Dealing with atypical noise sources: perceived difficulties when listening to dysphonic voices

Annelies Bockstael, Ingrid Verduyckt
École d'orthophonie et d'audiologie, Université de Montréal, Montreal, Canada.

Annelies Vandevelde, Dick Botteldooren
WAVES INTEC, Ghent University, Ghent, Belgium.

Summary
Voice disorders or dysphonia are clearly more prevalent in teachers compared to the general population. Voice disorders not only have negative effects on teachers' health and well-being, also for the students, information processing appears to be more cumbersome when listening to a dysphonic voice. The question is how listening to a dysphonic voice fits within the existing knowledge on the cognitive effects of background noise in general. In addition, most of the effects of dysphonic voices are studied for (primary school) children, it is unclear to what extent they hold true for adults. In this, adult students might be less affected by the dysphonic voices, as they have more (cognitive) resources to deal with adverse listening conditions. Finally, it is unclear whether different types of voice disorders, with different perceptual characteristics, affect information processing differently. This study investigates information processing by adults in conditions that are thought the be very challenging: speech in multitalker babble, speech produced by dysphonic voices without additional background noise, and dysphonic speech within multitalker babble. As a reference condition, speech has been produced by a healthy voice without additional background noise. For the dysphonic voices, three different voice disorders have been included. Participants were asked to report perceived difficulty of information processing. Results show that reported difficulty clearly varies depending on the conditions. Compared to the reference condition, reported difficulty is clearly higher for information presented in multitalker babble and information presented by a dysphonic voice. Remarkably, within multitalker babble no differences are seen between dysphonic voices and a healthy voice. Finally, no differences in rating were seen in-between the different voice disorders included in this study.

PACS no. 43.50.Qp

1. Introduction

Processing information that is presented orally in background noise requires focusing on the target signals while suppressing the irrelevant sounds. Whether this can be done successfully depends on a complex interplay between features of signal and noise, tasks at hand, and individual characteristics of the listener [1],[2],[3].

Irrelevant speech is often shown to be particularly disturbing, especially when it is intelligible and meaningful [1]. Besides intelligibility, mere phonological similarity between target and masking signal might also increase the influence of the masker [1].

Voice disorders can be regarded as a particular type of noise because here the noise is actually part of the target signal. Dysphonia is defined as a speech disorder 'characterized by the abnormal production and/or absences of vocal quality, pitch, loudness, resonance, and/or duration, which is inappropriate for an individual's age and/or sex.' (ASHA). The effect of dysphonia on learning is a very pertinent question, as dysphonia is often reported in teachers.

Dysphonic voices have been shown to affect information processing and language comprehension in children [4] as well as in adults [5]. This is in line with the generally known effects of irrelevant noise on task performance. It is less clear how important the effects of dysphonic voices are, compared to other external noise sources. In addition, it has been little investigated how dysphonic noise and background noise might interact, and what their combined effect on information processing would be. This is a pertinent question as teaching, also for teachers suffering from dysphonia, usually takes place in a certain level of (classroom) background noise.

Furthermore, especially for a population of young adults who supposedly have more cognitive resources to deal with adverse listening conditions as compared to school-aged children, it is not clear whether background noise and dysphonic voices have similar effects or not.

Finally, no relationship has been found between the severity of the dysphonic voice and the degree of decrease in information processing [6]. It remains to be seen whether voice disorders that sound differently, for example a creaky versus an overly breathy voice, have different effects on information processing.

In this work, the effect on information processing is investigated for (1) dysphonic voices versus multitalker babble background noise, and versus non-speech background noise with a spectrum similar to dysphonic noise, for (2) the combined effect of dysphonic voices and multitalker babble, and for (3) voice disorders with different perceptual characteristics. Self-reported ease of information processing is used as outcome.

2. Method

2.1. Participants and protocol

Forty-nine volunteers between 18 and 30 years old (average: 21.1 years) participated. Participants were instructed to listen carefully to 10 different lectures on various topics. Information was presented in layman terms to be comprehensible by young adults. The topics themselves were somewhat obscure to ensure that participants had little or no a-priori knowledge on the content.

Each lecture took about 5 minutes. After each lecture, participants had to write down up-to five key elements they had retained from the lecture, as well as to answer six true/false questions. After all lectures had been played back, participants were asked to order the 10 fragments in terms of how easy it was to follow the content of the lecture, with the easiest fragment first and the hardest at the last place.

2.2. Listening conditions

All lectures were read by a 40-year female speech therapist. Play-back of the lectures was done with different voice characteristics and different fragments of background noise.

For the voice characteristics, three different dysphonic voices were simulated using the software TC Helicon VoiceOne. A panel of three voice experts and five non-expert listeners judged the voice quality of the simulations. The selected simulations were judged as clearly dysphonic and could not be distinguished from natural (non-simulated) voices by the non-expert listeners. One healthy voice condition was added to the three simulated dysphonic voices, so in total four different voice conditions were included.

Two different background noise conditions were included: no additional background noise and multitalker babble. All four voice conditions were presented twice, i.e. once in multitalker babble and once without additional background noise.

Participants listened to the fragments through headphones, lectures were played at 68 dB calibrated with the Head And Torso Simular (HATS) type 4128C from Brüel & Kjær. Multitalker babble and dysphonic noise were played at 63 dB.

2.3. Self-reported effort

After all fragments had been presented, participants had to rank them on a ten-point scale between 1 and 10, 1 being the fragment that was most easy to retain information from, and 10 the hardest. Results are statistically analyzed using mixed model regression with the self-reported ranks as outcome variable. Participant is included as random factor. Two independent fixed effects are included in the model, voice condition (four levels: three types of dysphonic voices and one healthy voice condition), and background condition (two levels: no additional noise and multitalker babble).The interaction effect between voice and background condition is also considered, accounting for all eight listening conditions. Tukey post-hoc testing has been carried out to compare the ranking of the different listening condition pairwise.

3. Results

In Figure 1, the self-reported ease of following the content of the lectures is shown as a function of voice condition and background noise condition. Statistically, a strong interaction effect is seen between voice and background condition on self-reported ease of information processing ($p < 0.0001$).

Pairwise Tukey post-hoc testing reveals that multitalker babble stands out: information processing is systematically reported to be more difficult in multitalker babble compared to no additional background noise ($p < 0.01$ for all pairwise comparisons).

In the conditions without additional noise, the dysphonic voices appear to be systematically more difficult than the healthy voice ($p < 0.01$ for all relevant pairwise comparisons). Contrary, with multitalker babble noise present, no difference could be found between reported ease of information processing for dysphonic voices versus a healthy voice ($p > 0.1$ for all relevant pairwise comparisons). Finally, in-between the different types of dysphonic voices, no difference in reported ease of processing could be seen ($p > 0.1$ for all relevant pairwise comparisons).

4. Discussion

Both multitalker babble and dysphonic voices are rated less favorably compared to a healthy voice with-

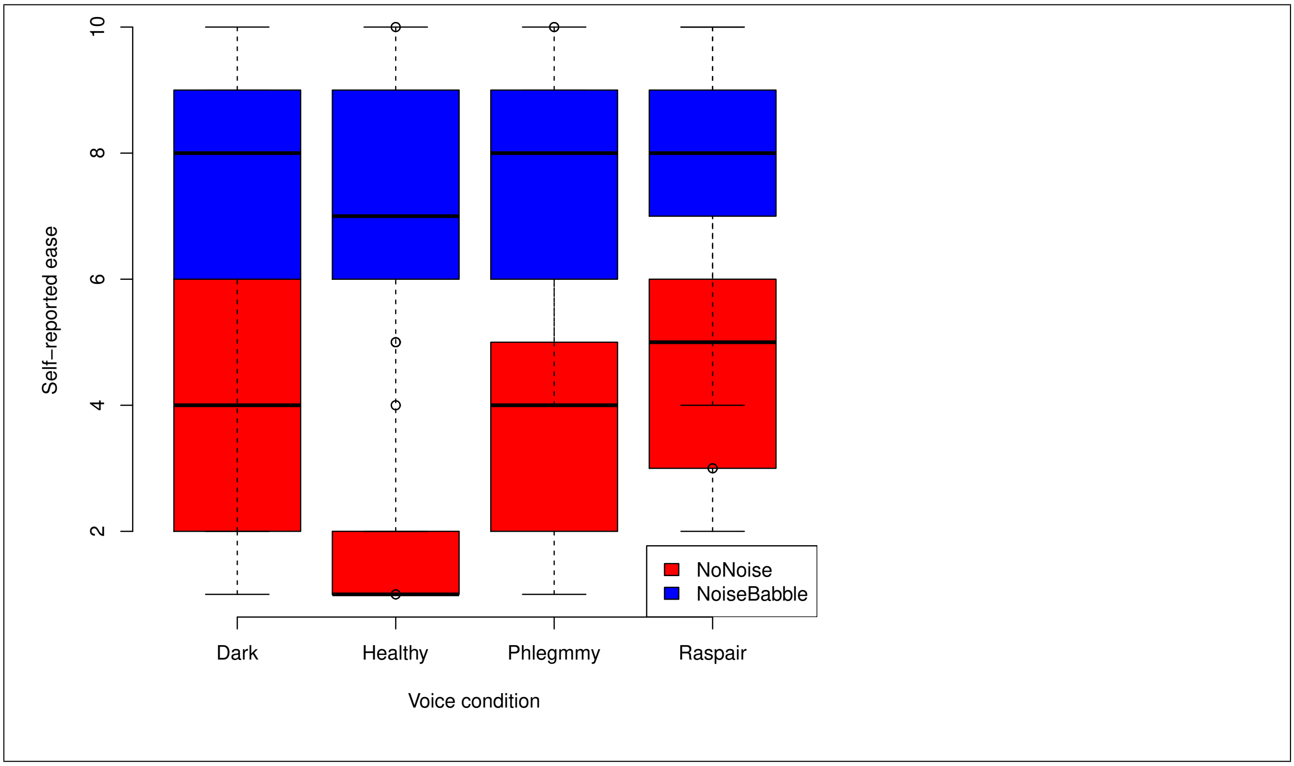

Figure 1. Self-reported ranking of ease of information processing as a function of voice condition (on the x-axis: dysphonic voices dark, phlegmmy and raspair, and healthy voice) and background condition (color coded: no addional background noise NoNoise and multitalker babble NoiseBabble)

out additional background noise. Speech sounds (multitalker babble) are known to be likely to draw the listener's attention, whereas the dysphonic sounds might be difficult to separate perceptually from the target signal as it is inherently part of it.

In this experiment, the multitalker babble has clearly been recognized as a stronger interfering noise source. The fact that the dysphonic characteristics are inherently part of the signal, as they are produced by the speaker, appears to be less important than the characteristics of multitalker babble. Within multitalker babble, listeners no longer make a distinction between a healthy voice and a dysphonic noise, so clearly the effects of background noise are more dominant than the effects of dysphonic voices. This is also suggested by the finding that participants do not perceptually distinguish between different types of dysphonic voices.

In this, it should be noted that the participants were young adults, often with a strong educational background. Effects might have been more distinguished in a population with less cognitive resources to deal with adverse listening condition, such as school-aged children.

5. Conclusions

The reported difficulty to process orally presented information clearly increased when lectures are presented in multitalker babble. Dysphonic voices have also a negative, albeit less strong, effect. Different dysphonic voices do not appear to lead to distinguishable effects, and within background noise, dysphonic voices do not lead to further increase in reported difficulty compared to the healthy voice.

References

[1] M. Rudner: Cognitive spare capacity as an index of listening effort. Ear Hear., 2016. 37, 69S–76S.

[2] K. S. Rhebergen, N. J. Versfeld, W. A. Dreschler: Release from informational masking by time reversal of native and non-native interfering speech. J. Acoust. Soc. Am., 2005. 118(3), 1274–1277.

[3] W. Wei, A. Bockstael, B. De Coensel, D. Botteldooren: Interference of speech and interior noise of chinese high-speed trains with task performance. Acta Acust. United Acust., 2012. 98(5), 790–799. Role: co-supervision of research quesions and experimental design, statistical analyses, interpretation, writing of specific sections, proofreading.

[4] V. Lyberg-Åhlander, M. Haake, J. Brännström, S. Schötz, B. Sahlén: Does the speaker's voice quality influence children's performance on a language comprehension test? International journal of speech-language pathology, 2015. 17(1), 63–73.

[5] M. Imhof, T.-R. Välikoski, A.-M. Laukkanen, K. Orlob: Cognition and interpersonal communication: The effect of voice quality on information processing and person perception. Studies in Communication Sciences, 2014. 14(1), 37–44.

[6] J. Rogerson, B. Dodd: Is there an effect of dysphonic teachers' voices on children's processing of spoken language? J. Voice, 2005. 19(1), 47–60.