et al., 2020; Schiller, Morsomme, Kob & Remacle, 2021).

Regarding Research Question 4 (i.e., the effect of moderators), children's processing of dysphonic speech may be moderated by the cognitive demands a listening task places on the child. Evidence suggests that listening to an impaired voice becomes significantly harder for children – both in terms of performance and listening effort – when the task induces a high processing load (Lyberg-Åhlander, Haake, et al., 2015; Sahlén et al., 2017). The reason may be that children "allocate capacity to the processing of the voice signal at the expense of listening comprehension, particularly when the linguistic difficulty is of borderline complexity for the child" (Lyberg-Åhlander, Brännström, et al., 2015, p.2). Note that, in a recent study, Schiller et al. (2021) also observed a significant interaction between cognitive demands and the speaker's voice quality on children's listening comprehension. However, results pointed in a different direction: children's comprehension of medium and difficult sentences did not vary with respect to the speaker's voice quality, while their comprehension of easy sentences was significantly better in the impaired-voice condition than in the normal-voice condition. The interaction between cognitive demands and the speaker's voice quality should be investigated in future works.

The impact of a speaker's impaired voice on children's spoken language processing may also be moderated by children's gender. However, this finding was based on a single study, in which dysphonic speech prolonged response times in girls but not boys (Sahlén et al., 2017). In another reviewed study that assessed children's answer accuracy as a function of the speaker's voice quality, no interaction with gender was found (Morsomme et al., 2011). Future studies should take into account potential gender effects.

Surprisingly, there was no indication of the degree of dysphonia moderating children's spoken language processing. This aspect was investigated in two of the studies included in this review (Chui & Ma, 2018; Rogerson & Dodd, 2005). In both studies, a speaker's impaired voice quality impeded children's listening comprehension, but this effect was not stronger with a higher degree of dysphonia. This has important implications for the educational setting. Even if a teacher only has a minor voice problem, this may degrade the speech signal and consequently disturb children's classroom listening.

Combined Effect of Noise and Impaired Voice on Children's Spoken Language Processing

In this systematic review and meta-analysis, we were not only interested in the isolated effects on noise and a speaker's impaired voice but also intrigued by whether these two types of acoustic degradations might interact (Research Question 3). Intuitively, listening to speech in noise should be particularly challenging when the speaker's voice quality was impaired. To our surprise, the two reviewed studies that addressed this question did not find an interaction between noise and a speaker's voice quality (Brännstöm, von Lochow et al., 2018; von Lochow et al., 2018). In fact, there was not even a main effect of voice quality. With respect to the notion that, during classroom listening, children are often exposed to noise and a teacher's impaired voice at the same time, the interplay between these two factors deserves further investigation. Schiller et al. (2020) recently picked up on that research topic and showed that 6-year-olds were significantly more disturbed by noise when the speaker's voice was impaired rather than normal.

Integrating Qualitative and Quantitative Findings Regarding the Effects of Noise and Impaired Voice

A visual summary of the effects of noise and a speaker's impaired voice on children's spoken language processing is provided in the form of a schematic presentation of the SPADE framework (Figure 3). The scheme highlights that both noise and a speaker's impaired voice can affect spoken language processing, although there is no proof of an interaction. The effects of noise and impaired voice may be moderated by other factors that relate to the listener, the task, the environment, and the exposure.

Our findings, as synthesized by means of the SPADE framework, are in line with earlier narrative reviews indicating that noise and a speaker's impaired voice disrupt children's listening efficiency (Gheller et al., 2020; Rodrigues et al., 2017; Shield & Dockrell, 2003; Klatte et al., 2013). It also expands on these findings, by quantifying these effects along different dimensions of spoken language processing and by identifying moderators. Applied to classroom listening, the sum of our findings suggests that children have trouble *hearing* their teacher, *understanding* the content of their speech, and *memorizing* information if noise levels are too high or if the teacher's voice quality is poor. This would be particularly true for speech and classroom noise and for children who are in the lower grades. The more capacity is needed to process the content of speech, the more likely it is that a teacher's impaired voice causes disruptions. Monitoring a teacher's vocal health in addition to reducing classroom noise is therefore essential.

Limitations

In the following, we discuss limitations of this systematic review and meta-analysis, which relate to the quality of the included studies and to our methodological choices.

First, regarding study quality, we could not always evaluate whether critical confounding factors (i.e., variables that may also affect the outcome variable but were not accounted for) were taken into consideration. For example, especially among noise studies, participants' language skills were often not adequately assessed. Thus, some of the investigated children might have had specific language impairments, which might have influenced the results. In many cases, inclusion and exclusion criteria were either not clearly specified or not rigorously reported.

Second, we did not find a quality assessment for interventional studies that entirely matched our needs. After carefully comparing different tools, we eventually opted for the NHLBI assessment tool for Observational Cohort and Cross-Sectional Studies (NHLBI, 2019). However, we had to remove four questions that were irrelevant to our studies in focus.

A third limitation of this study is the substantial heterogeneity we found between the study outcomes in the meta-analyses. We can only speculate on what caused this heterogeneity. A likely reason is that this was due to the different scales with which accuracy was measured and the different tasks and techniques applied to assess RT. Publication bias could have also contributed to the heterogeneity. However, no publication bias analysis was performed because each meta-analysis featured fewer than 10 studies.

Fourth, there may be further outcomes not considered in this systematic review and meta-analysis, which may help evaluate children's spoken language processing under acoustically challenging conditions. Examples are physiological outcomes, such as pupil dilation (McGarrigle et al., 2017), behavioural outcomes, such as SNR required for a certain performance level (Leibold et al., 2013), and neuro-psychological measures. Investigating more outcomes might have provided more extensive results (e.g., identifying neural markers of listening effort) but would have resulted in an unmanageable amount of eligible studies.

Finally, regarding the SPADE framework, it should be acknowledged that some of the moderators of the effects of noise and a speaker’s impaired voice were investigated in only a few of the included studies. An example is reverberation, which is generally known to increase the negative effect of noise on speech intelligibility (Neuman et al., 2010; Valente et al., 2012; Wróblewski et al., 2012). There may also be further moderators which were not investigated in any of the 31 reviewed studies, such as the speaker’s speech rate (Schiller, Morsomme, Kob, and Remacle, 2019), Lombard speech (Garnier & Henrich, 2014), and the presence of visual cues (Garnier et al., 2018). It will also be interesting to investigate whether non-typically developing children might be more vulnerable to acoustic degradations than typically developing children.

Recommendations

Listening to a dysphonic teacher in the presence of background noise is probably more difficult than listening to a healthy teacher in noise. Still, there is little research on potential interactions between noise and a speaker’s voice quality on children’s spoken language processing. The two studies that looked at this issue were conducted in the dimension of listening comprehension and did not suggest firm conclusions. Future studies should develop interventional designs incorporating both factors in an aim to close this research gap. In particular, speech perception and auditory working memory deserve more attention in this respect.

More research is needed to enhance and expand the SPADE framework, which targets the shadow areas in this research field. Further studies are necessary to confirm the impact of the identified moderators and explore their strength of moderation. Other factors which might influence children’s spoken language processing under adverse listening conditions should also be assessed. For example, it is still uncertain to which degree children’s executive functioning skills may predict children's ability to process a speaker's impaired voice (Lyberg-Åhlander, Holm, et al., 2015). Moreover, it is important to understand if children's processing of impaired voice may be moderated by the type of dysphonia, reverberation time, children's age, or their language proficiency. Regarding noise, more clarity is needed concerning the influence of reverberation (i.e., does reverberation rather lead to interaction effects or additive effects?). We hope that researchers will set out to investigate these aspects and complete the SPADE framework.

The SPADE framework may be useful to researchers, but also policy makers in the field of education. It may help psycholinguists, speech pathologists, and pedagogues to better understand spoken language processing areas susceptible to acoustic degradations and to gain an overview of listening tasks conducted in the past. Educational policy makers may use the framework to develop strategies to improve listening conditions in classrooms. For example, with regard to noise, a strategy could be to ensure that the youngest pupils are taught in the quietest classrooms.

Conclusion

In this systematic review and meta-analysis, we presented evidence for the adverse effects of noise and a speaker's impaired voice on children's spoken language processing. An evaluation of findings from listening tasks revealed that children make more processing errors and tend to take more time to process speech when the speech signal is acoustically degraded. By synthesizing and classifying results for the included studies in the SPADE framework, we illustrated that impediments may affect speech perception, listening comprehension, and auditory working memory, and are moderated by other factors such as a listener's age or their language proficiency. Due to a lack of studies, we were unable to determine the combined effect of noise and a speaker's impaired voice on children's spoken language processing. Future research in this domain is necessary to predict the challenges faced by school-aged children when listening and communicating in classrooms and to identify benchmarks to ensure quality communication.

Acknowledgements

This work was supported by a PhD grant of the University of Liège (N°: RD/DIR–vdu/2016.7166). The funding sources had no involvement in the preparation of this systematic review and in the decision to submit the article for publication. We thank the authors who responded to our e-mails and provided additional data on their studies. We also thank Charlotte Beaudart for her help with the statistical analyses.

References

References marked with an asterisk were included in the systematic review.

American National Standards Institute (ANSI). (2010). *Acoustical performance criteria, design requirements, and guidelines for schools, part 1: Permanent Schools*. ANSI S12.60-2010/Part 1. New York, NY: Acoustical Society of America.

*Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *Journal of the Acoustical Society of America*, *123*, 2078–2086. doi: doi: https://doi.org/10.1121/1.2839285

*Brännström, K. J., Kastberg, T., von Lochow, H., Haake, M., Sahlén, B., & Lyberg-Åhlander, V. (2018). The influence of voice quality on sentence processing and recall performance in school-age children with normal hearing. *Speech, Language and Hearing*, *21*, 1–9. https://doi.org/10.1080/2050571X.2017.1309787

*Brännström, K. J., von Lochow, H., Lyberg-Åhlander, V., & Sahlén, B. (2018). Immediate passage comprehension and encoding of information into long-term memory in children with normal hearing: The effect of voice quality and multitalker babble in noise. *American Journal of Audiology*, *27*, 231–237. https://doi.org/10.1044/2018_AJA-17-0061

*Chui, J. C. H., & Ma, E. P. M. (2018). The impact of dysphonic voices on children's comprehension of spoken language. *Journal of Voice*, *33*, 801.e7–801.e16. https://doi.org/10.1016/j.jvoice.2018.03.004

*Crandell, C. C., & Smaldino, J. J. (1996). Speech perception in noise by children for whom English is a second language. *American Journal of Audiology*, *5*, 47–51. https://doi.org/10.1044/1059-0889.0503.47

De Bodt, M., Van den Steen, L., Mertens, F., Raes, J., Van Bel, L., Heylen, L., Pattyn, J., Gordts, F., & Van de Heyning, P. (2015). Characteristics of a dysphonic population referred for voice assessment and/or voice therapy. *Folia Phoniatrica et Logopaedica*, *67*, 178–186. https://doi.org/10.1159/000369339

Fitzpatrick, J., & Wheeldon, L. (2000). Phonology and phonetics in psycholinguistic models of speech perception. In N. Burton-Roberts, P. Carr, & G. Docherty (Eds.), *Phonological Knowledge: Conceptual and Empirican Issues* (pp. 131–160). Oxford, UK: Oxford University Press.

Garnier, M., & Henrich, N. (2014). Speaking in noise: How does the Lombard effect improve acoustic contrasts between speech and ambient noise? *Computer Speech & Language, 28*, 580–597. https://doi.org/10.1016/j.csl.2013.07.005

Garnier, M., Ménard, L., & Alexandre, B. (2018). Hyper-articulation in Lombard speech: An active communicative strategy to enhance visible speech cues? *The Journal of the Acoustical Society of America*, *144*(2), 1059–1074. https://doi.org/10.1121/1.5051321

Geay, C., McNally, S., & Telhaj, S. (2013). Non-native speakers of English in the classroom: what are the effects on pupil performance?. *The Economic Journal*, *123*, F281-F307. https://doi.org/10.1111/ecoj.12054

Gheller, F., Lovo, E., Arsie, A., & Bovo, R. (2020). Classroom acoustics: Listening problems in children. *Building Acoustics*, *27*(1), 47–59. https://doi.org/10.1177/1351010X19886035

Higgins, J. P., Thompson, S. G., Deeks, J. J., & Altman, D. G. (2003). Measuring inconsistency in meta-analyses. *BMJ*, *327,* Article 7414, 557–560. https://doi.org/10.1136/bmj.327.7414.557

Holt, L. L., & Lotto, A. J. (2010). Speech perception as categorization. *Attention, Perception & Psychophysics*, *72*, 1218–1227. https://doi.org/10.3758/APP.72.5.1218

*Howard, C. S., Munro, K. J., & Plack, C. J. (2010). Listening effort at signal-to-noise ratios that are typical of the school classroom. *International Journal of Audiology*, *49*, 928–932. https://doi.org/10.3109/14992027.2010.520036

Hozo, S. P., Djulbegovic, B., & Hozo, I. (2005). Estimating the mean and variance from the median, range, and the size of a sample. *BMC medical research methodology*, *5*(1), Article 13. https://doi.org/10.1186/1471-2288-5-13

*Hurtig, A., Keus van de Poll, M., Pekkola, E. P., Hygge, S., Ljung, R., & Sörqvist, P. (2016). Children’s recall of words spoken in their first and second language: Effects of signal-to-noise ratio and reverberation time. *Frontiers in Psychology*, *6,* Article 2029. https://doi.org/10.3389/fpsyg.2015.02029

Ishikawa, K., Nudelman, C., Park, S., & Ketring, C. (in press). Perception and acoustic studies of vowel intelligibility in dysphonic speech. *Journal of Voice*. doi:10.1016/j.jvoice.2019.12.022

ISO (2008). *ISO 3382-2 – Acoustics – Measurement of room acoustic parameters – Part 2: Reverberation time in ordinary rooms*. International Organization for Standardization (ISO).

*Jamieson, D. G., Kranjc, G., Yu, K., & Hodgetts, W. E. (2004). Speech intelligibility of young school-aged children in the presence of real-life classroom noise. *Journal of the American Academy of Audiology*, *517*, 508–517. https://doi.org/10.3766/jaaa.15.7.5

Klatte, M., Bergström, K., & Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Frontiers in Psychology*, *4,* 578. https://doi.org/10.3389/fpsyg.2013.00578

Klatte, M., Lachmann, T., & Meis, M. (2010). Effects of noise and reverberation on speech perception and listening comprehension of children and adults in a classroom-like setting. *Noise & Health*, *12*, 270–282. https://doi.org/10.4103/1463-1741.70506

*Klatte, M., Meis, M., Sukowski, H., & Schick, A. (2007). Effects of irrelevant speech and traffic noise on speech perception and cognitive performance in elementary school children. *Noise and Health*, *9*, 64–74.

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 643–680. https://doi.org/10.3102/0034654315626799

Kraft, M. A., Blazar, D., & Hogan, D. (2018). The effect of teacher coaching on instruction and achievement: A meta-analysis of the causal evidence. *Review of Educational Research*, *88*(4), 547–588. https://doi.org/10.3102/0034654318759268

Leibold, L. J., & Buss, E. (2013). Children's identification of consonants in a speech-shaped noise or a two-talker masker. *Journal of Speech, Language, and Hearing Research*, *56*(4), 1144–1155. https://doi.org/10.1044/1092-4388(2012/12-0011)

Leibold, L. J., Hillock-Dunn, A., Duncan, N., Roush, P. A., & Buss, E. (2013). Influence of hearing loss on children's identification of spondee words in a speech-shaped noise or a two-talker masker. *Ear and Hearing*, *34*, 575–584. https://doi.org/10.1097/AUD.0b013e3182857742

Lyberg-Åhlander, V., Brännström, K. J., & Sahlén, B. S. (2015). On the interaction of speakers' voice quality, ambient noise and task complexity with children's listening comprehension and cognition. *Frontiers in Psychology*, *6*, Article 871. https://doi.org/10.3389/fpsyg.2015.00871

*Lyberg-Åhlander, V., Haake, M., Brännström, J., Schötz, S., Sahlén, B. (2015). Does the speaker's voice quality influence children's performance on a language comprehension test? *International Journal of Speech-Language Pathology*, *17*, 63–73. https://doi.org/10.3109/17549507.2014.898098

*Lyberg-Åhlander, V., Holm, L., Kastberg, T., Haake, M., Brännström, K. J., & Sahlén, B. (2015). Are children with stronger cognitive capacity more or less disturbed by classroom noise and dysphonic teachers? *International Journal of Speech-Language Pathology*, *17*, 577–588. https://doi.org/10.3109/17549507.2015.1024172

Martins, R. H. G., Pereira, E. R. B. N., Hidalgo, C. B., & Tavares, E. L. M. (2014). Voice disorders in teachers. A review. *Journal of Voice*, *28*, 716–724. https://doi.org/10.1016/j.jvoice.2014.02.008

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology*, *59*, 203–243. https://doi.org/10.1016/j.cogpsych.2009.04.001

Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, *27*, 953–978. https://doi.org/10.1080/01690965.2012.705006

*McCreery, R. W., & Stelmachowicz, P. G. (2013). The effects of limited bandwidth and noise on verbal processing time and word recall in normal-hearing children. *Ear and Hearing*, *34*, 585–591. https://doi.org/10.1097/AUD.0b013e31828576e2

*McGarrigle, R., Dawes, P., Stewart, A. J., Kuchinsky, S. E., & Munro, K. J. (2017). Measuring listening-related effort and fatigue in school-aged children using pupillometry. *Journal of Experimental Child Psychology*, *161*, 95–112. https://doi.org/10.1016/j.jecp.2017.04.006

Mealings, K. (2016). Classroom acoustic conditions: Understanding what is suitable through a review of national and international standards, recommendations, and live classroom measurements. In I. Hillock and D. Mee (Eds.), *Proceedings of Acoustics 2016 - The Second Australasian Acoustical Societies Conference* (pp. 145–172). Brisbane, Australia: The Australian Acoustical Society.

Medwetsky, L. (2011). Spoken Language Processing Model: Bridging auditory and language processing to guide assessment and intervention. *Language, Speech, and Hearing Services in Schools*, *42*, 286–296. https://doi.org/10.1044/0161-1461(2011/10-0036)

Moher, D., Liberati, A, Tetzlaff, J., Altman, D. G., & The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA statement. *PLOS Medicine*, *6*(7), Article e1000097.

*Morsomme, D., Minell, L., & Verduyckt, I. (2011). Impact of teachers' voice quality on children's language processing skills. *VOCOLOGIE: Stem En Stemstoornissen*, 9–15.

*Morton, V., & Watson, D. R. (2001). The impact of impaired vocal quality on children's ability to process spoken language. *Logopedics Phoniatrics Vocology*, *26*, 17–25. https://doi.org/10.1080/14015430118232

*Nakeva von Mentzer, C., Sundström, M., Enqvist, K., & Hällgren, M. (2018). Assessing speech perception in Swedish school-aged children: Preliminary data on the Listen-Say test. *Logopedics Phoniatrics Vocology*, *43*, 106–119. https://doi.org/10.1080/14015439.2017.1380076

National Heart, Lung and Blood Institute (NHLBI). Study quality assessment tools: Quality

Assessment Tool for Observational Cohort and Cross-Sectional Studies. (2019). https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools/

*Nelson, P., Kohnert, K., Sabur, S., & Shaw, D. (2005). Classroom noise and children learning through a second language: Double jeopardy? *Language, Speech, and Hearing Services in Schools*, *36*, 219–229. https://doi.org/10.1044/0161-1461(2005/022)

Neuman, A. C., Wroblewski, M., Hajicek, J., & Rubinstein, A. (2010). Combined effects of noise and reverberation on speech recognition performance of normal-hearing children and adults. *Ear and Hearing*, *31*(3), 336–344. https://doi.org/10.1097/AUD.0b013e3181d3d514

*Nirme, J., Haake, M., Lyberg Åhlander, V., Brännström, J., & Sahlén, B. (2019). A virtual speaker in noisy classroom conditions: Supporting or disrupting children's listening comprehension? *Logopedics Phoniatrics Vocology 44*, 79–86. https://doi.org/10.1080/14015439.2018.1455894

O'Malley, M., Uhl Chamot, A., & Küpper, L. (1989). Listening comprehension strategies in second language acquisition. *Applied Linguistics*, *10*(4), 418–437.

*Osman, H., & Sullivan, J. R. (2014). Children's auditory working memory performance in degraded listening conditions. *Journal of Speech, Language, and Hearing Research*, *57*, 1503–1511. https://doi.org/10.1044/2014_JSLHR-H-13-0286

*Peng, J., & Jiang, P. (2016). Chinese word identification and sentence intelligibility in primary school classrooms. *Archives of Acoustics*, *41*(2), 213–219. https://doi.org/10.1515/aoa-2016-0021

*Peng, J., Zhang, H., & Yan, N. (2016). Effect of different types of noises on Chinese speech intelligibility of children in elementary school classrooms. *Acta Acustica United with*

Acustica, *102*(5), 938–944. https://doi.org/10.3813/AAA.919008

*Peng, J., & Wu, S. (2018). Chinese speech intelligibility of children in noisy and reverberant environments. *Indoor and Built Environment*, *27*(10), 1357–1363. https://doi.org/10.1177/1420326X17716236

Pichora-Fuller, M. K., Kramer, S. E., Eckert, M. A., Edwards, B., Hornsby, B. W. Y., Humes, L. E., Larry, E., Lemke, U., Lunner, T., Matthen, M., Mackersie, C., Naylor, G., Phillips, N., Richter, M., Rudner, M., Sommers, M., Tremblay, K., & Wingfield, A. (2016). Hearing impairment and cognitive energy: The Framework for Understanding Effortful Listening (FUEL). *Ear and Hearing*, 37, 5S–27S. https://doi.org/10.1097/AUD.0000000000000312

*Picou, E. M., Bean, B. N., Marcrum, S. C., Hornsby, B. W., & Ricketts, T. A. (2019). Moderate reverberation does not increase subjective fatigue, subjective listening effort, or behavioral listening effort in school-aged children. *Frontiers in Psychology*, *10*, Article 1749. https://doi.org/10.3389/fpsyg.2019.01749

*Prodi, N., Visentin, C., Borella, E., Mammarella, I. C., & Di Domenico, A. (2019). Noise, age and gender effects on speech intelligibility and sentence comprehension for 11- to 13-year-old children in real classrooms. *Frontiers in Psychology*, *10*, Article 2166. https://doi.org/10.3389/fpsyg.2019.02166

*Prodi, N., Visentin, C., Peretti, A., Griguolo, J., & Bartolucci, G. B. (2019). Investigating listening effort in classrooms for 5- to 7-year-old children. *Language, Speech, and Hearing Services in Schools*, *50*, 196–210. doi:10.1044/2018_LSHSS-18-0039

RStudio Team (2019). *RStudio: Integrated Development for R* [Computer software]. www.rstudio.com/

Rodrigues, Medeiros, & Teixeira (2017). Impact of the teacher's voice in the classroom: A literature review. *Distúrb. Comun*, *29*(1), 2–9. https://doi.org/10.23925/2176-2724.2017v29i1p2-9

*Rogerson, J., & Dodd, B. (2005). Is there an effect of dysphonic teachers' voices on children's processing of spoken language? *Journal of Voice*, *19*, 47–60. https://doi.org/10.1016/j.jvoice.2004.02.007

Rönnberg, J., Lunner, T., Zekveld, A., Sörqvist, P., Danielsson, H., Lyxell, B., Dahlström, Ö., Signoret, C., Stenfelt, S., Pichora-Fuller, M. K., & Rudner, M. (2013). The Ease of Language Understanding (ELU) model: Theoretical, empirical, and clinical advances. *Frontiers in Systems Neuroscience*, *7*, Article 31. https://doi.org/10.3389/fnsys.2013.00031

Rueschemeyer, S.-A. & Gaskell, M., G. (Eds.). (2018). *The Oxford Handbook of Psycholinguistics: Second edition*. New York: Oxford University Press.

*Sahlén, B., Haake, M., von Lochow, H., Holm, L., Kastberg, T., Brännström, K. J., & Lyberg-Åhlander, V. (2017). Is children's listening effort in background noise influenced by the speaker's voice quality? *Logopedics Phoniatrics Vocology*, *43*, 47–55. https://doi.org/10.1080/14015439.2017.1324914

Schiller, I. S., Morsomme, D., Kob, M., & Remacle, A. (2021). Listening to a dysphonic speaker in noise may impede children's spoken language processing in a realistic classroom setting. *Language, Speech, and Hearing Services in Schools*, *52*(1), 396–408. https://doi.org/10.1044/2020_LSHSS-20-00078

Schiller, I. S., Morsomme, D., Kob, M., & Remacle, A. (2020). Noise and a speaker's impaired voice quality disrupt spoken language processing in school-aged children: Evidence from performance and response time measures. *Journal of Speech, Language, and Hearing Research*, *63*(7), 2115–2131. https://doi.org/10.1044/2020_JSLHR-19-

00348

Schiller, I., Morsomme, D., Kob, M., & Remacle, A. (2019). Children's perception of degraded speech at normal vs. fast speech rate. In M. Ochmann, M. Vorländer, & J. Fels (Eds.), *Proceedings of the 23rd International Congress on Acoustics, integrating 4th EAA Euroregio 2019* (pp. 5961–5967). Berlin, Germany: Deutsche Gesellschaft für Akustik eV (DEGA).

Schiller, I. S., Remacle, A., & Morsomme, D. (2019). Imitating dysphonic voice: A suitable technique to create speech stimuli for spoken language processing tasks? *Logopedics Phoniatrics Vocology*. https://doi.org/10.1080/14015439.2019.1659410

Shield, B. M., & Dockrell, J. E. (2003). The effects of noise on children at school: A review. *Building Acoustics*, *10*, 97–116. https://doi.org/10.1260/135101003768965960

*Sullivan, J. R., Osman, H., & Schafer, E. C. (2015). The effect of noise on the relationship between auditory working memory and comprehension in school-age children. *Journal of Speech, Language, and Hearing Research*, *58*, 1043–1051. https://doi.org/10.1044/2015_JSLHR-H-14-0204

Van Houtte, E., Claeys, S., Wuyts, F., & Van Lierde, K. (2011). The impact of voice disorders among teachers: Vocal complaints, treatment-seeking behavior, knowledge of vocal care, and voice-related absenteeism. *Journal of Voice*, *25*(5), 570–575. https://doi.org/10.1016/j.jvoice.2010.04.008

Valente, D. L., Plevinsky, H. M., Franco, J. M., Heinrichs-Graham, E. C., & Lewis, D. E. (2012). Experimental investigation of the effects of the acoustical conditions in a simulated classroom on speech recognition and learning in children. *The Journal of the Acoustical Society of America*, *131*(1), 232–246. https://doi.org/10.1121/1.3662059

Vandenbroucke, L., Spilt, J., Verschueren, K., Piccinin, C., & Baeyens, D. (2018). The classroom as a developmental context for cognitive development: A meta-analysis on the importance of teacher–student interactions for children's executive functions. *Review of Educational Research*, *88*(1), 125–164. https://doi.org/10.3102/0034654317743200

*von Lochow, H., Lyberg-Åhlander, V., Sahlén, B., Kastberg, T., & Brännström, K. J. (2018). The effect of voice quality and competing speakers in a passage comprehension task: Performance in relation to cognitive functioning in children with normal hearing. *Logopedics Phoniatrics Vocology*, *43*, 32–41. https://doi.org/10.1080/14015439.2017.1298835

Wróblewski, M., Lewis, D. E., Valente, D. L., & Stelmachowicz, P. G. (2012). Effects of reverberation on speech recognition in stationary and modulated noise by school-aged children and young adults. *Ear and Hearing*, *33*(6), 731–744. https://doi.org/10.1097/AUD.0b013e31825aecad

*Yacullo, W. S., & Hawkins, D. B. (1987). Speech recognition in noise and reverberation by school-age children. *Audiology*, *26*, 235–246.

*Zhang, D., Tenpierik, M., & Bluyssen, P. M. (2019). Interaction effect of background sound type and sound pressure level on children of primary schools in the Netherlands. *Applied Acoustics*, *154*, 161–169. https://doi.org/10.1016/j.apacoust.2019.05.007

Tables and Figures

Table 1: Characteristics of the 31 Studies Included in this Systematic Review

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Bradley & Sato (2008)	*N* = 840 Age = 6-11	Word-picture matching	Speech perception	Group testing in a classroom (stimuli presented via loudspeakers)	Classroom noise	Various SNRs ranging from –10 dB to 30 dB	Answer accuracy	Sign. performance decrease as SNR decreased. Noise was particularly detrimental for younger children.
Crandell & Smaldino (1996)	*N* = 40 Age = 8-10 (*M* = 8;9)	Sentence repetition	Speech perception	Individual testing in a sound treated room (stimuli presented via headphones)	Multi-talker babble	No added noise +6 dB SNR +3 dB SNR 0 dB SNR –3 dB SNR –6 dB SNR	Answer accuracy	Descriptively, performance decreased as SNR decreased. Speech-in-noise perception sign. poorer for stimuli presented in listeners' L2.
Howard et al. (2010)	*N* = 30 (17 ♀) Age = 9-12 (*M* = 10;8)	Word repetition[a]	Speech perception	Individual testing in a quiet room at school (stimuli presented via headphones)	Multi-talker babble	No added noise +4 dB SNR 0 dB SNR –4 dB SNR	Answer accuracy	Sign. performance decreased as SNR decreased.

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Hurtig et al. (2016)	*N* = 72 Age = 10	Word recall	Auditory working memory	Group testing in a classroom (stimuli presented via loudspeakers)	Broadband noise	+12 dB SNR +3 dB SNR	Answer accuracy	Sign. performance decrease. Sign. interaction between SNR and reverberation time (0.3 s vs. 1.2 s): at +3 dB SNR, shorter reverberation time impeded performance, at +12 dB SNR, longer reverberation time impeded performance.
Jamieson et al. (2004)	*N* = 30 Age = 6-8	Word-picture matching	Speech perception	Individual testing in a classroom (stimuli presented via headphones)	Classroom noise	No added noise 0 dB SNR –6 dB SNR –12 dB SNR	Answer accuracy	Descriptively, performance decreased as SNR decreased. Noise was particularly detrimental for younger children.

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Klatte et al. (2007)	*N* = 46 (6-8, *M* = 7;1) *N* = 22 (7-8, *M* = 8;5)	Word-picture matching Phonological discrimination Execution of oral instructions Odd one out	Speech perception Speech perception Listening compre-hension Auditory working memory	Group testing in a sound-proof booth (stimuli presented via headphones)	Back-ground speech Train noise	Low-noise control +5 dB SNR (speech) 0 dB SNR (train)	Answer accuracy	Word-picture-matching: sign. performance decrease for background speech and particularly for train noise. Discrimination task, execution of oral instructions, and odd-one-out task: sign. performance decrease for background speech but not for train noise.
McCreery & Stelmachowicz (2013)	*N* = 17 Age = 6-12 (*M* = 9;3)	Non-word repetition	Speech perception	Individual testing in a sound-treated room (stimuli presented via headphones)	Speech-shaped noise	+9 dB SNR +3 dB SNR	Answer accuracy and RT	Sign.performance decrease and RT increase.
McGarrigle et al. (2017)	*N* = 41 (23 ♀) Age = 8-11 (*M* = 9;3)	Word recall**	Auditory working memory	Individual testing in sound-proof booth (stimuli presented via headphones)	Multi-talker babble	+15 dB SNR –2 dB SNR	Answer accuracy and RT	No sign. performance decrease or RT increase.
Nakeva von Mentzer et al. (2018)	*N* = 27 (11 ♀) Age = 7-9 (*M* = 8;0)	Phonological discrimination	Speech perception	Individual testing, no room specifications (stimuli presented via headphones)	Multi-talker babble	No added noise +5 dB SNR	Answer accuracy and RT	Discrimination task: sign. performance decrease and RT decrease

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Nelson et al. (2005)	*N* = 22 Age = 7	Word-picture matching	Speech perception	Group testing in a classroom (stimuli presented via sound field amplification system)	Multi-talker babble	No added noise +10 dB SNR	Answer accuracy	Sign. performance decrease. Speech-in-noise perception sign. poorer for stimuli presented in listeners' L2.
Nirme et al. (2019)	*N* = 55 (34 ♀) Age = 8-9 (*M* = 8;6)	Passage comprehension	Listening comprehension	Individual testing in a quiet room at school (stimuli presented via headphones)	Multi-talker babble	No added noise +10 dB SNR (speech presented by dysphonic speaker)	Answer accuracy	Sign. performance decrease for content questions, but not inference questions.
Osman & Sullivan (2014)	*N* = 20 (11 ♀) Age = 8-10 (*M* = 9;2)	Forward digit recall Backward digit recall Word recall Veracity judgement	Auditory working memory Auditory working memory Auditory working memory Listening comprehension	Individual testing in a sound-proof booth (stimuli presented via loudspeakers)	Multi-talker babble	No added noise 0 dB SNR –5 dB SNR	Answer accuracy	Sign. performance decrease across tasks as SNR decreased. No interaction between task and SNR.

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Peng & Jiang (2016)	*N* = 30 (14 ♀) Age = 8-11	Word-word matching Sentence repetition	Speech perception Speech perception	Group testing in a classroom (stimuli presented via loudspeakers)	Speech-shaped noise	No added noise +20 dB SNR +15 dB SNR +10 dB SNR +5 dB SNR 0 dB SNR	Answer accuracy	Descriptively, performance in word identification and sentence repetition decreased as SNR decreased, particularly for younger children.
Peng et al. (2016)	*N* = 60 (23 ♀) Age = 7-12	Word-word matching	Speech perception	Group testing in a classroom (stimuli presented via loudspeakers)	Impact noise Traffic noise Fan noise Multi-talker babble Speech-shaped noise	+10 dB SNR 0 dB SNR	Answer accuracy	Sign. performance decrease at lower SNR, particularly for younger children. Speech-shaped noise and babble noise were more detrimental than impact, fan, or traffic noise.
Peng & Wu (2018)	*N* = 480 Age = 7-12	Word-word matching	Speech perception	Group testing in a classroom (stimuli presented via loudspeakers)	Speech-shaped noise	Various SNRs ranging from –6 dB to +17 dB	Answer accuracy	Sign. performance decrease as SNR decreased. Descriptively, noise was particularly detrimental for younger children.

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Picou et al. (2019)	*N* = 20 (15 ♀) Age = 10-17 (*M* = 13;3)	Word repetition [a]	Speech perception	Individual testing in a sound proof booth and room with moderate reverberation times (stimuli presented via loudspeakers)	Multi-talker babble	No added noise –5 dB SNR[c] –9 dB SNR[c]	Answer accuracy and RT	Sign. performance decrease as SNR decreased. Sign. RT increase from no added noise to –5 dB SNR. No interaction between SNR and reverberation time on performance or RT.
Prodi, Visentin, Borella, et al. (2019)	*N* = 159 (75 ♀) Age = 11-13 (*M* = 12)	Sentence repetition (nonverbal response mode) Sentence-picture matching	Speech perception Listening compre-hension	Group testing in a classroom (stimuli presented via loudspeakers)	Classroom noise Traffic noise	No added noise 0 dB SNR	Answer accuracy and RT	Sentence repetition: sign. performance decrease. Classroom noise more impeding than traffic noise. Regarding RT: 11- and 12- but not 13 year-olds slower in classroom noise than traffic or no noise. Sentence-picture matching: No sign. effect on performance. Sign. RT increase in classroom noise but not traffic noise.

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Prodi, Visentin, Peretti, et al. (2019)	*N* = 94 (45 ♀) Age = 6-7	Word-picture matching (with test-repetition shortly after the first test)	Speech perception	Group testing in a classroom (stimuli presented via loudspeakers)	Speech-shaped noise[d]	No added noise 0 dB SNR	Answer accuracy and RT	Sign. performance decrease and RT increase. Sign. interaction regarding RT: Repeating the test increased RTs in noise but not in the no-added-noise condition.
Sullivan et al. (2015)	*N* = 20 (11 ♀) Age = 8-10	Backward digit recall Word recall Veracity judgement Passage comprehension	Auditory working memory Auditory working memory Listening comprehension Listening comprehension	Individual testing, no room specifications (stimuli presented via loudspeakers)	Classroom noise	No added noise –5 dB SNR	Answer accuracy	Sign. performance decreases in all tasks. In passage comprehension, noise had the most negative impact on the reasoning subtest, followed by vocabulary.
Yacullo & Hawkins (1987)	*N* = 32 (19 ♀) Age = 8-10 (*M* = 9;3)	Sentence repetition	Speech perception	Individual testing in a sound-proof booth and classroom (stimuli presented via loudspeakers)	Multi-talker babble	+6 dB SNR +2 dB SNR	Answer accuracy	Sign. performance decrease at lower SNR.

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Zhang et al. (2019)	*N* = 290 (145 ♀) Age = 9-13 (*M* = 10;6)	Odd one out (identifying the odd word in a list of 3)	Speech perception	Group testing in an acoustically treated vs. untreated chamber (stimuli presented via loudspeakers)	Traffic noise Multi-talker babble Music	No added noise 45 dB(A) noise SPL 60 dB(A) noise SPL 33 dB(A) noise SPL 48 dB(A) noise SPL	Answer accuracy	Sign. interaction between noise SPL and noise source: at 45 dB(A), multi-talker babble was more impeding than traffic noise; at 33 dB(A), multi-talker babble was more impeding than traffic noise and music; at 48 dB(A), multi-talker babble was more impeding than music.
Bränn-ström, Kastberg, et al. (2018)	*N* = 57 (25 ♀) Age = 8-9 (*M* = 8;7)	Acceptability judgement Word recall	Listening compre-hension Auditory working memory	Individual testing in a quiet room at school (stimuli presented via headphones)	Provoked dysphonia	Healthy voice Mildly / moderately impaired voice	Answer accuracy	Acceptability judgement: sign. performance de-crease. Word recall: no effect
Chui & Ma (2018)	*N* = 134 (70 ♀) Age = 8-10	Passage comprehen-sion	Listening compre-hension	Group testing in a classroom (stimuli presented via loudspeakers)	Imitated dysphonia	Healthy voice Mildly impaired voice Severely impaired voice	Answer accuracy	Sign. performance decrease for impaired voice, but no effect of severity.

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Lyberg-Åhlander, Haake, et al. (2015)	*N* = 86 (43 ♀) Age = 7-9	Sentence-picture matching	Listening compre-hension	Individual testing in a quiet room at school (stimuli presented via loudspeakers)	Imitated dysphonia	Healthy voice Mildly / moderately impaired voice	Answer accuracy	No overall effect, but sign. performance decrease for more difficult sentences in dysphonic voice.
Lyberg-Åhlander, Holm, et al. (2015)	*N* = 93 (52 ♀) Age = 8-9	Sentence-picture matching	Listening compre-hension	Individual testing in a quiet room at school (stimuli presented via headphones)	Provoked dysphonia	Healthy voice Moderately impaired voice (speech merged with multi-talker babble at +10 dB SNR)	Answer accuracy	No overall effect, but in the impaired voice condition, children with better working memory capacity performed better for easier items, while in the normal voice condition, they performed better for more difficult items.
Morsomme et al. (2011)	*N* = 68 (34 ♀) Age = 8 (*M* = 8;5)	Phonological discrimination Passage comprehen-sion	Speech perception Listening compre-hension	Group testing in a classroom (stimuli presented via loudspeakers)	Imitated dysphonia	Healthy voice Moderately / severely impaired voice	Answer accuracy	Sign. performance decrease for impaired voice, particularly for discrimination task.

943

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Morton & Watson (2001)	*N* = 24 Age = 11	Passage comprehension Word recall[b]	Listening comprehension Auditory working memory	Individual testing in a quiet room at school (stimuli presented via headphones)	Real dysphonia	Healthy voice Moderately / severely impaired voice	Answer accuracy	Word recall: sign. performance decrease Passage comprehension: no effect.
Rogerson & Dodd (2005)	*N* = 107 (52 ♀) Age = 9-10	Passage comprehension	Listening comprehension	Group testing in a classroom (stimuli presented via loudspeakers)	Imitated dysphonia	Healthy voice Mildly impaired voice Severely impaired voice (speech merged with classroom noise at unknown SNR)	Answer accuracy	Sign. performance decrease for impaired voice, but no effect of severity.
Sahlén et al. (2017)	*N* = 93 (52 ♀) Age = 8-9	Sentence-picture matching	Listening comprehension	Individual testing, no room specifications (stimuli presented via headphones)	Provoked dysphonia	Healthy voice Moderately impaired voice (speech merged with babble noise at +10 dB SNR)	RT	Sign. RT increase, but only for girls.

Study	Study population *N* = sample size; Age (in years)	Task type	SPADE dimension	Interventional setting	Type of exposure	Listening conditions	Out-come	Finding
Brännström, von Lochow, et al. (2018)	*N* = 18 (14 ♀) Age = 9-12 (*M* = 10;1)	Passage comprehension	Listening comprehension	Individual testing in a quiet room at school (stimuli presented via headphones)	Provoked dysphonia Multi-talker babble	Healthy voice, no added noise Healthy voice, +10 dB SNR Mildly / moderately impaired voice, no added noise Mildly / moderately impaired voice, +10 dB SNR	Answer accuracy	Sign. performance decrease for noise. No effect of impaired voice. No interaction.
Von Lochow et al. (2018)	*N* = 49 (27 ♀) Age = 7-12 (*M* = 8;1)	Passage comprehension	Listening comprehension	No information	Provoked dysphonia Background speech Multi-talker babble	Healthy voice, no added noise Healthy voice, background speech Healthy voice, +5 dB SNR multi-talker babble Mildly / moderately impaired voice, no added noise	Answer accuracy	Inconclusive results for noise. No effect of impaired voice. No interaction.

	Mildly / moderately impaired voice, +5 dB SNR background speech	
	Mildly / moderately impaired voice, +5 dB SNR multi-talker babble	

Note. Color codes were used to highlight noise studies (no color), voice studies (light grey), and studies on the combined effect of noise and impaired voice (dark grey). Speech-shaped noise = steady-state signal that matches the spectral characteristics of speech; Multi-talker babble = babble noise created from several talkers (usually not intelligible); SNR = signal-to-noise ratio; RT = response time.

[a] Word repetition measured in dual-task design. The secondary task was a visual task.

[b] After children listened to passages in noise, children were presented with images depicting target and distractor words. Word recall was tested by asking the children whether or not these words were presented in the passage.

[c] SNRs represent estimated values due to the method by which they were obtained.

946 ***Table 2: Effects of Noise on Children's Answer Accuracy in Listening Tasks as a Function of SNR***

			Main effects			Heterogeneity test	
Predictor	Studies *N*	Children *N*	Cohen's *d* [95% CI]	*z*-value	*p*-value	I^2	*p*-value
Noise (SNR bins)							
+6 to +10 dB	5	130	–0.67 [–0.92; –0.42]	–5.24	< 0.01**	0%	0.93
+1 to +5 dB	5	156	–1.20 [–2.00; –0.40]	–2.94	< 0.01**	90%	< 0.01**
0 dB	7	371	–1.74 [–2.60; –0.88]	–3.96	< 0.01**	95%	< 0.01**
–1 to –5 dB	5	131	–2.24 [–3.82; –0.65]	–2.77	< 0.01**	96%	< 0.01**
–6 to –12 dB	3	70	–2.65 [–4.10; –1.21]	–3.60	< 0.01**	89%	< 0.01**

947

Table 3: Effects of Impaired Voice on Children's Answer Accuracy in Listening Comprehension Tasks and Auditory Working Memory Tasks

			Main effects			Heterogeneity test	
Predictor	Studies *N*	Children *N*	Cohen's *d* [95% CI]	*z*-value	*p*-value	I^2	*p*-value
Impaired voice							
Listening comprehension	9	545	–0.35 [–0.59; –0.11]	–2.82	< 0.01**	73%	< 0.01**
Auditory working memory	2	81	–0.13 [–0.72; 0.46]	–0.42	0.67	67%	0.08

Note. The dimension of speech perception is not featured as it was assessed in only one study (Morsomme et al., 2011).

Figure 1. PRISMA Flow Chart Depicting the Study Selection Process

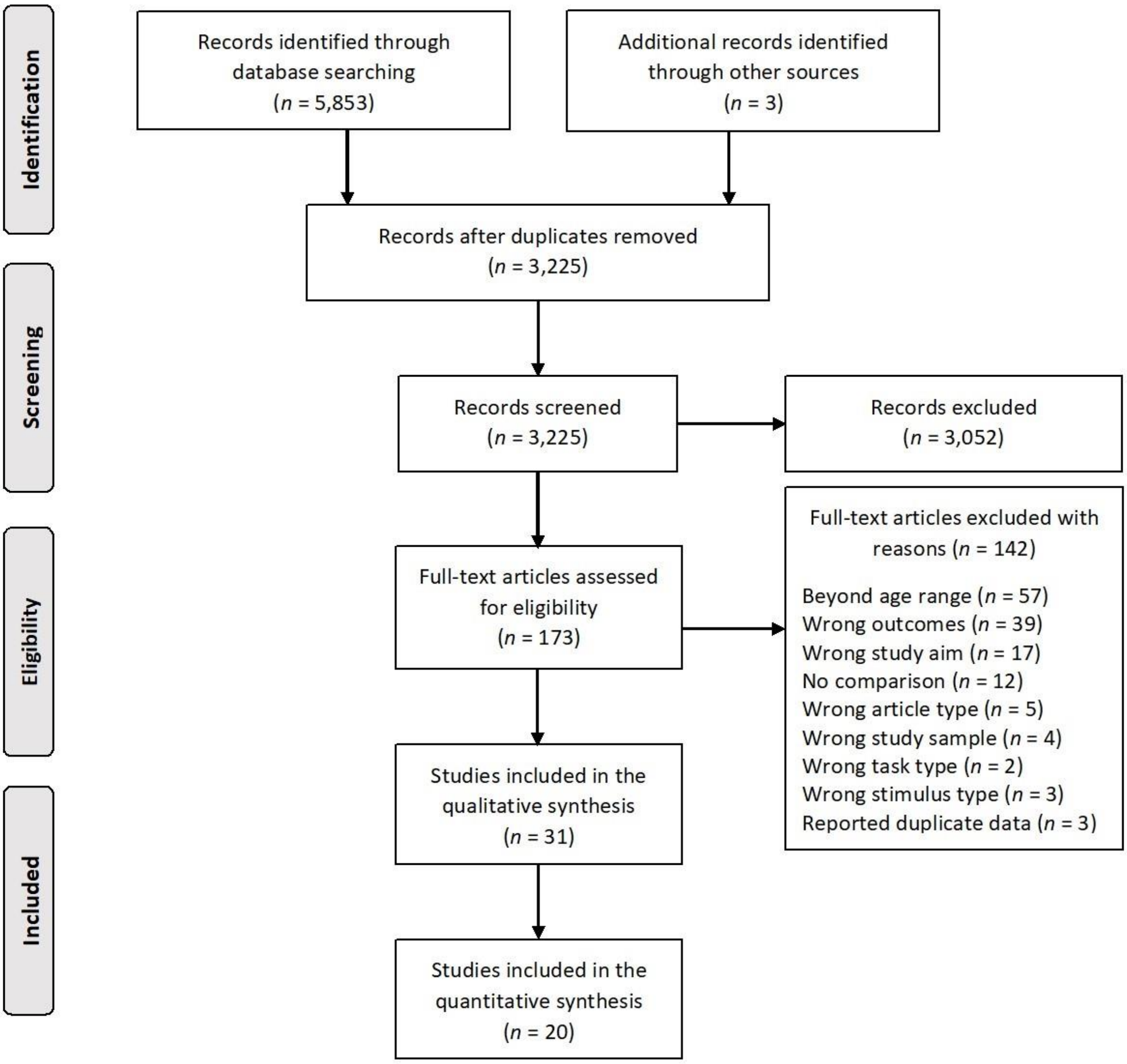
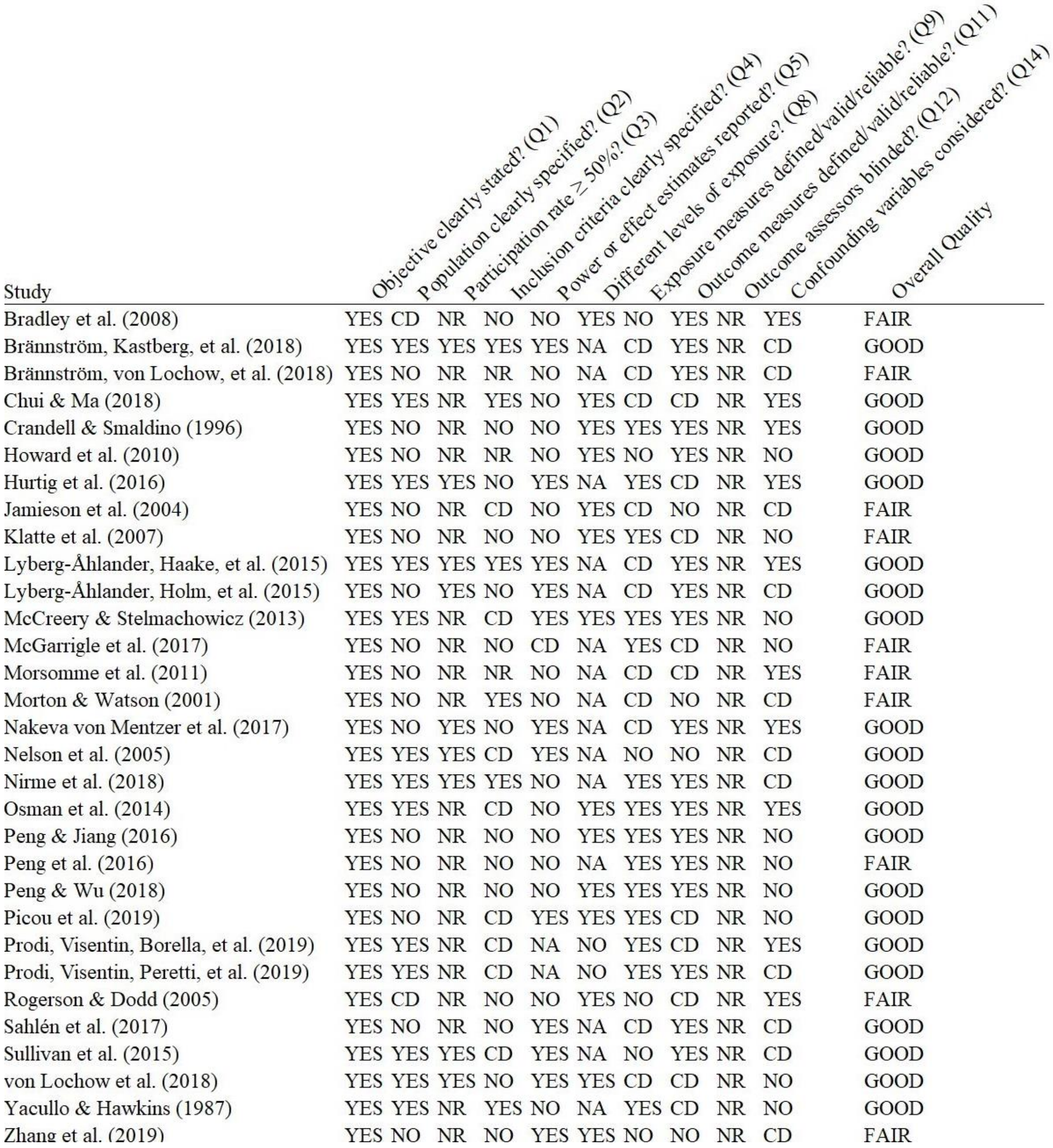


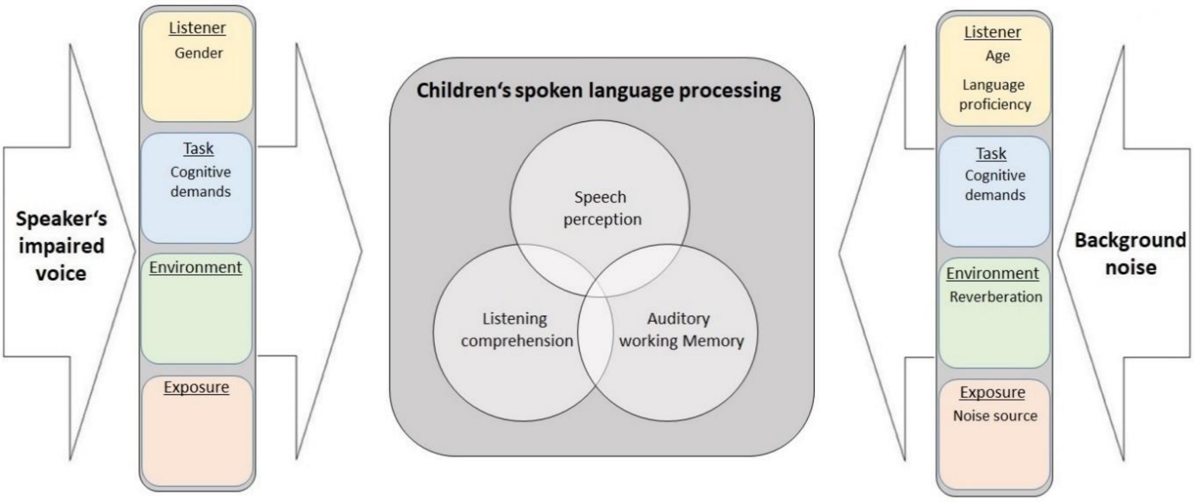
Figure 2. Study Quality Assessment Using the Quality Assessment Tool for Observational Cohort and Cross-Sectional Studies (NHLBI, 2019)



Study	Objective clearly stated? (Q1)	Population clearly specified? (Q2)	Participation rate ≥ 50%? (Q3)	Inclusion criteria clearly specified? (Q4)	Power or effect estimates reported? (Q5)	Different levels of exposure? (Q8)	Exposure measures defined/valid/reliable? (Q9)	Outcome measures defined/valid/reliable? (Q11)	Outcome assessors blinded? (Q12)	Confounding variables considered? (Q14)	Overall Quality
Bradley et al. (2008)	YES	CD	NR	NO	NO	YES	NO	YES	NR	YES	FAIR
Brännström, Kastberg, et al. (2018)	YES	YES	YES	YES	YES	NA	CD	YES	NR	CD	GOOD
Brännström, von Lochow, et al. (2018)	YES	NO	NR	NR	NO	NA	CD	YES	NR	CD	FAIR
Chui & Ma (2018)	YES	YES	NR	YES	NO	YES	CD	CD	NR	YES	GOOD
Crandell & Smaldino (1996)	YES	NO	NR	NO	NO	YES	YES	YES	NR	YES	GOOD
Howard et al. (2010)	YES	NO	NR	NR	NO	YES	NO	YES	NR	NO	GOOD
Hurtig et al. (2016)	YES	YES	YES	NO	YES	NA	YES	CD	NR	YES	GOOD
Jamieson et al. (2004)	YES	NO	NR	CD	NO	YES	CD	NO	NR	CD	FAIR
Klatte et al. (2007)	YES	NO	NR	NO	NO	YES	YES	CD	NR	NO	FAIR
Lyberg-Åhlander, Haake, et al. (2015)	YES	YES	YES	YES	YES	NA	CD	YES	NR	YES	GOOD
Lyberg-Åhlander, Holm, et al. (2015)	YES	NO	YES	NO	YES	NA	CD	YES	NR	CD	GOOD
McCreery & Stelmachowicz (2013)	YES	YES	NR	CD	YES	YES	YES	YES	NR	NO	GOOD
McGarrigle et al. (2017)	YES	NO	NR	NO	CD	NA	YES	CD	NR	NO	FAIR
Morsomme et al. (2011)	YES	NO	NR	NR	NO	NA	CD	CD	NR	YES	FAIR
Morton & Watson (2001)	YES	NO	NR	YES	NO	NA	CD	NO	NR	CD	FAIR
Nakeva von Mentzer et al. (2017)	YES	NO	YES	NO	YES	NA	CD	YES	NR	YES	GOOD
Nelson et al. (2005)	YES	YES	YES	CD	YES	NA	NO	NO	NR	CD	GOOD
Nirme et al. (2018)	YES	YES	YES	YES	NO	NA	YES	YES	NR	CD	GOOD
Osman et al. (2014)	YES	YES	NR	CD	NO	YES	YES	YES	NR	YES	GOOD
Peng & Jiang (2016)	YES	NO	NR	NO	NO	YES	YES	YES	NR	NO	GOOD
Peng et al. (2016)	YES	NO	NR	NO	NO	NA	YES	YES	NR	NO	FAIR
Peng & Wu (2018)	YES	NO	NR	NO	NO	YES	YES	YES	NR	NO	GOOD
Picou et al. (2019)	YES	NO	NR	CD	YES	YES	YES	CD	NR	NO	GOOD
Prodi, Visentin, Borella, et al. (2019)	YES	YES	NR	CD	NA	NO	YES	CD	NR	YES	GOOD
Prodi, Visentin, Peretti, et al. (2019)	YES	YES	NR	CD	NA	NO	YES	YES	NR	CD	GOOD
Rogerson & Dodd (2005)	YES	CD	NR	NO	NO	YES	NO	CD	NR	YES	FAIR
Sahlén et al. (2017)	YES	NO	NR	NO	YES	NA	CD	YES	NR	CD	GOOD
Sullivan et al. (2015)	YES	YES	YES	CD	YES	NA	NO	YES	NR	CD	GOOD
von Lochow et al. (2018)	YES	YES	YES	NO	YES	YES	CD	CD	NR	NO	GOOD
Yacullo & Hawkins (1987)	YES	YES	NR	YES	NO	NA	YES	CD	NR	NO	GOOD
Zhang et al. (2019)	YES	NO	NR	NO	YES	YES	NO	NO	NR	CD	FAIR

Note. Questions 6, 7, 10, and 13 were not applicable for our studies of interest and removed.

Figure 3. The Speech Processing under Acoustic DEgradations (SPADE) Framework



Note. The two types of acoustic degradations, *speaker's impaired voice* (arrow on the far left) and *background noise* (arrow on the far right), may disrupt *children's spoken language processing* (grey box in the center) in terms of speech perception, listening comprehension, and auditory working memory (overlapping circles). Moderators of these effects are presented in the two vertical squares. They refer to the listener (yellow), the task (blue), the environment (green), and the exposure (orange). The list of moderators is solely based on significant interactions revealed across the 31 studies included in this systematic review. There are likely more factors that can act as moderators, such as type of dysphonia (e.g., roughness vs. breathiness). Moreover, certain factors that moderated the effect of noise might also moderate the effect of a speaker's impaired voice (e.g., children's age or reverberation). Future works are required to complete the SPADE framework.