

Reducing the impact of wind noise on cochlear implant processors with two microphones

Kostas Kokkinakis^{a)}

*Speech-Language-Hearing: Sciences and Disorders, University of Kansas,
1000 Sunnyside Avenue, Lawrence, Kansas 66045
kokkinak@ku.edu*

Casey Cox

*Department of Hearing and Speech, University of Kansas Medical Center,
3901 Rainbow Boulevard, Kansas City, Kansas 66160
ccox5@kumc.edu*

Abstract: Behind-the-ear (BTE) processors of cochlear implant (CI) devices offer little to almost no protection from wind noise in most incidence angles. To assess speech intelligibility, eight CI recipients were tested in 3 and 9 m/s wind. Results indicated that speech intelligibility decreased substantially when the wind velocity, and in turn the wind sound pressure level, increased. A two-microphone wind noise suppression strategy was developed. Scores obtained with this strategy indicated substantial gains in speech intelligibility over other conventional noise reduction strategies tested.

© 2014 Acoustical Society of America

PACS numbers: 43.71.Ky, 43.71.Gv [SGS]

Date Received: January 16, 2014 **Date Accepted:** March 27, 2014

1. Introduction

Despite the success of cochlear implants (CIs) in partially restoring auditory sensation to individuals with profound hearing loss, noise induced by wind remains a major source of dissatisfaction for many CI users. Verbal communication in windy environments can become problematic for both the hearing aid (HA) and cochlear implant populations. This has been known to clinicians and engineers for some time (Dillon *et al.*, 1999; Thompson, 2002). For the CI population, wind noise can adversely affect sentence recognition and overall sound quality preference (Chung and McKibben, 2011). Typically, wind noise lowers the signal-to-noise ratio at the microphone of the device significantly, consequently reducing the intelligibility of speech.

Wind noise is the result of turbulent air flow interacting with the diaphragm of the microphone located in the speech processor. This turbulent air flow contains fluctuations in pressure that are converted to electrical fluctuations by the microphones. Wind noise can become a serious encumbrance for CI users because it can: (1) introduce discomfort especially in high intensity levels, (2) overload the microphones in the device, which produces unwanted sound distortions and clipping, and (3) further mask speech signals embedded in wind noise. BTE processors of HA instruments and CI devices typically offer little to almost no protection from wind noise. This is especially true when the angle of incidence is 0°. *In situ* measurements have shown that the amount of turbulent air flow at a 0° angle of incidence will almost always be the greatest in the area concentrated above the pinna (Dillon *et al.*, 1999). In BTE hearing instruments, the area where the intensity of the wind turbulence is typically the greatest is in the vicinity of the microphone ports (Thompson, 2002). Zakis (2011) measured one-third-octave wind noise levels at the front and rear microphones of two BTE devices at wind speeds of 3, 6, and 12 m/s. Those were found to be equal to 76, 97, and 114 dB sound pressure level (SPL),

^{a)} Author to whom correspondence should be addressed.

respectively, at the rear microphone of the BTE device when averaged across azimuths spanning the -135° to 180° range. Wind noise levels obtained at the front microphone of the same BTE device for the same wind speeds were 72, 91, and 108 dB SPL for the identical range of azimuths.

Chung and McKibben (2011) examined the effects of different microphone modes on speech understanding by CI listeners in wind. Flow noise recordings were obtained using a BTE digital HA placed inside a quiet wind tunnel. Wind flow velocities of 4.5 and 13.5 m/s were chosen. The speech stimuli were first recorded in an anechoic chamber and were subsequently mixed with wind flow noise to generate the speech testing materials. The results obtained from sixteen CI listeners indicated that omnidirectional microphones produce less wind noise and will almost always be preferred over directional microphones. For all the CI listeners tested, higher intelligibility scores were observed when relying on the omni-directional microphone mode than when using the directional microphone mode at 13.5 m/s. The scores between the two different microphone modes were not statistically different at 4.5 m/s wind flow speeds.

Chung and McKibben (2011) also found that a substantial reduction in wind noise levels could be achieved by switching from directional to omni-directional processing. The combination of omni-directional microphone processing in the low-frequency region and directional microphone processing in high-frequency bands was shown to be a viable wind noise reduction option (Chung and McKibben, 2011; Chung, 2012). An automatic switching algorithm exists in many commercially available HAs, whereby the device automatically switches to an omni-directional mode upon detection of wind. Wind is detected when the correlation between the two microphones is sufficiently low and the input level is greater than 75 dB SPL. Other alternatives to reduce wind include: (1) adopting the microphone mode with the lowest output and rely on level differences between the two microphones (Chung, 2012) and (2) combining the omni-directional microphones in addition mode (instead of subtraction mode) (Kates, 2008).

The aim of this study is to evaluate the efficiency of conventional noise reduction algorithms in reducing wind noise. This evaluation focuses on processing strategies (e.g., noise programs) currently offered in commercially available cochlear implant processors with two microphones (e.g., Nucleus 5 CP810). A two-microphone wind noise suppression strategy is also described and evaluated with CI users.

2. Strategies

Noise reduction strategies were implemented in MATLAB and verified in SIMULINK using the Nucleus 5 CP810 spatial simulator block. In the Nucleus 5 CP810, these strategies can be configured through commercial fitting software when programming the device.

2.1 ZOOM processing

In a directional design microphones are sensitive toward sounds emanating from the front and attenuate sounds originating outside the desired pickup area (Chung, 2004; Chung *et al.*, 2006). The Nucleus 5 CP810 device uses two closely spaced and matched omnidirectional microphones (front and rear). In the default mode, the Nucleus 5 CP810 relies on a cardioid pickup pattern. This pattern is implemented electronically by delaying the input to the rear omnidirectional microphone and then subtracting it from the input to the front omnidirectional microphone. In the ZOOM setting, the Nucleus 5 CP810 produces a supercardioid polar pattern instead.

2.2 BEAM processing

The BEAM strategy is a two-stage adaptive beamformer. A variant of the BEAM was first introduced as a noise program in Nucleus Freedom processors (Spriet *et al.*, 2007). The BEAM in the Nucleus 5 CP810 relies on a two-microphone system that combines the two outputs from the front and rear omnidirectional microphones. The goal of the BEAM is to focus on signals arriving from in front of the user and to automatically change the position of the null at the sides or behind the user to provide

attenuation for noise signals moving throughout the rear hemisphere. Because of the endfire placement of the microphones, the combination from the two microphones produces a main lobe that has its maximum at 0° . This increases sensitivity to sounds arriving from the front while suppressing sounds originating from the back and the sides.

2.3 2MWS processing

In dual-microphone devices, the spatial correlation of the two inputs can be used as a wind detection statistic. By applying a short-time discrete Fourier transform (DFT) to the inputs, the front microphone signal becomes $X_F(\omega, k)$, and the input to the rear microphone becomes $X_R(\omega, k)$, both expressed in the frequency-domain for frame index k and frequency bin ω . A wind noise detection statistic can be formed by resorting to the magnitude coherence function

$$|\Gamma_{FR}(\omega, k)| = \frac{|\Phi_{FR}(\omega, k)|}{\sqrt{\Phi_{FF}(\omega, k)\Phi_{RR}(\omega, k)}}, \quad (1)$$

where $\Phi_{FR}(\omega, k)$ denotes the cross-power spectral density equal to $\Phi_{FR}(\omega, k) = E[X_F(\omega, k)X_R^*(\omega, k)]$ and $\Phi_{FF}(\omega, k)$ and $\Phi_{RR}(\omega, k)$ are the auto-power spectral densities calculated at each microphone (front and rear) and are equal to $\Phi_{FF}(\omega, k) = E[|X_F(\omega, k)|^2]$ and $\Phi_{RR}(\omega, k) = E[|X_R(\omega, k)|^2]$, respectively. When the signals at the two microphones are uncorrelated (e.g., when wind is present), one can use the coherence function to distinguish between sound pressure caused by turbulence and sound pressure caused by a propagating sound wave. The signal correlation calculated in the time domain has been used previously as a detection statistic for wind noise (Kates, 2008). The coherence assumes values close to 1 if the input signals at the two microphones are correlated and holds values close to 0 if they are strongly uncorrelated (speech corrupted with wind). To suppress regions dominated by wind, we propose the use of the function

$$G(\omega, k) = \frac{1}{1 + \exp[-\gamma(\Gamma_{FR}(\omega, k) - 0.5)]}, \quad (2)$$

where γ is a parameter that controls the slope of the sigmoidal function and thus determines the degree (or aggressiveness) of suppression. Here, $\gamma = 10$ was chosen for optimal performance. The magnitude coherence will assume values in the entire $[0, 1]$ range for 3 m/s wind, while it will hold values much closer to 0 when speech is corrupted with 9 m/s wind noise. The proposed function will attenuate most frequency components when the magnitude of the coherence function approaches 0 (wind is dominant). It will also allow most frequency components to pass without attenuation as coherence approaches 1 (speech is dominant). In the proposed two-microphone wind suppression strategy (2MWS), the signals collected at the two microphones are first processed in 30 ms frames with a Hanning window and with a 50% overlap between successive frames. The auto- and cross-power spectral densities are computed through a first-order smoothing recursion. The actual suppression is performed in the frequency domain by multiplying the Fourier transform of the signal obtained after subtracting the delayed input to the rear microphone from the front microphone in the time domain with the suppression function $G(\omega, k)$. To reconstruct the enhanced signal, we apply an inverse FFT and synthesize with overlap-add.

3. Methods

3.1 Participants

Eight post-lingually deafened adult CI listeners with a fully inserted (long electrode array) implant on one side took part in this study. All eight participants were native speakers of American English and had acquired at least 12 months of experience post-implantation. More detailed demographic information is provided in Table 1. All

Table 1. Participant demographic information.

Subject	Gender	Age at testing	Months of experience with cochlear implant	Etiology of hearing loss	Cochlear implant type	Consonant-nucleus-consonant word-recognition score in quiet (% Correct)
1	M	54	38	Hereditary	Nucleus 5	88
2	M	49	18	Noise exposure	Nucleus 5	92
3	F	22	195	Unknown	Freedom	94
4	F	62	65	Hereditary	Freedom	72
5	F	57	24	Unknown	Nucleus 5	82
6	M	67	20	Noise exposure	Nucleus 5	90
7	M	56	68	Noise exposure	Nucleus 5	75
8	M	37	24	Meniere's	Nucleus 5	81
MEAN		50.5	56.5			84
SD		14.6	59.3			8

participants were paid an hourly wage. This study was approved by the Human Subjects Committee of the University of Kansas in Lawrence (HSCL). All participants gave written informed consent prior to testing, and a case history interview was conducted with each subject to determine eligibility. Only CI users who scored greater than 70% on the consonant-nucleus-consonant (CNC) test were included.

3.2 Speech stimuli

The speech stimuli were sentences from the IEEE database (IEEE, 1969). Two Knowles FG-3329 omni-directional microphones were mounted in a BTE hearing aid shell with approximately 12 mm spacing between the two ports. The dummy shell was placed behind the right ear of a Knowles Electronics Manikin for Acoustic Research (KEMAR). IEEE sentences were presented inside a double-walled IAC sound-attenuating booth from a single-cone loudspeaker (Tannoy Reveal 501 A) at 75 dB SPL. The speaker was at a 0° azimuth and 0.6 m away from the KEMAR. Front and rear digital recordings of the omni-directional microphone signals were made for each IEEE sentence. Stimuli were recorded at the sampling rate of 16 kHz.

3.3 Wind stimuli

The same two Knowles FG-3329 microphones were used to obtain wind noise recordings. The KEMAR was placed outdoors on top of an adapter. On two separate occasions, two wind noise samples with a total duration of 10 s each were recorded outdoors. The first recording was made at a speed of 3 m/s, which corresponds to a rating of 2 (gentle breeze) on the 13-point Beaufort wind force scale. The second recording was performed at a wind speed equal to 9 m/s that corresponds to a rating of 5 (strong breeze) on the 13-point Beaufort wind force scale. Wind speed was measured using a La Crosse Technology EA-3010U handheld anemometer securely placed near the top of the pinna of the KEMAR. The average levels of the recorded wind noise samples summed across all frequency bins were 65 and 80 dB SPL in the 3 and 9 m/s wind conditions, respectively. For each wind flow recording, the two microphones were connected through preamplifiers to a Roland R-26 six-channel, 24-bit portable digital recorder. The inputs were sampled at the rate of 16 kHz. Special attention was given during the recording process to avoid peak clipping. Front and rear digital recordings of the omni-directional microphone signals were made for each of the two wind velocity conditions. The two-channel wind noise recordings were saved as 16-bit WAV files.

3.4 Procedure

To generate the corrupted stimuli, each sentence was mixed with a different random segment of each of the two 10 s wind noise recordings (3 and 9 m/s) recorded at the

front microphone (channel 1) and the rear microphone (channel 2). Subjects participated in a total of five different conditions per wind flow velocity. The conditions tested were the following: (1) OMNI, (2) DIR, (3) ZOOM, (4) BEAM, and (5) 2MWS. The DIR condition was obtained by delaying the input to the rear omni-directional microphone and then subtracting it from the input to the front omni-directional microphone. There were 10 different conditions ($=2$ wind flow velocities $\times 5$ processing strategies). The stimuli were delivered to the CI device via direct audio input, and both microphones remained deactivated. All stimuli were presented at the subject's most comfortable level. Two IEEE lists (20 sentences) were used per condition. None of the lists was repeated across different conditions. To minimize order effects, the order of the test conditions was randomized across subjects. The participants were instructed to type as many of the words as they could identify via a computer keyboard. The responses of each individual were stored in a written sentence transcript and scored off-line based on the number of words correctly identified. All words were scored in each sentence. The percentage correct scores for each condition were calculated by dividing the number of words correctly identified by the total number of words in the particular sentence list.

4. Results

Individual speech intelligibility scores of eight CI listeners are plotted in Fig. 1. The different strategies are shown along the abscissa. All percent correct scores were rationalized arcsine unit-transformed prior to the analyses. A Kolmogorov–Smirnov test confirmed the normality of the transformed percent correct scores. A value equal to 0.05 was used as the significance level. Two-way analysis of variance (ANOVA) (with repeated measures) indicated a significant effect of the wind velocity ($F[1,7]=65.3$, $p < 0.05$), a significant effect of the processing strategy ($F[4,28]=13.01$, $p < 0.05$), and a significant interaction ($F[4,28]=20.05$, $p = 0.01$). *Post hoc* comparisons using Fisher's least significant difference were run to assess significant differences between different

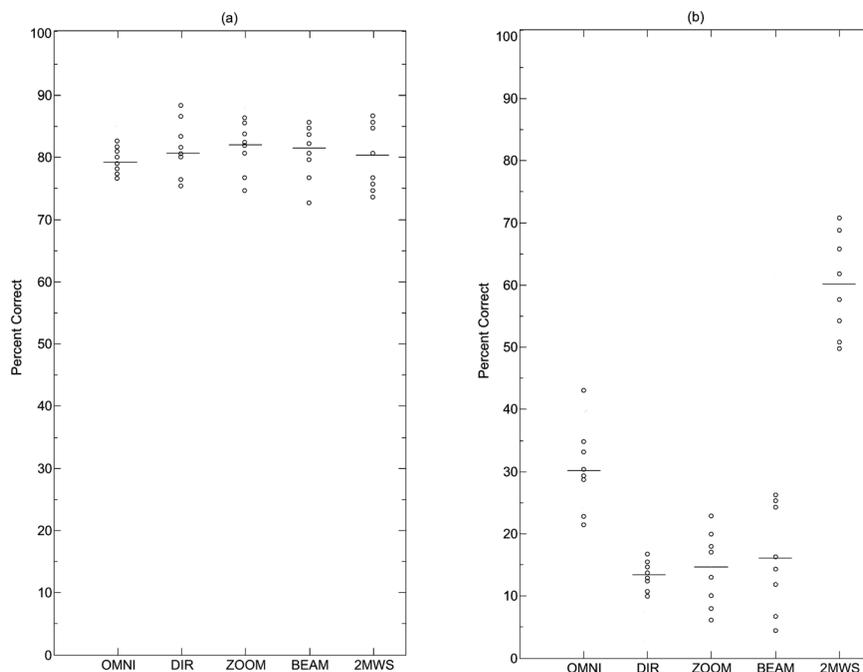


Fig. 1. Individual percent correct scores of eight CI listeners tested on sentences corrupted with (a) 3 m/s wind flow and (b) 9 m/s wind flow. The different processing strategies are shown along the abscissa. Each dot represents the score of a single participant. Mean scores for each condition are plotted as horizontal lines.

conditions. The scores obtained with the front omni-directional microphone (OMNI) did not differ significantly from the scores obtained with the directional microphone (DIR) in the 3 m/s condition ($p=0.724$) but were significantly higher in the 9 m/s wind flow condition ($p<0.05$). The scores obtained with the ZOOM did not differ significantly from the scores obtained with the BEAM in either the 3 m/s ($p=0.319$) or the 9 m/s wind conditions ($p=0.688$). In the 9 m/s condition, the scores obtained with the 2MWS strategy were significantly higher than the scores obtained in the OMNI ($p=0.018$), DIR ($p=0.007$), ZOOM ($p=0.007$), and BEAM ($p=0.026$) processing modes.

5. Summary and Discussion

In 3 m/s wind flow, all conditions tested yielded similar speech recognition scores. In the 9 m/s condition, our findings indicated that a directional microphone pattern is more susceptible to wind noise than an omni-directional pattern. The scores obtained with the ZOOM strategy did not differ significantly from the scores obtained with the BEAM strategy. For both strategies, the average scores in 9 m/s wind were about 15 percentage points lower than the scores obtained in the OMNI condition. Both the ZOOM and BEAM strategies operate under the assumption that the noise sources are strongly correlated (e.g., coherent). Naturally, coherent noise fields occur in cases when noise signals propagate to the microphones without undergoing any form of reflection, dispersion or dissipation and when wind turbulence effects are minimal. Here the wind noise captured by the microphones was not generated through acoustic propagation. Instead when wind flow was present, distinct air vortexes were created or sensed at each microphone of the device. Because each of these vortexes was random and unique, so was the signal at each microphone. As a consequence, the two wind signatures captured by the omni-directional microphones were spatially uncorrelated (e.g.,

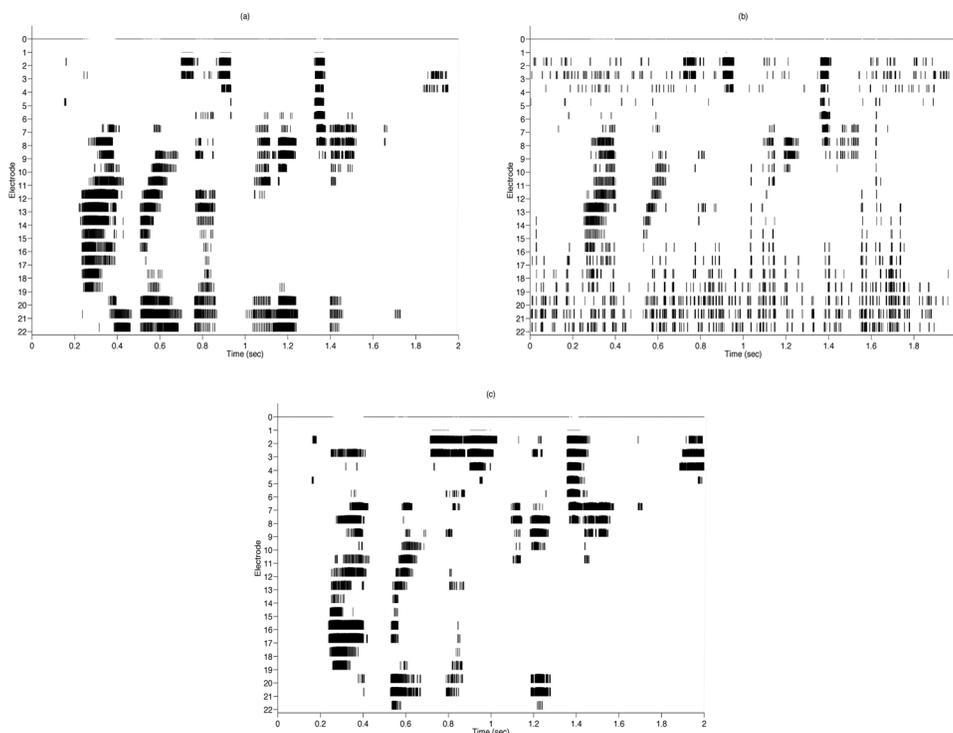


Fig. 2. (a) Electrodogram of unmodified (uncorrupted) sentence, (b) electrodogram of the same sentence corrupted with 9 m/s wind noise, (c) electrodogram of the same sentence processed with the 2MWS strategy. Time is shown along the abscissa and the electrode number is along the ordinate.

incoherent). In cases whereby incoherent (e.g., uncorrelated) noise sources are combined, the overall signal power will increase irrespective of the total phase of the two signals (Kates, 2008). Because both the ZOOM and BEAM strategies rely on first forming either the sum or the difference of the inputs to the two microphones, the overall level of wind noise was greater than the level observed at each individual omni-directional microphone. In contrast, the 2MWS strategy yielded a considerable benefit in all conditions tested. Figure 2 illustrates example stimulus output patterns (electrograms) of an IEEE sentence processed with the default advanced combination encoder (ACE) strategy. ACE is the default strategy used in Nucleus devices and operates by selecting only a subset of envelopes (typically around 8–12) for stimulation at each cycle. The principle underlying the use of ACE is that speech can be understood well even if only the peaks in the short-term spectrum are transmitted. As shown in Fig. 2(b), where speech is corrupted by wind noise recorded at 9 m/s, the masking effects of wind noise become more dominant in the low-frequency spectral channels represented in the more apical electrodes (e.g., 15–22). Wind noise flow smears the temporal envelope, blurs spectral cues, and also fills the gaps in unvoiced and silent segments. Strong wind bursts also increase high-frequency energy, which produces masking of higher speech frequencies (e.g., see consonant in the 1.8 s mark). When processing with ACE, the bands with the highest spectra will be the ones selected to drive the electrodes. Intense low- and mid-frequency energy present in wind noise could capture all the available electrode-driving time slots. Thus none will remain available to respond to the weaker high-frequency energy contained in the speech input. In such a scenario, the ACE processing strategy will mistakenly rely on the channels that contain mostly wind energy to drive the electrodes. As illustrated in Fig. 2(c), low-frequency energy due to wind noise can be attenuated considerably (especially in the apical electrodes) after processing with the 2MWS strategy. The less intense additive energy concentrated in the high-frequency range can be also suppressed to a large extent as is evident from examining the basal electrodes (e.g., 1–5). In essence, both low-frequency information contained in voiced segments (e.g., vowels) as well as high-frequency content present in obstruent sounds (e.g., fricatives, affricates, and stops) become more prominent.

Acknowledgment

This work was supported by NIH/NIDCD R03-DC-008882 to K.K.

References and links

- Chung, K. (2004). "Challenges and recent developments in hearing aids: Part I. Speech understanding in noise, microphone technologies and noise reduction algorithms," *Trends Amplif.* **8**, 83–124.
- Chung, K. (2012). "Wind noise in hearing aids: II. Effects of microphone directionality," *Int. J. Audiol.* **51**, 29–42.
- Chung, K., and McKibben, N. (2011). "Microphone directionality, pre-emphasis filter, and wind noise in cochlear implants," *J. Am. Acad. Audiol.* **22**, 586–600.
- Chung, K., Zeng F.-G., and Acker, K. N. (2006). "Effects of directional microphone and adaptive multichannel noise reduction algorithm on cochlear implant performance," *J. Acoust. Soc. Am.* **120**, 2216–2227.
- Dillon, H., Roe, I., and Katsch, R. (1999). "Wind noise in hearing aids," presented at *Hearing Aid Amplification for the New Millennium*, Sydney, Australia. It may be viewed at [http://www.nal.gov.au/pdf/wind noise.ppt](http://www.nal.gov.au/pdf/wind%20noise.ppt) (Last viewed November 8, 2013).
- IEEE (1969). "IEEE recommended practice speech quality measurements," *IEEE Trans. Audio Electroacoust.* **AU17**, 225–246.
- Kates, J. M. (2008). *Digital Hearing Aids* (Plural Publishing, Plymouth, UK), pp. 147–173.
- Spriet, A., Van Deun, L., Eftaxiadis, K., Laneau, J., Moonen, M., Van Dijk, B., Van Wieringen, A., and Wouters, J. (2007). "Speech understanding in background noise with the two-microphone adaptive beamformer BEAM in the Nucleus Freedom cochlear implant system," *Ear Hear.* **28**, 62–72.
- Thompson, S. C. (2002). "Microphone, telecoil, and receiver options: Past, present, and future," in *Hearing Aids: Standards, Options, and Limitations*, edited by M. Valente (Thieme Medical Publishers, New York), pp. 64–100.
- Zakis, J. A. (2011). "Wind noise at microphones within and across hearing aids at wind speeds below and above microphone saturation," *J. Acoust. Soc. Am.* **129**, 3897–3907.