Running Head: PNEUMOTACHOGRAPH MASKS MEASUREMENTS COMPARISON

Comparison of Two Pneumotachograph Masks for Sound Pressure Measurements in Speech

Hao Zhang, Xiao Chen and Stephen A. Zahorian

SUNY Binghamton University

Abstract

Ideally, respiratory masks, when used to make measurements of speech and singing, should not interfere with what is being measured. Unfortunately this is not always the case. In this paper, two masks intended for acoustic sound pressure measurements are experimentally compared. One is the circumferentially-vented (CV) pneumotachograph mask manufactured by Glottal Enterprises. The other is an unvented mask manufactured by KayPENTAX. Distortion measures, such as formant frequency shifts, and mask transfer functions, indicate that the unvented mask interferes with speech characteristics much more than does the CV mask. It is hypothesized that sound pressure level measurements are also likely to be adversely affected by the unvented mask.

Comparison of Two Pneumotachograph Masks for Sound Pressure Measurements in Speech

There are two primary types of pneumotachograph masks in use today for measuring the airflow from the mouth and/or nose during speech and singing. One is a circumferentially vented (CV) mask (Rothenberg, 1973 and Rothenberg & Nezelek, 1991), produced by Glottal Enterprises Inc. The second type is the unvented mask like the Phonatory Aerodynamic System (PAS) mask, produced by KayPENTAX Inc. (Stemple, 2008). Another unvented design is (Ghio & Teston, 2004) but the commonly used PAS mask is chosen to represent unvented masks. Figure 1 shows the two types of masks in measuring position.

The CV mask along is evaluated in detail by (Badin, Hertegard & Karlsson , 1990) with artificial sound source. In order to experimentally compare it with the PAS mask, with respect to possible interference with speech, an adult male and adult female, both native speakers of English, made recordings with and without a mask, using a microphone placed 10cm from the corner of the lips. Formant shifts caused by the masks and transfer functions of the two masks were estimated from recorded speech data.

Method

Participants

One adult female and one adult male native English-speaking subject

Apparatus

The subjects spoke into a headset mic (AV-LEADER AVL-605) while wearing attenuating headsets (Sigtronics SH-20 Headphone) to test CV mask (circumferentially vented mask) and PAS mask (Phonatory Aerodynamic System (PAS): Model 6600).

Procedure

One adult female and one adult male native English-speaking subject recorded 10 words of the form "bVd" ("bead," "bid," etc.) whose vowel portions included all nonglided English vowel phonemes (/i/, /ɪ/, /ɛ/, /æ/, /ɒ/, /ɔ/, /ʊ/, /u/, /ʌ/ and /ɜ/). Each subject also uttered a sentence: "Now is the time for all good Americans to come to the aid of their country." The words and sentence were repeated with no mask, with the CV mask and with the PAS mask in place. An AKG 520L omnidirectional microphone was located outside the mask, 10cm left (i.e. 90 degree angle to the subjects' anterior –posterior axis) of the corner of the lips for recording.

Results

Transfer functions for the masks were estimated by taking the ratio of the 1/3 octave-band spectral power with and without the mask in place. Subjects were wearing attenuating earphones during the whole session, in order not to be affected by audible feedback due to different masks. There is some variability in speaking when subjects are wearing different masks or not wearing a mask at all. However, the results of experimental results for different conditions are much larger than would be expected from these variations in speaking. Artificial sources were not used because experiments with human subjects can better show the performance of masks for real-world conditions.

The open source 1/3 octave band analysis algorithm (Couvreur, 1997) was used to estimate the power spectrum for words and complete sentences. Estimated transfer functions based on a complete sentence are shown in Figure 2 (Spectra of female and male speech are averaged.). Qualitatively similar transfer functions were obtained using whole words or vowels.

From Figure 2, it is clear that the PAS mask transfer function has dips (one at about 500 Hz and one at about 1800 Hz) or an overall low-pass filtering effect, whereas the CV mask transfer function is more nearly all-pass. This filtering effect is also illustrated by the three narrowband spectrograms shown in Figure 3. For this example, it is clear that the energy in the 2nd and 3rd formant regions is substantially lowered by the PAS mask (rightmost), but was changed very little by the CV mask. Observations were similar for most vowels. Qualitatively, the speech produced with the PAS mask in place sounded muffled (as per first two authors of this paper) whereas there was no apparent muffling caused by the CV mask.

Praat (Boersma, 2001) was used to track formants of the target word for each recording. Average values of the formats for a 20 ms steady state vowel portion of each word (20 ms portion manually determined) for each of the 10 vowels were computed and compared, as given in Figure 4 and Table 1.

Figure 4 and Table 1, show that the CV mask causes substantially less change of the 2nd and 3rd formant than does the PAS mask. For some vowels, the formant values computed from the of PAS mask speech differ by large amounts from those of the no mask case, especially for the 2nd formant. The reason may be that the corresponding formants are so weakened that the Praat tracker reverts to default values (e.g. Male - F2 - /æ /, /ɛ / and /ɜ /; Female – F3 - / ɜ / and / u /).

Discussion

Sound pressure level (SPL) only makes sense as a measure of the strength of the voice if it is related to the level of the sound at the lips. The Kay-Pentax mask appears to have a complex frequency selective acoustic system between the mouth and the microphone. The SPL measured by the Glottal Enterprises system is the amplitude of the derivative of the volume velocity at the mouth, which is the acoustic pressure at the mouth, modified only by the frequency limitations of the mask system.

In this study, the PAS mask and CV mask were compared as to their interference with speech. It was shown that the PAS mask changes the 2nd formant much more than does the CV mask. From comparing the transfer functions of the two masks, it can be seen that the PAS mask has characteristics of a low-pass filter, whereas the transfer function of the CV mask is more nearly uniform all pass. A preliminary version of this paper was presented as a poster at the 169th Meeting of the Acoustical Society of America.

References

Badin, P., Herteg ård, S., & Karlsson, I. (1990). Notes on the Rothenberg mask. *Speech Transmission Laboratory, Quarterly Progress and Status Report*, *1*, 1–7.

Boersma, P. (2001). Praat, a system for doing phonetics by computer. *Glot International*, *5:9/10*, 341-345.

Couvreur, C. (1998 – last update). Implementation of a one-third-octave filter bank in MATLAB. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi¼10.1.1.57.5728 [October 14, 2016].

Ghio, A., & Teston, B. (2004). Evaluation of the acoustic and aerodynamic constraints of a pneumotachograph for speech and voice studies. *Proceedings of international conference on voice physiology and biomechanics*, 55–58.

Rothenberg, M. (1973). A new inverse-filtering technique for deriving the glottal air flow waveform during voicing. *The Journal of the Acoustical Society of America*, *53.6*, 1632-1645.

Rothenberg, M., & Nezelek, K. (1991). Airflow-based analysis of vocal function. *Vocal Fold Physiology: Acoustic, Perceptual, and Physiological Aspects of Voice Mechanisms*, J. Gauffin and B. Hammerberg, Eds., Singular Publishing Group, San Diego, 139-148,

Stemple, J., Weinrich, B., & Brehm, S.B. (2008). Phonatory aerodynamic system: A clinical manual. KayPENTAX Corp., Lincoln Park, NJ;

Table 1

Means of absolute values of errors (mask versus no mask) for each formant and each speaker, averaged over all vowels (in Hz).

Male						Female					
CV			PAS			CV			PAS		
F1	F2	F3	F1	F2	F3	F1	F2	F3	F1	F2	F3
70	185	253	256	447	505	34	116	207	41	402	277

Figure Captions

Figure 1. CV mask (left) and PAS mask (right)

Figure Caption

Figure 2. Estimated transfer functions of CV mask and PAS mask. Two complete sentences (one spoken by female subject, the other male) were averaged.

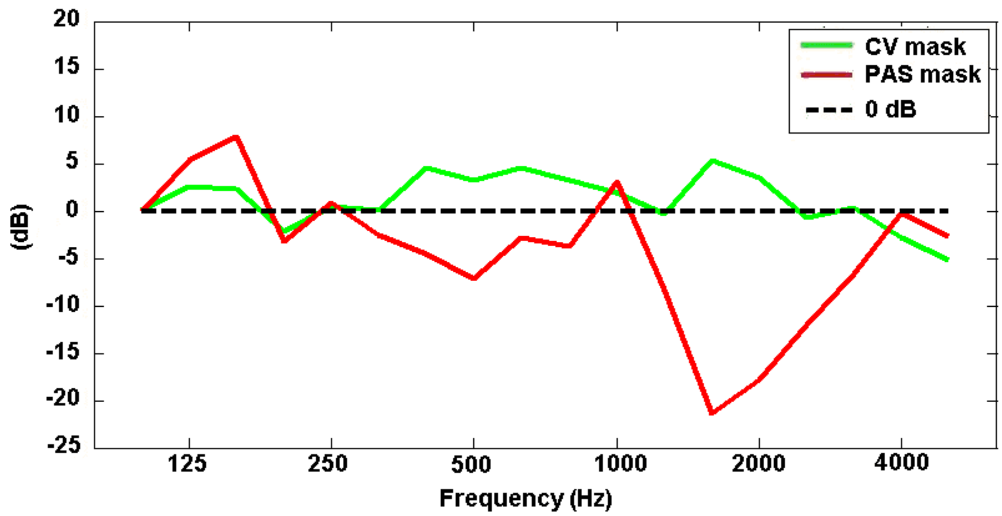

20
15
10
5
0
-5
-10
-15
-20
-25
(dB)
CV mask
PAS mask
0 dB
125
250
500
1000
2000
4000
Frequency (Hz)

Figure Caption

Figure 3. Spectrograms of female speaker for 250ms of “bed” for three cases

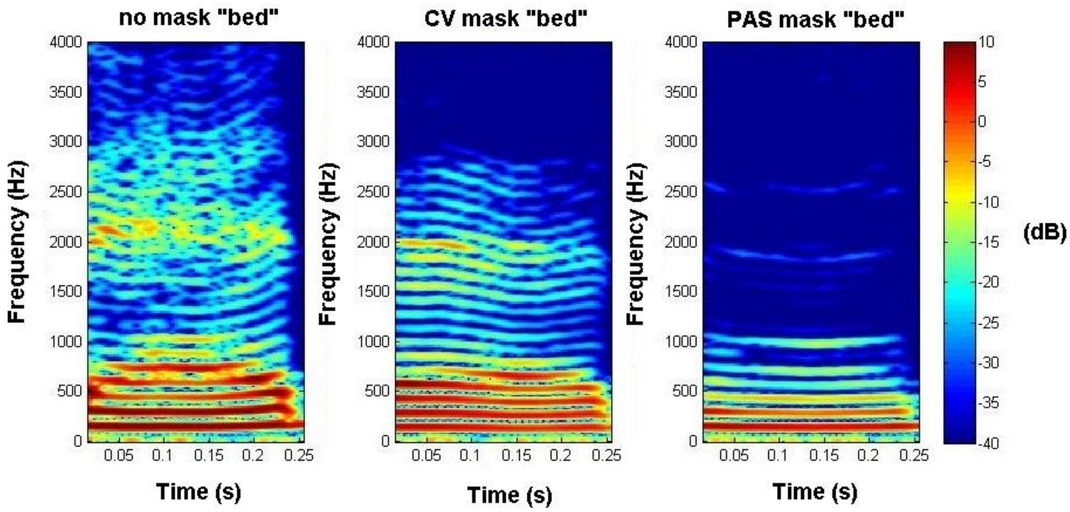
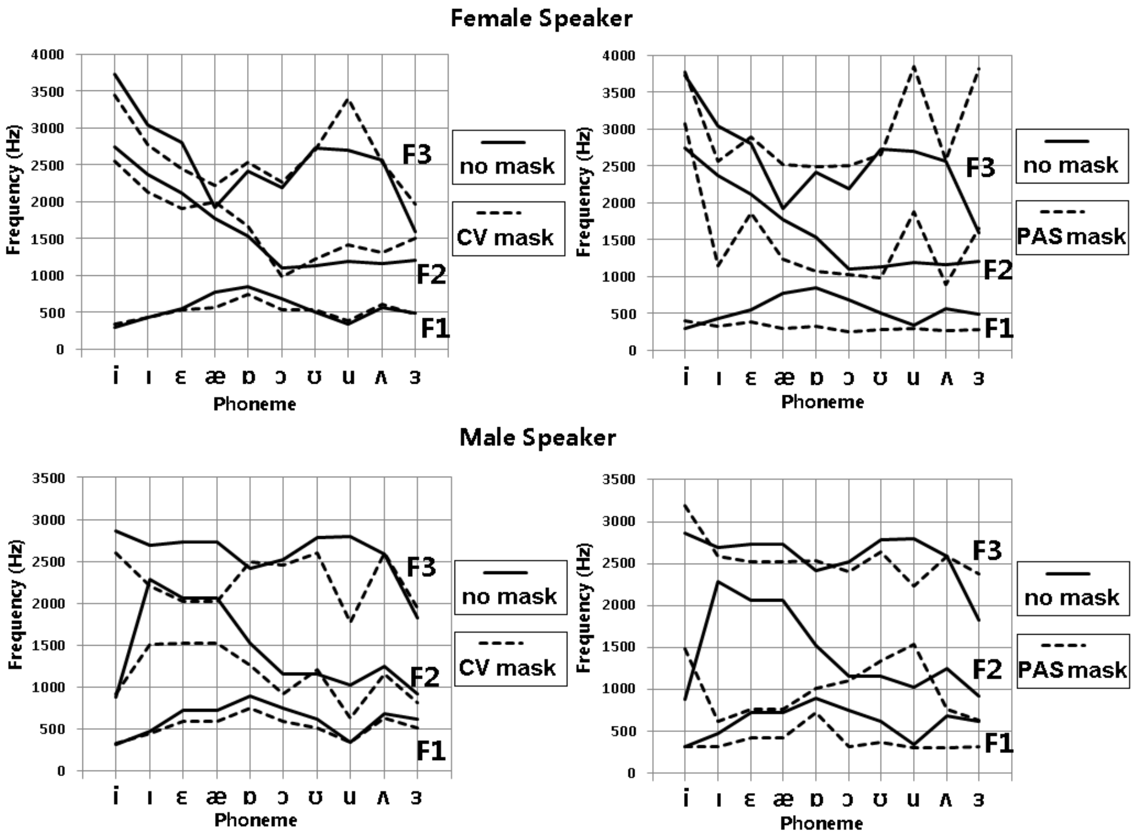

no mask "bed"
CV mask "bed"
PAS mask "bed"
Frequency (Hz)
Time (s)
(dB)
0
500
1000
1500
2000
2500
3000
3500
4000
0.05
0.1
0.15
0.2
0.25
10
5
0
-5
-10
-15
-20
-25
-30
-35
-40

Figure Captions

Figure 4. Female and male speakers' formant values for steady state vowels.


Female Speaker
Male Speaker
Frequency (Hz)
Phoneme
i ɪ ɛ æ ɒ ɔ ʊ u ʌ ɜ
F1
F2
F3
no mask
CV mask
PAS mask