IMPACT OF AUDITORY ATTENTION ON THE EFFERENT AUDITORY SYSTEM IN THE ABSENCE OF REAL AUDITORY TARGETS

Wing Yiu Stephanie Wong [*1], **Aravind Kumar Namasivayam, Ph.D.** [†1,2] **and Pascal van Lieshout, Ph.D.** [‡1,2,3,4,5]

[1]Oral Dynamics Lab, Department of Speech-Language Pathology, University of Toronto, 160-500 University Ave., ON, M5G1V7, Canada

[2] Toronto Rehabilitation Institute (TRI), Toronto, ON, Canada M5G 2A2

[3] Institute for Biomaterials and Biomedical Engineering (IBBME), University of Toronto, Toronto, ON, M5S 3G9, Canada

[4] Department of Psychology, Human Communications Lab, University of Toronto, Mississauga, ON, L5L 1C6, Canada

[5] Graduate Department of Rehabilitation Science, University of Toronto, Toronto, ON, M5G 1V7, Canada

Abstract

Previous studies have compared visual and auditory attention to no-task conditions and have demonstrated an attention-driven modulation of the efferent auditory system (De Boer & Thornton, 2007; Maison, Micheyl, & Collet, 2001). However, it is unclear whether these effects are modality-specific or a result of generalized attentional processes. In the present study, 16 young adults observed facial speech gestures related to productions of vowels /a/ and /u/ in the presence of contralateral broad band noise (BBN) under two instructions: (a) visual attention: visually count the number of /a/ productions and ignore BBN and (b) sham condition/ auditory attention: these trials did not have any vowels embedded in BBN, but participants were made to believe that there were sounds embedded and instructed to count the number of /a/ productions. These "sham" trials investigated the effect of auditory attention in the absence of real auditory targets. The influence of visual and auditory attention on the efferent auditory system was indirectly assessed by examining their effects on contralateral inhibition of click-evoked otoacoustic emissions (CS-CEOAE paradigm; Collet, Chanel, & Morgon, 1990). The mean inhibition from baseline for visual attention and auditory attention were 2.19 and 1.88 dB SPL, respectively. Cohen's *d* for the mean difference between the two conditions yielded a moderate positive effect size = 0.52. Twelve out of sixteen participants (75%; exact binomial test significant at one tailed $p = 0.03$) demonstrated a greater inhibition of CEOAEs amplitudes (mean difference = 0.31 dB SPL) in the visual attention condition relative to the auditory attention condition. Our results show that these effects are obtainable even in the absence of real auditory targets (i.e. without stimulus confound). Overall, finding a difference in inhibition of CEOAEs for visual and auditory attention conditions provide preliminary evidence for a modality-specific rather than a generalized attentional modulation in the efferent auditory system.

Keywords: Auditory attention, visual attention, contralateral inhibition, Transient-evoked otoacoustic emissions, efferent

Résumé

La comparaison de l'attention visuelle et auditive à des conditions sans-tâches a démontré une modulation du système efférent auditif dépendante de l'attention (De Boer & Thornton, 2007; Maison, Micheyl, & Collet, 2001). Cependant, il reste à déterminer si ces effets résultent de processus attentionnels généralisés ou de modalités. Dans cette étude, 16 jeunes adultes ont observé les mouvements du visage lors de la parole liés à la production des voyelles /a/ et /u/ en présence de bruit à bande large (BBN) controlatérale sous deux directives: (a) comptage visuel du nombre de production du /a/ en ignorant le BBN (attention visuelle) et (b) écoute soigneuse et comptage des sons cibles /a/ intégrés dans le BBN (condition feinte; attention auditive). Ces essais « feints » n'avaient pas de cibles acoustiques et reflètent l'effet de l'attention auditive en absence de véritables cibles auditives. L'influence de l'attention visuelle et auditive sur le système efférent auditif est mesurée par la inhibition controlatérale des otoémissions acoustiques provoquées (OEAP; Collet, Chanel, & Morgon, 1990). Les changements moyens du niveau de base pour l'attention visuelle et pour l'attention auditive sont respectivement de 2.19 et 1.88 dB SPL. La différence moyenne entre les deux conditions entraîne un effet positif modéré avec un d de Cohen de 0.52. Douze des seize participants (75%; valeur *p* du test binomial (unilatéral)= 0.03*) ont démontré une inhibition plus grande des amplitudes d'OEAPs (différence moyenne = 0.31 dB SPL) en condition d'attention visuelle qu'en condition d'attention auditive. Nos résultats démontrent que ces effets peuvent être obtenus même en absence de véritables cibles auditives. En résumé, l'observation d'une différence dans la inhibition de OEAPs entre les conditions d'attention visuelle et auditive fournit des preuves préliminaires soutenant une modulation attentionnelle spécifique plutôt qu'une modulation attentionnelle généralisée dans le système efférent auditif.

Mots clefs : attention auditive, attention visuelle, inhibition controlatérale, otoémissions acoustiques provoquées, efférent

*. wystephanie.wong@utoronto.ca

†. a.namasivayam@utoronto.ca

‡. p.vanlieshout@utoronto.ca

1 Introduction

In our day-to-day life, selective attention helps us tune in to relevant stimuli and ignore distractors as we try to make sense of the world around us. Research suggests that selective attention may be related to mechanisms that enhance relevant information or suppress irrelevant information [1]. Further, several studies have reported that attentional processes modulate the peripheral cochlear mechanisms [2,3,4], which are modulated by the efferent (descending) auditory pathway, specifically the medial olivocochlear (MOC) tracts. The MOC fibres are the only known descending connection between the corticofugal tracts originating from the auditory cortex and the cochlea, allowing top-down corticofugal modulation of the auditory system on a peripheral level [5,6]. Several animal studies have indicated that MOC tracts emerge from the superior olivary complex (SOC), and innervate the outer hair cells (OHCs) of the contralateral (75%) and ipsilateral (25%) cochlea [7,8]. The effects of corticofugal modulation of the peripheral auditory system can be indirectly assessed by examining their impact on the contralateral inhibition of evoked otoacoustic emission (OAE). OAEs are a byproduct of the cochlear amplifier and normal function of outer hair cells (OHC). In healthy ears, they can be recorded in the ear canal either spontaneously or in response to acoustic stimulation [9].

It has been reported that both visual and auditory attention leads to changes in OAEs, signifying a top-down modulation of the peripheral auditory system. For the visual system, attending to visual tasks (such as counting visual events) leads to an increase in contralateral inhibition (decrease OAE amplitude) relative to non-attending tasks [10, cf. 11]. In terms of auditory attention, attending to stimuli in the contralateral ear has also been shown to decrease contralateral inhibition compared to non-attending tasks [12]. However, given that both visual and auditory attention impact OAE amplitudes, it remains unclear whether these effects are modality-specific or a result of generalized attentional processes.

In the present study, we explored whether auditory attention, compared to visual attention, differentially modulates activity in the efferent auditory system. We investigated this using a well-reported procedure for assessing efferent auditory system modulation, which involves the presentation of broad band noise (BBN) in the contralateral ear and measuring OAE in the ipsilateral ear (CS-OAE paradigm; [13]). In this procedure, contralateral BBN is presumed to stimulate ipsilateral SOC via crossed efferent pathways; this in turn activates descending ipsilateral MOC fibres. Given that MOC fibres terminate at OHCs, it is assumed that they are in a position to modify the actions of OHCs and hence, modulate the gain of the cochlear amplifier and OAEs [7,14]. However, the resulting changes in OAEs may be a result of both active (OHC electromotility) and passive mechanisms (linear reflection along the cochlear partition) [15]. We hypothesize that cortically mediated release from MOC activity (i.e. level of contralateral OAE inhibition) at the level of cochlea would differ between tasks involving visual attention vs. auditory attention even when physical stimuli are identical. Such a differential response, if found, will support the influence of a modality-specific attentional process, as opposed to a more generalized attentional mechanism.

2 Method

2.1 Subjects

Sixteen young healthy adults (Mean age (S.D.) = 22.0 (3.16) years; Males= 4, Females= 12) participated in the study. All participants were right-handed, native English speakers, with no history of speech, language, learning, neurological, or otological issues, or noise exposure in the last 24 hours prior to the experiment. All participants met the following otological criteria: (a) normal tympanic membrane/ ear canal appearance on otoscopic examination, (b) bilateral audiometric thresholds between 500 Hz to 4000 Hz at 20 dB HL or lower, (c) normal middle ear function, exhibiting ear canal pressure values between -100 and +50 daPa, middle-ear compliance values between 0.3 and 1.6mL, and acoustic reflex thresholds ≥ 65 dB SPL. All participants were reimbursed at a standard fee of $10 CDN/hour. The study was approved by the University of Toronto's Health Sciences Research Ethics Board and participants provided informed consent prior to the start of the study.

2.2 Stimuli and Procedures

We used click-evoked OAEs (CEOAEs) elicited with clicks presented in a linear mode (same polarity) with the amplitude of 60 dB peak SPL (click duration of 80μs, click interval of 21.12 ms.) The responses were collected by averaging among 260 stimuli trains (1040 clicks), which was stored in two buffers (A and B) for a total of 2080 clicks. Whole wave reproducibility (WWR) was calculated as the Pearson correlation coefficient between the two obtained waveforms (A and B) and multiplied by 100. WWR is considered a quality index of the recorded OAEs; in the present study, WWR was set at > 70% as suggested by previous research [16,17]. The responses elicited were high and low pass filtered between 750 and 6000 Hz, respectively, with a recording window between 2.5 to 20.0 ms. CEOAEs at 2kHz centre frequency were recorded via the Vivosonic Integrity 4.5.3 system, with artifact rejection threshold of 45 dB SPL. A Signal-to-Noise Ratio (SNR) of > 6 dB was used as a criterion of CEOAE detection [18]. In the current study, we only analyzed CEOAEs centred around 2 kHz for three reasons: (1) this frequency yielded the largest SNR ratios across all participants, (2) contralateral inhibition effects are not strong above 3 kHz [19] and (3) most typical frequencies related to speech perception are < 3 kHz [20].

The study was conducted in a standard sound attenuated booth with a two-way observation window separating the control room and test room. The experimenter in the control room provided all instructions, presented different task conditions and controlled the stimuli presentation via a

Microsoft PowerPoint presentation on a laptop computer. The second experimenter sat next to the participant and carried out all CEOAE recordings, including probe fit monitoring on a trial-by-trial basis.

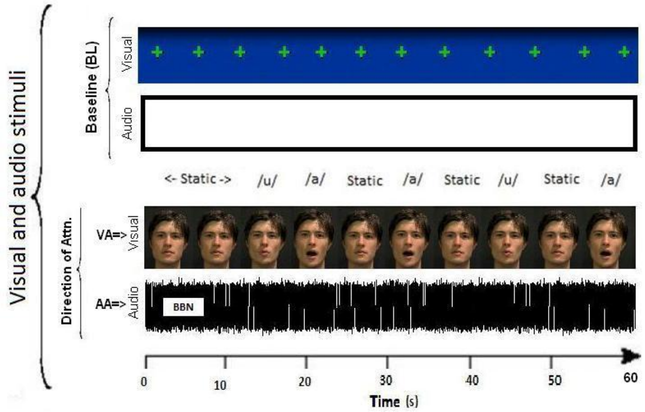

Figure 1. Task conditions: (a) Baseline (BL) condition: no contralateral BBN, (b) VA condition: contralateral BBN + attention (attn.) directed to visually observing speech gestures related to productions of vowels /a/ and /u/ (c) AA condition: contralateral BBN + attention directed to auditory stimuli (sham-condition). X-axis represents time in seconds.

CEOAEs were recorded from the right ear under 3 task conditions (Figure 1). The first condition was the baseline (BL) condition, in which participants focused their attention on a "+" symbol displayed on a computer monitor without any contralateral BBN. In the other two conditions, the participants were presented with continuous contralateral BBN, generated by a Grason-Stadler 61 (GSI-61) audiometer and delivered in the left ear at 55 dB HL via an ER-3A insert earphone. Real-ear or "in-situ" responses were measured (using a probe microphone real-ear measurement system; Audioscan RM500) for such BBN levels at the eardrum, and were found to be equivalent to 63-73 dB SPL (for frequencies between 750 to 4000 Hz) with roll offs at the higher and lower frequencies [18]. This noise level is the highest level of BBN that could be presented without eliciting acoustic reflexes [11, 12, 21]. While BBN was delivered, participants were also presented with a video of a man producing facial speech gestures related to productions of vowels /a/ and /u/ in both task conditions. In the visual attention condition (VA), the participants were instructed to mentally count the number of times they saw the person's face produce an /a/ speech gesture and ignore BBN. Prior to the start of the VA condition, participants were given two practice trials to familiarize themselves with the task condition. In the auditory attention (sham) condition (AA), we presented the subjects with a "practice" trial in which /a/ and /u/ sounds were embedded in BBN in different SNR (i.e. +10, +5, 0, -5 and -10). The participants were instructed to listen carefully to detect and mentally count the number of target sound /a/ embedded in BBN Importantly, the "sham" trials differed from the "practice" trial in that they did not have any real acoustic stimuli embedded in BBN. Furthermore, the "practice" trials were also used as random catch trials throughout the study to convince participants that there were vowel targets embedded in the BBN in the sham trials; OAEs from these catch trials were not recorded. In fact, participants were presented with identical visual and auditory stimuli in both the VA and the AA trials, and the only difference between the conditions was the information channel (visual/ auditory) to which they were instructed to direct their attention. This controlled for stimulus confound and probed the effect of auditory attention even when there was no real acoustic target. Notably, all participants reported "hearing" at least one embedded target in the "sham" trials, indicating that they were indeed paying auditory attention. There were 5 trials per block: the first block was always BL trials, followed by VA or AA trials, with the order of the latter two counterbalanced across participants. Trials within each block were also randomized; each trial lasted approximately 60 seconds, and was matched for both number of productions and movement duration of each /a/ or /u/ production (as timed with a metronome). Interstimulus interval (ISI) between any two visual speech gestures ranged from 1s to 6s, wherein all speech gesture presentation began at about 15s after the onset of BBN.

3 Results

The means and standard deviations (in parenthesis) for CEOAE amplitude (in dB SPL) across the 2kHz frequency band are depicted in Table 1. The mean of the VA block (or AA block) was subtracted from the mean of the BL block within a participant to derive a score representing change from baseline (ΔVA and ΔAA).

Table 1. Means (standard deviation) in dB SPL across 16 participants for 2kHz CEOAE test frequency (see text for more details).

CEOAE Frequency Band	BL	VA	AA	ΔVA	ΔAA
2kHz	3.36 (6.66)	1.17 (6.27)	1.48 (6.34)	2.19 (1.98)	1.88 (1.82)

The mean difference between the two conditions yielded a moderate positive effect size (Cohen's *d* adjusted for repeated measures = 0.52) [22, 23]. 75% of the participants tested (12 out of 16 participants; exact binomial test significant at one-tailed $p = 0.03$) exhibited an increase in inhibition of 0.31 dB SPL in the visual attention (VA) task relative to the auditory attention (AA) task.

4 Discussion

The current study investigated whether visual and auditory attention differentially modulates the peripheral auditory system. Overall, the presence of contralateral BBN inhibited CEOAE amplitude responses in the test ear across both attentional conditions, relative to baseline. The amounts of inhibition (see Table 1), as indicated as change from baseline, in the attentional conditions were 2.19 dB SPL (for

VA) and 1.88 dB SPL (for AA). Notably, our results show that auditory attentional effects are obtainable even in the absence of real auditory targets (i.e. without stimulus confound). Further, despite the identical physical stimuli presentation of the two conditions, a significant increase in inhibition of about .31 dB SPL was observed for the VA task, relative to the AA task (Table 1).

Such small differences (~0.35dB) found across attentional task conditions are not unusual and have been reported in other studies [2, 18]. Small changes in the amount of inhibition between conditions have larger implications if one takes into account the presumed role of MOC fibres and the efferent pathway. MOC fibre activity is assumed to have an inhibitory effect on OHC's electromotility, which is reflected in OAE inhibition. Evidence in the literature suggests that even small changes in the cochlear mechanics are able to alter target-specific input gain in the peripheral auditory system, resulting in increased signal amplitudes in the ascending auditory nerve fibres [21, 24].

Previous studies have also reported an increase in OAE inhibition from baseline during visual attention tasks [2, 3]; however, since the methodologies of these studies involved different stimuli during auditory attention and visual attention tasks, it was unclear whether selective attention was the only variable manipulated. In the current study's paradigm, given that VA and AA conditions employed the same stimuli and only differed in instructions of directing either visual or auditory attention, the differences observed between the conditions suggest a modality-specific rather than a generalized attentional modulation in the efferent auditory system. Alternatively, these effects may also be explained in terms of differences in neuronal bandwidths, wherein BBN and auditory attention may share the same neuronal bandwidth while visual attention may have access to additional bandwidth, either anatomically or functionally (e.g. [25]).

A potential limitation in the study is that the instrument we utilized (Vivosonic Integrity 4.5.3) does not allow for the time-locked recording of OAE with stimuli presentation. Thus, artifact rejection was not synchronized with the presentation of stimuli (and hence our blocked presentation approach). However, to ensure that there were no systematic differences in artifact rejection that could have biased the data towards a specific condition, we carried out a within participant post-hoc analysis on artifact rejection ratio (AAR%) across conditions. The results of this analysis did not reveal any systematic differences in AAR% across conditions within a participant. Thus, the condition effects in the present study are less likely due to differences in artifact rejection.

Another potential limitation is that, given the study's design, it is not possible to separate the sole influence of BBN from the effects of attention. However, since the aim of the study was to explore differences in modulation of OAE as a function of the direction of attention, the test conditions (VA and AA) have BBN as a common factor for we do not expect the influence of BBN across test conditions to be different.

Building on previous findings of both visual and auditory attention having an impact on OAE amplitude, the current results seem to indicate that the channel through which attention is directed may have the potential to differentially modulate efferent cochlear mechanisms.

Acknowledgments

The authors would like to thank Dr. Marshall Chasin for his assistance with real-ear noise measurements. The authors acknowledge the support of Toronto Rehabilitation Institute, which receives funding under the Provincial Rehabilitation Research Program from the Ministry of Health and Long-Term Care in Ontario. The views expressed do not necessarily reflect those of the Ministry. This research was undertaken, in part, thanks to funding from the Canada Research Chairs program, awarded to the last author.

References

[1] J. W. Couperus, and G. R. Mangun. Signal enhancement and inhibition during visual–spatial selective attention. *Brain Research*, 1359: 155-177, 2010.

[2] C. Meric, and L. Collet. Visual attention and evoked otoacoustic emissions: a slight but real effect. *International Journal of Psychophysiology*, 12(3): 233-235, 1992.

[3] P, Froehlich, L. Collet, and A. Morgon. Transiently evoked otoacoustic emission amplitudes change with changes of directed attention. *Physiology & Behavior*, 53(4): 679-82, 1993.

[4] S. Maison, C. Micheyl, and L. Collet. Influence of focused auditory attention on cochlear activity in humans. *Psychophysiology*, 38(01): 35-40, 2001.

[5] X. Perrot, P. Ryvlin, J. Isnard, M. Guénot, H. Catenoix, C. Fischer, and L. Collet. Evidence for corticofugal modulation of peripheral auditory activity in humans. *Cerebral Cortex*, 16(7): 941-948, 2006.

[6] S. Srinivasan, A. Keil, K. Stratis, K. L. Woodruff Carr, and D.W. Smith. Effects of cross-modal selective attention on the sensory periphery: Cochlear sensitivity is altered by selective attention. *Neuroscience*, 223 : 325-332, 2012.

[7] G. L. Rasmussen. The olivary peduncle and other fiber projections of the superior olivary complex. *Journal of Comparative Neurology*, 84(2): 141-219, 1946.

[8] W. B. Warr, and J. J. Guinan Jr. Efferent innervation of the organ of Corti: two separate systems. *Brain Research*, 173(1): 152-155, 1979.

[9] J. J. Guinan Jr. Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. *Ear and Hearing*, *27*(6), 589-607, 2006

[10] P. Froehlich, L. Collet, J. M. Chanal, and A. Morgon. Variability of the influence of a visual task on the active micromechanical properties of the cochlea. *Brain Research*, 508(2): 286-288, 1990.

[11] J. De Boer, and A. R. D. Thornton. Effect of subject task on contralateral inhibition of click evoked otoacoustic emissions. *Hearing Research*, 233(1): 117-123, 2007.

[12] A. W. Harkrider, and C. D. Bowers. Evidence for a cortically mediated release from inhibition in the human cochlea. *Journal of the American Academy of Audiology*, 20(3) : 208-215, 2009.

[13] L. Collet, D. T. Kemp, E. Veuillet, R. Duclaux, A. Moulin, and A. Morgon. Effect of contralateral auditory stimuli on active cochlear micro-mechanical properties in human subjects. *Hearing Research*, 43(2): 251-261, 1990.

[14] J. J. Guinan Jr. Cochlear efferent innervation and function. *Current Opinion in Otolaryngology & Head and Neck Surgery*, 18(5): 447, 2010.
[15] D. T. Kemp. Otoacoustic emissions, their origin in cochlear function, and use. *British Medical Bulletin*, *63*(1), 223-241, 2002.
[16] M. S. Robinette, and T. J. Glattke. *Otoacoustic Emissions: Clinical Applications*. Thieme, 2007.
[17] E. Berninger, E. Characteristics of normal newborn transient-evoked otoacoustic emissions: Ear asymmetries and sex effects. *International Journal of Audiology*, *46*(11), 661-669, 2007.
[18] A. K. Namasivayam, W. Y. S. Wong, D. Sharma, and P. van Lieshout. Visual speech gestures modulate efferent auditory system. *Journal of Integrative Neuroscience*, 14 (1), 73–83, 2015.
[19] A. Moulin, L. Collet, E. Veuillet, and A. Morgen. Interrelations between transiently evoked otoacoustic emissions, spontaneous otoacoustic emissions and acoustic distortion products in normally hearing subjects. *Hearing Research*, *65*(1), 216-233, 1993.
[20] P. Ladefoged. *A Course in Phonetics* (5th ed., international student ed.). Boston: Thomson Wadsworth, 2006.
[21] A. C. Garinis, T. Glattke, and B. K. Cone. The MOC reflex during active listening to speech. *Journal of Speech, Language, and Hearing Research*, 54(5): 1464-1476, 2011.
[22] J. Cohen. *Statistical Power Analysis for the Behavioural Sciences.* Routledge Academic, 2013.
[23] S. B. Morris, and R. P. DeShon. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychological Methods*, 7(1): 105, 2002.
[24] S. Puria, J. J. Guinan Jr., and M. C. Liberman. Olivocochlear reflex assays: effects of contralateral sound on compound action potentials versus ear‐canal distortion products. *The Journal of the Acoustical Society of America*, 99(1): 500-507, 1996.
[25] M. R. DeWeese, T. Hromádka, and A. M. Zador. Reliability and representational bandwidth in the auditory cortex. *Neuron*, *48*(3), 479-488, 2005.