LSHSS

Research Article

Listening to a Dysphonic Speaker in Noise May Impede Children's Spoken Language Processing in a Realistic Classroom Setting

Isabel S. Schiller,[a] Dominique Morsomme,[a] Malte Kob,[b] and Angélique Remacle[a,c]

Purpose: The aim of this study was to investigate children's processing of dysphonic speech in a realistic classroom setting, under the influence of added classroom noise.
Method: Typically developing 6-year-old primary school children performed two listening tasks in their regular classrooms: a phoneme discrimination task to assess speech perception and a sentence–picture matching task to assess listening comprehension. Speech stimuli were played back in either a typical or an impaired voice quality. Children performed the tasks in the presence of induced classroom noise at signal-to-noise ratios between +2 and +9 dB.
Results: Children's performance in the phoneme discrimination task decreased significantly when the speaker's voice was impaired. The effect of voice quality on sentence–picture matching depended on task demands: Easy sentences were processed more accurately in the impaired-voice condition than in the typical-voice condition. Signal-to-noise ratio effects are discussed in light of methodological constraints.
Conclusions: Listening to a dysphonic teacher in a noisy classroom may impede children's perception of speech, particularly when phonological discrimination is needed to disambiguate the speech input. Future research regarding the interaction of voice quality and task demands is necessary.

A classroom is an environment in which children spend a considerable amount of time listening to their teacher (Mealings, 2016). In doing so, they acquire knowledge and expand on that knowledge as they progress through school. However, various factors may interfere with classroom listening, two of them being a teacher's impaired voice quality (i.e., dysphonia) and background noise. In this field study, we explored children's perception and comprehension of dysphonic speech in classroom noise at classroom-typical signal-to-noise ratios (SNRs).

[a]Faculty of Psychology, Speech Therapy and Educational Sciences, University of Liège, Belgium
[b]Erich Thienhaus Institute, Detmold University of Music, Germany
[c]National Fund for Scientific Research (F.R.S.–FNRS), Brussels, Belgium

Correspondence to Isabel S. Schiller: isabel.schiller@uliege.be

Editor-in-Chief: Holly L. Storkel
Editor: Dawna E. Lewis

Received June 12, 2020
Revision received August 31, 2020
Accepted September 21, 2020
https://doi.org/10.1044/2020_LSHSS-20-00078

Voice Impairments Among Teachers

Voice impairments are a prevalent phenomenon among teachers. Every second teacher develops voice problems during their career (Roy et al., 2004). Although the etiology is not yet fully understood, underlying causes are thought to include vocal misuse or overuse in response to heavy vocal demands. Teachers with voice impairments show symptoms such as vocal fatigue, throat ache, roughness, and dysphonia (Martins et al., 2014). Although their voice is their primary tool for work, only about 50% of concerned teachers seek medical treatment for voice problems (Van Houtte et al., 2011). It can therefore be assumed that many children are taught by dysphonic teachers. This is problematic, because the dysphonic voice is characterized by acoustic disruptions (e.g., increased frequency perturbations [jitter], amplitude perturbations [shimmer], or a low harmonics-to-noise ratio [HNR]; Teixeira & Fernandes, 2015), which may be perceived similarly to noise. Consequently, dysphonic teachers may be less intelligible and children may find their voice unpleasant (Morsomme et al., 2011).

Disclosure: *The authors have declared that no competing interests existed at the time of publication.*

Classroom Noise and Room Acoustics

Background noise and poor room acoustics pose an additional challenge for classroom listening. In addition to high noise levels and low SNRs, classroom acoustics are commonly evaluated based on reverberation time and Speech Transmission Index (STI). Reverberation time is the time a sound takes to decay by 60 dB in a closed room. The STI gives an indication of the quality of speech signal transmission (Steeneken & Houtgast, 1980) and ranges between 0 and 1—the higher the value, the better the speech intelligibility.

The American National Standards Institute (ANSI, 2010) recommends maximum noise levels of 35 dBA and maximum reverberation times of 0.6 s for unoccupied classrooms. Mealings (2016) suggested that, for primary school children, who are more vulnerable to acoustic interference than older peers, "good" classroom conditions apply when the following criteria are met: unoccupied noise levels < 30 dBA, SNR > +15 dB, reverberation time < 0.4 s, and STI > 0.75. Unfortunately, real-world conditions often depart from these recommendations. Unoccupied noise levels have been reported to vary between 41 and 51 dBA (Crandell & Smaldino, 2000). SNRs typically range between –7 and +11 dB (Bradley & Sato, 2008; Crandell & Smaldino, 2000). Reverberation times range from 0.4 to 1.2 s (Crandell & Smaldino, 2000). STI values range between 0.33 and 0.88, but are often below 0.75 (Mealings, 2016). A listening scenario characterized by such noise interference and poor room acoustics is not ideal for classroom learning.

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

The effects of a speaker's impaired voice and noise on children's spoken language processing were recently investigated in a systematic review (Schiller, Remacle, & Morsomme, 2020). The authors proposed a classification of impaired-voice and noise effects along three processing dimensions: speech perception (referring to the initial stages of spoken language processing), listening comprehension (referring to higher linguistic processing stages), and auditory working memory (referring to information storage, manipulation, and recall). Below, we summarize the main findings.

Along the dimension of speech perception, impaired voice and noise may disrupt children's processing at an auditory–perceptual level and reduce intelligibility (e.g., Bradley & Sato, 2008; Howard et al., 2010; Morsomme et al., 2011; Peng & Jiang, 2016; Peng et al., 2016). Along the dimension of listening comprehension, impaired voice and noise may impede spoken language processing in terms of semantic and syntactic integration (e.g., Brännström, Kastberg, et al., 2018; Prodi, Visentin, Borella, et al., 2019). Finally, along the dimension of auditory working memory, impaired voice and noise may interfere with the storage, manipulation, and retrieval of speech-encoded information (Morton & Watson, 2001; Sullivan et al., 2015).

Regarding the dimension of listening comprehension, two laboratory studies suggested that the effect of impaired voice might be mediated by task demands (or cognitive demands related to solving a listening task; Lyberg-Åhlander, Haake, et al., 2015; Lyberg-Åhlander, Holm, et al., 2015). Task demands in spoken language processing tasks depend on a combination of different factors, most of which are linguistic in nature. They include lexical and semantic aspects, word or sentence length, syntactic structure, and even visual aspects related to response images. Lyberg-Åhlander, Haake, et al. (2015) found that children's performance in a sentence–picture matching task decreased significantly when listening to a dysphonic speaker, but only in the case of grammatically difficult sentences. In the study by Lyberg-Åhlander, Holm, et al. (2015), children with strong working memory skills had less trouble comprehending a dysphonic speaker than children with weaker skills, but only in the case of grammatically easy sentences. The nature of the interaction between task demands and a speaker's voice quality remains unclear and has never been investigated in a field experiment. Thus, this study takes a closer look at the influence of task demands on children's comprehension of dysphonic speech.

Methodological Considerations: Laboratory Versus Field Experiments

The traditional approach to explore the effects of acoustically degraded speech on children's spoken language processing is by means of laboratory experiments (e.g., Brännström, von Lochow, et al., 2018; Lyberg-Åhlander, Haake, et al., 2015; Lyberg-Åhlander, Holm, et al., 2015; Sullivan et al., 2015). In these experiments, children typically perform listening tasks in quiet rooms at school or in laboratories; they are tested individually or in small groups and listen to speech stimuli via earphones. Laboratory experiments offer a high degree of internal validity. Controlling for confounding factors, such as reverberation time or unwanted sounds, is relatively easy. A drawback is the limited generalizability of the results, due to the artificial setup.

Field experiments offer greater ecological validity because they are carried out under more authentic conditions (e.g., Bradley & Sato, 2008; Peng & Jiang, 2016; Peng et al., 2016; Prodi, Visentin, Borella, et al., 2019). By field experiments, we mean listening experiments conducted in a naturalistic setting (preferably in children's habitual classrooms), with children tested in groups (preferably together with their classmates), and with speech stimuli presented in a diffuse field (via loudspeakers). The drawbacks of field experiments are that the internal validity is lower and the effects of interest may be superimposed by confounding factors. Moreover, in most cases, it may not be possible to collect response times.

To bridge the gap between internal and ecological validity, this field experiment builds on a design that we previously applied in a laboratory experiment (Schiller, Morsomme, et al., 2020), where we investigated the effects of noise and a speaker's impaired voice in a highly controlled setting. Typically developing 6-year-old children performed a phonological discrimination task (to assess speech perception) and a sentence–picture matching task (to assess listening comprehension). They were tested in quiet rooms at school. Speech stimuli were presented via earphones in four conditions: typical voice in quiet, impaired voice in quiet, typical voice in noise, and impaired voice in noise (speech-shaped noise at 0 dB SNR). The results revealed that impaired voice and noise lowered performance and slowed down children's responses in the discrimination task. As for sentence–picture matching, there was an interaction between noise and voice quality: Noise disrupted children's performance when the speaker's voice was impaired, but not when it was not impaired. These findings provided a first indication that a teacher's impaired voice and noise might be detrimental for classroom listening. Whether these results hold true under more realistic circumstances was the starting basis of this work.

The aim of this study was to investigate the effects of a speaker's impaired voice and noise (at classroom-typical SNRs) on children's spoken language processing in a real classroom setting. A secondary aim was to document the acoustic conditions in the classrooms and take into account their potential effects on children's results in listening tasks. We used the same listening tasks as in Schiller, Morsomme, et al. (2020), measuring children's performance (but not response times) under different listening conditions. The participants were a new set of typically developing 6-year-old children. Children were examined in their habitual classrooms, together with their peers, and during regular school hours. Three hypotheses were tested:

- H1: Listening to an impaired voice will reduce children's performance in the speech perception task.
- H2: Listening to an impaired voice will reduce children's performance in the listening comprehension task, and this effect may interact with task demands.
- H3: Children's performance in classroom noise will drop with decreasing SNR, particularly when listening to an impaired voice.

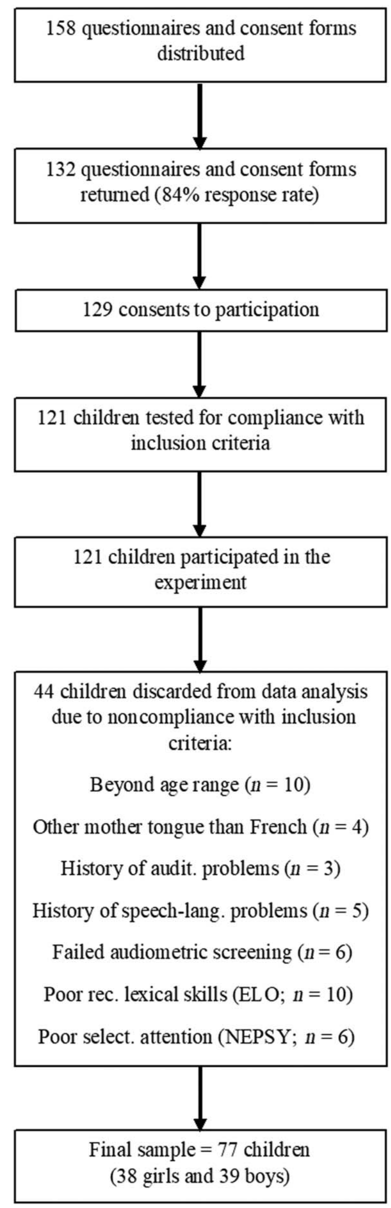
Method

Participants

The participant selection procedure is depicted in Figure 1. Participants were first graders recruited from eight primary schools in the French-speaking region of Belgium. From a total of 121 children who participated in the experiment, we discarded the data of 44 children due to noncompliance with the inclusion criteria presented below. Statistical analyses were run on a final sample of 77 children (38 girls, 39 boys) with a mean age of 6;6 (years;months; *SD* = 3 months).

Children were required to meet the following criteria: (a) 5–6 years old, (b) French as the mother tongue, (c) typical auditory development, (d) typical speech-language development, (e) hearing threshold ≤ 25 dB HL at octave frequencies between 500 and 4000 kHz, (f) normal or above-normal receptive lexical skills (i.e., score ≥ 25th percentile in the Lexique en Réception [Lexicon in Reception] subtest of the Évaluation du Langage Oral [ELO; Oral Language Assessment]; Khomsi, 2001), and (g) normal or above-normal auditory selective attention (i.e., score ≥ 25th percentile in the Auditory Attention subtest of the Bilan Neuropsychologique de L'Enfant–Seconde Edition [Developmental Neuropsychological Assessment]; Korkman et al., 2007).

Figure 1. Flowchart presenting the recruitment of participants and selection of the final sample. Eight children who consented to participate were absent on the days when we assessed children's compliance with the inclusion criteria. audit. = auditory; ELO = Évaluation du Langage Oral; NEPSY = Bilan Neuropsychologique de L'Enfant; rec. = receptive; select. = selective; speech-lang. = speech-language.



Compliance with criteria (a)–(d) was evaluated based on parental report, using a self-administered questionnaire. Compliance with criteria (e)–(g) was based on the results of pretests. In these pretests, children individually underwent a pure-tone audiometric screening (MAICO MA 50 audiometer with DD45 earphones) and performed the receptive lexical task (the LexR subtask is a subtask of the ELO; Khomsi, 2001) and the auditory selective attention task (the auditory selective attention task is a subtask of the NEPSY; Korkman et al., 2007).

Oral informed consent was obtained from the participants, and written informed consent was obtained from their parents. This study was approved by the ethics committee of the Faculty of Psychology, Speech Therapy and Educational Sciences (University of Liège, Belgium; File No. 1617-54).

Tasks

Children performed two listening tasks. Speech perception was assessed with the Épreuve Lilloise de Discrimination Phonologique (ELDP; Lille Phonological Discrimination Test; Macchi et al., 2012), and listening comprehension was assessed with the Comprehension2 subtest from the ELO (Khomsi, 2001). For the purpose of this study, we created pen–paper versions of both tasks and used speech stimuli recorded for this research project (available from the NODYS [NOrmophonic and DYsphonic Speech samples] database; Schiller et al., 2019).

Speech Perception

The ELDP task (Macchi et al., 2012) is a phonological discrimination task. Children listen to pairs of pseudowords (i.e., nonexistent words that comply with the phonotactic rules of French) and have to decide whether the two words sounded the same or different. We used List 1 of the ELDP task, developed for 5- to 6-year-old children. This list includes 36 speech items (pseudoword pairs). Half of them consist of two identical pseudowords and the other half of two slightly different pseudowords, such as /paʀum/–/pamuʀ/ (structural opposition) or /muko/–/luko/ (phonemic opposition). In the original task, children respond by pointing to response images of either two identical-looking planets (words sounded the same) or two different-looking planets (words sounded different). In our version of the task, participants circled the planet images in their answer booklets. Correct responses were coded as 1, and incorrect responses were coded as 0.

Listening Comprehension

The Comprehension2 subtest from the ELO (Khomsi, 2001) is a sentence–picture matching task, designed for 5- to 10-year-old children. The children's task is to listen to a sentence and match it to the corresponding picture. Each target picture is presented along with three distractors, which are morphosyntactically or semantically similar. The task contains a total of 32 sentence items of varying complexity but can be stopped after Item 21. We chose this option due to our participants' young age and because they had to perform the speech perception task in the same session. To account for the varying complexity, we classified the items into three levels of task demand, based on the ELO norm data. Items closest to the median performance level of 65% were classified as medium items (n = 7). Items with higher and lower performance levels were respectively classified as easy (n = 7) and difficult (n = 7) items. In the original task, children respond by pointing. In our version of the task, they circled the corresponding pictures in their answer booklets. Correct responses were encoded as 1, and incorrect responses were coded as 0.

Listening Conditions

Children performed the speech perception task and the listening comprehension task in their normal classrooms. We manipulated the speaker's voice quality and the background noise condition. As for voice quality, items were randomly presented in a typical voice or an impaired voice. Concerning noise, we played back classroom noise throughout the entire experiment. SNRs varied between +2 and +9 dB (range: 8 dB), as is typical for teaching situations (Bradley & Sato, 2008; Crandell & Smaldino, 2000). This SNR range is narrow considering that the just noticeable difference in SNR has been claimed to be around 3 dB (McShefferty et al., 2015). However, past studies have shown that even small differences of 3–4 dB SNR may affect children's performance in speech perception (Howard et al., 2010) and listening comprehension (Valente et al., 2012) tasks. In the following sections, we provide more information on the speech and noise signals and on the experimental setup.

Speech Signals

Speech items for both listening tasks were recorded in two voice quality conditions. The speaker was a female speech therapist, who first read out all items in her normal voice and then while mimicking dysphonia. We followed the recording guidelines outlined in Barsties and De Bodt (2015). Schiller, Remacle, and Morsomme (2020) described the characteristics of the two voice qualities. The acoustic analysis included the Acoustic Voice Quality Index (AVQI; Maryn et al., 2010), based on connected speech and sustained vowels, as well as jitter, shimmer, and HNR measures on sustained vowels. The perceptual analysis included a GRBAS (Grade, Roughness, Breathiness, Asthenia, and Strain) rating (Hirano, 1981) on connected speech and sustained vowels, as well as consistency and authenticity ratings of the voice qualities. Acoustic and perceptual analyses confirmed that (a) the speaker's normal voice was free of a voice disorder (AVQI = 2.53; jitter [local] = 0.31%; shimmer [local] = 1.39%; HNR = 25 dB; $G_0R_0B_0A_0S_0$), (b) the speaker's imitated impaired voice was moderately to severely dysphonic and characterized by a high degree of roughness and asthenia (AVQI = 6.89; jitter [local] = 2.77%; shimmer [local] = 9.18%; HNR = 11 dB; $G_3R_3B_2A_3S_1$), and (c) the speaker's imitated impaired voice showed a consistent quality throughout the recordings and was perceived as reasonably authentic. Note that the same speech stimuli were used in our laboratory experiment (Schiller, Morsomme, et al., 2020), which allows for a direct comparison.

Classroom Noise

The noise signal was classroom noise, recorded during a mathematics class in a fourth-grade primary school classroom. Our rationale was to use a realistic noise source that children would actually encounter during regular classroom listening. Therefore, we decided not to use speech-shaped noise as we did in Schiller, Morsomme, et al. (2020). For the recording, we used a binaural headset (BHS II, HEAD acoustics). Signal processing was conducted in Praat (Version 6.0.29; Boersma & Weenink, 2017). We cut out all intelligible speech segments from the recording, as well as the most prominent noise bursts visually detected in the spectrum. The resulting signal contained typical ambient noise found in a classroom (i.e., children clearing their throat, opening pencil cases, moving chairs, rustling paper, and occasionally whispering). The root-mean-square level was normalized to 50 dB SPL, with a dynamic range of 30 dB (32–62 dB). Finally, we looped and time-shifted the signal to create two 45-min noise chains (Noise A and Noise B), identical in spectral and temporal characteristics but with different starting points. In the listening experiment, we simultaneously played back these noise chains from diagonally aligned loudspeakers to create a realistic listening experience.

Experimental Setup, Calibration, and Acoustic Measurements

The listening experiment was conducted in eight primary school classrooms. Table 1 lists information regarding the experimental context for each classroom. Figure 2 shows a typical classroom setup. All classrooms were prepared in the same way. In each corner of the room, we positioned one loudspeaker (Neumann KH 120 A) to broadcast the classroom noise. In front of the class, where the teacher would normally stand, we positioned a fifth loudspeaker (Neumann KH 120 A) to broadcast the speech signals. The loudspeakers were connected to and controlled from a Dell laptop via an audio interface (RME Babyface Pro). Chairs were arranged in four rows of seats (from R1 to R4; see Figure 2). Between the two middle rows (R2 and R3), we defined a central measurement position (CMP). In that position, we calibrated speech and noise presentation levels. SNRs were measured at the CMP and each seat row.

Presentation levels were calibrated in unoccupied classrooms. We leveled speech and noise signals in the CMP to approximately 70 and 65 dBA (fast, A-weighted sound levels), respectively, as measured with a calibrated Class 2 sound-level meter (NL-21, Rion), which was positioned on a microphone stand. Calibration was done based on quasistationary speech-shaped noise (same root-mean-square level as speech and noise signals). First, we broadcast the calibration signal from the speech loudspeaker and adjusted the volume until the sound-level meter in the CMP steadily showed approximately 70 dBA. The same procedure was applied for the noise loudspeakers, to yield a sound level of approximately 65 dBA. After calibration, we used the sound-level meter to measure SNRs per seat row by moving the microphone stand to the seating positions in the center of each row. The resulting +5 dB SNR in the CMP, as well as the subsequently measured SNRs in each seat row (see Table 1), should be regarded as best estimated fits, not exact or constant ratios. Uncertainties arise from the calibrated accuracy of the sound-level meter (±2 dB), natural intensity fluctuations of speech and noise signals across time, and additional noise caused by the presence of children in the room.

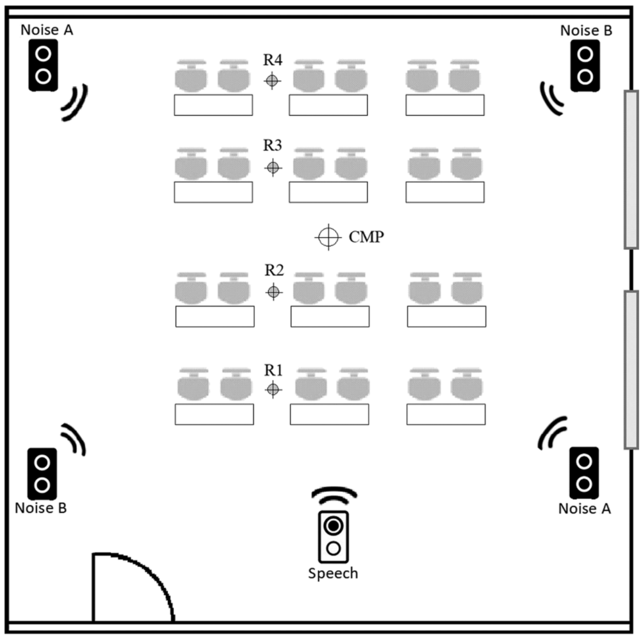
In each classroom, we also assessed the inherent acoustic conditions. This evaluation was based on reverberation time, STI, unoccupied noise levels, and occupied noise levels. Reverberation time (T_{30}) and STI were derived from room impulse responses in octave bands from 60 Hz to 4 kHz. For this purpose, we used WinMF measurement software (Four Audio, 2018). The unoccupied classroom was excited with a sine sweep signal radiated from the four noise loudspeakers, which were directed toward the CMP. The receiver was an

Table 1. Information regarding the experimental setting and the artificially induced signal-to-noise ratios (SNRs) in the eight classrooms.

Classroom ID	Room volume (m^3)	Children present during experiment (*n*)	Distance between speech source and seat rows (R1–R4) (m)				SNR per seat row[a] (R1–R4) (dB)			
			R1	R2	R3	R4	R1	R2	R3	R4
1	214	16	1.2	2.3	3.0	4.2	9	8	5	3
2	129	20	1.2	2.3	3.2	4.3	8	7	5	4
3	213	15	1.1	2.3	3.6	4.9	9	6	4	3
4	124	20	0.7	1.7	2.9	4.2	7	6	2	0
5	121	19	2.1	3.3	4.5	5.7	6	5	4	3
6	168	19	1.3	2.4	3.9	4.9	7	6	5	4
7	59	14	0.9	1.9	2.5		6	5	4	
8	118	18	1.2	2.2	3.4	4.6	9	5	3	2
M	143	18	1.2	2.3	3.4	4.7	8	6	4	3

[a]These SNRs are based on the calibrated presentation levels of the speech and noise signals.

Figure 2. Diagram of the typical experimental setup in each classroom. Noise A and Noise B refer to the same chain of classroom noise, which was time-shifted (i.e., different starting points). CMP = central measurement position; R = measurement points in each seat row (R1–R4).



omnidirectional MM 1 microphone (Beyerdynamic) located in the CMP. Due to time restrictions, we did not vary receiver positions. Impulse responses were digitized and later used for calculating reverberation time and STI. Noise levels were measured using the NL-21 sound-level meter, which was located in the CMP. Unoccupied noise levels (L_{aeq}, 5 min, in dBA) were measured in empty classrooms. Occupied noise levels (L_{aeq}, 1 min, in dBA) were measured in the presence of all participants, who were instructed to sit silently at their desks.

Procedure

We conducted a pilot study with a group of seven children aged 6 years old. They were tested in a meeting room at the University of Liège. This pilot study helped us to determine appropriate presentation levels for speech and noise signals, improve the clarity of the task instructions and answer booklets, and estimate how much time would be required for the experimental setup, calibration, and acoustic measurements (about 45 min); to run the experiment (about 35 min); and to remove the material (about 15 min).

The main experiment was carried out between December 2018 and March 2019 in eight Belgian primary schools. During the 2 days that preceded the experiment in each school, children were assessed for compliance with the inclusion criteria. On the day of the experiment, while the school was still closed, three experimenters set up the material in the participants' habitual classroom. One experimenter calibrated the speech and noise presentation levels and took the acoustic measurements (except occupied noise levels). The experiment was then conducted in the first hour of the morning. As children entered the room, they were assigned random seating positions. Tables were equipped with screens (to prevent copying), answer booklets, and pens (see Appendix).

After ensuring that all children were quietly seated, we measured occupied noise levels. Then, the experiment was explained, and the instructions for the first task (speech perception task) were read out: "You will listen to pairs of fantasy words. After each pair, your task is to decide whether the two words sounded the same or different. If they sounded the same, circle the picture of the planets that look exactly the same. If they sounded different, circle the image with the different-looking planets. Sometimes, it will be difficult to understand the speaker, because her voice sounds a bit rough. There will also be noise in the background. Just try to focus on the task and answer as best you can." The task began with four practice items, followed by the 36 test items. Response time was restricted to 8 s per item, based on the maximum response times in Schiller, Morsomme, et al. (2020). Speech items were randomly presented in a typical versus an impaired voice quality. SNRs varied depending on where participants were seated (i.e., children in the back rows performed the task under poorer SNRs than children in the front rows; see Table 1).

The speech perception task was directly followed by the listening comprehension task. The experimenter explained: "In this task, you will listen to sentences. Each sentence is accompanied by four pictures that you can see in your answer booklet. Your task is to circle the picture that matches the sentence you have heard. Again, understanding the speaker might be difficult, so listen carefully, focus on your task, and answer as best you can." The task began with four practice items, followed by the 21 test items, which were played randomly in a typical or an impaired voice. SNRs remained the same as in the speech perception task. Response time was limited to 12 s per item, based on maximum response times in Schiller, Morsomme, et al. (2020). After the experiment, we collected the response booklets and removed the material.

Statistical Analysis

To statistically analyze the listening task data, we fitted generalized linear mixed-effects models (GLMMs) using R software (Version 3.6.1; R Core Team, 2019). This was done with the glmer function of the lme4 package (Version 1.1-15; Bates et al., 2015). The assumed significance level was α = .05. We modeled our data with GLMMs, because GLMMs do not require a prior transformation of binary data (Lo & Andrews, 2015). Furthermore, our study design included repeated measures, which may be accounted for in GLMMs by introducing random effects.

We built different models for the speech perception task and the listening comprehension task. GLMMs were specified with a binomial distribution and logit link function as in Schiller, Morsomme, et al. (2020). A forward procedure was used for model selection (Prodi, Visentin, Peretti, et al., 2019). Using R's anova function, models were compared based on the Akaike information criterion (Akaike, 1974). Significant effects were further investigated in pairwise comparisons using the lsmeans package (Lenth, 2016), with Tukey's honestly significant difference test accounting for multiple comparisons.

The final speech perception model predicted children's performance as a function of the fixed factors *voice quality* (typical vs. impaired) and *SNR* (continuous variable ranging from +2 to +9 dB). Our rationale for treating *SNR* as a continuous variable was related to the narrow range of SNR values (i.e., from +2 to +9 dBA) resulting from the presentation level calibration that was conducted within each of the eight classrooms. The GLMM included random intercepts for effects of participant (n = 77), item (n = 36), discrimination target (same vs. different), trial (n = 36), and school (n = 8). The final listening comprehension model predicted performance as a function of the *Voice Quality* × *Task Demands* interaction (easy vs. medium vs. difficult) and *SNR,* considering the random effects of participant and item.

Results

In the following sections, we will first report on the acoustic conditions in the eight classrooms in which the experiments were conducted and whether they affected children's listening performance. Then, we present the results regarding children's performance in the speech perception task and the listening comprehension task.

Classroom Acoustics

To reduce the impact of varying classroom acoustics on the results, we normalized speech and noise presentation levels in each classroom by means of calibration. As this does not cancel out all room-related differences, we further considered the following acoustic parameters in our statistical analyses: reverberation time, STI, unoccupied noise levels, and occupied noise levels. Table 2 shows the respective measurement results. Unoccupied noise levels varied between 37 and 45 dBA. Occupied noise levels varied between 43 and 50 dBA. Note that the highest occupied noise levels were measured in Classroom 8, although this classroom exhibited the lowest unoccupied noise levels. Reverberation times varied between 0.4 and 0.8 s. Finally, STI values ranged from 0.69 to 0.89.

The potential influence of these acoustic parameters on children's performance was assessed by treating them as random effects in the GLMMs of both tasks. Other random effects assessed in the GLMMs were children's age and gender. None of these random effects resulted in a statistically significant improvement of the model fits, so they were dropped from the final GLMMs. A reason for a factor's incapacity to improve the model fits could be a poor predictive value with regard to the dependent variable or the fact that including this factor would have resulted in overfitting.

The Effect of Voice Quality

Figure 3 illustrates children's performance in the two listening tasks as a function of voice quality. Results from the GLMMs revealed that, in the speech perception task, children's performance was statistically significantly impeded by a speaker's impaired voice, $\chi^2(1) = 10.3$, $p = .001$. Figure 3 shows the performance drop from a proportion-correct level of 0.79 (SE = 0.13, 95% CI [0.45, 0.94]) in the typical-voice condition to 0.73 (SE = 0.15, 95% CI [0.37, 0.92]) in the impaired-voice condition, indicating that children discriminated phonemes in pseudowords with about 8% lower accuracy. There was no main effect of voice quality on children's performance in the listening comprehension task, $\chi^2(1) = 0.2$, $p = .62$.

Interaction Between Voice Quality and Task Demands

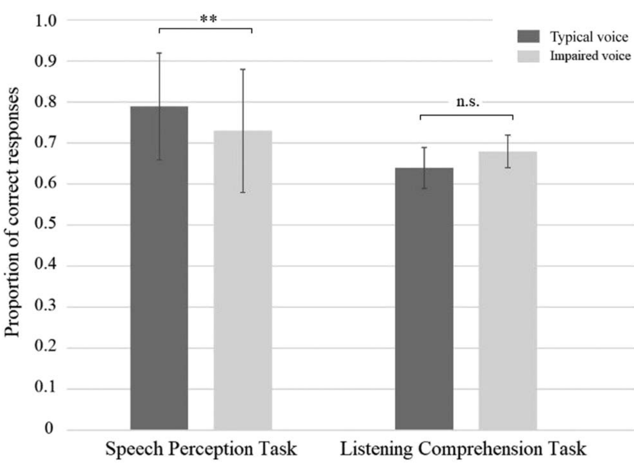
While voice quality alone had no statistically significant effect on children's listening comprehension, GLMM results revealed a statistically significant interaction between voice quality and task demands, $\chi^2(2) = 11.07$, $p = .004$. This interaction is depicted in Figure 4. Contrary to our

Table 2. Descriptive results from the acoustic measurements taken in the eight classrooms.

Classroom ID	Unoccupied noise level (L_{aeq} in dB)	Occupied noise level (L_{aeq} in dB)	RT (T_{30} in s)	STI
1	45	49	0.52	0.76
2	38	42	0.79	0.67
3	38	49	0.45	0.78
4	40	41	0.36	0.89
5	39	47	0.73	0.69
6	43	49	0.72	0.70
7	37	43	0.60	0.73
8	37	50	0.52	0.76
M	40	46	0.59	0.75

Note. RT = reverberation time; STI = Speech Transmission Index.

Figure 3. Mean task performance as a function of voice quality in the speech perception task (left) and the listening comprehension task (right). Error bars represent *SE*. n.s. = not significant.

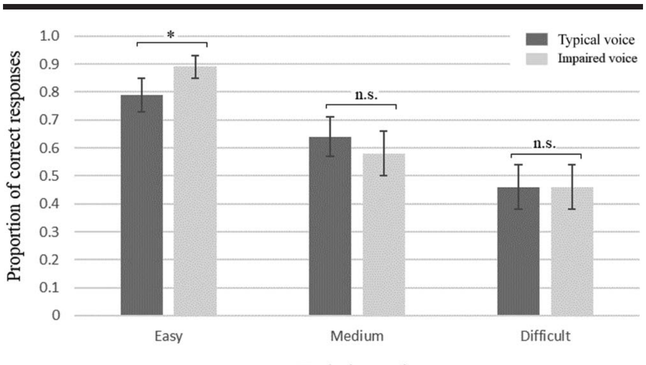


expectations, pairwise comparisons by means of Tukey's honestly significant difference test showed a statistically significantly weaker performance for the typical voice than for the impaired voice, when children listened to easy sentences (z = 3.0, p = .03). Under this condition, the GLMM estimated proportion-correct levels of 0.78 (*SE* = 0.06, 95% CI [0.63, 0.88]) for the typical voice and 0.88 (*SE* = 0.04, 95% CI [0.78, 0.94]) for the impaired voice. When sentences were of medium difficulty, performance was slightly but not statistically significantly better in the typical-voice condition (z = –1.54, p = .64). In the case of difficult sentences, performance in the typical- and impaired-voice conditions did not differ (z = –0.18, p = 1.0).

The Effect of Classroom Noise

The effect of classroom noise was assessed in terms of the numeric variable SNR. GLMM results revealed a statistically significant effect of SNR on children's performance in the speech perception task (β = .07, z = 2.1, p = .03), suggesting that, with a decreasing SNR, children discriminated phonemes less accurately. However, when plotting the proportion of correct responses for each estimated SNR unit (ranging from +2 to +9 dB), this effect appears negligible (see the left-hand graph in Figure 5). Visual inspection of the data shows considerable variance, as indicated by the large standard errors. Finally, no statistically significant interaction between SNR and voice quality, $\chi^2(1) = 0.14$, p = .71, was found. Regarding listening comprehension, GLMM results revealed neither a statistically significant effect of SNR (β = .02, z = 0.55, p = .58) nor a statistically significant interaction between SNR and voice quality, $\chi^2(1) = 0.32$, p = .57. The right-hand graph in Figure 5 shows the proportion of correct responses in the listening task for each of the estimated SNR units (ranging from +2 to +9 dB).

Figure 4. Mean task performance as a function of voice quality and task demands in the listening comprehension task. Error bars represent *SE*. n.s. = not significant.



Discussion

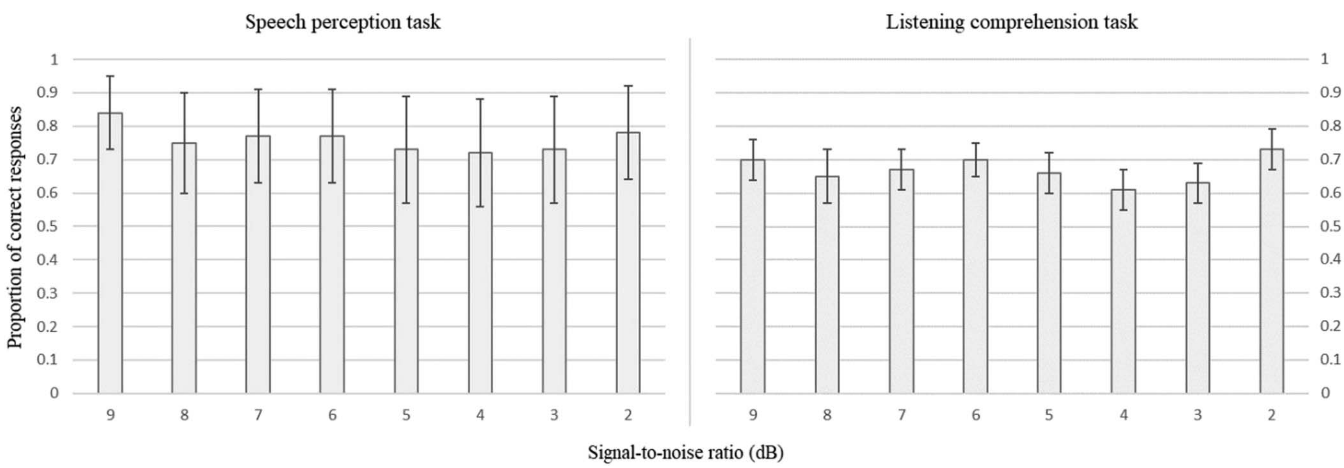
In classrooms, pupils may frequently be required to listen to dysphonic teachers and deal with high noise levels (Crandell & Smaldino, 2000; Martins et al., 2014; Mealings, 2016; Roy et al., 2004; Van Houtte et al., 2011). This prompted us to carry out in-depth investigations into the effects of impaired voice and noise on spoken language processing in typically developing children. In our previous works, we reviewed the literature regarding these effects and provided a classification along different processing dimensions (Schiller, Remacle, & Morsomme, 2020). In a laboratory experiment, we showed that speech-shaped noise and a speaker's impaired voice disrupt spoken language processing in 6-year-olds (Schiller, Morsomme, et al., 2020). The results from the present field experiment confirmed that these findings largely hold true under more realistic circumstances. Beyond that, they suggested that children's processing of dysphonic speech may vary with respect to task demands. These findings will be discussed in light of the previous literature.

The Effect of Voice Quality

We hypothesized that listening to a dysphonic voice would significantly impair children's performance in a speech perception task (H1). Our results confirmed this hypothesis and aligned with findings from our systematic review (Schiller, Remacle, & Morsomme, 2020), our laboratory experiment (Schiller, Morsomme, et al., 2020), and another field experiment (Morsomme et al., 2011). We interpreted the negative effect of impaired voice on speech perception as an indication that dysphonic speech was less intelligible. This is probably related to the increased proportion of noise components in the spectrum, as indicated by the low HNR of 11 dB compared to 25 dB HNR in the typical voice. Discriminating phonemes in a dysphonic speech stream may be significantly more difficult for children when they cannot deduce a word meaning from the context.

Interestingly, the performance drop from the typical-voice condition to the impaired-voice condition was about

Figure 5. Mean task performance as a function of estimated signal-to-noise ratio unit in the speech perception task (left) and the listening comprehension task (right). Error bars represent *SE*.



9% weaker than in the speech-in-noise conditions of our laboratory experiment (Schiller, Morsomme, et al., 2020). We speculate that the speech-shaped noise used in Schiller, Morsomme, et al. (2020) induced greater energetic masking effects (i.e., greater physical overlapping of physical characteristics with the speech signal; Mattys et al., 2009) on the impaired voice than the real classroom noise. The collection of response times in this study would have allowed a more fine-grained comparison, especially because we previously showed that children's speech-in-noise perception was not only less accurate but also slower when the speaker's voice was impaired (Schiller, Morsomme, et al., 2020). Future studies are needed for an in-depth investigation of the interaction between a speaker's voice quality and noise source on speech perception.

Our second hypothesis (H2) stated that listening to an impaired voice would reduce children's performance in the listening comprehension task and that this effect might interact with task demands (easy, medium, difficult). Taken together, our results showed no negative effect of impaired voice on children's listening comprehension. This is in line with earlier findings by Morton and Watson (2001) and Schiller, Morsomme, et al. (2020). However, it diverges from the prevailing assumption that listening to an impaired voice (in noise) increases children's processing load, thereby leaving less resources available for comprehending the spoken message (Brännström, Kastberg, et al., 2018; Lyberg-Åhlander, Haake, et al., 2015; Lyberg-Åhlander, Holm, et al., 2015). We assume that increased processing load might instead manifest in prolonged response times rather than in reduced task performance.

Interaction Between Voice Quality and Task Demands

We observed an interesting interaction between voice quality and task demands. Recall that task demands refer to the degree of difficulty of the 21 sentence items as derived from the ELO norm data (Khomsi, 2001). These demands are thought to result predominantly from sentence length, word familiarity, syntactic complexity, and semantic distance between target and distractor pictures. When task demands were low (i.e., when an item results in high performance levels, according to the ELO norm data), children performed statistically significantly better in the impaired-voice condition than in the typical-voice condition. No such difference was found regarding medium or high task demands. We suspect that two opposing effects may explain the observed interaction, as explained below.

On the one hand, listening to an atypical voice might have attracted children's attention back to the task in a situation when their overall concentration was fading (recall that the listening comprehension task was presented after the speech perception task). In other words, the impaired voice might have had a standout effect, as it sounded quite different to the speech children would normally encounter. In the case of easy sentences, this standout effect might have generated a performance advantage by increasing children's alertness. On the other hand, in the case of more difficult sentences, the increased processing demands might have outweighed the standout effect. This might explain why no effect of impaired voice quality was seen for moderately and very difficult sentences.

Our theory of the counteracting effects would also explain why Lyberg-Åhlander, Haake, et al. (2015) found disruptive effects of impaired voice on children's processing of difficult sentences but not of easier sentences. Note that this study included only children with normal and above-normal auditory selective attention skills. These children might have had better abilities to process dysphonic speech, which might explain why their processing of difficult sentences was not impeded by the impaired voice. Lyberg-Åhlander, Holm, et al. (2015) had previously provided indications that children with strong cognitive skills may be less affected by a speaker's impaired voice than their peers.

Future research is needed to validate statements regarding the interaction between a speaker's voice quality and task demands, as well as children's ability to respond to these demands.

The Effect of Classroom Noise

Our third hypothesis (H3) stated that children's task performance in classroom noise would decline with decreasing SNR, particularly when the speaker's voice was impaired. This was not confirmed by our results. Regarding the speech perception task, the effect of SNR was statistically significant, concurrent with previous results from laboratory experiments (Howard et al., 2010; Schiller, Morsomme, et al., 2020; Sullivan et al., 2015) and field experiments (Bradley & Sato, 2008; Peng & Jiang, 2016). However, a visual inspection of the performance data per SNR failed to show a clear downward trend in performance with decreasing SNR (see Figure 5). This likely relates to the small SNR range combined with potential confounding factors, as is further discussed below. Regarding the listening comprehension task, our statistical analysis showed no significant effect of SNR. This result was similar to our previous findings (Schiller, Morsomme, et al., 2020) but diverged from Valente et al.'s (2012) finding that children's performance in a listening comprehension task significantly decreased as the SNR dropped from +10 to +7 dBA (SNR is treated as a categorical variable). Finally, no statistically significant interaction between SNR and the speaker's voice quality on children's performance in either task was found.

For several reasons, these results should be interpreted cautiously: (a) The SNR range was narrow (i.e., 8 dBA). Although even small SNR decreases may disrupt children's spoken language processing (Howard et al., 2010; Valente et al., 2012), a broader SNR range would have certainly made the detection of noise-induced performance changes more likely. (b) SNR values were positive (i.e., varying between +2 and +8 dBA). Particularly in the case of the listening comprehension task, in which children could rely on context cues for sentence interpretation, the level of classroom noise might have been too low to impede performance. Response time measures might have revealed more subtle effects with regard to listening effort. (c) SNR values provide only an average estimate, because speech and noise signals fluctuated and SNRs were measured before children entered the classroom. Finally, (d) further uncertainties may result from the study design (e.g., varying group dynamics, individual differences) and the measurement material (e.g., ±2 dB accuracy of the sound-level meter).

In the context of listening comprehension, the lack of a main effect of SNR or of a significant SNR × Voice Quality interaction on performance could also relate to practice and/or habituation effects. Because the children performed the listening tasks in classroom-typical SNRs, it is possible that they were adept at processing speech under such conditions due to daily exposure. The fact that speech-in-noise training can generally improve children's processing of speech in noise was confirmed by Millward et al. (2011). The extent to which daily-life situations, such as listening in a noisy classroom or living in a noisy household, may result in similar training effects remains to be discovered (e.g., by increasingly integrating questionnaire data in experimental studies). Given that noise was present during the entire experiment, which lasted about 35 min, it is also possible that children became less disturbed by it over time. To date, little is known about children's habituation to noise in listening tasks. However, a study in which adults had to perform a working memory task in noise showed that noise habituation may be possible (Röer et al., 2014). More research on this interesting topic is needed.

Considerations on the Acoustic Conditions Within Classrooms

A subordinate aim of this article was to evaluate the acoustic conditions of the classrooms in which the listening experiments were performed. Classroom acoustics may influence children's listening conditions and, therefore, need to be considered in field studies. In this study, reverberation time, STI, unoccupied noise levels, and occupied noise levels did not significantly affect children's listening task performance. Importantly, however, the unoccupied noise levels we measured (i.e., 37–45 dBA) consistently surpassed the recommended maximum thresholds of between 30 dBA (Mealings, 2016) and 35 dBA (ANSI, 2010). Occupied noise levels varied between 41 and 50 dBA, with the highest measure (i.e., 50 dBA) obtained in Classroom 8—a peculiar finding, because Classroom 8 also showed the lowest unoccupied noise level (i.e., 37 dBA). This variation might be due to different agitation levels of the children in relation to the short measurement time of 1 min. Reverberation times varied between 0.4 and 0.8 s, with the mean of 0.59 s falling barely below the recommended maximum of 0.6 s (ANSI, 2010), but still surpassing Mealings' (2016) proposed threshold of 0.4 s. STI values varied between 0.69 and 0.89, with the mean of 0.75 suggesting appropriate conditions for speech transmission (Steeneken & Houtgast, 1980). Given the alarming classroom acoustic measures reported in the literature (Crandell & Smaldino, 2000; Mealings, 2016), the conditions we measured across the eight classrooms can be regarded as fair, but they could definitely be improved.

Limitations and Future Directions

We presented and discussed the results of a field experiment that arose from a previous laboratory experiment (Schiller, Morsomme, et al., 2020). Our adaptation of the study design allowed us to test the ecological validity of our previous findings in a more naturalistic setting. Nevertheless, there are some limitations that should be acknowledged and future directions that must be discussed.

One limitation was the difficulty of ruling out the effects of varying classroom characteristics on the results. Because we sought to test children under the most realistic circumstances possible, the experiment was performed in various classrooms with different shapes and acoustic conditions. To address this problem, we calibrated the sound presentation levels in order to equalize listening conditions, and we included various acoustic variables in our statistical models. Nevertheless, there might be other confounding factors we did not control for (e.g., different group dynamics or the duration of each individual experiment). Moreover, our procedure resulted in a narrow SNR range, which might have made it difficult to detect noise effects.

Another limitation is that the tasks presented to the children were different from tasks they would encounter during normal lessons. During lessons, children might be required to listen for a sustained period of time. Tasks might require them to switch back and forth between speech perception and production. We did not use such tasks, as they have their own drawbacks. Prolonged speech-in-noise listening tasks preclude the assessment of low-level speech perception. Moreover, standardized test material is rarely available. It would be interesting to build on the concept of passage comprehension tasks, by creating a task in which children listen to and answer questions about even longer texts.

The effects of impaired voice and noise should increasingly be investigated in relation to fatigue resulting from sustained listening effort. Children might tire sooner when listening to a dysphonic teacher in noise. However, the opposite effect—an adaptation to impaired voice or noise—is also possible. More research is needed to understand the effect of prolonged exposure to impaired voice. Whenever possible, the collection of response times is recommended and may allow deeper insight into children's listening effort.

Conclusions

This study was the first to assess the combined effect of a speaker's voice quality and noise on school-aged children's spoken language processing in a realistic classroom setting. When the speaker's voice was impaired, children had more problems processing speech in noise, as indicated by the results of a phoneme discrimination task. On the level of complex listening comprehension, however, no main effect of impaired voice was detected. Response time measurements might have provided more subtle information regarding this question. An interesting finding was that, when sentences induced few processing demands, exposure to an impaired voice appeared to improve performance, possibly because it increased children's arousal. Regarding the effect of classroom noise, the results precluded firm conclusions, mainly as a consequence of a narrow SNR range.

Our findings indicated that, even at the very beginning of primary school, children possess a certain competency to restore acoustically degraded speech based on linguistic context. This should not, however, tempt us to assume they are unaffected by classroom noise or by a teacher's dysphonic voice. Disruptions during low-level speech perception might carry over to high-level listening comprehension and make listening more effortful.

Finally, in terms of classroom acoustics, we showed that none of the eight primary school classrooms in which the listening tasks were carried out provided optimal listening and learning conditions. Concurrently with what has been observed in international noise surveys, noise levels, reverberation times, and STI values mostly deviated from the recommended standards. It is still important to tackle this problem to support children's academic performance and make both teaching and learning pleasant experiences.

Acknowledgments

This article was published with support from the Belgian University Foundation. This study was also supported by University of Liège PhD Grant RD/DIR-vdu/2016.7166, awarded to Isabel S. Schiller, and by a grant from the National Fund for Scientific Research (F.R.S.–FNRS), awarded to Angélique Remacle.

The authors would like to thank the participating schools for their assistance during the experiment. We also acknowledge the help of Justine Jonlet, Cloé Deflandre, Sacha Pszenica, and Elsa Poncet in collecting the data. We also thank Sysmex Belgium N.V. for providing the NL-21 sound-level meter used in this study. Finally, we thank the University of Music in Detmold for providing the loudspeakers, measurement equipment, and audio interface.

References

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

American National Standards Institute. (2010). *ANSI/ASA S12.60-2010/Part 1: American National Standard: Acoustical performance criteria, design requirements, and guidelines for schools, Part 1: Permanent schools.*

Barsties, B., & De Bodt, M. (2015). Assessment of voice quality: Current state-of-the-art. *Auris Nasus Larynx, 42*(3), 183–188. https://doi.org/10.1016/j.anl.2014.11.001

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Boersma, P., & Weenink, D. (2017). *Praat* (Version 6.0.29) [Computer software]. http://www.fon.hum.uva.nl/praat/

Bradley, J. S., & Sato, H. (2008). The intelligibility of speech in elementary school classrooms. *The Journal of the Acoustical Society of America, 123*(4), 2078–2086. 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), 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 noise. *American Journal of Audiology, 27*(2), 231–237. https://doi.org/10.1044/2018_AJA-17-0061

Crandell, C. C., & Smaldino, J. J. (2000). Classroom acoustics for children with normal hearing and with hearing impairment. *Language, Speech, and Hearing Services in Schools, 31*(4), 362–370. https://doi.org/10.1044/0161-1461.3104.362

Four Audio. (2018). *WinMF measurement software* [Computer software].

Hirano, M. (1981). Psycho-acoustic evaluation of voice. In G. Arnold, F. Winckel, & B. Wyke (Eds.), *Disorders of human communication* (pp. 81–84). Springer.

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*(12), 928–932. https://doi.org/10.3109/14992027.2010.520036

Khomsi, A. (2001). *ELO: Évaluation du Langage Oral* (Oral Language Evaluation) [Measurement instrument]. ECPA Pearson.

Korkman, M., Kirk, U., & Kemp, S. L. (2007). *NEPSY-II* (2nd ed.) [Measurement instrument]. Harcourt Assessment.

Lenth, R. V. (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software, 69*(1), 1–33. https://doi.org/10.18637/jss.v069.i01

Lo, S., & Andrews, S. (2015). To transform or not to transform: Using generalized linear mixed models to analyse reaction time data. *Frontiers in Psychology, 6,* 1171. https://doi.org/10.3389/fpsyg.2015.01171

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*(1), 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*(6), 577–588. https://doi.org/10.3109/17549507.2015.1024172

Macchi, L., Descours, C., Girard, É., Guitton, É., Morel, C., Timmermans, N., & Boidein, F. (2012). *ELDP: Épreuve Lilloise de Discrimination Phonologique* (The Lille Test of Phonological Discrimination) [Measurement instrument]. http://orthophonie.univ-lille2.fr/stocks/stock-contents/epreuve-lilloise-de-discrimination-phonologique.html

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*(6), 716–724. https://doi.org/10.1016/j.jvoice.2014.02.008

Maryn, Y., Corthals, P., Van Cauwenberge, P., Roy, N., & De Bodt, M. (2010). Toward improved ecological validity in the acoustic measurement of overall voice quality: Combining continuous speech and sustained vowels. *Journal of Voice, 24*(5), 540–555. https://doi.org/10.1016/j.jvoice.2008.12.014

Mattys, S. L., Brooks, J., & Cooke, M. (2009). Recognizing speech under a processing load: Dissociating energetic from informational factors. *Cognitive Psychology, 59*(3), 203–243. https://doi.org/10.1016/j.cogpsych.2009.04.001

McShefferty, D., Whitmer, W. M., & Akeroyd, M. A. (2015). The just-noticeable difference in speech-to-noise ratio. *Trends in Hearing, 19,* 1–9. https://doi.org/10.1177/2331216515572316

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. D. M. Hillock & D. J. Mee (Eds.), *Proceedings of Acoustics 2016: The Second Australasian Acoustical Societies Conference* (pp. 1047–1056). The Australian Acoustical Society.

Millward, K. E., Hall, R. L., Ferguson, M. A., & Moore, D. R. (2011). Training speech-in-noise perception in mainstream school children. *International Journal of Pediatric Otorhinolaryngology, 75*(11), 1408–1417. https://doi.org/10.1016/j.ijporl.2011.08.003

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*(1), 17–25. https://doi.org/10.1080/14015430118232

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

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,* 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*(2), 196–210. https://doi.org/10.1044/2018_LSHSS-18-0039

R Core Team. (2019). *R: A language and environment for statistical computing* [Computer software]. https://www.R-project.org/

Roy, N., Merrill, R. M., Thibeault, S., Parsa, R. A., Gray, S. D., & Smith, E. M. (2004). Prevalence of voice disorders in teachers and the general population. *Journal of Speech, Language, and Hearing Research, 47*(2), 281–293. https://doi.org/10.1044/1092-4388(2004/023)

Röer, J. P., Bell, R., & Buchner, A. (2014). Evidence for habituation of the irrelevant-sound effect on serial recall. *Memory & Cognition, 42*(4), 609–621. https://doi.org/10.3758/s13421-013-0381-y

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. S., Remacle, A., Durieux, N., & Morsomme, D. (2020, January, 9–10). *The effects of noise and impaired voice quality on spoken language processing in school-aged children: Preliminary results of a systematic review* [Poster presentation]. The 12th Speech in Noise Workshop, Toulouse, France.

Schiller, I. S., Remacle, A., & Morsomme, D. (2019). *NODYS: NOrmophonic and DYsphonic Speech samples, Mendeley Data* (Version 1) [Database]. Mendeley. https://doi.org/10.17632/g2fmkw8t85.1.

Schiller, I. S., Remacle, A., & Morsomme, D. (2020). Imitating dysphonic voice: A suitable technique to create speech stimuli for spoken language processing tasks. *Logopedics Phoniatrics Vocology, 45*(4), 1651–2022. https://doi.org/10.1080/14015439.2019.1659410

Steeneken, H. J. M., & Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *The Journal of the Acoustical Society of America, 67*(1), 318–326. https://doi.org/10.1121/1.384464

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*(3), 1043–1051. https://doi.org/10.1044/2015_JSLHR-H-14-0204

Teixeira, J. P., & Fernandes, P. O. (2015). Acoustic analysis of vocal dysphonia. *Procedia Computer Science, 64,* 466–473. https://doi.org/10.1016/j.procs.2015.08.544

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

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

Appendix

Picture of the Table Setup

Note. The purpose of the screens was to prevent children from copying their neighbors' answers. Each child received an answer booklet and a pen.