A hypothetical mechanism of auditory processing for extraction of directional cues. Integration with oculomotor function

V.A. Gorelik
Neuronix
7003 Sand Road
Savannah, GA 31410

***Abstract* – A hypothetical mechanism involving delayed forced oscillations of Outer Hair Cells (OHC) in conjunction with the Tectorial membrane is introduced here to explain noise filtering, high dynamic range and efficient extraction of directional cues in the mammalian auditory system. It will be shown how an ensemble of individually slow but highly synchronized neurons can produce a phase-correlated high frequency oscillation in a row of OHCs at a given tonotopic location. We will also suggest how superposition of two forces – acoustical pressure difference between scala tympany and scala vestibuli and neurally-induced oscillations - produce synchronized deflection of a row of OHCs and lead to precisely controlled phase selectivity, and how such phase selectivity can be attributed to sound localization and noise suppression. Finally we will show how a mechanism for extraction of auditory directional cues is linked to the oculomotor function. It is also suggested that functional and anatomical proximity of these two systems assist each other to warrant low-level auditory and visual situation awareness.**

I. INTRODUCTION

Attempts to explain the ability of the mammalian auditory tract to extract directional cues from an incoming acoustic signal [1] are facing some difficulties due to an inconsistency between the auditory frequency spectrum and neurons' maximum firing rate. Mechanisms based on phase-sensitivity were previously ruled out [2] as physiologically impossible. To contradict this belief, a new physiologically plausible mechanism that attempts to explain phase sensitivity of the organ of Corti in conjunction with its neural organization is proposed here. Even though an individual neuron cannot fire at a frequency much above several hundred Hertz [3], the proposed hypothesis provides the means for cooperative neural activation of OHCs that greatly exceeds the highest firing rate of a single neuron. Additionally, it will be shown that the characteristic frequency of a given tonotopic location is determined by the motility of the OHCs [4], their number per tonotopic location, the number of efferent neurons innervating each row of outer hair cells, but not by the dynamics of an individual neuron. The suggested mechanism allows selective frequency and phase tuning of each tonotopically-defined group of outer and inner hair cells. This is accomplished in such a way that acoustic stimuli with specifically defined phase shifts between left and right cochleae are amplified, while others are attenuated, hence providing tightly controlled directional localization of the source of sound.

Some mammals possessing echolocation capabilities respond to frequencies in excess of 60KHz. Bats can extract auditory directional cues at even higher frequencies [5] – behavior that is very difficult to explain without invoking phase-sensitive techniques. Several algorithms [6] were proposed previously to work around this problem; yet another neural mechanism is suggested here. It is based on the assumption that a row of outer hair cells in a mammalian cochlea can be forced into sustained high frequency harmonic oscillations by cooperative action of an ensemble of relatively slow neurons. In addition to traveling wave action in the cochlea's tectorial membrane, this mechanism explains some previously puzzling behaviors of the cochlea. It will be shown here that under some conditions OHCs can actively drive the tectorial membrane into oscillations at a given tonotopic frequency, which may explain some functional intricacies of the cochlea's behavior. It explains how a human cochlea can respond to frequencies far in excess of the neuron's maximum "firing" rate; it also gives additional explanation of the nature of otoacoustic emission [7], [8] and sound localization in the horizontal plane.

The fundamentals of the mechanism responsible for the high-frequency oscillations in the tectorial membrane will be described in section II. The neural structure providing support for this mechanism is introduced in section III. Section IV correlates the preceding discussion with some factual and relevant information from neurophysiology and neuroanatomy of corresponding structures in the midbrain, pons and medulla. Conclusions and some additional remarks are made in section V.

II. FREQUENCY AND PHASE SENSITIVITY AS THE RESULT OF OHC-INDUCED OSCILLATIONS

Certain properties of the mammalian cochlea and its outer hair cells allow some assumptions to be made about their functionality. Among these properties are the cochlea's otoacoustic emission and OHCs' motility. Additionally, anatomical evidence suggests that the OHCs' length and mechanical stiffness are proportional to their tonotopic position along the basilar membrane [9]; lower frequency locations correspond to longer cells. Higher frequency cells receive efferents from multiple neurons, while lower frequency cells are innervated by a single neuron [10]. It is also known that a single neuron cannot fire in excess of several hundred Hertz. This last fact imposed one of the most stringent conceptual limitations on previous attempts to model cochlear dynamics. It also left otoacoustic emission

without an adequate explanation. It appears that the motility (ability to oscillate) of any given OHC is much higher than the frequency response of a corresponding efferent [11]. Lastly, efferents innervating outer hair cells are unmyelinated and very thin, which makes them bad conductors with significant propagation delays.

Otoacoustic emission can be explained only by the existence of precisely controlled oscillations induced by the collective behavior of an ensemble of outer hair cells. The following is a hypothetical mechanism that allows highly organized arrays of low-frequency neurons to induce high-frequency oscillations in outer hair cells and thus in the tectorial membrane.

It is assumed here that 1) the tectorial membrane is a linear and non-dispersive medium, 2) an OHC and its stereocilia are capable of mechanical oscillations at or above the characteristic frequency of a given tonotopic location, 3) the impulse response of an OHC uncoupled from the tectorial membrane is a rapidly decaying high-frequency oscillation, 4) a displacement of the tectorial membrane by an OHC creates a transverse wave propagating in both directions from the point of excitation, 5) a delay is introduced as action potential propagates along an axon, and 6) an abrupt impedance change at the edges of the tectorial membrane is ignored. Also, considering that the tectorial membrane is a damping element at a "native" OHC's frequency, it "sees" this decaying oscillation as a delta function ($\delta(t)$).

Fig. 1 shows a single tonotopic location with three outer hair cells coupled to the tectorial membrane. Outer hair cells in a human's cochlea are organized in rows of three or four per tonotopic location. Additionally, each high frequency tonotopic location is innervated by a bundle of several neurons. Each axon in the bundle branches out to form a synapse with an OHC as it passes by it. A delay is introduced by each neural efferent as the action potential propagates along the thin unmyelinated axon from the first OHC in a row to the last. This delay forces the next OHC in the row to respond with a phase shift from the previous cell. For a three-cell tonotopic location the displacement of the tectorial membrane under the influence of a harmonic actuation from three OHCs can be described as:

$$X(t(k+1)) = F([\delta(t), \delta(t+\tau), \delta(t+2\tau)]) \quad (1)$$

where: $X(t)$ – is a harmonic function representing the transverse displacement of the tectorial membrane at the location of the inner hair cell, t - is the moment in time when the efferent's action potential arrives at the location of the first OHC in the row, τ- is the propagation delay between two consecutive OHCs, and k – is a positive constant. A harmonic envelope with a characteristic frequency, determined by τ, is produced by the apexes of the OHCs at a given tonotopic location.

A bundle of thin unmyelinated axons innervates each row of corresponding outer hair cells at the high-frequency tonotopic locations. It is assumed here that each neuron in the bundle is activated with a specific delay in reference to a previously active neuron. Fig. 1 also shows how a transport (propagation) delay along the axon of a single Auditory Efferent Master Neuron provides delayed presynaptic activation to five Auditory Efferent Slave Neurons. Each slave neuron "fires" at a relatively low frequency defined by the master neuron. However, multiple delayed slave neurons form synapses on each OHC in the row, resulting in a combined phase-shifted low frequency activation of the corresponding OHC and, for that matter, the entire tonotopic group.

Four timing diagrams are shown at the bottom of Fig. 1. The topmost diagram represents the activity of the Auditory Efferent Master Neuron at the branching point with the first (uppermost) slave neuron, which is identical to the activity of the first slave neuron. As the depolarization event propagates along the axon of an Auditory Efferent Master Neuron, it consequently "triggers" five slave neurons. The slave transport delay is the product of the spacing between the slave neurons' postsynaptic terminals along the axon of the Auditory Efferent Master Neuron and the speed of the action potential in the axon of the Master Neuron. Each slave neuron in the bundle "fires" at a frequency that is equal to the depolarization rate of the corresponding Auditory Efferent Master Neuron. However, depolarization of the next slave neuron in the bundle is shifted in time by the slave transport delay. The slave transport delay is a fraction of the period of oscillation at a given tonotopic location.

The lower three diagrams in Fig. 1 show resulting presynaptic activation that is "seen" by the first, second and third OHCs correspondingly. Each slave neuron innervates all three or four OHCs in the row with "OHC transport delay," defined as the product of spacing between the OHCs and the speed of the action potential in the axon of a slave neuron. OHCs receive at least one synapse from each slave neuron in the bundle. The resulting activation is a superposition of phase-shifted activations originating from a bundle of five slave neurons. It is clear from the diagram

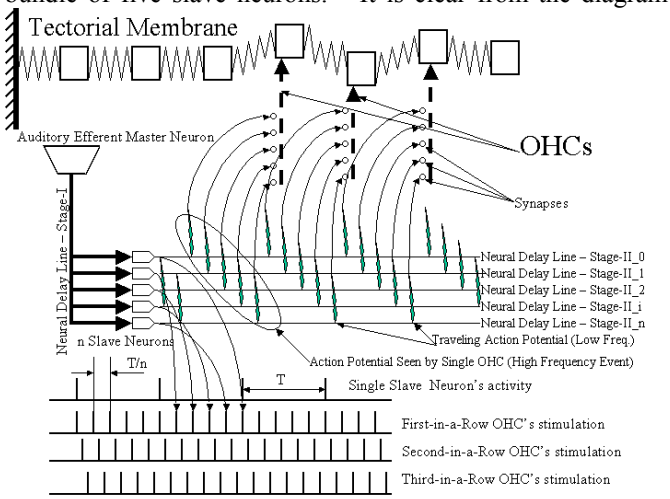

Fig. 1

that each OHC is activated at a frequency which exceeds the frequency of a single slave neuron by a factor equal to the number of slave neurons synapsing on each OHC. It is also clear from the diagram that each OHC receives its activation with a delay in reference to the previous cell.

Each OHC at a given tonotopic location (specific frequency) receives a presynaptic activation with a phase shift of 120 degrees for a three-OHCs-per-ensemble or 90 degrees for four-OHCs. As a result, a row of outer hair cells that is mechanically coupled to the tectorial membrane forces the membrane to flex into oscillation at a frequency much higher than would be possible if it were innervated by a single neuron.

Inner hair cells are embedded in the bony spiral lamina and thus their bases are immobile, while their stereocillia acts as an effective transducer of the shearing force from the tectorial membrane in response to external acoustical stimuli. Fig. 2 shows that unlike IHC, both bases and apexes of outer hair cells are embedded into the flexible basilar and tectorial membranes respectively, which introduces some additional complexity but will be ignored at this time. One end of the tectorial membrane is embedded into the spiral lamina; the other end is free to move and is subjected to the transverse displacing force from the corresponding outer hair cells and acoustic pressure, which propagates along the cochlea channel and displaces the basilar membrane.

The IHC responds with amplification if the displacing force induced by the OHC is in-phase with the acoustic stimuli or with attenuation if it is out-of-phase, as shown in Fig. 2. Another assumption that is made here is that IHCs act as signal rectifiers, whose responses are proportional to the amplitude of the local oscillations of the tectorial membrane, but not to the local frequency of oscillations. This mechanism allows phase-selective amplification and attenuation at each tonotopic location of the organ of Corti. It also accounts for the very large dynamic range of the cochlea's acoustic responsiveness to sound pressure.

The propagation speed of an action potential in a typical

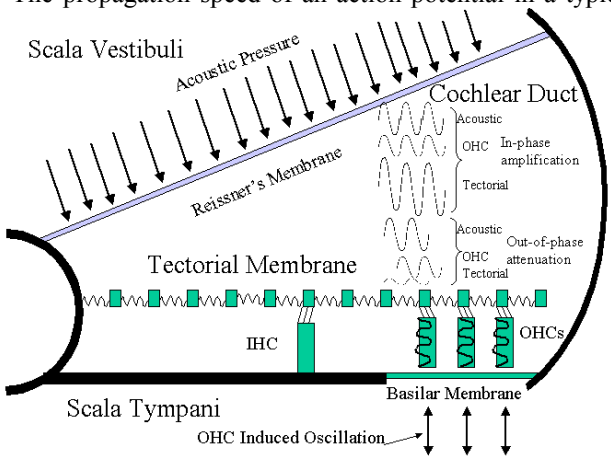

Fig. 2

and relatively thick myelinated axon is approximately 30m/sec or 30,000,000 μm/sec. It is reasonable to expect that this parameter for a thin unmyelinated efferent cochlear axon is in the range of 1m/sec. or 1,000,000 μm/sec [12], in which case approximately 5 neurons are required to sustain a long-term activation of a single 20KHz tonotopic location, assuming that there are four outer hair cells per row spaced apart by 10 μm, and a slave neuron requires 10,000 μsec to recover after each depolarization event.

Additional information is needed to prove the physiological validity of the proposed mechanism. One way to prove it is to show that low frequency tonotopic locations are indeed innervated by myelinated axons or employ another mechanism that can attribute to shortening of the propagation delay. It is known that high-frequency OHCs are innervated by multiple neurons and the higher the frequency the more neurons form synapses with the cell. Another proof would be to show that such cells are innervated by multiple time-shifted synapses with accumulated time shift over the population of neurons equal to or multiple of the period of a given tonotopic frequency.

III. Underlying Neural Organization

Fig. 3 shows the neural micro-circuitry and corresponding timing relationships that are hypothesized to be responsible for the behavior of the organ of Corti. The proposed mechanism can synchronize fast dynamics of outer hair cells with slow neurons to produce an adequate frequency response. However, afferent pathways from inner hair cells are not included in the current analysis as the related information is available in a large body of previous work.

This figure presents only a functional description of the hypothesized mechanism; for clarity and convenience we keep the names of various elements similar to the names of actual neuroanatomical structures. However, note that if the proposed mechanism does exist in the human brain, some or all of the components shown may not necessarily "map" directly onto the anatomical structures with same names. For example, components of the spiral ganglion in Fig. 3 may actually be located in one of the cochlear nuclei, etc.

Fig. 3 shows how two Auditory Efferent Master Neurons representing one tonotopic location in the left and right cochleae are integrated with a 2D cluster of Θ-F-Phase-Selecting Neurons, where Θ-F signifies phase-frequency. Each tonotopic location's frequency response is not sharply tuned to a specific auditory stimulus; instead it works as a bandpass filter and therefore may be tuned to a reasonably broad spectrum of frequencies centered on a specific median tone. To reiterate, a 2D cluster of Θ-F-Phase-Selecting Neurons is organized in rows where all neurons in the same row can fire only at one specific frequency. The number of rows per 2D array is roughly equal to the number of IHC-OHCs rows in the organ of Corti. There are also probably some lateral connections between rows to allow for tuning of each tonotopic location to a spectrum of frequencies, which

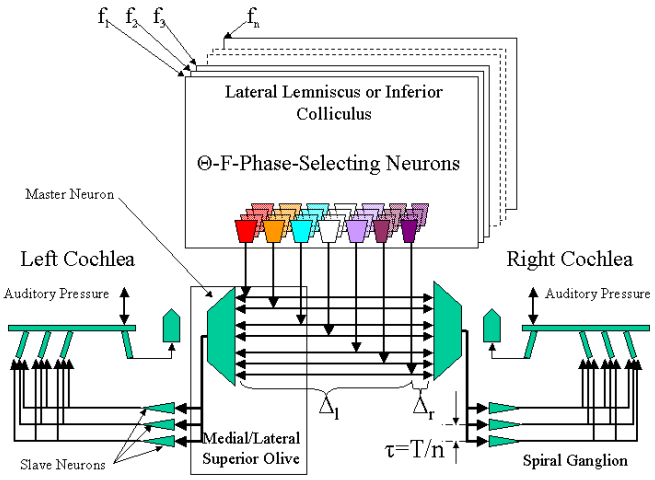

Fig. 3

means that several Θ-F-Phase-Selecting Neurons from adjacent rows but belonging to the same column may synapse on a single tonotopicaly specific slave bundle.

All Θ-F-Phase-Selecting Neurons belonging to the same row are identical. However, each of these neurons has an axon branching out in two directions - to the left and right Auditory Efferent Master Neurons corresponding to the left and right cochleae. Each of these neurons is responsible for innervating a specific tonotopically-defined group. Each of the Θ-F-Phase-Selecting Neurons in a row occupy a fixed spatial location in the brain, and although they all have the same characteristics, they need different lengths of the left and right axonal branches to reach their corresponding targets - Auditory Efferent Master Neurons at the left and right. They therefore provide uniquely delayed action potentials to the left and right cochleae. For example, to reach the left Auditory Efferent Master Neuron the action potential from the right-most Θ-F-Phase-Selecting Neuron in Fig. 3 will experience a delay equal to Δl, while requiring only Δr to reach the right Auditory Efferent Master Neuron. Activation of one of the Θ-F-Phase-Selecting Neurons in a row of the 2D array tunes both cochleae to a specific frequency and specific phase shift, effectively amplifying the auditory input emerging from a specific direction while attenuating the same frequencies emanating from all other directions. Such behavior explains why we cannot effectively tune our audition to recognize separate ongoing discussions from spatially separated groups during a cocktail party.

The maximum acoustic delay between the left and right cochleae separated by 10cm is ~300 μsec. It will be necessary to have at most only ~3.3mm separation between two Θ-F-Phase-Selecting Neurons in a row to achieve the required delay in unmyelinated axons; this distance is consistent with the sizes of several cochlea-related nuclei. The result of such anatomical intricacy is manifested in binaural phase sensitivity to incoming acoustic stimuli due to acoustic propagation delay for each orientation and matched axonal delay in corresponding Θ-F-Phase-Selecting neurons.

IV. Neuroanatomy of the Human Auditory Tract and its Correlation with Oculomotor Function

In this section we will analyze topographic arrangements of oculomotor and auditory nuclei to derive some of their intricate functional and behavioral relationships.

First – both the eyes and the cochleae in humans are roughly in the same coronal plane.

Second - eye motions are controlled by three groups of muscles, as seen in Fig. 4: superior/inferior oblique (rotation around optical axis), superior/inferior rectus (up/down) and lateral/medial rectus (left/right).

Third - three nerves originating from two relatively distant areas of the brain stem innervate these muscles.

Three of four rectus' muscles are controlled by oculomotor - Cranial Nerve III (CN-III), which originates in the upper midbrain [13], [14]. The remaining lateral rectus is innervated by CN-VI, which takes origin from the lower pons.

CN-III and CN-VI are of the most importance to our analysis: CN-III originates from the oculomotor nucleus at the level of the superior colliculus and in close proximity to the inferior colliculus and lateral lemniscus; it innervates four oculomotor muscles. Both left and right oculomotor nuclei that give origin to CN's-III are located far above the pons and almost on the central sagittal section of the brain stem, and thus are in a very close proximity to each other.

The only muscle that is not innervated by CN-III is the lateral rectus that rotates the eye away from the nose. CN-VI controls this muscle - it originates from the abducens nucleus in the caudal pons at the level of facial colliculus in a very close proximity to CN-VIII. CNs-VIII or the vestibulo-cochlear nerves originate from the vestibulocochlear nuclei and CN-VI and CN-VIII are located very far below all other oculomotor functions.

Unlike points of origin of both left and right CNs-III, which almost coincide at the brain's midsection, left and right

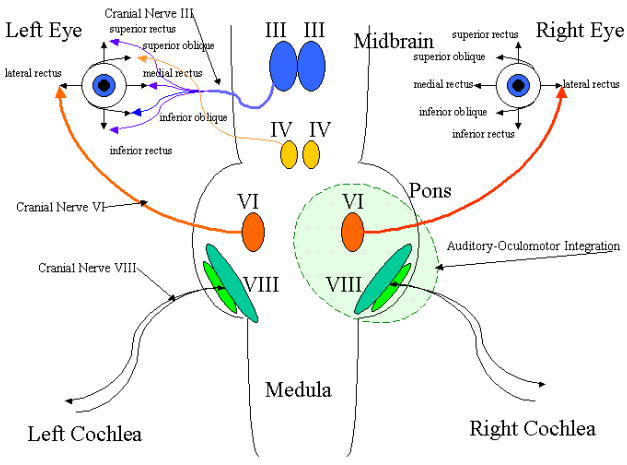

Fig. 4

CNs-VI are pulled apart and are located much more laterally. Such spatial separation of closely related oculomotor functions suggests that proximity of the cochlear nuclei to the nuclei controlling lateral recti' is preferred over tight integration with the oculomotor nuclei.

Specifically, rather than keeping neural connections short it was more important from an evolutionary standpoint to ensure fast switching of visual attention in response to rapidly changing auditory stimuli and vise-versa - Fig. 5.

A mechanism that controls convergence of the eyes and is similar to the one described above may exist in the vicinity of the abducent nuclei of the CN-VI. However, instead of controlling the phase of neural activation, it may modulate the neural firing rate that produces an appropriate contraction of left and right lateral recti, forcing the eyes to converge on a specific point in the horizontal plane. For example an auditory-to-visual loop may be activated as the result of auditory stimulation leading to activation of a number of monophased Θ-F-Phase-Selecting Neurons. Activation of a number of monophased Θ-F-Phase-Selecting Neurons means that there is a spatial location that is sourcing a polyphonic acoustic stimulus, which is probably associated with a physical object. The information about acoustically triangulated spatial location of a suspected physical object is conveyed to a structure associated with the CN-VI for conversion into the appropriate stimulation of left and right lateral recti' that leads to eye convergence on the suspected source of acoustic stimulus.

Similarly, in a "cocktail party environment" eye convergence on an object of interest produces visual triangulation that is conveyed to the auditory subsystem to tune the phases of both cochleae for noise filtering and better perception of a conversation.

We suspect that there is a different mechanism that is responsible for vertical localizations of auditory and visual stimuli. This is because both cochleae and eyes are located approximately in the same horizontal plane and thus a more intricate mechanism should be responsible for vertical localization.

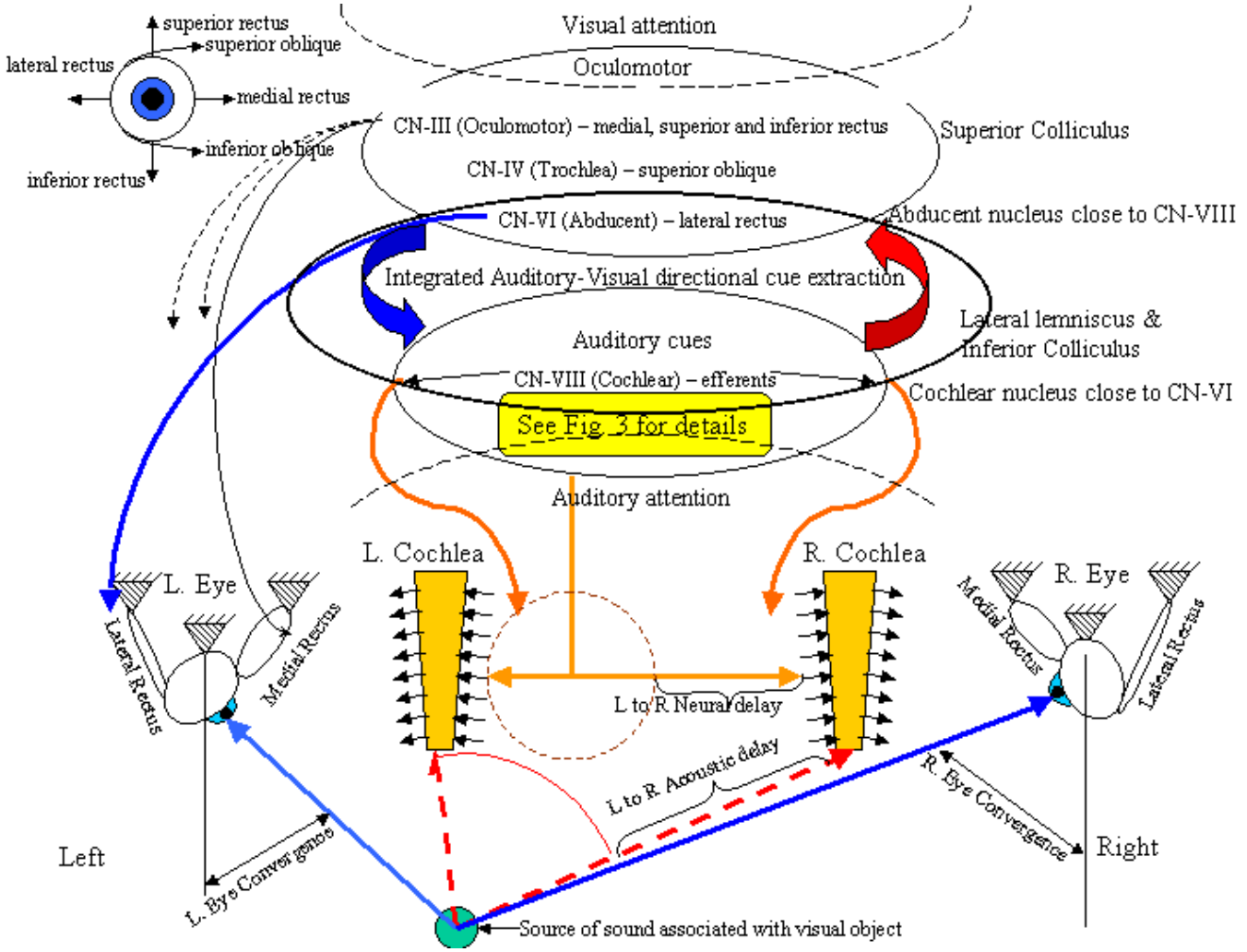

Fig. 5

V. CONCLUSIONS

The above-described mechanism, if found in the human brain, would allow filtering of acoustic information in a noisy environment and tuning to a specific acoustic source, as well as switching visual attention in response to acoustic stimulus and vise versa.

Auditory filtering is achieved through phase-selective amplification of in-phase and attenuation of out-of-phase signals. Visual fixation on an object results in "triangulation" of the source of sound associated with it. Each visually "triangulated" direction manifests itself in activation of a single or very few neurons that convey this information to auditory centers and activate the corresponding Phase Selecting neurons presumably in the vicinity of lateral lemniscus. When activated, a Phase Selecting neuron tunes the left and right organs of Corti by attenuating all frequencies that are not originating from the direction of the object of interest. A similar mechanism is responsible for the otherwise difficult to explain broad dynamic range of a mammalian ear.

Our ability to quickly switch visual attention may be attributed to the close proximity of the auditory and oculomotor centers. If not suppressed by high-level cortical processes, such proximity allows fast convergence of both eyes on a source of sound. If an acoustic source with significant auditory information is presented, the direction to the source is instantly "triangulated" by the auditory system and forwarded to the oculomotor center. Similarly to the described mechanism of auditory convergence, the oculomotor center may also contain a mechanism that forces both eyes to optically converge in the direction of the source of sound.

Such reciprocal arrangements between some auditory and oculomotor centers allow for competitive "push-pull" activity between these two subsystems that may lead to audio-visual sensory "awareness" on a subcortical level.

If the proposed mechanism survives the critics it may become reasonable to ask whether some animals without a developed vocal tract may actually have the capacity for complex acoustic communications via emission from their cochleae. And if so, it will be interesting to see how the evolution of Broca-like areas in such brains has brought up the emergence of Wernicke-like areas.

It should also be noted here that the proposed delay-based mechanism is generic and may exist in other areas of the brain, including the cortex. It is known, for example, that axons of mossy fibers spread over large distances in the molecular layer of the cortex in the direction parallel to the brain's surface. Along their way, these axons innervate multiple Purkinje cells with different time delays. Such feedback loops may produce oscillatory behavior in the cortical structures and may also contribute to the functionality of central pattern generators to produce coordinated motor responses.

VI. REFERENCES

[1] Yost, W. A., and G. Gourevitch, Eds. (1987). Directional Hearing (Springer-Verlag, New York).

[2] Lloyd Watts, (1993), "Cochlear Mechanics: Analysis and Analog VLSI," PhD Thesis, CalTech

[3] Sridhar Kalluri, and Bertrand Delgutte. (2002), "Characteristics of Cochlear Nucleus Onset Units Studied with a Model", Computational Models of Auditory Function, IOS Press

[4] He, D.Z.Z. (1997). Relationship between the development of outer hair cell electromotility and efferent innervation: A study in cultured organ of Corti of neonatal gerbils. J. Neurosci. 17, 3634-3643.

[5] Moss, C.F. and Carr, C.E. Comparative Audition. In of Psychology, Wiley, 2002, in press.

[6] Simon JZ, CE Carr, SA Shamma, (1999), "A dendritic model of coincidence detectin in the avian brainstem," Neurocomputing 26-7:263-269

[7] Thierry Morlet, Kelly Rose, Linda J. Hood, and Charles I. Berlin in http://www2.unife.it/oae/editorials/August_september_2002.html

[8] Nakajima, H.H., Hubbard, A.E. and Mountain D.C. (2000) A physiologically-based nonlinear active feedback model of the cochlea. In: Symposium on Recent Developments in Auditory Mechanics, H. Wada, T. Takasaka, K. Ikeda, K. Ohyama, T. Koike eds. World Scientific Publishing, Singapore, pp. 202-208.

[9] Ian Russell and Cornelia Schauz, (1995), "Salicylate Ototoxicity: Effects on the Stiffness and Electromotility of Outer Hair Cells Isolated from the GuineaPigCochlea", http://www.northwestern.edu/auditory-neuroscience/vol.1/no.3/v1n3p309.html;

[10] Fabio A. Thiers, Barbara J. Burgess, and Joseph B. Nadol, "Reciprocal Innervation of Outer Hair Cells in a Human Infant," Journal of the Association for Research in Otolaryngology DOI: 10.1007/s101620020024 © 2002 Springer-Verlag New York, Inc.

[11] Serena Preyer, Sabine Renz, Werner Hemmert, Hans-Peter Zenner, and Anthony W. Gummer, "Receptor Potential of Outer Hair Cells Isolated from Base to Apex of the Adult Guinea-Pig Cochlea: Implications for Cochlear Tuning Mechanisms," http://www.northwestern.edu/auditory-neuroscience/vol.2/no.2/v2n2p145.html

[12] Christof Koch, (1999), "Boiphysics of Computation", Oxford University Press

[13] http://info.med.yale.edu/caim/cnerves

[14] Stephen J. DeArmond, Madeline M. Fusco and Maynard M. Dewey, (1989), "Structure of the Human Brain, a photographic atlas", third edition, New York, Oxford University Press