

Iowa State University

From the Selected Works of Jonathan W. Kelly

2007

Accurate vocal compensation for sound intensity loss with increasing distance in natural environments

Pavel Zahorik, *University of Louisville*

Jonathan W. Kelly, *Vanderbilt University*



Available at: https://works.bepress.com/jonathan_kelly/8/

Accurate vocal compensation for sound intensity loss with increasing distance in natural environments

Pavel Zahorik

*Department of Psychological and Brain Sciences, University of Louisville, Louisville, Kentucky 40292
pavel.zahorik@louisville.edu*

Jonathan W. Kelly

*Department of Psychology, Vanderbilt University, Nashville, Tennessee 37212
jonathan.kelly@vanderbilt.edu*

Abstract: Human abilities to adjust vocal output to compensate for intensity losses due to sound propagation over distance were investigated. Ten normally hearing adult participants were able to compensate for propagation losses ranging from -1.8 to -6.4 dB/doubling source distance over a range of distances from 1 to 8 m. The compensation was performed to within 1.2 dB of accuracy on average across all participants, distances, and propagation loss conditions with no practice or explicit training. These results suggest that natural vocal communication processes of humans may incorporate tacit knowledge of physical sound propagation properties more sophisticated than previously supposed.

© 2007 Acoustical Society of America

PACS numbers: 43.70.Mn, 43.66.Qp [QJF]

Date Received: July 12, 2007 **Date Accepted:** August 20, 2007

1. Introduction

Under ideal conditions sound intensity obeys an inverse square law with distance: Each doubling of sound source distance decreases sound intensity by 6 dB. Previous research has demonstrated that humans increase their vocal output in order to compensate for these sound propagation losses (Healey, Jones, and Berky, 1997; Johnson *et al.*, 1981; Markel, Prebor, and Brandt, 1972; Michael, Siegel, and Pick, 1995; Warren, 1968). Although the compensation appears to be performed naturally to facilitate effective communication over varying distances between talker and listener and is evident in children as young as 3 years of age (Johnson *et al.*, 1981), there is considerable variability in the amount of compensation reported in the literature. Warren reports increases in vocal output level of 6 dB per doubling distance, suggesting that talkers may perhaps have internalized the ideal inverse-square law relationship for sound propagation loss (Warren, 1968; Warren, 1981). Other studies have reported considerably less level compensation, ranging from 5 to less than 1 dB/doubling (Healey *et al.*, 1997; Johnson *et al.*, 1981; Markel *et al.*, 1972; Michael *et al.*, 1995).

One potential source for this variability is the extent to which the listening environments in which these past experiments were conducted approximated a free-field environment with an ideal 6 dB/doubling propagation loss. Departure from this ideal, such as in rooms with sound reflecting surfaces, in general results in propagation losses less than 6 dB/doubling. Accurate compensation in such environments would, therefore, require less than a 6 dB/doubling increase in vocal output level. This may explain why talkers did not increase their vocal output levels by a full 6 dB/doubling distance in a number of past studies (Michael *et al.*, 1995), although it is important to note that most past studies did not report physical sound propagation losses in the testing environments. For one study that did report propagation losses, there does seem to be a relationship (Michael *et al.*, 1995). A propagation loss of around 2 dB/doubling corresponded to a vocal compensation increase of around 2 dB/doubling for certain condi-

tions. Other results show a much less clear relationship: 4–6 dB/doubling propagation losses in the testing environment of another study corresponded to vocal compensation ranging from less than 2 dB/doubling for adults to more than 35 dB/doubling for children (Johnson *et al.*, 1981). Still unknown, however, is the extent to which talkers may be able to adjust their amount of vocal compensation to different environments with different propagation losses. An additional issue relates to the range of distances over which vocal compensation abilities have been evaluated. Valid tests of a relationship between propagation loss and vocal compensation amounts will require evaluation at multiple distances.

The current study seeks to determine the extent to which vocal compensation depends on the specific propagation losses present in the listening environment over a wide range of distances, and whether talkers can accurately modulate their amounts of compensation to match widely varying amounts of physical propagation loss across different natural listening situations. To the extent that these vocal compensation abilities are used in everyday vocal communication, one might expect normal adult talkers to have developed considerable skill and accuracy in these abilities.

2. Methods

2.1 Propagation loss and source directionality measurements

Estimates of the physical sound level decay with increasing source distance were made for each of four listening conditions used in the experiment: Two acoustic environments and two source orientations. The two acoustic environments, one outdoor and one indoor, had widely different reverberant properties. The outdoor environment was chosen to approximate a free-field listening situation. It was a grassy field approximately 80 m × 40 m, with closest non-ground sound-reflecting surfaces at least 20 m from the measurement locations. The indoor environment was a reverberant hallway, approximately 20 × 3.5 × 3 m (L × W × H) with hard walls, hard floor, and an absorptive ceiling. Source orientation was either directly facing the measurement location (0°), or else rotated 180° in the horizontal plane.

All decay measurements were made using a high-quality omni-directional microphone (Sennheiser KE4-211-2) mounted on a movable tripod. A small (12.7 cm full-range driver, 17.8 × 15.2 × 13.0 cm cabinet) high-output loudspeaker (MicroSpot Monitor, Galaxy Audio, Inc.) with high-quality amplification (D-75, Crown, Inc.) mounted on a tripod with a rotating head at a fixed location served as the sound source. Both microphone and loudspeaker were positioned 1.5 m above the ground surface. In the indoor environment, the source location was approximately 5 m from one end of the hallway, and approximately midway between the side walls. The measurement signal was spectrally-shaped noise (10 s duration), with flat spectrum between 0.1 and 1.5 kHz, decreasing at 60 dB/octave below 0.1 kHz and 20 dB/octave above 1.5 kHz. The spectral shape of this signal was chosen to roughly approximate the spectra of the speech signals used in subsequent portions of the experiment: Male and female talkers producing the vowel /a/. This signal was processed digitally using Matlab software (Mathworks, Inc.) and stored to a standard audio compact-disc for later presentation during measurement conditions. The output level of the measurement signal was fixed for all measurements and corresponded to 90 dBA at 1 m, 0° orientation, in the outdoor environment as measured via a calibrated sound-level meter (Realistic 33-2059, calibrated with a B&K piston-phone, model 4228). Background noise levels were approximately 48 and 37 dBA for the outdoor and indoor environments. Broadband (0.2–4 kHz) reverberation time, T_{60} , for the indoor reverberant environment was approximately 0.7 s, measured with the loudspeaker in the 180° orientation using an energy integration technique (Schroeder, 1965). Decay measurement results were represented in dB relative to the observed microphone output voltage (RMS) at 1 m for each of the two measurement environments and two source orientations.

Occasional nonstationary noise disturbances did occur during both measurement and later experimental sessions. During these occurrences the experimenters suspended the measurement and/or experimental session until the noise disturbance had subsided and discarded any potentially noise-contaminated data.

The extent to which the level decay measurements made using a loudspeaker sound source are valid for making inferences regarding level decay of vocal sound sources depends critically on the two sources having similar directional responses. This is a particularly important issue in acoustically reflective environments where source directivity can strongly affect sound propagation. Measurements were therefore made to estimate the directional responses characteristics of four representative talkers (two male, two female) producing the vowel /a/ and the measurement loudspeaker, using methods fundamentally similar to those described by [Studebaker \(1985\)](#). These measurements were conducted in a second quiet outdoor environment also chosen to approximate free-field conditions: A large grassy field free from sound reflecting surfaces other than the ground. Average noise level in this environment was approximately 40 dBA. All sound level measurements were conducted relative to the measurement location, which was at fixed distance of 1 m. Loudspeaker response measurements were made at 0° and 180° angular orientations in the horizontal plane at a distance of 1 m, using the same source material and microphone used in the level decay measurements. Loudspeaker output level was fixed for all measurements: 90 dBA at 1 m, 0° orientation. Vocal response measurements were also made at 0° and 180° orientations, although two matched measurement microphones (Sennheiser KE4-211-2) were used: One at distance of 1 m, and one fixed to the talker's head at a distance of approximately 10 cm from the mouth. Talkers were instructed to use "conversational" vocal output levels and to produce the required vocalization for at least 2 s at each of the measurement orientations. Decibel levels were computed in 1/3-octave bands (0.25–5 kHz) for all measurements. For the voice measurements, the decibel difference in each frequency band between close and far microphone measurements was computed. This representation of relative output level allowed control for differences in absolute source output levels from measurement to measurement, and was used for all subsequent analyses.

2.2 Vocal output level compensation

Participants: Ten adult volunteers (seven females and three males, ages 20–29 years) were paid for their participation. All reported having normal hearing, normal or corrected-to-normal vision, and normal vocal abilities.

Design: In a completely within-participants design, there were two test environments (indoor or outdoor), two participant orientations (facing toward or away from the target location, referred to as 0° and 180° orientations, respectively), and four participant-target distances (1, 2, 4, and 8 m). Test environment order was counterbalanced. Within each test environment, participant orientation was blocked and order was fixed (0°, then 180°). Within each such block, the order of participant-target distances was also fixed, from nearest to farthest.

Stimuli and procedures: Participants attempted to compensate for the physical sound level losses associated with increasing distance by adjusting the output level of their own voice. The testing procedure was as follows: Participants were lead to the reference location (0 m) in the listening environment (either indoor or outdoor) where they remained for the duration of testing at a given source orientation. The experimenter instructed the participant to produce the vowel /a/ for approximately 3 s and adjust their vocal output level such that the level reaching the experimenter (distal level) remained constant at each of the four target distances. At each target location, starting from the 1 m location and successively increasing in distance for subsequent locations, the experimenter measured and recorded the sound level using a hand-held sound level meter (Realistic 33-2059), which served as the target. Participants were instructed to make initial /a/ productions at conversational levels for the 1 m target distance. For both orientations (0° and 180°), participants were instructed to look at the experimenter prior to vocal production in order to provide the participant with visual distance information. For the 180° orientation, this required participants to turn, look, and then return to the appropriate orientation prior to vocal production. All instructions were provided verbally to each participant at a fixed distance of approximately 1 m (0° orientation) between experimenter and participant prior to vocal compensation testing at different distances. Instructions were provided in detail in the initial environment and then repeated in capsule form in the second environment. No feed-

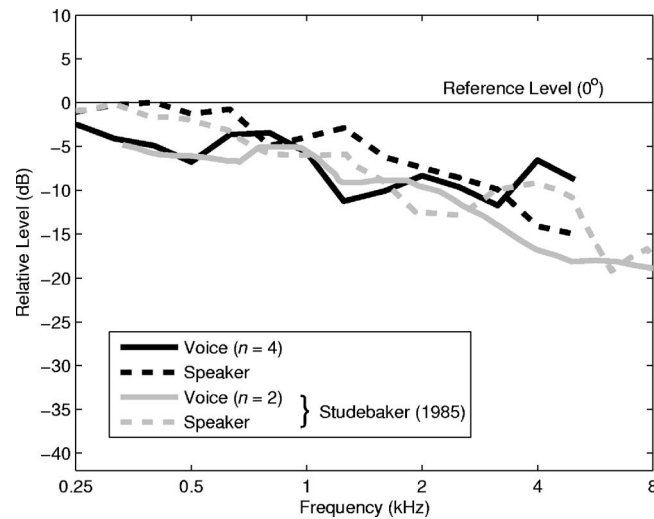


Fig. 1. Signal levels observed in 1/3-octave frequency bands at 180° in the horizontal plane relative to the levels at 0° azimuth for human voices and small loudspeakers measured in free-field environments. Voice levels are mean data, where n indicates the number of voices contributing to the mean values (see text for details).

back or explicit training related to sound propagation loss with distance was provided to the participants at any time during the experiment. Participant 103 was not tested at the 180° orientation in either listening environment for unforeseen logistical reasons.

3. Results

3.1 Source directionality

Figure 1 displays the frequency-dependent source directionality results in which levels measured at 180° are compared relative to those measured at 0° (reference level). Small loudspeaker ($17.9 \times 11.1 \times 10.5$ cm cabinet) and vocal (continuous discourse averaged from one male and one female talker) directional response results from a previous study (Studebaker, 1985) are also displayed for comparison purposes. Small loudspeaker directionality is, in general, quite similar to that of the human voice. Both become more directional with increasing frequency at roughly the same rates. Verification of this similarity is particularly important when sound pressure levels averaged across frequency are considered, as in the case of the vocal output measurements made using a sound level meter. The mean level difference between loudspeaker and voice signals (speaker–voice) across 1/3 octave bands from 0.25–5 kHz was 1.5 dB for the current study, which is a slightly better match than the 2.8 dB difference that results from the comparison measurement data (Studebaker, 1985). Overall, this relatively close match in directional responses suggests that valid inferences regarding vocal source propagation may be made from loudspeaker-based sound propagation measurements when averaging across frequency.

3.2 Propagation loss

Sound propagation measurement results for all source orientations and measurement environments are displayed in Figs. 2(a)–2(d). Rates of sound level loss (dB) per doubling of source distance are also shown for each condition, determined via exponential fits to the data using a least-squares criterion. The fitted functions were adequate descriptions of the data in all cases. The RMS error between predicted and measured values was less than 0.3 dB in all conditions except indoors at 0° , where RMS error was still less than 1.5 dB. The outside 0° condition had a measured propagation loss quite similar to that predicted by the inverse-square law for free-field sources. The departure from this ideal 6 dB/doubling loss for the 180° orientation outside

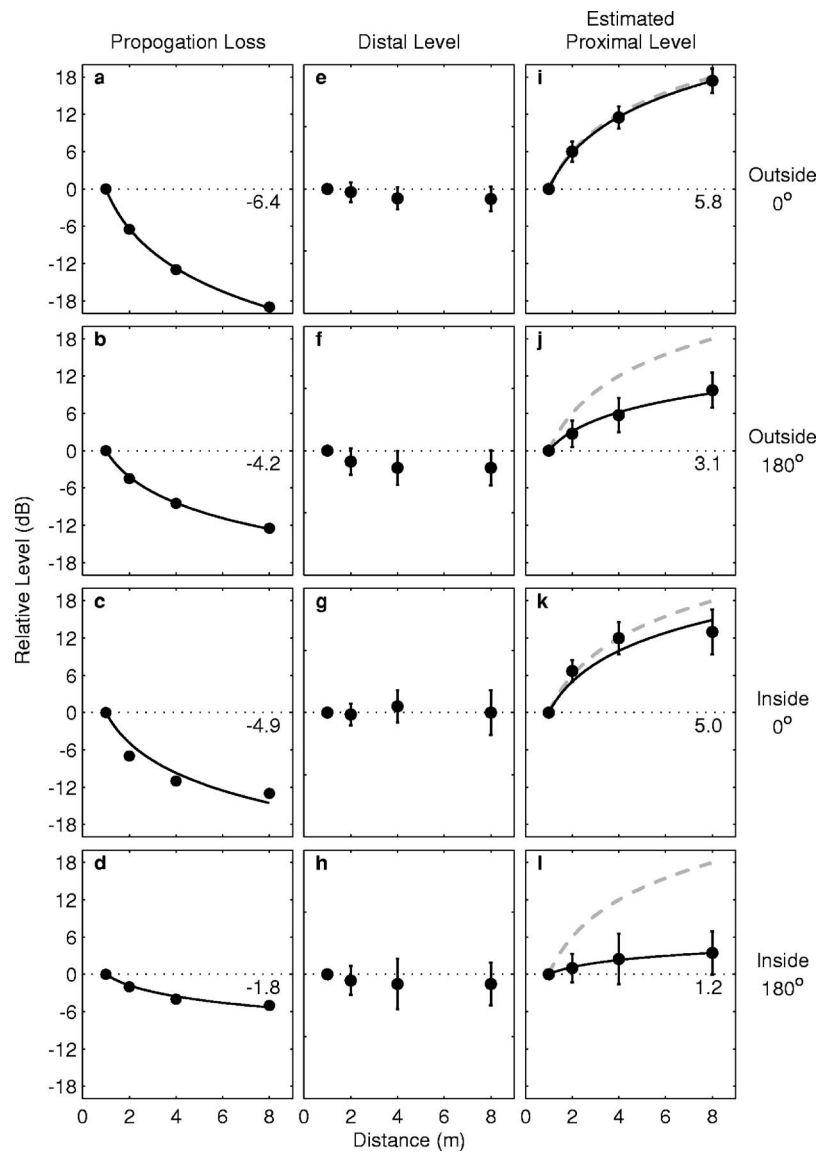


Fig. 2. Propagation loss and vocal compensation results for each of four acoustic conditions (inside and outside environments, 0° and 180° source orientations). All levels are expressed in dB relative to 1 m. Propagation losses are shown in the left column (a–d), and are well described by exponential fits to the data (solid curves). Slopes (dB/distance doubling) are displayed for each fit. Mean levels (dB) at each measurement distance for vocal sources when talkers are instructed to compensate for propagation losses are shown in the center column (e–h). Bars indicated 99% confidence intervals. Mean estimated levels (dB) at the talker’s location based on the measured propagation losses are shown in the right column (i–l). Bars again indicate 99% confidence intervals. Solid curves show exponential fits to the data, with slopes indicated (dB/doubling) indicated for each fit. The gray dashed line displays a 6 dB/doubling increase for reference.

is not well understood, but likely resulted from acoustically reflective surfaces that did exist in the outdoor environment in the direction opposite of the measurement microphone, but at a distance of at least 20 m. As expected, sound reflecting surfaces in the indoor environment also resulted in propagation losses less than 6 dB/doubling. The loss for the 0° orientation was similar to that reported in a previous study in which the listening environment had a similar

reverberation time. The indoor 180° orientation had the least propagation loss, given that the energy reaching the measurement microphone in this case was mostly reverberant energy, which is relatively independent of source distance. Propagation loss in this case was also similar to the reverberant-only energy loss observed in a previous study in a room with similar reverberation time. Overall, these four conditions resulted in a considerable range of physical sound propagation losses with distance to the sound source.

3.3 Vocal output compensation

“Conversational” levels at 1 m ranged from 57 to 68 dBA across talkers, with a median level of 62 dBA. Level compensation results are shown in Figs. 2(e) and 2(f), where mean distal levels relative to the levels measured at 1 m for each participant in each condition are displayed. Given that the levels remain within 3 dB of the level at 1 m (0 dB) and lie within the confidence intervals at all distances, it may be concluded that talkers on average accurately increase their vocal output to remain constant at each measurement distance. The mean level relative to 1 m was -1.2 dB across all distances (not including 1 m) and conditions. Because the physical propagation losses are different in the different measurement conditions, the amount of increase in vocal output to compensate for increased distance also differs across condition. Figures 2(i)–2(l) displays estimates of the increases in vocal output level (proximal) based on the measured propagation losses to produce the measured distal levels. All data points are mean levels relative to 1 m. For comparison purposes, a 6 dB/distance doubling is also indicated in these plots (dashed gray line). These estimates further suggest that the amount of level compensation as a function of distance differed dramatically, on average, across the four measurement conditions. It is also clear from these data that talkers are, on average, not simply applying a 6 dB/doubling increase to their vocal outputs in all conditions, although they do apply this rule where it is appropriate to the observed physical propagation loss (outside, 0°).

Although the data in Fig. 2 suggest that talkers can, on average, accurately compensate for propagation losses with distance under different loss conditions, it is important to determine the extent to which individual talkers are also capable of accurate compensation. Individual talker data was, therefore, analyzed via separate exponential function fits, which were found to adequately describe the data in all cases. RMS error between predicted and measured levels ranged from 0.3 to 2.6 dB across all listeners and conditions, with half of all RMS errors below 1.2 dB. Figure 3 displays slope values from the fitted functions for all participants and conditions using vocal signals. Slope values based on functions fit to the mean level data across all participants [e.g., Figs. 2(e)–2(h)] are also displayed. Sound propagation losses (dB/doubling) are also shown for each corresponding measurement condition. Although the slope estimates based on individual data were more variable than those based on mean data, 0 dB/doubling fell within the 99% confidence regions all but two slope values [Participant 102, Fig. 3(b); Participant 103, Fig. 3(c)]. The median slope across all participants and conditions was -0.9 dB/doubling, with half of all slopes falling within $+0.8$, -1.8 of 0 dB/doubling. Overall, these results indicate that individual participants can compensate for the variable sound propagation losses present in the acoustic conditions with considerable accuracy.

4. Discussion and conclusions

This study has demonstrated that talkers can adjust their vocal output to compensate with considerable accuracy for sound propagation losses ranging from approximately -1.8 to -6.4 dB/doubling distance. This suggests that humans may have tacit knowledge of sound propagation properties more sophisticated than previously thought, and may explain at least some of the previously unexplained variance in past reports of vocal compensation abilities for changing distance (Healey *et al.*, 1997; Johnson *et al.*, 1981; Markel *et al.*, 1972; Michael *et al.*, 1995; Warren, 1968). From the standpoint of vocal communications, the ability to adjust vocal output for sound propagation losses to a listener's position is clearly advantageous, and this advantage is potentially extended with an accurate match between physical propagation loss and vocal output increase. Applying a 6 dB/doubling increase of vocal output for all situations

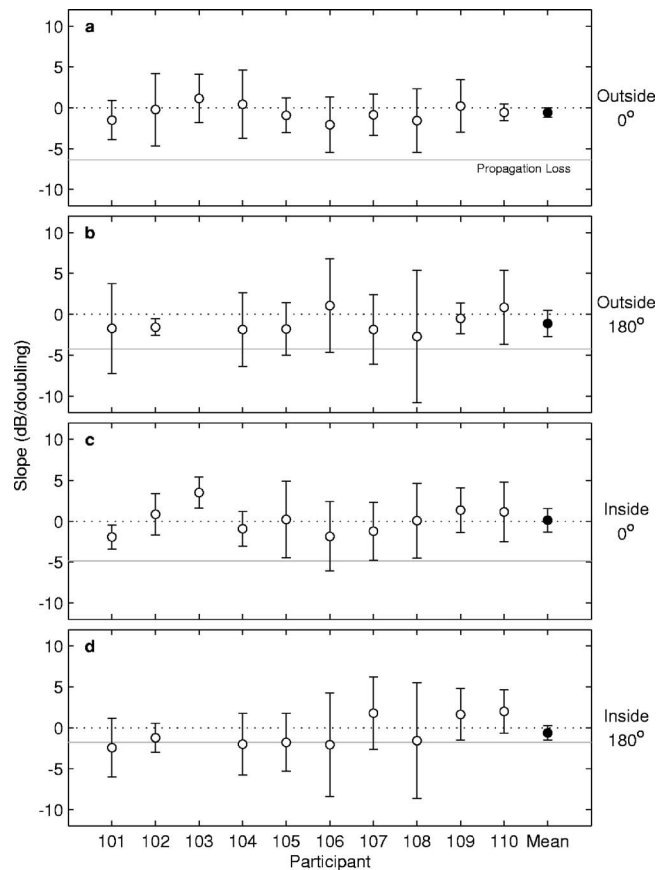


Fig. 3. Slopes of exponential fits to the distal level measurements of each participant and also to the mean levels across participants for four measurement conditions. Bars indicated 99% confidence intervals. Physical propagation losses (dB/doubling) are also shown for each condition.

would unnecessarily limit compensation distances in environments with less than a 6 dB/doubling loss. It is clear, however, that regardless of listening environment, vocal compensation abilities do have practical limits governed by various factors such as the effective dynamic range of the human voice and listening environment noise levels. Although neither of these factors was tested in this study, reasonably accurate vocal compensation was observed in quiet environments over a distance range of 1 to 8 m.

In certain respects, this vocal compensation ability is similar to another, more well-known, form of vocal compensation, known as the Lombard Reflex (Lombard, 1911), in which talkers increase their vocal output level when background noise level is increased. Both forms of compensation facilitate vocal communication by keeping signal-to-noise ratio constant at the listener's location, can be performed with considerable accuracy (i.e., signal-to-noise ratio held constant at the listeners location), and appear to be performed naturally as part of the vocal communication process. The Lombard Reflex has also been documented in other mammals (Scheifele *et al.*, 2005; Sinnott, Stebbins, and Moody, 1975) and songbirds (Cynx *et al.*, 1998), is remarkably robust to volitional control of human talkers (Pick *et al.*, 1989), and appears to depend critically on auditory feedback in its regulation of vocal output level (Siegel and Pick, 1974). Although informal observation suggests both that other species may at least roughly compensate for sound propagation losses, and that it may also be difficult for human talkers to

suppress distance compensation, the extent to which distance compensation depends on auditory feedback is unknown. Clearly additional scientific research is needed in all of these areas.

Results from this study may also have important implications for auditory distance perception, where systematic biases in distance estimates to sound sources have been documented in numerous studies using a wide range of stimulus conditions and psychophysical procedures (see Zahorik, Brungart, and Bronkhorst, 2005, for review). If the vocal compensation accuracy to changing distance observed here does represent tacit knowledge of sound propagation losses in the listening environment, then it is surprising that listeners are not able to use this knowledge to make accurate judgments of sound source distance. This seeming dissociation may perhaps be similar to the well-documented dissociation between accurate visually directed action and inaccurate conscious visual experience (Creem and Proffitt, 1998; Milner and Goodale, 1995), although at least one known deficit in auditorily directed action, dysarthria resulting from Parkinson's Disease, does not appear to affect talkers' compensation for sound propagation losses (Ho, Ianssek, and Bradshaw, 1999). This suggests that vocal compensation abilities are not purely action-based. Accurate vocal compensation abilities may instead depend on perceptual processes at least partially distinct from those underlying (inaccurate) conscious experience of sound source distance. Further research will be needed to more fully evaluate these potential relationships.

Acknowledgments

The authors wish to thank Dr. Jack Loomis for his helpful comments and for the use of the facilities in which this study was conducted. Work supported in part by grants from ONR (N00014-01-1-0098) and NIH (F32EY007010, R03DC005709, R01DC008168).

References and links

- Creem, S. H., and Proffitt, D. R. (1998). "Two memories for geographical slant: separation and interdependence of action and awareness," *Psychon. Bull. Rev.* **5**, 22–36.
- Cynx, J., Lewis, R., Tavel, B., and Tse, H. (1998). "Amplitude regulation of vocalizations in noise by a songbird, *Taeniopygia guttata*," *Anim. Behav.* **56**, 107–13.
- Healey, E. C., Jones, R., and Berky, R. (1997). "Effects of perceived listeners on speakers' vocal intensity," *J. Voice* **11**, 67–73.
- Ho, A. K., Ianssek, R., and Bradshaw, J. L. (1999). "Regulation of Parkinsonian speech volume: the effect of interlocuter distance," *J. Neurol., Neurosurg. Psychiatry* **67**, 199–202.
- Johnson, C. J., Pick, H. L., Jr., Siegel, G. M., Cicciarelli, A. W., and Garber, S. R. (1981). "Effects of interpersonal distance on children's vocal intensity," *Child Dev.* **52**, 721–723.
- Lombard, E. (1911). "Le Signe de l'Élévation de la Voix," *Ann. Maladies Oreille, Larynx, Nez, Pharynx* **37**, 101–119.
- Markel, N. N., Prebor, L. D., and Brandt, J. F. (1972). "Biosocial factors in dyadic communication," *J. Pers. Soc. Psychol.* **23**, 11–13.
- Michael, D. D., Siegel, G. M., and Pick, H. L., Jr. (1995). "Effects of distance on vocal intensity," *J. Speech Hear. Res.* **38**, 1176–1183.
- Milner, A. D., and Goodale, M. A. (1995). *The visual brain in action* (Oxford University Press, New York).
- Pick, H. L., Jr., Siegel, G. M., Fox, P. W., Garber, S. R., and Kearney, J. K. (1989). "Inhibiting the Lombard effect," *J. Acoust. Soc. Am.* **85**, 894–900.
- Scheifele, P. M., Andrew, S., Cooper, R. A., Darre, M., Musiek, F. E., and Max, L. (2005). "Indication of a Lombard vocal response in the St. Lawrence River Beluga," *J. Acoust. Soc. Am.* **117**, 1486–1492.
- Schroeder, M. R. (1965). "New method of measuring reverberation time," *J. Acoust. Soc. Am.* **37**, 409–412.
- Siegel, G. M., and Pick, H. L., Jr. (1974). "Auditory feedback in the regulation of voice," *J. Acoust. Soc. Am.* **56**, 1618–1624.
- Sinnott, J. M., Stebbins, W. C., and Moody, D. B. (1975). "Regulation of voice amplitude by the monkey," *J. Acoust. Soc. Am.* **58**, 412–414.
- Studebaker, G. A. (1985). "Directivity of the human vocal source in the horizontal plane," *Ear Hear.* **6**, 315–319.
- Warren, R. M. (1968). "Vocal compensation for change in distance," *Proceedings of the 6th International Congress on Acoustics (Tokyo)*, pp. 61–64.
- Warren, R. M. (1981). "Measurement of sensory intensity," *Behav. Brain Sci.* **4**, 175–223.
- Zahorik, P., Brungart, D. S., and Bronkhorst, A. W. (2005). "Auditory distance perception in humans: A summary of past and present research," *Acta Acust.* **91**, 409–420.