Is cochlear gain reduction related to speech-in-babble performance?

KRISTINA DEROY MILVAE*, JOSHUA M. ALEXANDER,
AND ELIZABETH A. STRICKLAND

Department of Speech, Language, and Hearing Sciences, Purdue University, West Lafayette, IN, USA

Noisy settings are difficult listening environments. With some effort, individuals with normal hearing are able to overcome this difficulty when perceiving speech, but the auditory mechanisms that help accomplish this are not well understood. One proposed mechanism is the medial olivocochlear reflex (MOCR), which reduces cochlear gain in response to sound. It is theorized that the MOCR could improve intelligibility by applying more gain reduction to the noise than to the speech, thereby enhancing the internal signal-to-noise ratio. To test this hypothesized relationship, the following measures were obtained from listeners with clinically normal hearing. Cochlear gain reduction was estimated psychoacoustically using a forward masking task. Speech-in-noise recognition was assessed using the QuickSIN test (Etymotic Research), which generates an estimate of the speech reception threshold (SRT) in background babble. Results were surprising because large reductions in cochlear gain were associated with large SRTs, which was the opposite of the hypothesized relationship. In addition, there was a large range for both cochlear gain reduction and SRT across listeners, with many individuals falling outside of the normal SRT range despite having normal hearing thresholds.

INTRODUCTION

We are able to navigate the world around us using sensorineural systems that give us a sense of touch, sight, smell, taste, and sound. These sensory systems work by detecting changes in our environment, such as the sound of a friend's voice above the noise of a restaurant. To detect the friend's voice, it would be helpful for our auditory system to have a differential response to the varying speech relative to the ongoing background noise. It is known that one function of cochlear outer hair cells is to provide gain to basilar membrane motion for low-level acoustic stimulation. If relatively less gain is applied to the steady noise, then acoustic changes associated with the speech can be detected more easily. One possible mechanism to accomplish this is the medial olivocochlear reflex (MOCR).

---

*Corresponding author: kderoy@purdue.edu

Proceedings of ISAAR 2015: Individual Hearing Loss – Characterization, Modelling, Compensation Strategies. 5th symposium on Auditory and Audiological Research. August 2015, Nyborg, Denmark. Edited by S. Santurette, T. Dau, J. C. Dalsgaard, L. Tranebjærg, and T. Andersen. ISBN: 978-87-990013-5-4. The Danavox Jubilee Foundation, 2015.

The MOCR is a bilateral reflex in the auditory system involving the inner ear and brainstem pathways. Once activated by acoustic stimulation along some place on the basilar membrane, the MOCR acts to reduce the outer hair cell gain at that place (Cooper and Guinan, 2006). This reflex takes about 25 ms to fully activate, making it a sluggish feature of the auditory system (Backus and Guinan, 2006; James *et al.*, 2005).

One hypothesis for the role of the MOCR is that it improves perception in noise. Auditory nerve fibers are able to better respond to changes in a signal embedded in noise when the MOCR is activated (Kawase *et al.*, 1993; Winslow and Sachs, 1987). In addition, the MOCR may improve the signal-to-noise ratio (SNR) for speech in noise, as shown in modelling studies (Ghitza, 1988; Messing *et al.*, 2009).

The relationship between physiological estimates of MOCR gain reduction strength (using contralateral suppression of otoacoustic emissions [OAEs]) and speech-in-noise performance has been measured in correlational studies. Results have been mixed. Some studies have found a positive correlation (Bidelman and Bhagat, 2015; de Boer and Thornton, 2008; Giraud *et al.*, 1997; Kumar and Vanaja, 2004), but another found a negative correlation (de Boer *et al.*, 2012). In addition, some work has found no correlation between the two measures (Wagner *et al.*, 2008). The reason for this variety of findings is not yet clear.

An alternative measure of cochlear gain reduction, likely related to MOCR activity, can be estimated using psychoacoustic measures (Krull and Strickland, 2008; Roverud and Strickland, 2010; Strickland, 2001; Yasin *et al.*, 2014). There are some advantages to the use of behavioral measures over OAEs. Behavioral measures allow for quantification of cochlear gain reduction in terms that may help us better understand functional consequences for perception. In addition, measures of cochlear gain reduction from ipsilateral stimulation can be easily measured with this technique, so that ipsilateral gain reduction can be compared to ipsilateral speech-in-noise performance. Studies investigating psychoacoustic measures of ipsilateral gain reduction have primarily investigated cochlear gain reduction at the 4-kHz place. However, because speech is a broadband signal and has more energy at lower frequencies, it is important to consider MOCR function at lower frequencies as well. With contralateral acoustic stimulation, psychoacoustic evidence of cochlear gain reduction has been found at frequencies as low as 500 Hz (Aguilar *et al.*, 2013). The present study will estimate cochlear gain reduction at both 2 and 4 kHz to improve our understanding of ipsilateral cochlear gain reduction across frequency.

We hypothesize that participants with relatively larger gain reduction estimates will perform better on a speech-in-noise task. This study builds on previous work in that we measure ipsilateral gain reduction at a frequency that is more relevant to speech perception than that traditionally measured. In addition, the perceptual measure of gain reduction is compared to performance on the QuickSIN (Etymotic Research), thereby allowing us to investigate the relationship between psychoacoustic measures of ipsilateral cochlear gain reduction and speech-in-noise performance.

METHOD

Participants

Twenty young adults (7 male, 13 female) between the ages of 18 and 28 years (median: 20 years) completed this experiment in exchange for modest monetary compensation. All participants reported English as a first language. Participants passed a hearing screening of 15 dB HL at 0.25, 0.5, 1, 2, 4, and 8 kHz in both ears. The hearing screening was completed in a sound-treated booth. One additional participant did not pass the hearing screening and testing was discontinued.

Stimuli and procedure

Speech-in-babble performance. Speech understanding in noise was measured using the QuickSIN (Speech-in-Noise) Test (Etymotic Research; Killion *et al.*, 2004). Participants listened to a recording of a woman's voice and background speech babble at various SNRs. Sentences were presented from 0-25 dB SNR in 5-dB steps and descending order for each list of six sentences (easiest to most difficult condition). Participants responded by repeating the target sentence at each SNR. A practice list was used to familiarize participants with the task. Next, four test lists (lists 1-4) were used. Sentences were scored according to the test instructions and were based on the number of keywords repeated correctly.

The QuickSIN measures speech reception threshold (SRT), which is the SNR required for 50%-correct performance. SRTs above +4 dB (the normative range for SNR loss plus the 2 dB reference for listeners with normal hearing) are considered outside the normal range (Killion *et al.*, 2004). The QuickSIN was presented at 70 dB HL to each participant's right ear via ER-3A insert earphones using a CD player routed to an audiometer (GSI-61).

Estimate of cochlear gain reduction. Estimates of cochlear gain reduction were measured in the right ear. The signal was a 2-kHz, 10-ms tone (5-ms $\cos^2$ ramps) or a 4-kHz, 6-ms tone (3-ms $\cos^2$ ramps). These durations were chosen to keep the signals as short as possible with minimal frequency spread. Participant thresholds for the tone alone were compared with those for the same tone preceded by a 50-ms, 60-dB SPL broadband noise precursor and a 20-ms silent gap. The precursor bandwidth was 0.25-8 kHz, and 5-ms $\cos^2$ ramping was used at onset and offset. High-pass noise was presented during each precursor interval to limit off-frequency listening (e.g., Nelson *et al.*, 2001). The noise began 50 ms before the first stimulus and ended 50 ms after the signal (5-ms $\cos^2$ ramps), and was presented at a spectrum level 50 dB below the signal level. The frequency content of the high-pass noise ranged from 1.2 times the signal frequency to 10 kHz.

This paradigm is based on the one used by Roverud and Strickland (2010), with silence replacing the off-frequency masker in an effort to isolate masking due to cochlear gain reduction from masking due to excitation. Previous research has provided evidence that preceding stimulation in this temporal paradigm is more consistent with cochlear gain reduction than temporal integration of sound (Jennings *et al.*, 2009; Roverud and Strickland, 2014).

Stimuli were generated using a custom Matlab (2012a, The Math Works, Natick, MA) program and delivered by a Lynx II XLR sound card. The sounds were passed through a headphone buffer (TDT HB6) and then delivered to insert earphones (ER-2). Adaptive tracking (Levitt, 1971) was implemented in the computer program to approximate the 70.7% correct threshold on the psychometric function, using a rule that increases the intensity of the signal after one incorrect response and decreases the intensity of the signal after two correct responses. Step sizes began at 5 dB and then decreased to 2 dB after the fourth reversal. The program continued testing until 12 reversals were completed. Threshold was defined as the average of the levels of the final 8 reversals.

Participants were instructed that they would hear three intervals. The task was to identify the interval containing the signal for each set of stimuli. Four thresholds were measured for each of the four conditions. Adaptive runs with a reversal point standard deviation greater than 5 dB were discarded, and additional runs were completed to obtain four estimates of threshold for each condition. However, due to a programming error, only 3 threshold estimates were obtained for one of the forward masking conditions in 4 participant data sets.

The order in which the conditions were completed always began with a signal-alone condition and ended with a signal-and-precursor condition. In addition, same-frequency conditions had no more than 1 condition separating them in time. This ordering was used to ensure that participants were familiarized with the signal before completing the forward masking task, and resulted in 4 groups of 5 participants who completed the task in the same order.

RESULTS

QuickSIN results from each list tested were averaged to estimate the SRT of each participant. All scores fell in the normal/near normal to mild SNR loss range as indicated by the scoring guide.

Estimates of cochlear gain reduction were calculated by subtracting the average threshold for the signal alone condition from the average threshold for the broadband noise condition for each signal frequency. One outlier with a gain reduction estimate at 2 kHz that was greater than 2 times the interquartile range (Tukey's criteria) was excluded from further analysis. A within-subjects ANOVA [$F(1,18) = 9.66$, $p = 0.006$] revealed that gain reduction was significantly greater at 4 kHz ($M = 11.61$, $SD = 4.73$) than at 2 kHz ($M = 8.17$, $SD = 3.22$).

A similar relationship was seen between SRTs and estimates of gain reduction at 2 and 4 kHz. Participants with better speech-in-noise performance (lower SRT) had smaller gain reduction estimates than those with poorer speech-in-noise performance. In fact, a linear relationship was found between these two variables for 2 kHz [$r(17) = 0.70$, $p = 0.001$], excluding one outlier. However, the relationship was not statistically significant at 4 kHz [$r(18) = 0.32$, $p = 0.174$].

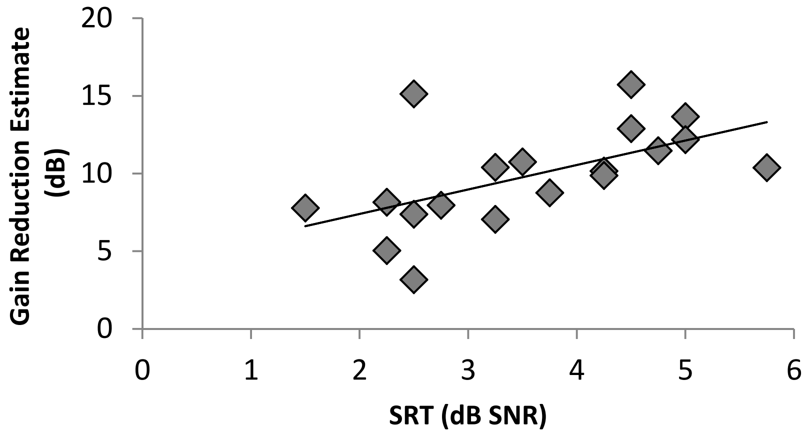

Fig. 1: Observed relationship between SRT and gain reduction estimates averaged across 2 and 4 kHz for 19 participants (one outlier excluded).

Figure 1 demonstrates the correlation between gain reduction estimates and speech-in-noise performance when gain reduction estimates were averaged across the two frequencies [$r(17) = 0.57$, $p = 0.010$]. One participant's data were again excluded according to Tukey's criteria for outliers.

DISCUSSION

The correlation showed that participants with better speech-in-noise performance had smaller gain reduction estimates than those with poorer speech-in-noise performance. The correlation was stronger when gain reduction was estimated at 2 kHz than when gain reduction was estimated at 4 kHz. This relationship between speech-in-noise performance and cochlear gain reduction is the opposite of that hypothesized.

This counterintuitive finding is similar to that found by de Boer *et al.* (2012), who used a different technique in an attempt to examine the same relationship. They used a consonant identification-in-noise task and compared those results to a reduction in OAE amplitude with contralateral stimulation. De Boer and colleagues found that participants with large contralateral suppression of OAEs performed poorer on the speech-in-noise task. They reasoned that the demand on attention is different between the two measures. It has been shown that the MOCR is under some attentional control (Delano *et al.*, 2007; Maison *et al.*, 2001). Since OAEs do not require the participant's attention, de Boer and colleagues (2012) hypothesized that differences in attentional control across the conditions could possibly explain their counterintuitive findings.

In our study, however, participants were actively engaged in both the measure of MOCR strength and the speech-in-noise task, which suggests that the attentional control explanation does not explain the observed relationship. Alternatively, perhaps there is something about the measure that explains this relationship. It is unlikely that activation of the MOCR leads to poorer perception in noise, given physiological data that suggests the opposite (Kawase *et al.*, 1993). Behavioral data also suggests that the MOCR improves perception in noise. For example, May *et al.* (2004) found that cats performed much better on a localization task in noise with their olivocochlear neurons intact.

The measure of speech perception in noise in this study involved an estimate of the SNR where performance was 50% correct. Because of this, each participant's SRT represented threshold performance at different SNRs. Although this is a valid way to measure a decrement in speech-in-noise performance, perhaps measurement at different SNRs is not the best choice to examine the relationship between MOCR strength and speech-in-noise performance. Research has shown that this method can confound data when an effect is SNR-dependent (Bernstein, 2012).

The MOCR may improve performance at certain SNRs and hinder performance at other SNRs. Kumar and Vanaja (2004) found that contralateral acoustic stimulation improved speech perception for ipsilateral SNRs of +10 and +15 dB, but not +20 dB. In hearing aid research, a parallel is the action of wide dynamic range compression (WDRC). When WDRC is activated, gain is provided by the hearing aid to the input sound, increasing the level presented to the ear. As the level of the input sound to the hearing aid rises, the hearing aid decreases the amount of gain provided. This variable gain has similarities to that provided by the outer hair cells in the inner ear. Research has shown that WDRC progressively decreases positive SNRs, especially for fast-acting multichannel compression and steady background noise (Alexander and Masterson, 2015). This body of research inspires the idea that a more systematic approach to measurement of speech perception in noise is preferable. By measuring performance at several SNRs, it will be possible to see if the relationship between ipsilateral cochlear gain reduction and speech-in-noise performance changes with SNR.

It is also possible that bilateral stimulation is needed to observe a beneficial relationship between cochlear gain reduction and speech-in-noise performance. The MOCR is, after all, a bilateral reflex. Natural listening situations such as cocktail parties, where MOCR activity could be beneficial, are situations where both ears are involved in listening to a target. There is possible interplay between cochlear feedback and localization cues.

The results of this experiment also bring to light individual differences. All participants passed a hearing screening at 15 dB HL, yet there was a range of both SRT and gain reduction estimates for these individuals. In the case of the SRT measurements, many participants with hearing thresholds in the normal range had SRTs outside of the normal range.

This study is the first to examine ipsilateral cochlear gain reduction with psychoacoustic methods at 2 kHz. This frequency may be more relevant to speech perception than 4 kHz, which is the frequency most frequently examined. This experiment is a first step in connecting psychoacoustic observations of cochlear gain reduction to speech perception, by showing that cochlear gain reduction is observed at a frequency with higher importance for speech intelligibility (Fletcher and Galt, 1950). This study is also the first to relate a psychoacoustic measure of cochlear gain reduction to speech-in-noise performance, allowing a comparison between two conditions where the ipsilateral MOCR pathway may be activated.

REFERENCES

Alexander, J.M. and Masterson, K.M. (**2015**). "Effects of WDRC release time and number of channels on output SNR and speech recognition," Ear Hearing, **36**, e35-e49.

Aguilar, E., Eustaquio-Martin, A., and Lopez-Poveda, E.A. (**2013**). "Contralateral efferent reflex effects on threshold and suprathreshold psychoacoustical tuning curves at low and high frequencies," J. Assoc. Res. Otolaryngol. **14**, 341-357.

Backus, B.C. and Guinan, J.J., Jr. (**2006**). "Time-course of the human medial olivocochlear reflex," J. Acoust. Soc. Am., **119**, 2889-2904.

Bernstein, J.G.W. (**2012**). "Controlling signal-to-noise ratio effects in the measurement of speech intelligibility in fluctuating maskers," Proc. ISAAR 2011: *Speech Perception and Auditory Disorders*. 3rd International Symposium on Auditory and Audiological Research, Nyborg, Denmark. Eds. T. Dau, M.L. Jepsen, T. Poulsen, and J.C. Dalsgaard. ISBN 978-87-990013-3-0 (The Danavox Jubilee Foundation, Copenhagen), pp. 33-44.

Bidelman, G.M. and Bhagat, S.P. (**2015**). "Right-ear advantage drives the link between olivocochlear efferent 'antimasking' and speech-in-noise listening benefits," NeuroReport, **26**, 483-487.

Cooper, N.P. and Guinan, J.J., Jr. (**2006**). "Efferent-mediated control of basilar membrane motion," J. Physiol., **576**, 49-54.

de Boer, J. and Thornton, A.R.D. (**2008**). "Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task," J. Neurosci., **28**, 4929-4937.

de Boer, J., Thornton, A.R.D., and Krumbholz, K. (**2012**). "What is the role of the medial olivocochlear system in speech-in-noise processing?" J. Neurophysiol., **107**, 1301-1312.

Delano, P.H., Elgueda, D., Hamame, C.M., and Robles, L. (**2007**). "Selective attention to visual stimuli reduces cochlear sensitivity in chinchillas," The J. Neurosci., **27**, 4146-4153.

Fletcher, H. and Galt, R.H. (**1950**). "The perception of speech and its relation to telephony," J. Acoust. Soc. Am., **22**, 89-151.

Ghitza, O. (**1988**). "Auditory neural feedback as a basis for speech processing," ICASSP, **88**, 91-94.

Giraud, A.L., Garnier, S., Micheyl, C., Lina, G., Chays, A., and Chery-Croze, S. (**1997**). "Auditory efferents involved in speech-in-noise intelligibility," NeuroReport, **8**, 1779-1783.

James, A.L., Harrison, R.V., Pienkowski, M., Dajani, H.R., and Mount, R.J. (**2005**). "Dynamics of real time DPOAE contralateral suppression in chinchillas and humans," Int. J. Audiol., **44**, 118-129.

Jennings, S.G., Strickland, E.A., and Heinz, M.G. (**2009**). "Precursor effects on behavioral estimates of frequency selectivity and gain in forward masking," J. Acoust. Soc. Am., **125**, 2172-2181.

Kawase, T., Delgutte, B., and Liberman, M.C. (**1993**). "Antimasking effects of the olivocochlear reflex. II. Enhancement of auditory-nerve response to masked tones," J. Neurophysiol., **70**, 2533-2549.

Killion, M.C., Niquette, P.A., Gudmundsen, G.I., Revit, L.J., and Banerjee, S. (**2004**). "Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am., **116**, 2395-2405.

Krull, V. and Strickland, E.A. (**2008**). "The effect of a precursor on growth of forward masking," J. Acoust. Soc. Am., **123**, 4352-4357.

Kumar, U.A. and Vanaja, C.S. (**2004**). "Functioning of olivocochlear bundle and speech perception in noise," Ear Hearing, **25**, 142-146.

Levitt, H. (**1971**). "Transformed up-down methods in psychoacoustics," J. Acoust. Soc. Am., **49**, 467-477.

Maison, S., Micheyl, C., and Collet, L. (**2001**). "Influence of focused auditory attention on cochlear activity in humans," Psychophysiology, **38**, 35-40.

May, B.J., Budelis, J., and Niparko, J.K. (**2004**). "Behavioral studies of the olivocochlear efferent system – Learning to listen in noise," Arch. Otolaryngol., **130**, 660-664.

Messing, D.P., Delhorne, L., Bruckert, E., Braida, L.D., and Ghitza, O. (**2009**). "A non-linear efferent-inspired model of the auditory system; matching human confusions in stationary noise," Speech Commun., **51**, 668-683.

Nelson, D.A., Schroder, A.C., and Wojtczak, M. (**2001**). "A new procedure for measuring peripheral compression in normal-hearing and hearing-impaired listeners," J. Acoust. Soc. Am., **110**, 2045-2064.

Roverud, E. and Strickland, E.A. (**2010**). "The time course of cochlear gain reduction measured using a more efficient psychophysical technique," J. Acoust. Soc. Am., **128**, 1203-1214.

Roverud, E. and Strickland, E.A. (**2014**). "Accounting for nonmonotonic precursor duration effects with gain reduction in the temporal window model," J. Acoust. Soc. Am., **135**, 1321-1334.

Strickland, E.A. (**2001**). "The relationship between frequency selectivity and overshoot," J. Acoust. Soc. Am., **109**, 2062-2073.

Wagner, W., Frey, K., Heppelmann, G., Plontke, S.K., and Zenner, H.-P. (**2008**). "Speech-in-noise intelligibility does not correlate with efferent olivocochlear reflex in humans with normal hearing," Acta Oto-Laryngol., **128**, 53-60.

Winslow, R.L. and Sachs, M.B. (**1987**). "Effect of electrical stimulation of the crossed olivocochlear bundle on auditory nerve response to tones in noise," J. Neurophysiol., **57**, 1002-1021.

Yasin, I., Drga, V., and Plack, C.J. (**2014**). "Effect of human auditory efferent feedback on cochlear gain and compression," J. Neurosci., **34**, 15319-15326.