Mapping the effect of body position: Voice quality differences in connected speech

Helena Engström[1], Marcin Włodarczak[1], Sten Ternström[2]
[1] Department of Linguistics, Stockholm University, Sweden
[2] Division of Speech, Music and Hearing, KTH Royal Institute of Technology, Sweden
[1]heen5613@student.su.se, [1]marcin.wlodarczak@ling.su.se, [2]stern@kth.se

Abstract

This work investigates the effect of body position on voice quality based on cepstral peak prominence (CPP) and spectrum balance (SB) metrics layered on a mapped speech range profile (SRP) across a sound pressure level (SPL) and fundamental frequency (f_o) plane. Eight participants were tested in an upright position, supine position at 0º and an inverted position at -10º. Findings show varied and small changes in voice quality in connected speech between positions and that effects may occur at specific SPL and f_o ranges among some participants.

Introduction

Body position has a marked impact on speech physiology due to changes in gravitational influence. Previous findings in respiratory kinematics indicate that the respiratory system is susceptible to changes in body position (Hoit, 1995). In an upright position the abdominal muscles are activated to counter the downward pull of the diaphragm and remain active during speech. In a supine position the abdominal muscles are inactive since the abdomen is held in place by gravitational load. This increases the mobility of the abdomen in a supine position, causing a difference in respiratory kinematics compared to an upright position. Additional electromyographic (EMG) findings in quiet breathing research by De Troyer (1983) showed that the upper abdomen was activated among some participants while placed in an inverted position at -45º, also suggesting a difference in respiratory kinematics between a supine and an inverted position.

The effect of body position on voice quality is less known. An abductive effect on the vocal folds at high lung volumes found in an upright position (Iwarsson et al., 1998) suggests that the voice source is affected by gravity due to tracheal pull. Tracheal pull entails that the lungs are pulled downwards in an upright position, causing a lengthening of the tracheal tube and widening of the glottis. Since tracheal pull occurs when the lungs are pulled in a downwards direction, the abductive effect on the vocal folds is likely to disappear in a supine or inverted position. Furthermore, in a magnetic resonance imaging (MRI) study by Traser et al. (2021) the larynx was found to be more cranially displaced during phonation among classically trained singers placed in a supine position, suggesting a loss of tracheal pull.

Speech data is typically procured from participants positioned upright, however, considering that current MRI technology predominantly requires participants to be placed in a supine position, it is possible that results may not be comparable between studies. Ternström and Pabon (2022) presented voice maps as a tool for comparative analysis of the individual voice. Based on synchronously obtained voice- and electroglottographic (EGG) signals mapped across a sound pressure level (SPL) and fundamental frequency (f_o) plane, voice maps can be generated and compared within the FonaDyn software (Ternström et al., 2018). Each map cell is one semitone wide and one decibel high and contains its own metric averages based on phonatory cycles observed at the corresponding SPL and f_o. The comparative function of voice maps, proposed as a way for assessing voice disorders before and after treatment, may also be of benefit for analysing voice quality differences between body positions by mapping a speech range profile (SRP).

As part of a master's thesis project by the first author, which aims to investigate the effects of body position on voice quality and respiratory kinematics using a mapped SRP, the present work highlights cepstral peak prominence (CPP) and spectrum balance (SB) findings between an upright position, a supine position at 0º and an inverted position at -10º.

Material and methods

Participants

Eight participants in total were tested. Three men, one non-binary person and four women within the age range of 24 to 35 years. All participants reported no pre-existing health risks that posed a problem with being tested in an inverted position. No compensation was offered.

Environment and equipment

The sessions were held at the Phonetics laboratory at Stockholm University.

A MAXXUS Gravity Pro inversion table modified with an angle chart was used to ensure that the supine and inverted position was comparable between participants. A head-mounted Sennheiser HSP2 microphone was used to record the voice signal and SPL was calibrated with a Galaxy Audio CM C200 calibrator. The EGG signal was acquired using a Glottal Enterprises EG2-PCX2 electroglottograph with two dual-channel electrodes. Additionally, respiratory inductance plethysmography (RIP) was used to record respiratory kinematics through two transducer bands connected to a RespTrack 2.4 processor (Heldner et al., 2019). The RIP signals were traced in a MORDAX DATA oscilloscope during testing. Calibration and usage of the RespTrack system is described in Heldner et al. (2019) and similar application of acquiring and analysing RIP signals across the same body positions used in the present work is described in Engström (2023). All signals were recorded with the FonaDyn software connected through an Expert Sleepers ES-9 audio interface. EGG and microphone signals were acquired at a 44100 Hz sampling frequency and RIP signals were acquired at a 100 Hz sampling frequency. The speech material consisted of three standard texts in Swedish that were modified to exclude any instances of voiced fricatives since these speech sounds affect the SB metric.

Experimental procedure

Each participant was recorded at separate occasions with one experimenter present (the first author). The participant signed a consent form and provided information about age, gender and past vocal training experience by answering a voluntary information questionnaire. Rib cage circumference, abdominal circumference and height was measured to obtain a body type estimate. Information about height was also necessary for adjustments of the inversion table. The participant was equipped with the microphone, electrodes and transducer bands. One transducer band was placed at axillary height and the other with the umbilicus as reference. To prevent the transducer bands from slipping between positions, the participant's upper garment was taped down onto the trousers. The microphone was SPL calibrated by following *Scenario A* inside the FonaDyn environment and placed at a 30 cm distance from the participant's mouth. The participant was instructed through the calibration process of the transducer bands (See Heldner et al., 2019) and had an initial vocal warm-up by familiarizing themselves with the speech material.

The body positions were always tested in the same order: The upright position, the supine position at 0º and lastly the inverted position at -10º. The electrode placement indicator on the electroglottograph was used to ensure that the electrodes were at level with the larynx in each position by having the participant phonate [aː]. The recording sequence of each position consisted of calibration of the transducer bands and phonation of an initial [aː] followed by reading aloud the three texts.

Data processing and analysis

A Praat script (Version 6.4.04; Boersma & Weenink, 2024) was used to extract the speech segments from each body position and mapped together with the EGG signal in the FonaDyn software. The voice threshold was set at ≥ 0.9 to exclude overly aperiodic EGG cycles and de-noise was set at 0.02 to filter out interfering EGG noise. Mapping the EGG and voice signal resulted in three csv files containing smoothed SRP maps for each participant: An upright map, a supine map and an inverted map. The maps were filtered to obtain representative values, similarly to Patel and Ternström (2021). In the present work, the cycle threshold for comparative analysis of the SRP maps was set at the median of the total number of cycles per cell. This means that the 50% of the most visited map cells were used for analysis. The lowest median value was used across each comparison. The three comparisons consisted of the upright and supine comparison (U_S), the upright and inverted comparison (U_I) and the supine and inverted comparison (S_I). The comparison of the CPP and SB metrics between positions was obtained by creating a difference map calculated by the formula:

$$\Delta MAP = MAP2 - MAP1 \qquad (1)$$

The SB map layer available in FonaDyn measures the difference in acoustic energy above 2000 Hz and below 1500 Hz, corresponding to increased vocal effort when SB is less negative. The CPP map layer indicate less periodicity in the acoustic signal when CPP is decreased, corresponding to increased breathiness.

Only the participants that show a marked change in ΔCPP and ΔSB distributions between positions are included in the results in the present work. Participants that are excluded show ΔCPP and ΔSB distributions that are close to zero between positions.

Results

Figures 1-4 visualize the data distribution of each comparison for selected participants.

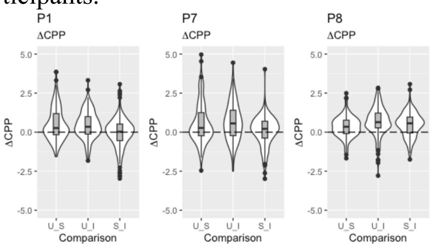

Figure 1. ΔCPP distribution of participants P1, P7, P8 in comparisons U_S, U_I, S_I.

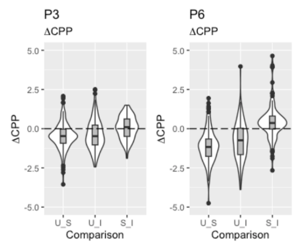

Figure 2. ΔCPP distribution of participants P3 and P6 in comparisons U_S, U_I and S_I.

As seen in Figure 1, the ΔCPP distribution in comparisons U_S and U_1 show predominantly small positive changes for participants P1, P7 and P8. CPP increased in the supine and the inverted position compared to the upright position. The S_I comparison show a less systematic pattern. The ΔCPP distribution of P1 and P7 is closer to zero, which indicates less change than P8 where there is a positive change.

In Figure 2, the ΔCPP distribution of participants P3 and P6 in the U_S and U_I comparison show that there is an overall CPP decrease in the supine and inverted positions compared to the upright position. The ΔCPP distribution is centred around zero in the S_I comparison for P3, whereas the ΔCPP distribution of P6 show a positive change from S to I. Participants P3 and P6 also shared similarities in ΔSB distribution, as seen in Figure 3. The U_S comparison show a negative change in the ΔSB distribution for both participants. This indicates that there was less energy at high frequencies in the supine position compared to the upright position.

Among the eight participants, only P8 show a positive change in ΔSB distribution across all comparisons, as seen in Figure 4. This points to more energy at high frequencies in the supine and inverted position compared to the upright position, as well as in the inverted position compared to the supine position.

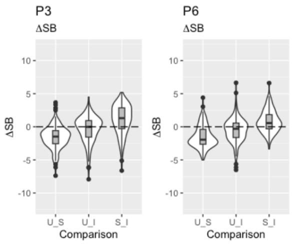

Figure 3. ΔSB distribution of participants P3 and P6 in comparisons U_S, U_I and S_I.

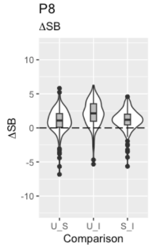

Figure 4. ΔSB distribution of participant P8 in comparisons U_S, U_I and S_I.

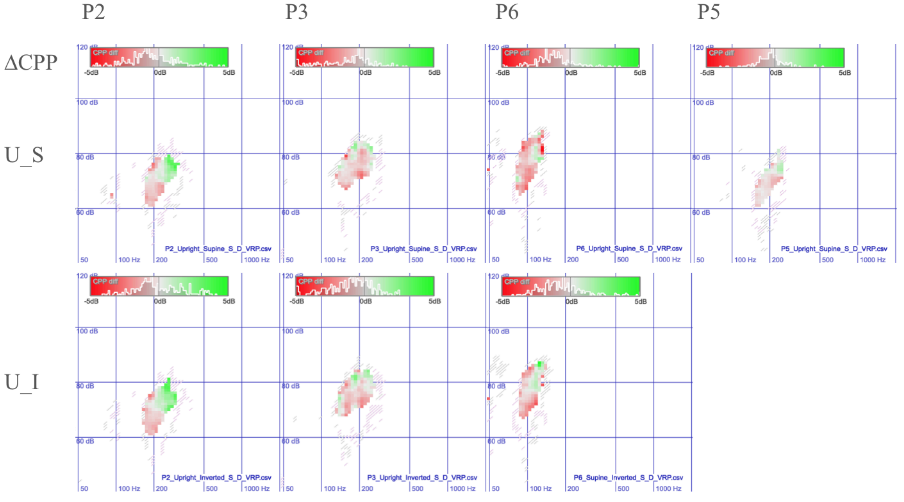

Figure 5. Filtered SRP maps visualizing similar pattern of ΔCPP in the U_S comparison (top row) and U_I comparison (bottom row) for participants P2, P3, P6 and P5.

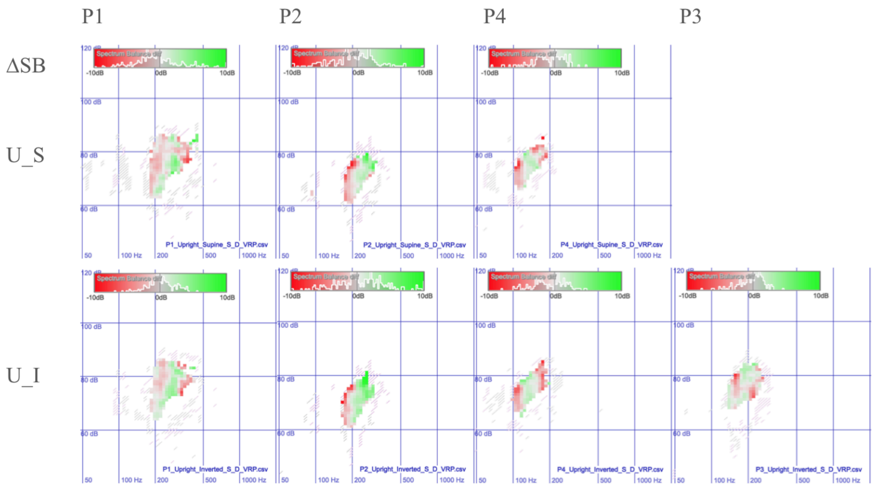

Figure 6. Filtered SRP maps visualizing similar pattern of ΔSB in the U_S comparison (top row) and U_I comparison (bottom row) for participants P1, P2, P4 and P3.

Despite less change in ΔCPP and ΔSB distribution between positions, other participants share similar patterns when visualized as filtered SRP difference maps. Figure 5 show overall a pattern of a negative change in ΔCPP at low SPL ranges and a positive change at high SPL ranges for participants P2, P3, P6 and P5 across U_S and U_I comparisons. As seen in Figure 6, there are overall negative changes in ΔSB at the lowest and highest f_0 ranges as well as a positive change in middle f_0 ranges for participants P1, P2, P4, and P3 across U_S and U_I comparisons.

Discussion

This investigation has shown that body position has influence on voice quality among some participants. However, among the participants that showed the most change in ΔCPP and ΔSB distributions included in the results, the changes were not necessarily similar. P3 and P6 showed an increase in breathiness and P8 showed less breathiness across the CPP and SB metrics in the supine and inverted position compared to the upright position. However, shared systematic patterns across SPL and f_0 ranges were also found among participants P1, P2, P4 and P5 when considering the visual output of the filtered SRP difference maps.

Iwarsson et al. (1998) found an abductive effect on the vocal folds at a high lung volume, however, reading and conversational speech is restricted to the middle lung volume range (Hixon et al., 1973). The present results therefore suggest that the effects of gravity on voice quality also occur at habitual lung volume ranges in speech, reflected as small and varied changes in CPP and SB metrics. However, other unexplored factors may influence these results, such as a potential increase in subglottal pressure caused by the abdominal contents pressing on the chest cavity in a supine and an inverted position.

Additionally, the advantage of using a filtered mapped SRP has been shown in the fact that the similarities observed among the participants when visualized in difference maps across an SPL and f_0 plane would have gone unnoticed if only value distributions had been considered. This indicates that voice mapping could benefit voice quality analyses in speech and phonation as well as being informative of where effects occur based on SPL and f_0 ranges.

Conclusions

This work has shown that differences in voice quality based on changes in body position can be observed among some participants, possibly indicating that the loss of tracheal pull in a supine and an inverted position may have a small impact on speech compared to an upright position. In addition, that future voice quality investigations may benefit from the usage of voice maps for analyses.

Acknowledgements

We thank all participants for their time and effort.

References

Boersma, P., & Weenink D. (2024). *Praat: doing phonetics by computer* (Version 6.4.04) [Computer software]. URL: https://www.fon.hum.uva.nl/praat/

De Troyer, A. (1983). Mechanical role of the abdominal muscles in relation to posture. *Respiration physiology, 53*(3), pp. 341-353. doi: https://doi.org/10.1016/0034-5687(83)90124-X

Engström, H. (2023). *The effect of body position on the relative contribution of the rib cage to speech breathing and voice quality*. [Bachelor's thesis]. Stockholm University. URL: https://www.diva-portal.org/smash/record.jsf?dswid=6218&pid=diva2%3A1768562

Heldner, M., Włodarczak, M., Branderud, P., & Stark, J. (2019). The RespTrack system. *In 1st International Seminar on the Foundations of Speech: BREATHING, PAUSING, AND THE VOICE, Sønderborg, Denmark, 1-3 December, 2019*, pp. 16-18. URL: https://www.diva-portal.org/smash/record.jsf?pid=diva2%3A1467649&dswid=6218

Hixon, T. J., Goldman, M. D., & Mead, J. (1973). Kinematics of the chest wall during speech production: Volume displacements of the rib cage, abdomen, and lung. *Journal of speech and hearing research, 16*(1), 78-115. doi: https://doi.org/10.1044/jshr.1601.78

Hoit, J. D. (1995). Influence of body position on breathing and its implications for the evaluation and treatment of speech and voice disorders. *Journal of Voice, 9*(4), pp. 341-347. doi: https://doi.org/10.1016/S0892-1997(05)80196-1

Iwarsson, J., Thomasson, M., & Sundberg, J. (1998). Effects of lung volume on the glottal voice source. *Journal of voice, 12*(4), pp. 424-433. doi: https://doi.org/10.1016/S0892-1997(98)80051-9

Patel, R. R., & Ternström, S. (2021). Quantitative and qualitative electroglottographic wave shape differences in children and adults using voice map–based analysis. *Journal of Speech, Language, and Hearing Research, 64*(8), 2977-2995. doi: https://doi.org/10.1044/2021_JSLHR-20-00717

Ternström, S., & Pabon, P. (2022). Voice Maps as a Tool for Understanding and Dealing with Variability in the Voice. *Applied Sciences, 12*(22), 11353. doi: https://doi.org/10.3390/app122211353

Ternström, S., Johansson, D., & Selamtzis, A. (2018). FonaDyn—A system for real-time analysis of the electroglottogram, over the voice range. *SoftwareX, 7*, pp. 74-80. doi: https://doi.org/10.1016/j.softx.2018.03.002

Traser, L., Schwab, C., Burk, F., Özen, A. C., Burdumy, M., Bock, M., ... & Echternach, M. (2021). The influence of gravity on respiratory kinematics during phonation measured by dynamic magnetic resonance imaging. *Scientific Reports, 11*(1), 22965. doi: https://doi.org/10.1038/s41598-021-02152-y