



LUND UNIVERSITY

Does the speaker's voice quality influence children's performance on a language comprehension test?

Lyberg Åhlander, Viveka; Haake, Magnus; Brännström, Jonas; Schötz, Susanne; Sahlén, Birgitta

Published in:
International Journal of Speech-Language Pathology

DOI:
[10.3109/17549507.2014.898098](https://doi.org/10.3109/17549507.2014.898098)

2015

[Link to publication](#)

Citation for published version (APA):

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>

Total number of authors:
5

General rights

Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: <https://creativecommons.org/licenses/>

Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Does the speaker's voice quality influence children's performance on a language comprehension test?

Lyberg-Åhlander V¹., Haake M²., Brännström K.J¹., Schötz S³., Sahlén B¹.

¹Department of Logopedics, Phoniatrics and Audiology, Lund University

²Department of philosophy, Cognitive science, Lund University

³Department of Linguistics, Lund University

Key words: Voice perception, voice quality, children, cognitive capacity, working memory, occupational voice, language comprehension

Running head: voice quality influence on childrens' performance

Abstract

A small number of studies have explored children's perception of speakers' voice quality and its possible influence on language comprehension. The aim of this explorative study was to investigate the relationship between the examiner's voice quality, the child's performance on a digital version of a language comprehension test, the Test for Reception of Grammar (TROG-2), and two measures of cognitive functioning. The participants were ($n=86$) mainstreamed 8-year old children with typical language development. Two groups of children ($n = 41/45$) were presented with the TROG-2 through recordings of one female speaker: one group was presented with a typical voice and the other with a simulated dysphonic voice. Significant associations were found between executive functioning and language comprehension. The results also showed that children listening to the dysphonic voice achieved significantly lower scores for more difficult sentences ("the man but not the horse jumps") and used more self-corrections on simpler sentences ("the girl is sitting"). This suggests that a dysphonic speaker's voice may force the child to allocate capacity to the processing of the voice signal at the expense of comprehension. Our findings have implications for clinical and research settings where standardized language tests are used.

Introduction

Language comprehension is multimodal – children grasp the meaning of an utterance based on a combination of perceptual/lower level visual and auditory cues in the utterance and their higher level language and cognitive skills. Speakers naturally vary considerably as for speech dynamics, voice quality, gaze and gesture behaviour. Listeners have to recognize words and sentences despite this variability and adapt dynamically to different speakers.

The influence of the speaker's speech rate on children's comprehension has previously been studied at our lab (Haake, Hansson, Gulz, Schötz, & Sahlén, 2013). Instructions in a language comprehension test (TROG-2, Bishop, 2003) phrased by an examiner with fast speech rate seemed to hinder comprehension, whereas instructions spoken with slow speech rate was facilitating for off-line language comprehension. Language comprehension tests often lack specific instructions to examiners about paralinguistic factors (speech rate; voice quality; intonation; gaze and gesture behaviour) and there is therefore reason to believe that the variation in test administration challenges the reliability of language comprehension test results. In the present study the influence of a speaker's voice quality on children's performance in a language comprehension test is explored. External factors as voice quality and the level of difficulty of the test item (length and grammatical complexity of sentences) are related to internal cognitive factors in the child such as working memory capacity and executive functioning.

Most research on the relationship between speech recognition and cognitive factors has been performed in the field of audiology, where speech recognition in noise has typically been studied in adults with hearing impairment and controls (for an overview see Akeroyd, 2008). Working memory capacity was an effective predictor of speech recognition in noise, more so than the degree of hearing impairment when audibility was restored. There is

therefore reason to believe that the younger and the less cognitively mature a child is, the stronger her/his dependency on lower level the processing will be. Low level deviances in the acoustic signal caused by background noise or deviant voice quality in a speaker may, to a higher extent, disrupt language processing in children with less developed language and cognitive skills than in those with more developed skills.

Comprehension and cognitive skills

Language comprehension is intimately related to working memory capacity (Hansson, Forsberg, Löfqvist, Maki-Torkko, & Sahlén, 2004; Sahlén, Reuterskiöld-Wagner, Nettelbladt, & Radeborg, 1999), which is one of the aspects of cognition that we focus on in the present study. The study takes its departure in one of several current working memory theories, namely in Just and Carpenter's (1992) *Capacity Theory of Comprehension*. According to this model, working memory is a unitary system responsible for the simultaneous processing and storing of information over a short period of time. Information storage and information processing thus share the same mental resources, and the partitioning of resources for different functions differ according to the character of a task. For instance, in dealing with a grammatically complex piece of information, resources may mainly be allocated for processing of the verbal content and the resources remaining for storing are diminished. Working memory capacity differs between individuals. According to Just and Carpenter (ibid), this is why individuals differ regarding the speed and accuracy with which they can process and understand language. For instance, individuals with more limited working memory capacity often have difficulties with the processing of complex and/or long sentences (Bishop, 1997). In this study, cognitive capacity is assessed by a measure of general working memory capacity, tapping in to the competing processing and recall of verbal content (the Competing Language Processing Test, Gaulin & Campbell, 1994; Swedish version, Pohjanen

& Sandberg, 1999) and a measure of executive functioning: Elithorn Mazes (EM) (in Wechsler, 2004).

Comprehension and external factors

Apart from the child's linguistic and cognitive capacity, a range of external factors influence language comprehension in children. Our knowledge is extremely sparse regarding the association between speakers' or examiners' voice quality and children's comprehension of language. Only a couple of studies so far have explored this relationship. Morton and Watson (2001), Rogerson and Dodd (2005) and a not peer reviewed, magazine article by Morsomme, Minell, & Verduyck, (2011) are to our knowledge the only studies within this area. In all three studies groups of children were presented to instructions or texts phrased by a female voice, either typical or hoarse (dysphonic). The studies all found adverse effects of the dysphonic voice on the included children's performance on different kinds of tasks (recall of words; drawing inferences (Morton & Watson, 2001) language comprehension tasks (Rogerson & Dodd, 2005; Morsomme, Minell and Verduyck, 2011) and discrimination tasks (Morsomme, Minell and Verduyck, 2011). In the study by Rogerson and Dodd (2005), the children were asked for their opinion about the speaker. The authors found that the children experienced even a slightly dysphonic voice as "not nice".

None of the aforementioned studies controlled for task difficulty. Alterations in speakers' speech and voice quality may have different impacts depending on the complexity of the task the child is asked to perform. In Haake et al. (2013), a facilitating effect of slow speech rate on comprehension appeared in more difficult, not already well mastered, tasks. The effect appeared more specifically in longer and more grammatically complex sentences. The combination of a speaker's deviant voice quality and increasing difficulty of tasks is therefore of interest to explore.

In summary, so far very little attention has been paid to the child's perception of a speaker's voice quality and to its possible consequences for comprehension and learning. The verbal content communicated by a dysphonic voice due to its deviant acoustic content may cause the child to allocate too much cognitive capacity to the processing of the voice signal, thus reducing the capacity to process the actual content of the message. The child's working memory capacity and executive functioning may be more important for performance in language comprehension tasks when the acoustic content of the examiner's voice is deviant. The novelty in the present study lies in the exploration of the combined effects of external factors (the speaker's voice quality and task difficulty) and internal (cognitive) factors in the child on a language comprehension task.

Purpose

Our overall aim is to study if and how a speaker's voice quality (normal vs dysphonic) influences language comprehension in 8-year-old children with typical language development as defined by parents and teachers.

Our specific research questions are: Is there an association between speakers' voice quality and the result on the TROG-2? Is there an association between performance on cognitive tasks (working memory capacity and executive functioning) and language comprehension measured as the performance on the TROG-2? We predict an association between voice quality and performance on the TROG-2, increasing with the degree of difficulty of test items. Further, we predict a stronger association between performance in the TROG-2 and cognitive tests in children presented to the dysphonic speaker's voice than in children presented to the typical voice.

Method

Design

Two groups of children were tested with a digital version of the TROG-2 (Bishop, 2003; Swedish version, 2009) using test instructions recorded with two different types of voice quality (typical voice and dysphonic voice) from the same female speaker. One group was presented with the typical voice quality and the other group with the dysphonic voice quality. The children were also tested for aspects of cognitive capacity using Elithorn's Mazes (EM, Wechsler, 2004) and The Competing Language Processing Task (CLPT, Gaulin & Campbell, 1994; Swedish version, Pohjanen & Sandberg, 1999). The data collection for the present study was performed by master students for a Master's thesis in Logopedics (in Swedish) at Lund University (Anderberg, Johnell & Halvardsson, 2012).

Participants

The participating children (8-9 years old in the 2nd grade of the Swedish school system) were recruited through a randomized selection of grade schools. Written information along with consent forms were distributed to the parents by the children's teachers. Of 250 distributed forms, 102 were returned, 94 with acceptance to participate (92%). The form also comprised relevant factors that might influence the results: multilingualism, hearing disorders/impairments, and any former or ongoing contact with an SLP. This information was considered necessary for the control of the distribution of the children within the two groups and also for possible exclusion (see below). During the test period, four children were reported ill and thus a total of 90 children participated in the tests. Before the testing started, each child was assigned to one of the two groups: every second child to the group presented with the typical voice and every second child to the group presented with the dysphonic voice.

To remain included, the participants had to understand and proceed with the test procedures for all three tests (EM, CLPT, and TROG-2) and perform above the lower minimum value of the normative data for TROG-2 (Johansson & Rutgersson, 2011). On the basis of this, a total of four children were removed due to reported test procedure problems in combination with: (i) hearing disorders (n = 2, one from each group), (ii) language problems (n = 1, typical voice group), and (iii) ongoing SLP contact (n = 1, typical voice group). The final data set included 86 children, see table 1 (that also presents distributions for voice condition, age, gender, multilingualism, and previous SLP contact).

The final data set included 86 children: 43 girls/43 boys mean age range 7:10--9:1 (y:m), see table 1 (that also presents distributions for voice condition, age, gender, multilingualism, and previous SLP contact).

Additionally four children ended up with test procedure problems resulting in single extreme data points: three children scoring zero on EM and one child making 16 self-corrections on TROG-2 (cf. descriptive statistics in table 2). No additional problems or annotated deviances were present for these four children and the corresponding extreme data points were treated as single data points with no available data (n/a).

[table 1 about here]

Procedure

Before the TROG-2 testing, all children were tested with EM (Wechsler, 2004) and CLPT (Gaulin & Campbell, 1994; Swedish version, Pohjanen & Sandberg, 1999). Three test leaders administrated the tests to the children who participated individually and in separate rooms. The test leaders were SLP students at their final semester of the SLP program (semester 8/8). All three had fulfilled their clinical training. The tests were presented to the children in the same sequence: EM, CLPT and TROG-2. Two test leaders administrated TROG-2 and one test leader EM and CLPT. The TROG-2 was installed on two lap top

computers and the children listened to the sentences via the computer's integrated loudspeakers at the level of the child's own choice. The test procedure was thoroughly checked for fidelity through comparisons of repeated video recordings of the three testers to secure similarity of procedure and processing of the tests. The children judged their subjective experience of the test situation and of the teacher's voice after the completion of TROG-2. The complete testing procedure lasted for 45 min.

Instruments

Digitalized version of the TROG-2

The TROG-2 is a picture selection test consisting of 80 sentences (Bishop, 2003). The child listens to a sentence and selects the picture (one out of four pictures) that corresponds to the sentence. The 80 sentences are organized into 20 blocks (A--T) with four items in each. Each block assesses one specific grammatical construction and the degree of difficulty increases with each block. The degree of difficulty depends on different factors, i.e., the grammatical construction involved, the length of the sentence, the semantic content and also on how the distracting pictures relate to the target picture. Thus, the response mirrors all levels of sentence processing; auditory, syntactic, semantic and, in this case, also visual processing.

With the digitalized version of the test it is possible to analyse additional quantitative aspects such as self-correction and response latency. For this study, the children's self-corrections were chosen for analysis. Normative data for response accuracy are available for a Swedish child population (Johansson & Rutgersson, 2011).

The child was presented with each sentence only once and was informed that he/she had unlimited time to deliver an answer and could change the response once. After listening to the sentence, the child indicated the chosen picture with the computer's mouse. The latest indicated picture was interpreted as the final answer. For each item within the block a correct answer received a score of 1 and an incorrect answer received a score of 0. This means that

the maximum score for each block was 4 and 80 for all items of the TROG-2. The result was computed on the total of correct responses and the sum was later transformed to percent for the purpose of the analysis.

TROG-2 Sentences: Recordings and Assessment of Voice Quality

The TROG-2 sentences were read aloud by one of the authors, female, 46 years, and recorded on two separate occasions, first recording with typical voice quality and second recording with simulated dysphonic quality. The recordings were made in an anechoic chamber. The sentences were recorded using the freeware application *Audacity* (version 1.3.12-beta) and a Neumann U87 microphone hanging from the centre of the ceiling at 44.1 kHz/16 bit sampling frequency. The sentences were later segmented and normalized for sound pressure level (SPL) using *Praat* (freeware version 5.2.46).

Three experienced SLPs assessed the dysphonic voice through the sound player in their own PCs, wearing headphones. The recordings of three randomly selected sentences were assessed with a 10 cm Visual Analogue Scale (0=no deviance, 10= max. deviance) using a voice assessment protocol ad modum SWEA (Hammarberg, 2000). They were allowed to listen to the recordings as many times as needed and in any sequence. The SLPs agreed on the prevalence of the parameters hyperfunction (M : 6.1 cm); vocal fry (M : 5.6 cm); reduced sonority (M : 2.9 cm) and hard glottal onsets (M : 1.6 cm). The mean for grade of voice disorder was 3.9 cm which is considered a mild to moderately deviant voice according to clinical praxis. The typical voice was not analysed since the SLPs considered it without deviations out of the typical range.

Assessment of Working Memory and Executive Functioning

The CLPT is a test used for assessment of simultaneous processing and storing of information, complex working memory capacity (Gaulin & Campell, 1994; Swedish version, Pohjanen & Sandberg, 1999). The CLPT is a child version of the Reading Span Test,

which was originally developed within the theoretical working memory framework “A capacity theory of comprehension” (Just & Carpenter, 1992). In the CLPT, the participant is first asked to judge the semantic acceptability of sentences (processing) and thereafter to repeat the final words of each sentence (storing). The test consists of 42 sentences in blocks of one to six sentences. Each sentence is orally presented to the child once. One point is given per correctly recalled word (maximum 42 points). The final words in a block can be repeated in any order (i.e., free recall).

EM is a subtest from the Wechsler Intelligence Scale for Children–A Processing Instrument, Fourth Edition (WISC–IV), (Wechsler, 2004) and can also be administered independently of the WHISC-IV (ibid). The EM aims at measuring executive functions such as organization, processing, planning and inhibition skills in children between 8-16 years. It is a timed pen and paper test and the objective is to successfully trace a path in ten different mazes with increasing complexity. Scores are given for speed and correctness with time bonus (EM) and without time bonus (EMN). The EMN turned out to have a lower explanatory value and was excluded accordingly in this study.

Statistics and ethical considerations

The statistical power has been calculated and with an assumed TROG-2 block score difference between two groups and a SD of 2.5 statistical power of 95% is reached with 34 subjects in each group. According to Altman (1991) power is defined as the probability of avoiding a Type II error and is suggested to be at least about 80%.

All analyses were conducted in R v3.0 (R Core Team, 2013). Independent student's *t*-tests were used for comparisons of TROG-2 block scores between voice conditions. Non-parametric Wilcoxon's rank-sum tests were used for analyses of grammatical difficulty and self-corrections showing non-normal data distributions.

To analyse the relation between TROG-2 block scores, EM, CLPT & TROG-2 self-corrections on block basis, a correlation matrix was calculated followed by a multiple linear regression analysis with AIC using stepwise regression (forward and backward).

The alpha level for all statistical analyses was set to .05.

Ethical Considerations

The information to parents and school staff and the testing of the children were performed in accordance with the Helsinki Declaration. The study was approved by the local committee for ethical approval.

Results

All analyses were conducted in R v3.0 (R Core Team, 2013).

Age, Gender, SLP contact and Multilingualism

As seen in table 1, the distributions for age, gender, previous contact with SLP, and multilingualism were similar between the two voice condition groups.

The EM and the CLPT

Table 2 presents descriptive statistics for executive functioning (EM) and working memory capacity (CLPT) separated on the two voice conditions (typical/dysphonic) for EM, and CLPT. Furthermore, the EM and CLPT distributions were tested positively for normality and homogeneity of variance and no significant difference was found between voice conditions using Student's *t*-tests for independent samples: CLPT ($t(84) = -1.11, p = 0.2$), EM ($t(81) = 1.93, p = 0.057$). A complementary nonparametric Wilcoxon's rank-sum test being more robust towards outliers supported the non-significant result for EM ($WS = 1033, p = 0.111$).

[table 2 about here]

TROG-2 Test Scores

The TROG-2 test scores were analysed on block basis as this aligns with the underlying construction and intended use of the TROG-2 test. The TROG-2 test scores (per participant) revealed skewed distributions together with some outliers for the typical voice condition (figure 1). Evaluations of the block scores on the two voice conditions displayed only a small difference between the means and Student's *t*-test for two independent samples displayed no significant effect. A complementary estimation using trimmed means (to compensate for outliers) and Mann-Whitney U test (not assuming normality and being more robust towards outliers) showed similar results, see table 3.

[figure 1 about here]

[table 3 about here]

TROG-2 Scores and Grammatical Difficulty

To pursue the investigation, the TROG-2 test scores (as percentage of participants with correct answers to adjust for differences in group size) were plotted for each block A to T, separated on the two voice conditions (figure 2). In this plot, the x-axis (blocks A-T) reflects the overall increase of the degree of difficulty in coherence with the construction of the TROG-2 test. According to the TROG-2 score plots in figure 1, block T exhibits floor effects; this task was evidently far too difficult for all children in both groups. Furthermore, block M seems to reveal a pronounced “reversed” effect with the group exposed to the dysphonic voice answering more correctly. A closer investigation showed that this effect can be explained by one specific item (M4), where there were ambiguities in the test.

[figure 2 about here]

A closer look at figure 2 suggests that the dysphonic voice condition affects the TROG 2 block results more negatively than the typical voice condition for the last five blocks (block O-S; block T disregarded due to floor effects).

The TROG-2 scores for the last five blocks (blocks O-S: block T excluded due to floor effects) was performed using a nonparametric test (Mann-Whitney's U test) as the distribution did not apply to normality. The results displayed a significant effect between the two voice conditions ($W_S = 1182.5$, $p = 0.018$), in that the dysphonic voice condition yielded significantly lower block scores on blocks O-S.

TROG-2 Self-corrections and task difficulty

Self-corrections on TROG-2 were assessed on block basis (one or more self-corrections within a block counted as one self-correction for that block), in coherence with the calculation of TROG-2 performance scores. Relating to the analysis of TROG-2 test scores regarding task difficulty, a corresponding evaluation of self-corrections on degree of difficulty gives a partly reversed picture. Figure 3 suggests that the dysphonic voice produces more self-corrections for the first two thirds of the blocks (block A-M) – whereas for the last third (block N-T) no such effect seems to be present. From figure 3 it is also evident that the amount of self-corrections increases in both groups with the degree of task difficulty.

[figure 3 about here]

A Mann-Whitney's U test for the first two thirds of blocks (block A-M) on the two voice conditions displayed a significant effect ($W_S = 46$, $p = 0.049$), whereas the upper third of blocks (block N-T) showed no significant effect ($W_S = 21.5$, $p = 0.748$).

TROG-2 Block Scores and Self-Corrections in Relation to EM and CLPT

The data was further analysed with regard to executive functioning (EM) and working memory capacity (CLPT) data.

Correlation. A correlation analysis using Pearson's r was performed on EM, CLPT, TROG-2 block scores and TROG-2 block wise self-corrections (table 4). In this analysis, data from both groups were used.

[table 4 about here]

The effect sizes of the correlations were interpreted using Cohen's (1988) conventions: *small* $\pm 0.1-0.3$; *moderate* $\pm 0.3-0.5$; *large* $\pm 0.5-1.0$. Thus, for EM, the correlation with TROG-2 block scores was moderate and the correlation with TROG-2 block wise self-corrections was negative and small. For CLPT, all the correlations were small and only marginally significant. Finally, for TROG-2 block scores, there was a small negative correlation with self-corrections.

These results suggest that decreasing results on the EM test, decreasing results on the TROG-2 block scores, and an increasing amount of self-corrections correlate to each other with approximately the same small to moderate effect size (Pearson's r around 0.3).

Regression analysis. A multilevel regression model with TROG-2 block scores predicted by EM, CLPT and self-corrections was established using stepwise back- and forward elimination with AIC (table 5).

[table 5 about here]

As already indicated by the correlation results (table 4), using ANOVA with Type III SS to evaluate the model (table 5) displayed a significant effect for EM on TROG-2 block scores [$F(1,78) = 5.07$, $p = 0.027$], whereas CLPT and self-corrections only showed non-significant trends, [CLPT: $F(1,78) = 2.71$, $p = 0.104$; self-corrections: $F(1,78) = 2.15$, $p = 0.15$].

In these analyses, the voice variable was excluded since it did not contribute to the model. Complementary evaluations adding the voice variable to the model did not reveal any significant main effects or interaction effects.

Summary of results

1. There was no significant overall difference between the voice conditions on the TROG-2 results.
2. The children in the group presented with the dysphonic voice scored significantly lower on the later and more difficult TROG-2 tasks (blocks O-S; block T excluded due to floor effects);
3. The children in the group presented with the dysphonic voice made significantly more self-corrections on the simpler TROG-2 tasks (blocks A-M).
4. The percentage of children making self-corrections increased throughout the test for both groups, although the difference between the voice conditions was significant only for the simpler TROG-2 tasks (blocks A-M).
5. There was a significant positive correlation of moderate effect size between the EM results and the number of correct answers on TROG-2.
6. There was a significant negative correlation of small (approaching moderate) effect size between the EM results and the number of self-corrections.
7. EM, CLPT, and self-corrections explained about 13% of the variance encountered for the TROG-2 block scores with EM as the dominant factor describing almost half of these 13%.

Discussion

In this study we wanted to investigate the relationship between the examiner's voice quality, the children's performance in a digital version of a test of language comprehension, the Test for Reception of Grammar (TROG-2, Bishop 2003, Sw. version 2009) and two measures of cognitive functioning. We found the dysphonic voice affected the children's performance, and that the performance varied depending on the difficulty of the task, only when the sentences were presented with a dysphonic voice. The children in the group that was

presented with the dysphonic voice performed significantly poorer on the more difficult items/blocks compared to the performance in the group that was presented with the typical voice. As for self-corrections, the frequency was significantly higher in the simpler blocks/items when the test sentences were presented with the dysphonic voice.

Bishop (1997) demonstrated that there is a close connection between working memory capacity and language comprehension in children. In this study we wanted to explore the role of cognitive functions (working memory capacity and executive functions) for language comprehension, as measured with the TROG-2. The results for the cognitive function tests in this study showed that executive functions were significantly associated to performance on the language comprehension test (TROG-2). Our expectations that the cognitive tests would predict the outcome on the TROG-2 more strongly when instructions were given by a dysphonic voice than with a normal voice, were not met for the entire test, only for the more difficult tasks. Therefore, the interpretation of the results is executive functions are of great importance for exclusion of irrelevant information or inhibition of disturbing aspects of the acoustic signal from a dysphonic voice when the difficulty of the task increase. The perceptual demands in listening to a dysphonic voice may be compared to the demands of listening in noise (Ljung et al. 2009; Ljung, Israelsson & Hygge, 2013).

The difference in occurrence of self-corrections for the simpler items between the groups is puzzling if self-corrections are interpreted as indications of uncertainty. Why did the frequency of self-corrections not differ more between voice conditions for the more difficult tasks? One possible explanation is that the children adapt to the dysphonic voice, i.e: learn to “normalize” the acoustic signal along the test in line with the concept of normalization (Palmeri, Goldinger and Pisoni, 1993).

Palmeri, Goldinger and Pisoni (1993) discuss *the concept of normalization*. This concept describes a process of perception making it possible for a listener to understand a

message unrelated to the source's signal and thus unrelated to the speaker, since the listener learns to normalize for acoustical variations. The findings of Palmeri et al. (1993) show that the normalization process does not discard the voice signal, instead the voice signal seems to be represented in long term memory along with the phonetic representations. This process challenge the listener's executive functioning (Ibid). In relation to our results, this could mean that the children presented with the dysphonic voice normalize the dysphonic characteristic during the testing to the point where the difficulty of the task "takes over" and the gives an incorrect response.

However, the registration of self-corrections needs some consideration. Our purpose was to capture processing aspects of performance during the comprehension task. We tracked self-corrections unrelated to the response being changed into the correct or incorrect answer. There might, of course, be different explanations to these behaviours, like poor motivation, poor inhibition, true uncertainty etc. A differentiation of different types of self-corrections is beyond the scope of this study and will be dealt with in a future study, which more thoroughly will explore different aspects of processing of tasks (response times and effectiveness in choices) in the language comprehension task.

Future studies are needed on the effects of a dysphonic teacher's voice on the child's everyday comprehension and learning in a school setting with poor acoustics and background noise and also with children of different age. The results from Creel and Jimenez (2012) show that children might get increasingly more skilled in processing dysphonic voices over the years. One of the factors that develops is the "voice lexicon": the memory representations of different voices. Small children, naturally, have not been exposed to a large variety of voices, and thus, there is a greater need for new memory representations to be constructed while listening to a dysphonic voice, demanding allocation of cognitive capacity to this process. It is also plausible that the "voice lexicon" develops over time (ibid). In such case, the acceptance

of a dysphonic voice increases and the older child can better handle the demands caused by the voice of a dysphonic teacher.

Present research in relation to previous studies

Our results align largely with the previous studies (Morton & Watson, 2001; Rogerson & Dodd, 2005, Morsomme et al., 2011) regarding the effect of dysphonic voice on language comprehension. There is; however, a range of methodological factors that can explain why results in the small number of previous studies (ibid) seem more clear-cut than in our study regarding the influence of a dysphonic voice on language comprehension

Materials and presentation of the recorded voices. In the present study, both voices were presented to the children individually, through the computer's loudspeakers, with 2-7 words per sentence. The child had to match the sentence to one out of four pictures. Hence, the children had a limited choice of pictures and visual support, which was not the case in the previous studies. Task demands were quite different in the other studies where the children were tested for their ability to recall words and to draw inferences from a series of short passages. The latter task may be more challenging for different reasons. The visual cues given in the TROG-2 test may require less working memory capacity than listening and drawing inferences from a text passage read aloud without visual support to the child, and also to children of a different age. This might contribute to the stronger effect of the dysphonic voice on performance in the previous studies. Further, the groups of children in our study were only exposed to one voice condition, whereas in the other studies, all children were exposed to two or three voice conditions, possibly reducing the possibility of normalization. Only in the study by Rogerson and Dodd (2005) the material was presented via videorecordings, thus also providing the children with visual cues of the speaker. The dysphonic voices in Rogerson and Dodd (2005) were simulated as in the present study. Rogerson and Dodd (2005) compared the

effect of three levels of dysphonia (mildly-, moderately- and severely dysphonic) and concluded that even a simulated mildly dysphonic voice influenced the children's performance.

The selection of the included children. The selection of children differed in all studies. In the present study the children were checked for a range of factors (SLP contact; hearing disorders; multilingualism) and thorough testing of working memory capacity and executive functioning was performed. The overall effects of the dysphonic voice found in the previous studies might be explained by the fact that important factors were not controlled for, apart for the exclusion of children with hearing impairments. Morton and Watson (2001) aimed at controlling the included children's cognitive level by making a selection of 24 top-students. Without a proper testing the included children's cognitive skills would remain unknown.

Furthermore, sociocultural factors may explain differences between our results and the results in previous studies. It is well known that individuals with deviant resonance or dysphonic voices are sometimes met by condescending attitudes from listeners (Lallh & Rochet, 2000; Verduyckt, 2013), attitudes that even are hard to sway with information about the underlying disorder (Lallh & Rochet, 2000). A negative attitude towards a certain voice may influence task performance. Little is yet known about if and how different cultures differ regarding attitudes towards dysphonic voices.

In the study by Morton and Watson (2001) the children expressed a clear dislike for the dysphonic voice. According to Lallh and Rochet (2000) a good executive functioning is needed to shut out the feelings that might be caused by a dysphonic voice.

As we wanted to explore the included children's likes or dislikes of the two voice conditions and the test situation, we developed a version of a "Smileys-test". In Swedish schools this instrument is commonly used for facilitating the younger children's self-

assessment of aspects of well-being and relations to staff and friends. The instrument is a straightforward test where the child is supposed to draw mouths on smileys-icons where only the eyes are pre- indicated. However, this part of the test-procedure was not included in the fidelity check and the responses to the questions were unclear. We therefore chose to exclude the test from further analysis. For coming research projects however, we want to improve and validate this kind of test.

Methodological Considerations and Future Directions

For the present study, we recorded the same female speaker for both voice conditions. This can be considered a strength since linguistic differences between speakers might constitute a confounder. However, one drawback is that the recordings were made on separate occasions. Notations were made in the written test instructions regarding emphatic stress and pausing but no systematic comparison was made for such factors. For future studies the recordings need to be controlled for such aspects. It is important to point out that there is hardly any research on, for example, the effects on prosody and emphatic stress by a dysphonic voice. Further, all testing was performed at the children's schools. This brought the advantage of making the children feel as they normally would. The disadvantage was the difference as to the sound environment, both regarding the specific room and surrounding distractors. The TROG-2 sentences were presented to the children via the computer's built-in loudspeakers, which may have affected the children's perception. The use of headphones would have been a better choice to exclude surrounding, and maybe distracting, noise sources. However, the children seemed to enjoy participation and no distractions were observed during the test sessions.

To account for normalization, we could have shifted back and forth between voice qualities within the test. From validity point of view we decided to use TROG-2 according to

practice, since we wanted to simulate a natural test situation where the same voice is used all throughout the test.

The children were consecutively assigned to the two groups before commencing the test-procedure. It was found that the groups did not differ significantly for age, gender or the previously annotated factors. Thus the groups turned out to be similar, although not matched beforehand. In future studies the statistical strength would increase if the matching is based on the distribution of deviant factors. To be able to further explore the relation between cognitive factors and language comprehension, children with different levels of cognitive capacity should be included.

The dysphonic voice in the present study was assessed and controlled by three experienced SLPs and judged as a “moderately dysphonic voice” according to clinical praxis. Voice judgements commonly rely on perceptual evaluations which might be considered a problem. Based on the results by e.g. (Holmberg, Hillman, Hammarberg, Södersten, & Doyle, 2001) skilled clinicians rely on “inner standards” of a healthy voice. Further, Holmberg et al. (2001) found a high agreement among her assessors on what could be considered a healthy voice and what was judged to be pathological. In the present study the assessment of the dysphonic voice relies on the inner standards of the expert panel and, hence, we conclude that the assessment is to be considered ecologically valid.

Clinical implications and implications for school settings

The current research, although being just a first step, has relevance, not only for clinical settings but also for school settings. The clinical assessment of language comprehension should be related to the purpose of the assessment. For SLPs, the purpose is often twofold: to diagnose, for example, language impairment and to generate implications for therapy. To reach the first purpose, a standardized procedure must be used; however, this may

strongly underestimate the child's capacity to reach a certain behaviour in less de-contextualized settings than the standardized test setting. For the second purpose, dynamic assessment is needed. There is always a risk that standardized and dynamic test administration overlap. The TROG-2 should be administered in a similar way by all clinicians, otherwise normative data cannot be used. Our study specifically showed that the result on the TROG-2 can vary depending on the examiner's voice. Children may underachieve if an examiner speaks with dysfunctional voice. The use of a digital version of the TROG-2 for clinical use would also increase the reliability of repeated testing. It would ensure that it is the one and same voice that occurs at different test occasions which is of importance based on the results by Pisoni (1997), who showed that comprehension is increased through familiarity of the speaker's voice.

School settings and teachers' voices. The sound environment in today's classroom is often noisy. Several experimental studies, although using adult participants, show that the ability to listen in noise is related to individual cognitive abilities interacting with sensory information (Akeroyd, 2008); to language development and language processing skills within the child (Redey-Nagy, 2009) and to reading and writing performance (Ibertsson, 2010; Savage, Cornish, Manly, & Hollis, 2006). Further, poor classroom acoustics and background noise may negatively influence comprehension (Kjellberg, Ljung, & Hallman, 2008; Ljung, Sörqvist, Kjellberg, & Green, 2009; Strukelj, Holmberg, Lindström, Mossberg, Brännström, & Holmqvist, 2012). The question is whether a speaker or teacher with a dysphonic voice is another source of distracting noise in the classroom, adding to the noise load!

The present study also has relevance for teachers in the classroom and teacher education: Classrooms of today seem to be built for the techniques of yesterday (Greenland & Shield, 2011; Shield & Dockrell, 2008). Teaching methods have developed dramatically

during the last decades, from strictly lectern teaching (listening to one speaker) to group based work and interactive teaching and learning (listening and interacting in the presence of multiple talkers). Further, the use of computers in learning increases and computer fans also add to the background noise. In all, this development challenges the sound design in school buildings (Bako-Biro, Kochhar, Clements-Croome, Awbi, & Williams, 2007) and the teachers' possibilities to successfully communicate in the classroom (Lyberg Åhlander et al., 2011). Poor classroom acoustics have also been shown to contribute to teachers' vocal load and voice problems (Lyberg Åhlander et al., 2010; Pelegrín García, Lyberg Åhlander, Rydell, Löfqvist, & Brunskog, 2010). This makes the present study relevant for SLPs working with voice disorders. Furthermore, the teacher's voice and communication may have an even greater importance in e. g., children struggling with language comprehension, such as many multilingual children, children with language impairment (LI), hearing impairment (HI) and attention deficit/hyperactivity disorder (ADHD). Further studies are required.

Hence, teachers' dysphonic voices are a potential problem in the classroom and in teaching (Morton & Watson, 2001; Yiu, 2002). As a consequence of a disordered voice, a dysphonic teacher's struggle to make her voice function properly can be misinterpreted negatively and influence the child-teacher relationship. There is a need for implementation of courses in the teacher education programs, teaching future teachers voice ergonomics, care of the professional voice and communication skills.

Conclusion

The present findings showed that the dysphonic voice affected the children's performance. The dysphonic voice yielded significantly lower scores in more difficult items (blocks O-S) and also more self-corrections on the simpler items (blocks A-M). For all

children together there were associations between the cognitive tests (executive functioning) and the comprehension test (TROG-2, both for accuracy and self-corrections). Our findings indicate that a dysphonic speaker's voice may force the child to allocate too much capacity to the processing of the voice signal at the expense of comprehension, particularly when the child is approaching her/his limits for mastering a comprehension task.

Acknowledgements

The authors gratefully acknowledge the master students, Linnéa Anderberg, Anna Johnell and Emilia Lucena Halvardsson, who collected the data; the Linneaus' environment Cognition, Communication and Learning at Lund University for financial support; Professor Agneta Gulz for valuable comments on the proof, and SLP Susanna Whitling for proof reading and style review. We also like to express our gratitude to the referees for very valuable comments on the manuscript.

References

- Akeroyd, M. A. (2008). Are individual differences in speech reception related to individual differences in individual differences in cognitive ability? A survey of twenty experimental studies with normal and hearing-impaired adults. *International Journal Audiology, 47*, 53-71.
- Altman D.G. 1991. *Practical statistics for medical research*. London: Chapman & Hall.
- Anderberg, L., Johnell, A., & Halvardsson, E. L. (2012). *Om sambandet mellan logopedens röstskvalitet, språkförståelse och kognitiv förmåga hos barnet. (Master's thesis). [The relationship between the Speech-Language Pathologist's quality of voice and the language comprehension and cognitive capacities in the child]* Retrieved from Lund University's student theses database:
<http://www.lunduniversity.lu.se/o.o.i.s?id=24965&postid=3233200>. Swedish
- Bako-Biro, Z., Kochhar, N., Clements-Croome, D., Awbi, H., & Williams, M. (2007). Ventilation rates in schools and learning performance. In *Proceedings of Clima 2007 "WellBeing Indoors"*. Helsinki, Finland. Retrieved from
<http://www.irbnet.de/daten/iconda/CIB7218.pdf>
- Bishop, D. (1997). *Uncommon understanding. Development and disorders of language comprehension in children*. East Sussex: Psychology Press Ltd.
- Bishop, D. (2003). Test for reception of grammar, version 2 (TROG-2). London, UK: Pearson Assessment.
- Bishop, D. (2009). Test for reception of grammar, version 2 (TROG-2). Swedish version. London, UK: Pearson Assessment.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

- Creel, S. C., & Jimenez, S. R. (2012). Differences in talker recognition by preschoolers and adults. *Journal of Experimental Child Psychology*, *113*, 487-509.
- Gaulin, C., & Campbell, T. (1994). Procedure for assessing verbal working memory in normal schoolage children: Some preliminary data. *Perceptual and Motor Skills*, *79*, 55-64.
- Greenland, E. E., & Shield, B. M. (2011). A survey of acoustic conditions in semi-open plan classrooms in the United Kingdom. *The Journal of the Acoustical Society of America*, *13*, 1399-1410.
- Haake, M., Hansson, K., Gulz, A., Schötz, S., & Sahlén, B. (2013). The slower the better? Does the speaker's speech rate influence children's performance on a language comprehension test? *Int J Speech Lang Pathol*. Early online:00:0, 1-10. Retrieved from:
<http://informahealthcare.com/action/showCitFormats?doi=10.3109%2F17549507.2013.845690>
- Hammarberg, B. (2000). Voice Research and Clinical Needs. *Folia phoniatica et logopaedica*, *52*, 93-102.
- Hansson, K., Forsberg, J., Lofqvist, A., Maki-Torkko, E., & Sahlén, B. (2004). Working memory and novel word learning in children with hearing impairment and children with specific language impairment. *Int J Lang Commun Disord*, *39*, 401-422.
- Holmberg, E. B., Hillman, R. E., Hammarberg, B., Södersten, M., & Doyle, P. (2001). Efficacy of a behaviorally based voice therapy protocol for vocal nodules. *J Voice*, *15*, 395-412.
- Ibertsson, T. (2010). *Cognition and communication in children/adolescents with cochlear implant* (Master's thesis). Retrieved from Lund University's student theses database:
<http://www.lunduniversity.lu.se/o.o.i.s?id=24965&postid=1488278>

- Johansson, H. G., & Rutgersson, K. (2011). Svensk normering av TROG-2 för åldrarna 8-10 år och sambandet mellan grammatisk språkförståelse och ordavkodning. [Swedish TROG-2 normative data for 8-10 year old children and the association between grammatical language understanding and word decoding] (Master's thesis). Retrieved from Gothenburg University Publications Electronic Archive (GUPEA):
https://gupea.ub.gu.se/bitstream/2077/26900/1/gupea_2077_26900_1.pdf
- Just, M. A., & Carpenter, P. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99, 122-149.
- Kjellberg, A., Ljung, R., & Hallman, D. (2008). Recall of words heard in noise. *Applied Cognitive Psychology*, 22, 1088-1098.
- Lallh, A. K., & Rochet, A. P. (2000). The effect of information on listeners' attitudes toward speakers with voice or resonance disorders. *Journal of Speech, Language, and Hearing Research*, 43, 782-795.
- Ljung, R., Sörqvist, P., Kjellberg, A., & Green, A.-M. (2009). Poor listening conditions impair memory for intelligible lectures: Implications for acoustic classroom standards. *Building Acoustic*, 16, 257-265.
- Ljung, R., Israelsson, K., Hygge, S. (2013). Speech Intelligibility and Recall of Spoken Material Heard at Different Signal-to-noise Ratios and the Role Played by Working Memory Capacity: The difference in memory & speech perception. *Applied Cognitive Psychology*, 27(2), 198.
- Lyberg Åhlander, V., Rydell, R., & Löfqvist, A. (2010). Speaker's comfort in teaching environments: Voice problems in Swedish teaching staff. *Journal of Voice*, 25(4), 430-440.

- Lyberg Åhlander, V., Rydell, R., & Löfqvist, A. (2011). How do teachers with self-reported voice problems differ from their peers with self-reported voice health? *Journal of Voice*, 26(4), 149-161.
- Morsomme D, Minell L, & Verduyck I. (2011). Impact of teachers' voice quality on children's clanguage processing skills *Logopedie* 9-15.
- Morton, V., & Watson, D. (1998). The teaching voice: Problems and perceptions. *Logopedics Phoniatrics Vocology*, 23, 133-139.
- Morton, V., & Watson, D. (2001). The impact of impaired vocal quality on children's ability to process spoken language. *Logopedics Phoniatrics Vocology*, 26, 17-25.
- Palmeri, T. J., Goldinger, S. D., & Pisoni, D. B. (1993). Episodic encoding of voice attributes and recognition memory for spoken words. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 309-328.
- Pelegrín García, D., Lyberg Åhlander, V., Rydell, R., Löfqvist, A., & Brunskog, J. (2010). *Influence of classroom acoustics on the voice levels of teachers with and without voice problems: a field study*. In *Proceedings of Meetings on Acoustics*, 11, 060001. Retrieved from <http://asadl.org/poma/resource/1/pmarcw/v11>
- Pisoni, D. B. (1997). Some thoughts on "normalization" in speech perception. In: K. Johnson, J. W. Mullennix (Eds.), *Talker Variability in Speech Processing* (pp. 9-32). San Diego, CA: Academic Press.
- Pohjanen, A., & Sandberg, M. (1999). *Arbetsminnet hos svenska fem-, sju- och nioåriga barn med normal språkutveckling. [Working memory capacity in Swedish five-, seven- and nine year-olds with typical language development]* (Master's thesis). Retrieved from Lund University's student theses database: <http://www.lunduniversity.lu.se/o.o.i.s?id=24965&postid=2968822>

- R Core Team (2013). R: A language and environment for statistical computing [Computer Software]. R Foundation for Statistical Computing, Vienna, Austria. Available from <http://www.R-project.org/>
- Redey-Nagy C. (2009). *Speech perception in noise, phonemic categorization and auditory attention in specific language impairment* (Doctoral thesis, University College, London, UK).
- Rogerson, J., & Dodd, B. (2005). Is there an effect of dysphonic teachers' voices on children's processing of spoken language? *Journal of Voice*, *19*, 47-60.
- Sahlén, B., Reuterskiöld-Wagner, C., Nettelbladt, U., & Radeborg, K. (1999). Non-word repetition in children with language impairment--pitfalls and possibilities. *Int J Lang Commun Disord*, *34*, 337-352.
- Savage, R., Cornish, K., Manly, T., & Hollis, C. (2006). Cognitive processes in children's reading and attention: The role of working memory, divided attention, and response inhibition. *British Journal of Psychology*, *97*, 365-385.
- Shield, B. M., & Dockrell, J. E. (2008). The effects of environmental and classroom noise on the academic attainments of primary school children. *Journal of Acoustical Society of America*, *123*, 133-144.
- Strukelj, A., Holmberg, N., Lindström, P., Mossberg, F., Brännström, K. J., & Holmqvist, K. (2012). Text comprehension during noise exposure: Effects on eye movements, galvanic skin responses and subjective performance. In *The Scandinavian Workshop on Applied Eye Tracking (SWAET 2012)*. Extended abstract retrieved from <http://www.eyelab.se/swaet/program/presentations?p=15>
- Verduyckt, I. (2013). *Perception de la personnalité du locuteur dysphonique sur base de sa qualité vocale [the perception of the speaker based on voice quality]*. (PhD Doctoral thesis), Université catholique de Lovain, Lovain, Belgium.

Wechsler, D. (2004). WISC-IV Integrated. Wechsler Intelligence Scale for Children – Fourth Edition – Integrated [Manual]. London, UK: Pearson Assessment.

Yiu, E. M. L. (2002). Impact and prevention of voice problems in the teaching profession – embracing the consumers' view. *Journal of Voice*, 16, 215-229.

Figure captions

Figure 1. Left: TROG-2 test results for total block score per participant separated on voice conditions; right: corresponding boxplot with means indicated by a solid circle.

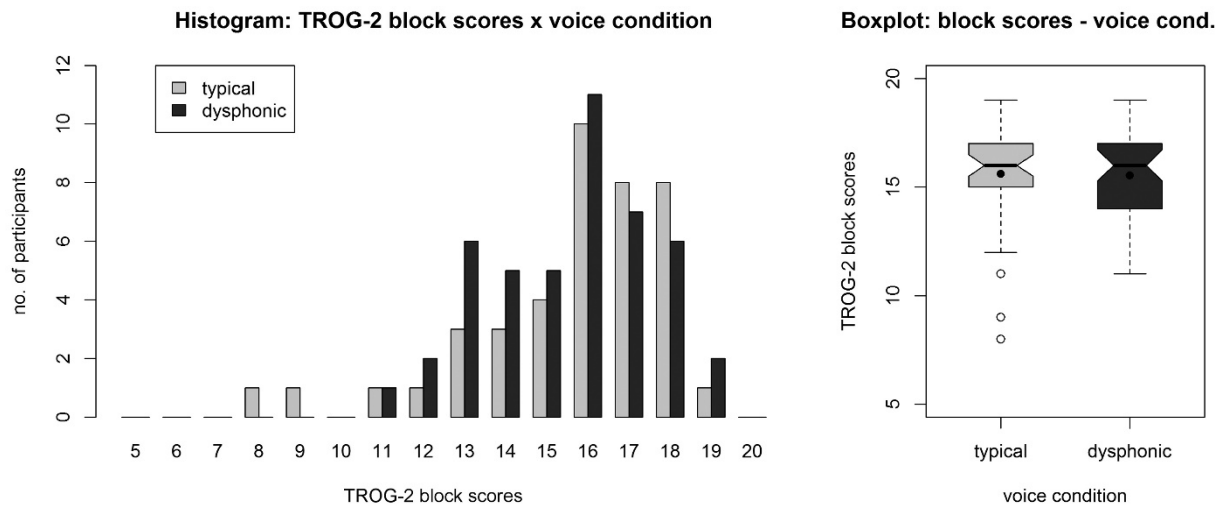


Figure 2. Distribution of $n= 86$ children’s TROG-2 block results (as per cent of participants correct) for the two speech quality conditions (typical and dysphonic voice). Dashed lines present lowess (locally weighted) smoothing curves.

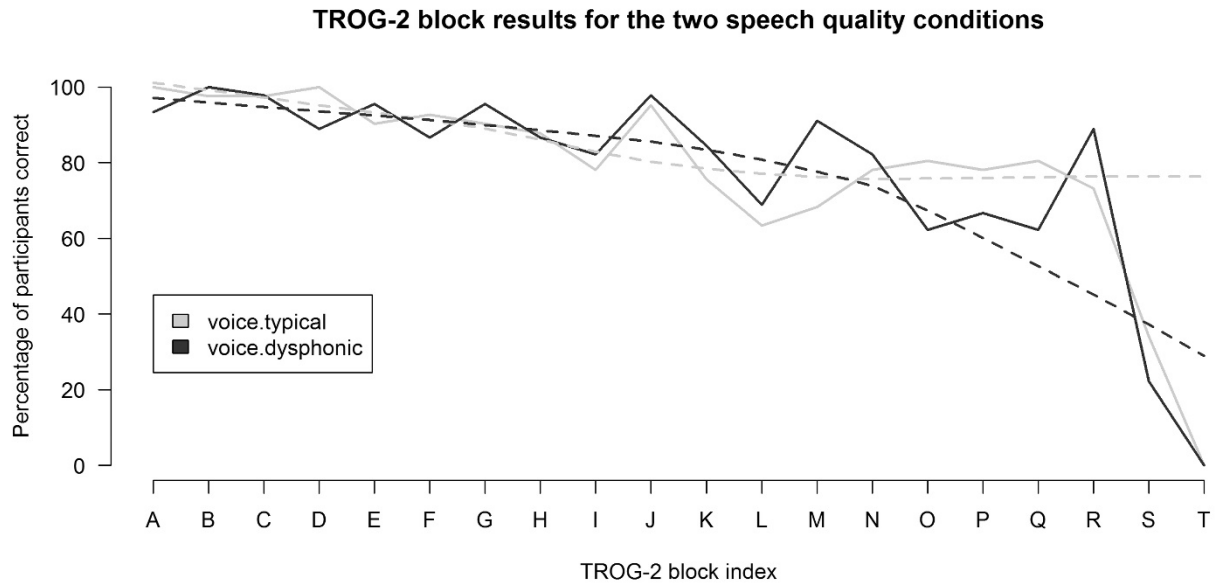


Figure 3. TROG 2 blockwise self-corrections (as percentage of participants making corrections) for TROG-2 block A-T. Dashed lines present lowest (locally weighted) smoothing curves.

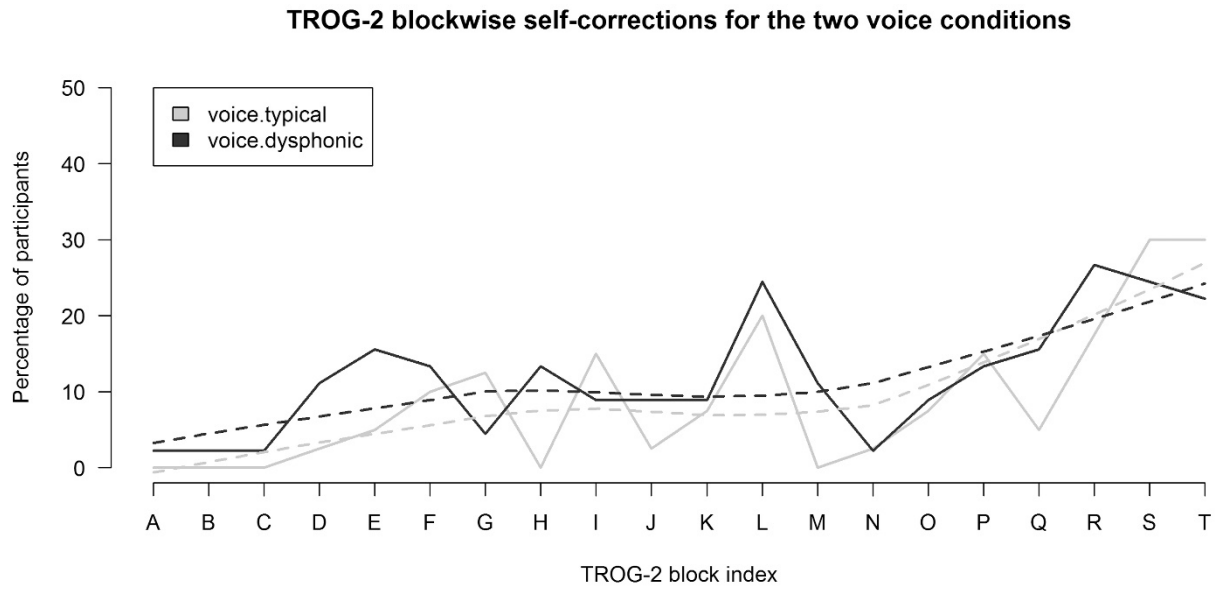


Table 1

Distribution of age, gender, multilingualism and previous contact with speech-language pathologists (SLP) for n = 86 children in two groups exposed to one of two conditions: "Typical voice" and "Dysphonic voice".

	Typical voice	Dysphonic voice
Test subjects (n)	41	45
Age: Mean (Range)	8.7 (7;10-9;1)	8.6 (8;1-9;0)
Girls/Boys	21/20	22/23
Multilingualism	6	8
Prev. contact SLP	7	6

Table 2

Test results for n = 86 children separated regarding the two voice quality conditions (typical voice = “typ”, dysphonic voice = “dys”): executive functioning (EM), working memory capacity (CLPT), TROG-2 block scores (block), and TROG-2 blockwise self-corrections (scrt).

Variable	n	Min	Max	Median	Mean	SD
EM(typ / dys) ¹	39 / 44	7 / 5	39 / 32	20 / 18.5	21.4 / 18.5	7.59 / 6.13
CLPT(typ/dys)	41 / 45	16 / 15	32 / 36	25 / 25	24.3 / 25.2	4.17 / 3.79
block:(typ/dys)	41 / 45	8 / 11	19 / 19	16 / 16	15.6 / 15.5	2.45 / 2.00
scrt(typ/dys) ²	40 / 45	0 / 0	7 / 7	1 / 2	1.73 / 2.16	1.74 / 1.83

Note:

¹ EM: three occurrences of extreme data points (see section “Participants” under “Method”).

² scrt: one occurrence of an extreme data point (see section “Participants” under “Method”).

Table 3

Descriptive statistics for TROG-2 block scores for the two voice conditions (typical and dysphonic voice): means (M), trimmed means excluding 5 % of highest and lowest scores (M), and standard deviation (SD), together with comparisons between the two voice conditions by means of Student's t-test (t(p)) and Mann-Whitney's U test (WS(p)).*

	<i>M</i>	<i>M*</i>	<i>SD</i>	<i>t (p)</i>	<i>WS (p)</i>
typical	15.6	15.8	2.45		
				0.159 (0.874)	998 (0.511)
dysphonic	15.5	15.6	2.00		

Table 4

Correlation matrix presenting Pearson's product-moment correlation coefficient (r) for executive functioning (EM), working memory capacity (CLPT), TROG-2 block scores (block), and TROG-2 blockwise self-corrections (s crt) in n=86 children.

	CLPT	block	s crt
EM	0.17	0.32 **	-0.28 *
CLPT		0.20 .	-0.20 .
block			-0.28 **

$p < 0.1$ * $p < 0.05$ ** $p < 0.001$

Table 5

Regression model for TROG-2 block scores (block) depending on executive functioning (EM), working memory capacity (CLPT) and TROG-2 blockwise self-corrections (scrt).

Model (AIC)	adj. R^2	model statistics	relative importance¹
block ~ em + clpt + scrt	13.4%	$F(3,78) = 5.20, p < 0.01$	EM: 46%, CLPT: 27%, scrt: 27%

¹ Relative importance by function calc.relimp (method LMG), R-package relaimpo v2.2 (Grömping, 2006).