Supplementary Material.

Description of the intervention for participants randomly assigned to the experimental group or to the control group of the randomized controlled trial. The description is based on the Template for Intervention Description and Replication (TIDieR) (Hoffmann et al., 2014).

Note that the detailed intervention protocol has been published previously by Remacle et al. (2022).

Items	Experimental group	Control group
Brief name	VirtuVox Program (virtual reality applied to voice behavior). Voice-related prevention for teachers based on training in oral communication skills in environmental and communicative situations that are close to the reality of teaching.	/
Why	Teachers represent a population at risk for voice disorders (Cantor Cutiva et al., 2013; Epstein et al., 2011; Martins et al., 2014; Phyland & Miles, 2019). The high ambient noise level in school environments represents a risk factor (Cutiva et al., 2013). A program for training communication skills to deal with background noise was created to acquire and retain healthy vocal motor behavior and ultimately prevent voice disorders in teachers (Remacle et al., 2022). This mixed program combines a direct intervention with an indirect intervention (Van Stan et al., 2015). It was inspired by cognitive behavioral therapy. The active ingredients are immersion in a virtual classroom and simulations of communicative situations in noise. The hypothesized mechanisms of action are that these real and virtual simulations will allow trainee teachers to train and automatize oral communication skills, promote immunogenic behaviors and reduce pathogenic behaviors.	/
What – Materials & procedures	Indirect intervention: A group information session on voice functioning and vocal hygiene (duration: 1 h). - Presentation of a video on preventing voice disorders in teachers (Jowaheer, 2019) (duration: 30 min). The video has the following sections: (1) voice production (explanations of voice functioning and anatomical and physiological concepts); (2) definition of a voice disorder; (3) importance of posture and breathing; (4) voice in different contexts: how we can adapt; (4) vocal fatigue: symptoms and measures to take; (6) advice on vocal hygiene; (7) the role of the environment; and (8) what to do about voice disorders. Link to video: https://www.youtube.com/watch?v=2wFFVM7ehRI - Demonstration of equipment (duration: 15 min): (1) presentation and handling of a tool using the traffic light system to give visual feedback on the noise level in a school environment (Vinco noise level traffic light, Ref: LD22744N); (2) presentation and handling of portable voice amplifiers (e.g., WAP-5, Monacor). - Answers to participants' questions (duration: 15 min).	
What – Materials & procedures	Direct intervention: 3 individual vocal training sessions including immersion in a virtual classroom and simulations of communicative situations in background noise. The purpose is to train and automatize	/

healthy vocal motor behavior. Each session addresses three specific skills with voice production activities with and without noise, inspired by cognitive behavioral therapy.

1. SESSION 1

1.1. Skills targeted:

(1) Adjust body posture in order to facilitate voice projection; (2) direct the gaze to the listener/interlocutor to encourage the speaker to actively use gaze as a clue to nonverbal communication; and (3) adjust general and local muscle tone with the goal of avoiding laryngeal tension and vocal effort.

1.2. Procedure and equipment:

1.2.1. Contextualization of proposed training (10 min).

Method: Have the participant discuss their previous experiences of communicating in noise, with the help of a semi-structured interview.

Script: "Can you identify a noisy situation in which you had to communicate? How long ago? Can you describe the kind of noise? Who were you talking to and how many of them were there? Can you explain what happened? What did you feel physically, in terms of voice quality, and in terms of your efficacy in this communicative situation? What was your strategy? Would it be an appropriate one to copy in future?"

1.2.2. Verification of theoretical comprehension of targeted skills (5 min)

Method: In a semi-structured interview, the participant is invited to explain their representations of the three skills covered in the session. The trainer provides feedback if necessary to correct or clarify matters.

1.2.3. Training of targeted skills (15 min)

Method: The trainer asks the participant to introduce themselves and to pay attention to (1) the position of their body (nape, neck, shoulders, jaw, pelvis, legs, knees, feet); (2) the orientation of their gaze; and (3) their muscle tone (face, nape, neck, shoulders, throat, abdomen, back, legs).

The participant verbalizes their feeling regarding each of the three target skills. The trainer gives feedback on the behaviors she observed and on the participant's verbalizations. She suggests relevant ways to improve the trained skills. The activities are practiced standing up in the following conditions:

(1) gaze on the trainer, no noise; (2) gaze on the trainer, cocktail-party-type noise played through a speaker at 60 dBA; and (3) gaze in other directions, with the same kind of noise.

NB: (1) The participant and the trainer can move around if they want; the initial distance between them is 2.5 m. (2) A mirror allows the participant to observe themselves if they feel the need.

Script: "I'm going to present three practical scenarios:

(1) "Standing up, try to focus your attention on the position of your head, shoulders, back, pelvis, legs, knees, feet. Do you feel tension in some places? Where are you looking? Can you introduce yourself while looking at me? While you talk, you'll pay attention to your body's position, your gaze and your muscle tone. What do you feel?

(2) "Can you describe one of your hobbies to me, while looking at me? Some noise will be played. You'll pay attention to your body's position, your gaze and your muscle tone. What do you feel?

(3) "Can you explain your motivations for studying to be a teacher? At the same time, noise will be played and you'll look in other directions (avoid looking at me). Also think about transmitting a clear message. You'll pay attention to your body's position, your gaze and your muscle tone. What do you feel? Do you perceive any differences between the situation without noise and with noise? Do you perceive any differences between the situation where you look at the other person and where you look in another direction? Following these scenarios, can you identify something you became aware of that you could use in your daily routine?"

1.2.4. Practicing skills with the help of virtual reality (15 min)

Method: The participant is encouraged to practice and automatize the target skills during immersion in a virtual classroom. They are asked to introduce themselves to the pupils in the virtual classroom in the context of their first contact with the children.

The VR environment used consists of an immersive classroom developed for the purpose of the VirtuVox project and validated in Remacle et al. (2021). The environment features an elementary school classroom with 16 pupils aged 9 to 12 years old animated with typical childlike actions. The auditory stimuli corresponding to background noise without understandable semantic content are spatialized and played through the speakers of an Oculus Rift™ head-mounted display. The trainer can adjust the level of three noise sources and the children's agitation, which allows for clinical flexibility. The noise level (playground, corridor, and classroom) is set at 70 dBA and the agitation level of pupils in the virtual class at 30%.

Script: "You're going to be immersed in a virtual elementary school class. You will see that these pupils behave like real children: they'll move around and make noise, but they won't answer your questions

if you ask any. Since this is the first time that you're meeting these pupils, we'd like you to introduce yourself to them for 5 or 10 minutes, as you would do the first time you meet pupils when you're a trainee teacher. It's important for these children to get to know you and learn why you're there, as if you were starting out as a trainee teacher. You can also call on them if you want. You can explain to them what you're going to do together. During this experience, you'll pay attention to your posture, your gaze and your muscle tone."

1.2.5. Semi-structured interview (15 min):
Method: After the VR immersion, the participant expresses what they felt concerning the three targeted communication skills; their perception of noise in the virtual classroom and the virtual pupils' agitation; and the use of the target skills during the VR simulation. The trainer provides feedback and asks the participant to (1) identify something they noticed during the session that they would like to use in future; and (2) for the next session, identify a noisy situation in which they had to communicate and how they felt in relation to the three target skills.
Script: "What did you feel regarding your body position, your gaze, your muscle tone (movements)? What was your perception of the noise in the virtual classroom and the pupils' agitation level? How did you use your voice in that environment? Do you perceive differences between your voice use in the virtual classroom and outside that environment? If you had to retain one thing you could use in the future, what would it be?
"For the next session, during the week to come, can you identify (1) a noisy situation in which you had to communicate; (2) what you felt physically in relation to that situation (posture, gaze and muscle tone), your voice quality, and your feeling of efficacy (did the listener understand your message?)."

2. SESSION 2
2.1. Skills targeted:
(1) Search for effective breathing patterns to support speech production and voice projection; (2) search for resonant voice patterns involving a configuration of the vocal tract that allows for improved supraglottic resonance; this technique aims to increase vocal level and spectral energy in the region where the human ear is most sensitive to SPL; and (3) adjust the speech rate to promote perception and understanding of the message by the listener/interlocutor.

2.2. Procedure and equipment:

2.2.1. Contextualization of proposed training (5 min).

Method: Have the participant discuss their previous experiences of communicating in noise, with the help of a semi-structured interview.

Script: "During the week, were you able to identify a noisy situation in which you had to communicate? What day was it? Can you describe the kind of noise? Who were you talking to and how many of them were there? Can you explain what happened? What did you feel physically, in terms of voice quality, and in terms of your efficacy in this communicative situation? What were your strategies? Were they appropriate and could they be used again in future?"

2.2.2. Verification of theoretical comprehension of targeted skills (5 min)

Method: In a semi-structured interview, the participant is invited to explain their representations of the three skills covered in the session. The trainer provides feedback if necessary to correct or clarify matters. A 3D anatomical representation of the larynx and a 2D model of the speech organs are used to illustrate the explanations.

2.2.3. Training of targeted skills (20 min)

Method: The trainer asks the participant to produce automatized series, the sustained phoneme /m/, a reading, and spontaneous speech, while paying attention to (1) breathing patterns, (2) supraglottic resonance, and (3) speech rate.

The participant verbalizes their feelings regarding each of the three target skills. The trainer gives feedback on the behaviors she observed and on the participant's verbalizations. She suggests relevant ways to improve the trained skills and refers to the concepts covered in the indirect intervention. The activities are practiced standing up in the following conditions: (1) paying attention to respiratory movements, without noise and then with speech-shaped noise played through a speaker at 60 dBA; (2) paying attention to the vibratory sensations in the face and supraglottal tract, without noise and then with school noise recorded in a playground and played through a speaker at 60 dBA; and (3) experimenting with different speech rates while reading a children's book, without noise and then with speech-shaped noise played through a speaker at 60 dBA.

NB: (1) The participant and the trainer can move around if they want; the initial distance between them is 2.5 m. (2) With the participant's agreement, the trainer can touch them to guide the position of the hands, or help them perceive respiratory movements or vibrations in the facial area. (3) A

mirror allows the participant to observe themselves if they feel the need. (4) The participant can refer to their student teaching session the previous year.

Script: "I'm going to present three practical scenarios, each one without noise and then with background noise:

(1) "Standing up, focus your attention on your respiratory movements by touching your chest and abdomen. Do you perceive movements? What kind? When (while breathing in or breathing out)?

"Visualize yourself in class, facing your pupils. Can you count up to 20 while you imagine having these pupils do a psychomotor exercise and pay attention to the movements of your chest and abdomen. What do you feel? Can you do the same thing in noise? What do you feel? Do you perceive any differences without noise and with noise?

(2) "Can you sustain the sound /m/ and focus your attention on the vibrations in your throat and face (nose, cheekbones, lips, cheeks, jaw). Produce this sound with your usual pitch, a high-pitched voice, and a low-pitched voice [if necessary, the trainer gives examples]. What do you feel?

"Still imagining that you're in class, can you say the months of the year while focusing your attention on the vibrations in your throat and face (nose, cheekbones, lips, cheeks, jaw)? A background noise will be played at the same time. What do you feel?

(3) "Imagine that you're in class, facing your pupils. Can you read this book, first at a slow rate, then fast and then medium? [Possible adaptation of the instructions based on the participant's normal speech rate.]

"You will focus your attention on what you feel in terms of breathing, the feeling of ease and control, your ability to bring the story to life (using intonation). What do you feel? Can you do the same exercise in noise? What do you feel? Do you perceive any differences between the condition without noise and with noise?

"Following these scenarios, can you identify something you became aware of that you could use in your daily routine?"

2.2.4. Practicing skills with the help of virtual reality (15 min)

Method: The participant is encouraged to practice and automatize the target skills during immersion in a virtual classroom. They are asked to explain an activity to students that they are going to do together.

The VR environment used is the same as in 1.2.4. The noise level (playground, corridor, and classroom) is set at 70 dBA and the agitation level of pupils in the virtual class at 50%.

Script: "You're going to be immersed in a virtual elementary school class. You will see that these pupils behave like real children: they'll move around and make noise, but they won't answer your questions if you ask any. Remembering your student teaching session last year, we suggest that you take 5 to 10 minutes to explain an activity that you're going to do with them. You can also call on them if you want. During this experience, you'll pay attention to your breathing patterns, vibrations in the face area, and speech rate."

Topics to present to the virtual class are suggested to the participant: a school celebration, a field trip, a tour (e.g., museum, city), a longer stay (e.g., open-air class/winter sports class/city exploration class). The participant can explain issues related to the context, the timetable (departure, return), travel (e.g., mode of transportation), things to bring (e.g., hat, sunblock, coat), meals, how much spending money to bring, activities on-site, etc.

2.2.5. Semi-structured interview (15 min):
Method: After the VR immersion, the participant expresses what they felt concerning the three targeted communication skills; their perception of noise in the virtual classroom and the virtual pupils' agitation; and the use of the target skills during the VR simulation. The trainer provides feedback and asks the participant to (1) identify something they noticed during the session that they would like to use in future; (2) for the next session, identify a noisy situation in which they had to communicate and how they felt in relation to the three target skills; and (3) prepare for the VR immersion in session 3 by planning a story to tell or a teaching activity they can present to the pupils in the virtual class.
Script: "What did you feel regarding your breathing patterns, vibrations in your face, and speech rate? What was your perception of the noise in the virtual classroom and the pupils' agitation level? How did you use your voice in that environment? Do you perceive differences between your voice use in the virtual classroom and outside that environment? If you had to retain one thing you could use in the future, what would it be?
"For the next session, during the week to come, can you identify (1) a noisy situation in which you had to communicate; (2) what you felt physically in relation to that situation (breathing, vibrations in the face area, and speech rate), your voice quality, and your feeling of efficacy (did the listener understand your message?). Also, (3) prepare for the VR immersion in session 3 by planning a story to tell or a teaching activity you can present to the pupils in your virtual class."

3. SESSION 3
3.1. Skills targeted:

(1) Search for variations in pitch and vocal intensity to favor temporal modulations in f_0 and SPL; (2) search for more precise and higher-amplitude articulatory movements to favor auditory and visual perception of speech; and (3) search for visual communication cues such as facial expressions, gaze and articulatory movements.

3.2. Procedure and equipment:
3.2.1. Contextualization of proposed training (5 min).
Method: Have the participant discuss their previous experiences of communicating in noise, with the help of a semi-structured interview.
Script: "During the week, were you able to identify a noisy situation in which you had to communicate? What day was it? Can you describe the kind of noise? Who were you talking to and how many of them were there? Can you explain what happened? What did you feel physically, in terms of voice quality, and in terms of your efficacy in this communicative situation? What were your strategies? Were they appropriate and could they be used again in future?"

3.2.2. Verification of theoretical comprehension of targeted skills (5 min)
Method: In a semi-structured interview, the participant is asked to recall the competences covered in sessions 1 and 2. The participant is then invited to explain their representations of the three skills covered in session 3. The trainer provides feedback if necessary to correct or clarify matters.

3.2.3. Training of targeted skills (15 min)
Method: The trainer asks the participant to read aloud, while paying attention to (1) variations in pitch and intensity; (2) the articulation of speech sounds; and (3) visual communication cues. The trainer explains the importance of multimodal perception for pupils (visual and auditory aspects), and the fact that some pupils are more sensitive to the auditory channel and others to the visual channel, which is why it is important for teachers to use both channels to communicate in ways appropriate to each pupil, especially those who have learning difficulties.
The participant verbalizes their feelings regarding each of the three target skills. The trainer gives feedback on the behaviors she observed and on the participant's verbalizations. She suggests relevant ways to improve the trained skills and refers to the concepts covered in the indirect intervention. The following activities are practiced in a standing position:
(1) Reading without variations in pitch and intensity by the trainer, followed by verbalization by the participant; reading with variations in pitch and intensity by the trainer, followed by verbalization by

the participant; reading without variations by the participant, followed by verbalization and feedback by the trainer; and reading with variations by the participant, followed by verbalization and feedback by the trainer.

(2) Reading with hypoarticulation by the trainer, followed by verbalization by the participant; reading with hyperarticulation by the trainer, followed by verbalization by the participant; reading with hypoarticulation by the participant, followed by verbalization and feedback by the trainer; reading with hyperarticulation by the participant, followed by verbalization and feedback by the trainer.

(3) Reading with a hidden face by the trainer, followed by verbalization by the participant; reading with a visible face by the trainer, followed by verbalization by the participant.

(4) Reading by the participant, paying attention to the three target skills, without noise and then with speech-shaped noise played through a speaker at 60 dBA.

NB: The participant and the trainer can move around if they want; the initial distance between them is 2.5 m.

Script: "I'm going to present four practical scenarios, with noise for the last one.

(1) "I'm going to start reading the book in two different ways (the trainer reads without and then with variations in pitch and intensity). What are your impressions? What differences do you perceive? Can you do the same thing: first without and then with variations in your voice? What do you feel? What differences do you perceive?

(2) "I'll continue reading in two different ways (the trainer reads while hypoarticulating and then while hyperarticulating). What are your impressions? What differences do you perceive visually and auditorily? Can you do the same thing: first without articulating and then articulating (making lip and tongue movements more precise and increasing the movement amplitude)? What do you feel? What differences do you perceive? (lips, tongue, jaw, vertical mouth opening)

(3) "I'll finish by reading in two different ways (face hidden, then face visible). What are your impressions? What differences do you perceive?

(4) "Can you continue reading, without noise and then with noise, producing variations in your voice and hyperarticulating. You will first focus your attention on the voice variations, then on articulation, then on both at once. Do you perceive any differences with and without noise?

"Following these scenarios, can you identify something you became aware of that you could use in your daily routine?"

3.2.4. Practicing skills with the help of virtual reality (15 min)

	Method: The participant is encouraged to practice and automatize the target skills during immersion in a virtual classroom. They are asked to tell a story or present a teaching activity to the pupils in the virtual class. The VR environment used is the same as in 1.2.4. The noise level (playground, corridor, and classroom) is set at 70 dBA and the agitation level of pupils in the virtual class at 70%. Script: "You're going to be immersed in a virtual elementary school class. You will see that these pupils behave like real children: they'll move around and make noise, but they won't answer your questions if you ask any. We suggest that you take 5 to 10 minutes to tell them a story or present a teaching activity (the scenario we asked you to prepare at the end of the last session). During this experience, you'll pay attention to the variations in your pitch and intensity, the articulation of speech sounds, and visual communication cues." 3.2.5. Semi-structured interview (20 min): Method: After the VR immersion, the participant expresses what they felt concerning the three targeted communication skills; their perception of noise in the virtual classroom and the virtual pupils' agitation; and the use of the target skills during the VR simulation. The trainer provides feedback and asks the participant to identify a key factor to remember in future. At the end of the intervention, the trainer asks the participant how they plan to implement what they have seen in practice: in their next student teaching session (identification of a short-term goal) and in their first year of employment as an elementary schoolteacher (identification of a long-term goal). Script: "What did you feel regarding the variations in the pitch and intensity of your voice, your articulation of speech sounds, and the visible aspects of communication? What was your perception of the noise in the virtual classroom and the pupils' agitation level? How did you use your voice in that environment? Do you perceive differences between your voice use in the virtual classroom and outside that environment? If you had to retain one thing you could use in the future, what would it be? "What did you learn in this program that you plan to put in practice in your next student teaching session? How? "What did you learn in this program that you plan to put in practice when you start working as a teacher? How?"	
Who provided	A female speech therapist with a specialty in voice (20 years' experience), who was trained in advance in the technical aspects of VR and in the implementation of VR in cognitive behavioral therapy.	Two female speech therapists with a specialty in voice
How	Indirect intervention: A group session given in person to all participants in the experimental and control groups (group size = 101).	

	Direct intervention: 3 sessions given in person individually to participants in the experimental group.	/	
Where	Indirect intervention: The session took place in the participants' school (in French-speaking Belgium), in a 200-seat auditorium.		
	Direct intervention: The sessions took place in the participants' school (in French-speaking Belgium), in a room measuring 10.3 x 5.5 x 4.5 m, with a mean sound level of 41 dB LAeq. The participant and the trainer were standing face to face, 2.5 m apart.	/	
When and how much	All sessions took place on participants' regular schooldays.		
	Direct intervention: The 3 individual training sessions (duration = 1 h) were spaced one week apart. The 3 sessions took place over a total duration of 19 ± 5 days. Indirect intervention: The group information session on voice functioning and vocal hygiene (duration = 1 h) took place between direct sessions 1 and 2.	Indirect intervention: The group information session on voice functioning and vocal hygiene (duration = 1 h) took place at the same time as for the experimental group.	
Tailoring	Indirect intervention: Participants were encouraged to ask questions and the trainer answered them. Direct intervention: The protocol was designed to allow for personalization during the semi-structured interviews and question periods. The level of difficulty of activities and feedback were tailored based on each participant's skills.		
Modifications	No modification was applied during the course of the study.		
How well – Planned	To ensure that the protocol was applied faithfully, the project leader (first author) supervised the trainer during a pilot phase with 4 participants before the study. The project leader then gave the trainer feedback with the goal of fine-tuning the standardization of the intervention. As for participants' adherence, those who wanted them were sent reminders by SMS or email the day before each session.		
How well – Actual	Of the 34 participants enrolled in the experimental group, 27 attended the pre-intervention assessment; 26 attended session 1 of the direct intervention; 22 attended the indirect intervention session and session 2 of the direct intervention; 21 attended session 3 of the direct intervention; and 21 attended the post-intervention assessment. The analyses and results presented in the study include only participants who completed the entire program (n=21).	Of the 35 participants enrolled in the control group, 28 attended the pre-intervention assessment and the indirect intervention session, and 20 attended the post-intervention assessment. The analyses and results presented in the study include only participants who completed the entire program (n=20).	

References

Cantor Cutiva, L. C., Vogel, I., & Burdof, A. (2013). Voice disorders in teachers and their associations with work-related factors: A systematic review. *Journal of Communication Disorders, 46,* 143–155.

Epstein, R., Remacle, A., & Morsomme, D. (2011). From reactive intervention to proactive prevention: The evolution of occupational dysphonia. *Perspectives on Voice and Voice Disorders, 21*, 48–55.

Hoffmann, T. C., Glasziou, P. P., Boutron, I., Milne, R., Perera, R., Moher, D., Altman, D. G., Barbour, V., Macdonald, H., Johnston, M., Lamb, S. E., Dixon-Woods, M., McCulloch, P., Wyatt, J. C., Chan, A.-W., & Michie, S. (2014). Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *BMJ, 348,* Article g1687. http://dx.doi.org.ezproxy1.library.usyd.edu.au/10.1136/bmj.g1687

Jowaheer, R. (2019). *Evaluation des connaissances des troubles de la voix chez un groupe d'enseignants de maternelle et primaire par la création d'une vidéo de prévention : Etude pilote* ["Assessment of knowledge of voice disorders in a group of kindergarten and elementary schoolteachers by creating a prevention video: Pilot study"] [Unpublished master's thesis]. Université catholique de Louvain. https://www.youtube.com/watch?v=2wFFVM7ehRI

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

Phyland, D. & Miles, A. (2019). Occupational voice is a work in progress: Active risk management, habilitation and rehabilitation. *Current Opinion in Otolaryngology and Head and Neck Surgery, 27*(6), 439–447. https://doi.org.10.1097/MOO.0000000000000584

Remacle, A., Ancion, V., & Morsomme, D. (2022). Protocole pour l'entraînement des compétences de communication orale des enseignants dans un objectif de prévention vocale : Description du programme VirtuVox ["Protocol for training teachers in oral communication skills for the purpose of voice protection: Description of the VirtuVox program"], *Langue(s) & Parole, 7,* 23–46.

Remacle, A., Bouchard, S., Etienne, A.-M., Rivard, M.-C., & Morsomme, D. (2021). A virtual classroom can elicit teachers' speech characteristics: Evidence from acoustic measurements during in vivo and in virtuo lessons, compared to a free speech control situation. *Virtual Reality, 25*(4), 935–944.

Van Stan, J., Roy, N., Awan, S., Stemple, J., & Hillman, R. (2015). A taxonomy of voice therapy. *American Journal of Speech-Language Pathology, 24*(2), 101–125.