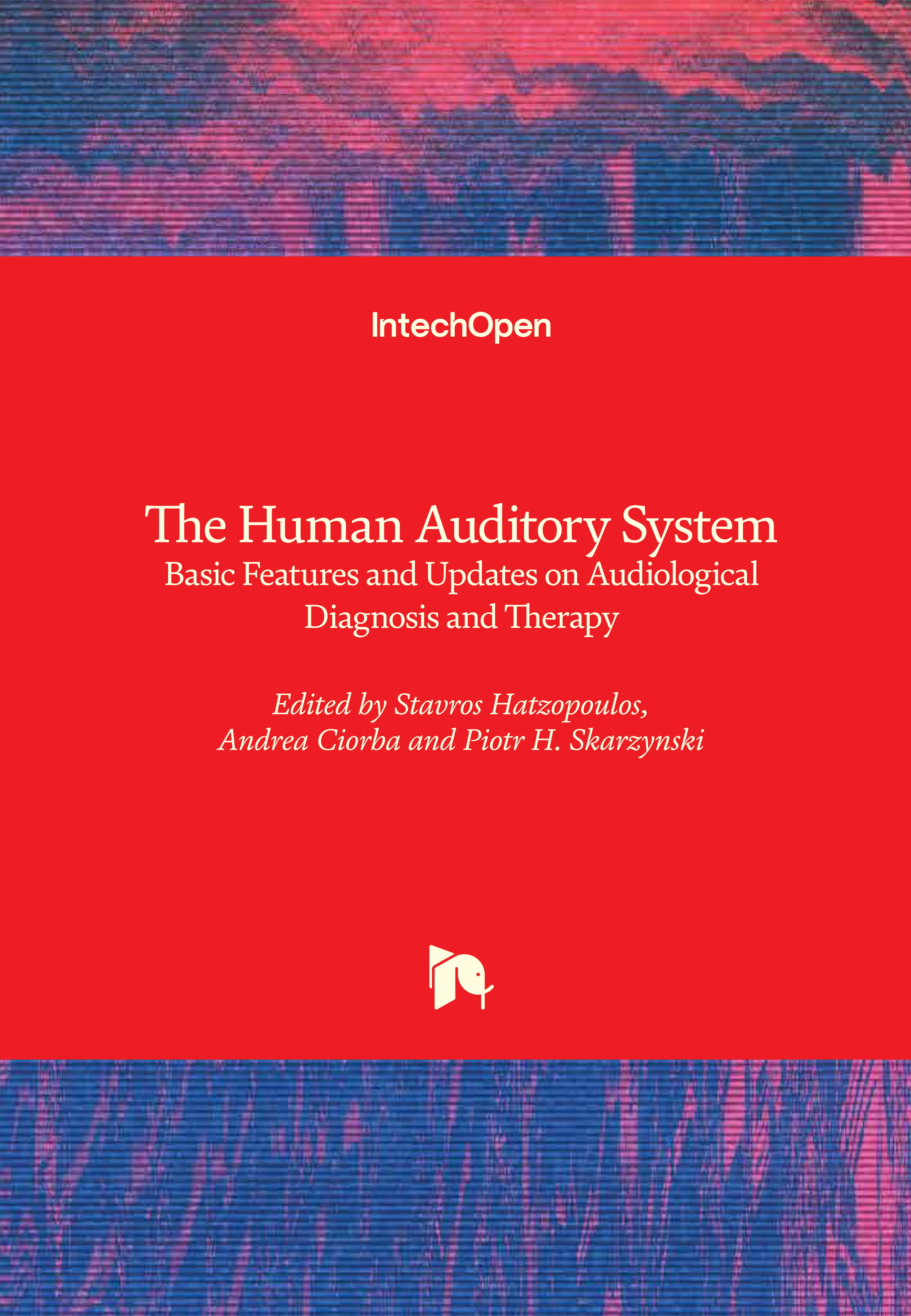

IntechOpen

The Human Auditory System

Basic Features and Updates on Audiological Diagnosis and Therapy

Edited by Stavros Hatzopoulos,
Andrea Ciorba and Piotr H. Skarzynski

The Human Auditory System - Basic Features and Updates on Audiological Diagnosis and Therapy

Edited by Stavros Hatzopoulos, Andrea Ciorba and Piotr H. Skarzynski

Published in London, United Kingdom







IntechOpen






Supporting open minds since 2005

The Human Auditory System - Basic Features and Updates on Audiological Diagnosis and Therapy
http://dx.doi.org/10.5772/intechopen.77713
Edited by Stavros Hatzopoulos, Andrea Ciorba and Piotr H. Skarzynski

Contributors
Milaine Dominici Sanfins, Caroline Donadon, Stavros Hatzopoulos, Chiung Yao Chen, Brian Barton, Alyssa A Brewer, Luca Oscar Redaelli De Zinis, Nader Nassif, Patrick Boyle, Piotr H. Skarzynski, Maria Francisca Colella-Santos, Magdalena Skarzynska

© The Editor(s) and the Author(s) 2020
The rights of the editor(s) and the author(s) have been asserted in accordance with the Copyright, Designs and Patents Act 1988. All rights to the book as a whole are reserved by INTECHOPEN LIMITED. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECHOPEN LIMITED's written permission. Enquiries concerning the use of the book should be directed to INTECHOPEN LIMITED rights and permissions department (permissions@intechopen.com).
Violations are liable to prosecution under the governing Copyright Law.

Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 3.0 Unported License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at http://www.intechopen.com/copyright-policy.html.

Notice
Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in London, United Kingdom, 2020 by IntechOpen
IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, 7th floor, 10 Lower Thames Street, London, EC3R 6AF, United Kingdom
Printed in Croatia

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from orders@intechopen.com

The Human Auditory System - Basic Features and Updates on Audiological Diagnosis and Therapy
Edited by Stavros Hatzopoulos, Andrea Ciorba and Piotr H. Skarzynski
p. cm.
Print ISBN 978-1-78923-937-9
Online ISBN 978-1-78923-938-6
eBook (PDF) ISBN 978-1-83968-453-1

Meet the editors

Dr. Stavros Hatzopoulos has been a faculty member of the Audiology & ENT Clinic, University of Ferrara, Italy, since 1994. His background is in biomedical, engineering, audiological engineering, and hearing science (BSc degree from University of Southern California, MSc from Texas A&M, PhD from Worcester Polytechnic Institute in Massachusetts). He received his habilitation as an associate professor in 2014. He is the author of more than 195 book chapters, peer review papers, and congress presentations. He has participated in numerous European Concerted Action projects in the areas of otoacoustic emissions, genetics, and nanotechnologies. He currently serves as the editor of the Portal on Otoacoustic Emissions (www.otoemissions.org) and as the audiology section editor in the Journal of Hearing Science.

Andrea Ciorba, MD, PhD, has been a member of the Audiology & ENT Clinic, University of Ferrara, Italy, since 2004. After earning his degree in Medicine in 1999 and a specialization in Otolaryngology and Head and Neck Surgery in October 2003, he obtained a PhD in Audiology in March 2008. He received his habilitation as an associate professor in 2014 and he is presently a member of the Audiology & ENT Clinic, University of Ferrara, Faculty Board. He is the author of approximately 130 peer-reviewed papers and book chapters and serves as a member of the editorial board or as a reviewer for a number of scientific journals. His clinical field of interest is audiology and vestibology, and he is also involved in research activities in the field of otoacoustic emissions and inner ear electrophysiology.

Prof. Piotr H. Skarzynski, MD, PhD, MS is occupied with the implementation of scientific, clinical and surgical work at the World Hearing Center, Medical University of Warsaw and Institute of Sensory Organs in Warsaw, Poland. He is a member of 29 scientific societies. He has been a Representative Board Member (until 2019) and Vice-President and Institutional Representative (since 2019) of ISfTeH; Vice-President of the Hearing Group; Member of Hearing Committee of AAO–HNS; an auditor of EFAS; a member of the Consultant Committee of International Experts of CPAM-VBMS (by special invitation); an Honorary Member of ORL Danube Society; and a member of the Roster of Experts on Digital Health of WHO. He also participated in the 3rd Stakeholders Consultation meeting during which the World Hearing Forum of WHO was announced. He has been an author and co-author of over 1767 documents (invited speaker presentations – 115, round table presentations – 118, as an instructor – 118 courses) and 843 publications (IF – 166,554, IC – 44450,31). He is a reviewer of 30 national and international scientific journals. He is currently involved with numerous international projects, connected with hearing screening and telemedicine in European, Asian, and African countries.

Contents

Preface

The objective of this book is to present information on the basic features of the auditory system, as well as hearing loss, which is among the most prevalent chronic disabilities worldwide. Nowadays, it is clear that the identification and the rehabilitation of hearing impairment, when possible, has to be adequately and promptly managed, since hearing loss can seriously interfere with psychosocial development, family dynamics, and social interactions. This book presents a number of basic characteristics of the auditory system and a number of novel approaches to issues pertaining to hearing deficits.

This book was made possible through the substantial contribution of several authors.

The format of this volume targets graduate courses in audiology, speech pathology and hearing science, neurosciences, and basic graduate courses in otolaryngology.

The volume is divided in two main sections as follows:

1. **Part I:** Anatomical and Pathophysiological Features. In these chapters, several aspects of the auditory system are discussed, particularly focusing on the pathophysiology of the auditory pathways.

2. **Part II:** Updates on Diagnosis and Therapy. This part offers an update on several innovative approaches for a modern evaluation and therapy of the ear, such as endoscopic ear surgery.

Acknowledgements

We would like to thank all the contributing authors, and the IntechOpen personnel for the continuous and generous assistance during the preparation of the volume.

Stavros Hatzopoulos PhD and Dr. Andrea Ciorba
Clinic of Audiology and ENT,
University of Ferrara,
Ferrara, Italy

Piotr H. Skarzynski MD PhD
Department of Teleaudiology and Screening,
World Hearing Center,
Institute of Physiology and Pathology of Hearing,
Warsaw, Poland

Section 1

Anatomical and Pathophysiological Features of the Auditory Pathways

Chapter 1

Attention and Working Memory in Human Auditory Cortex

Brian Barton and Alyssa A. Brewer

Abstract

Human sensory systems are organized into processing hierarchies within cortex, such that incoming sensory information is analyzed and compiled into our vivid sensory experiences. Computations that are common to these sensory systems include the abilities to maintain enhanced focus on particular aspects of incoming sensory information (i.e., attention) and to retain sensory information in a short-term memory store after such sensory information is no longer available (i.e., working memory). In at least the auditory and visual systems, the necessary computational steps to create these experiences take place in cloverleaf clusters of cortical field maps (CFMs). The human auditory CFMs represent the spectral (i.e., tones) and temporal (i.e., period) aspects of sound, which are represented along the cortical surface as two orderly gradients that are physically orthogonal to one another: tonotopy and periodotopy, respectively. Knowledge of the properties of such CFMs is the foundation for understanding the specific sensory computations carried out in particular cortical regions. This chapter reviews current research into auditory nonverbal attention, auditory working memory, and auditory CFMs, and introduces the next steps to measure the effects of attention and working memory across the known auditory CFMs in human cortex using functional MRI.

Keywords: human auditory cortex, fMRI, tonotopy, periodotopy, cloverleaf cluster, cortical field maps, attention, working memory

1. Introduction

Mammalian sensory systems are composed in cortex of many functionally specialized areas organized into hierarchical networks [1–6]. The most fundamental sensory information is embodied by the organization of the sensory receptors, which is maintained throughout most of the cortical hierarchy of sensory regions with repeating representations of this topography in cortical field maps (CFMs) [5, 7–13]. Accordingly neurons with receptive fields situated next to one another in sensory feature space are positioned next to one another in cortex within a CFM.

In auditory cortex, auditory field maps (AFMs) are identified by two orthogonal sensory representations: tonotopic gradients from the spectral aspects of sound (i.e., tones), and periodotopic gradients from the temporal aspects of sound (i.e., period or temporal envelope) [5, 10, 14]. On a larger scale across cortex, AFMs are grouped into cloverleaf clusters, another fundamental organizational structure also common to visual cortex [8, 10, 15–20]. CFMs within clusters tend to share properties such as receptive field distribution, cortical magnification, and processing specialization (e.g., [18, 19, 21]).

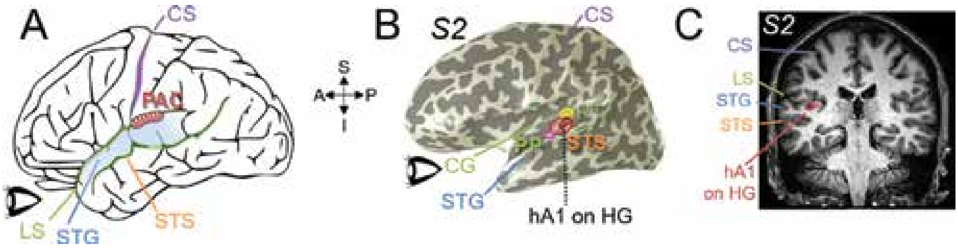

Figure 1.
Primary auditory cortex. (A) The lateral view of the left hemisphere is shown in the schematic. Major sulci are marked by black lines. The approximate position of primary auditory cortex (PAC) is shown with the red overlay inside the black dotted line. The white dotted line within the red region indicates the extension of PAC into the lateral sulcus (LS) along Heschl's gyrus (HG; hidden within the sulcus in this view). Inset refers to anatomical directions as A: anterior; P: posterior; S: superior; I: inferior. PAC: primary auditory cortex (red); LS: lateral sulcus (green; also known as the lateral fissure or Sylvian fissure); CS: central sulcus (purple); STG: superior temporal gyrus (blue); STS: superior temporal sulcus (orange). (B) The cortical surface of the left hemisphere of one subject (S2) is displayed as a typical inflated 3-D rendering created from high-resolution, anatomical MRI measurements. Light gray regions denote gyri; dark gray regions denote sulci. The exact location of this subject's hA1 auditory field map is shown in red within the black dotted lines. Note that HG in S2 is composed of a double peak, seen here as two light gray stripes, rather than the more common single gyrus. The locations of the three cloverleaf clusters composed of the core and belt AFMs are shown along HG by three colored overlays as yellow: hCM/hCL cluster; red: HG cluster including hA1, hR, hRM, hMM, hML, hAL; and magenta: hRTM/hRT/hRTL cluster (cite?). Additional cloverleaf clusters are under investigation along PP, PT, STG, and the STS. Green-labeled anatomical regions are sections within the lateral sulcus—CG: Circular gyrus (green); PP: planum polare (green); PT: planum temporale (green). (C) This single T1 image shows a coronal view of hA1 on HG (red within dotted white line). Adapted from Refs. [5, 12].

Across the cortical hierarchy, there is generally a progressive increase in the complexity of sensory computations from simple sensory stimulus features (e.g., frequency content) to higher levels of cognition (e.g., attention and working memory) [6, 13, 22]. CFM organization likely serves as a framework for integrating bottom-up inputs from sensory receptors with top-down attentional processing [12, 17]. With the recent ability to measure AFMs in the core and belt regions of human auditory cortex along Heschl's gyrus (HG) using high-resolution functional magnetic resonance imaging (fMRI), the stage is now set for investigation into this integration of basic auditory processing with higher-order auditory attention and working memory within human AFMs (**Figure 1**) [5, 12, 15, 23].

This chapter first provides a brief history of research into models of auditory nonverbal attention and working memory, with comparisons to their visual counterparts. Next, we discuss the current state of research into AFMs within human auditory cortex. Finally, we propose directions of future research investigating auditory attention and working memory within these AFMs to illuminate how these higher-order cognitive processes interact with low-level auditory processing.

2. Attention and working memory in human audition

2.1 Models of attention and working memory

Attention, the ability to select and attend to aspects of the sensory environment while simultaneously ignoring or inhibiting others, is a fundamental aspect of human sensory systems (for reviews, see [24–27]). Given the limited resources of the human brain, attention allows for greater resources to be allocated to processing of important incoming sensory stimuli by diverting precious resources from currently unimportant stimuli. Such allocation can be controlled cognitively, in what is generally referred to as 'top-down' attentional control in models of attention, in reference to the higher-order cognitive processes controlling attention from the 'top'

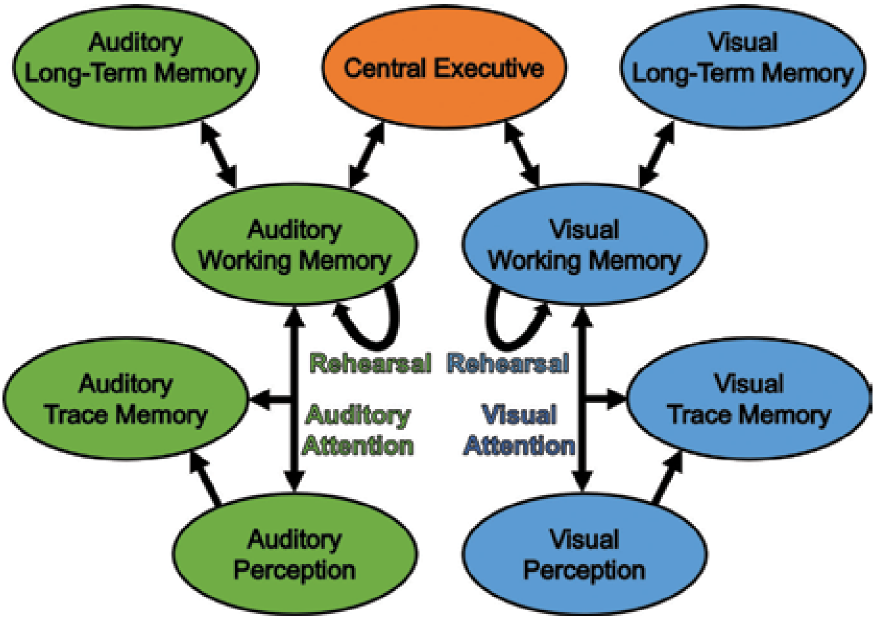

Figure 2.
Attention and working-memory model. A model of the interactions between perception, trace memory, attention, working memory, and long-term memory in the visual and auditory systems, as well as the central executive. Ovals represent neural systems. Arrows represent actions of one system on another. Attention is the term for the action of perception and trace memory on working memory and vice versa. Rehearsal is the term for maintaining information in working memory. This model is not intended to indicate that these systems are discrete or independent; within each sense, they are in fact highly integrated.

of the sensory-processing hierarchy and acting 'down' on the lower levels (**Figure 2**) [24, 28–31]. Despite lower priority being assigned to the currently unimportant stimulus locations, change is constant, so the resource diversion to attended stimuli is not absolute, allowing for the sensory environment to continue to be monitored. If, instead, processing resources were evenly distributed throughout the sensory field, without regard to salience, more resources would be wasted on unimportant aspects of the field. If something in the unattended sensory field should become important, the system requires a mechanism to reorient attention to that aspect of the field. Such stimulus-driven attentional control is referred to as 'bottom-up', referring to the ability of incoming sensory input at the bottom of the hierarchy to orient the higher-order attention system. This broad framework of attentional models is common at least to the senses most commonly studied, vision and audition [25, 27, 31, 32].

In the effort to elucidate the parameters of auditory attention, researchers have taken a myriad of approaches in numerous contexts. Researchers have attempted to decipher at what level of the sensory-processing hierarchy stimulus-driven attention occurs (after which sensory-processing steps does attention act) [24, 30, 31, 33–35], how attention can be deployed (to locations in space or particular sensory features) [36–40], and how can attention be distributed (to how many 'objects' or 'streams' can attention be simultaneously deployed) [41–44]. Many studies have narrowed the range of possibilities without precisely answering these questions, and so remain active areas of research. Modern models of attention generally agree that stimuli are processed to some degree before attention acts, accounting for the stimulus-driven 'bottom-up' attentional shifts, though it is unclear to precisely which degree [24, 30, 33]. Neuroscientific evidence suggests that attention acts throughout sensory-processing hierarchies, so the idea of attention being located at a particular 'height' in the hierarchy may not be a particularly useful insight for identifying the cortical locus of attentional control [45, 46]. Modern attentional models also generally agree that attention can be deployed to locations in or features of sensory space, both of which are fundamental aspects to the sensory-processing hierarchy [24, 35]. Finally, modern models of attention agree that attention is very limited, but not about precisely how it is limited. Some models are still fundamentally 'spotlight' models [25, 44], in which attention is limited to a single location or feature set, while others posit that attention can be divided between a small number of locations or features [41, 47]. Based on related working-memory research, the latter theory is gaining prominence as likely correct.

Working memory (i.e., a more accurate term for 'short-term memory') is the ability to maintain and manipulate information within the focus of attention over a short period of time after the stimulus is no longer perceptible (for reviews, see [48–51]). Without explicit maintenance, this retention period is approximately 1–2 s, but is theoretically indefinite with explicit maintenance. Working memory should not to be confused with 'sensory memory', also known as 'iconic memory' in vision and 'echoic memory' in audition [52]. Sensory memory is a fundamental aspect of sensory systems in which a sensory trace available to attention and working-memory systems persists for less than ~100 ms after stimuli are no longer perceptible. Models of working memory are nearly indistinguishable from models of attention; the key difference is that working memory is a 'memory' of previously perceptible stimuli, whereas attention is thought to act on perceptible stimuli or sensory traces thereof. Working-memory models posit, by definition, that working memory acts after perception processing has occurred (**Figure 2**; for review, see [53]). However, it has been difficult to isolate exactly where working-memory control resides along the cortical hierarchy of sensory processing, likely because low-level perceptual cortex is recruited at least for visual working memory and attention [40, 46, 54, 55].

Like attention, working-memory models also posit that working memory is a highly limited resource, in which a small set of locations or objects (e.g., 3–4 items on average) can be simultaneously maintained [42, 49]. In fact, some modern measures of attention and working memory are nearly identical. The change-detection task is a ubiquitous one in which subjects are asked to view a sensory array, then compare that sensory array to a second one in which some aspect of the array may have changed, and indicate whether a change has occurred (**Figure 3**) [56–60]. A short delay period (i.e., retention interval) is included during each array, which may include a neutral presentation or, if desired, a mask of the sensory stimuli to prevent the use of 'sensory memory'. The length of the delay period can be then be altered to either measure attention or working memory. If the delay period is on the order of ~0–200 ms, it is considered an attentional task; if it is longer, on the order of 1–2 s, it is considered a working-memory task [53]. Therefore, attention and working-memory systems are at a minimum heavily intertwined and very likely the same system studied in slightly different contexts, with attention being a component of a larger working-memory framework.

With the relatively recent invention of fMRI, researchers have been able to begin to localize these models of attention and working memory to their cortical underpinnings (e.g., [6, 37, 40, 50, 55, 61, 62]). FMRI, through its exquisite ability to localize blood oxygenation-level dependent (BOLD) signals (and thus the underlying neural activity) to just a couple of millimeters is the best technology available for such research [63, 64]. Two broad approaches have been employed for studying these high-order cognitive processes: model-based and perception-based. Model-based investigations tend to use tasks based on behavioral investigations into attention and working memory, adapt them to the strict parameters required of fMRI, and compare activity in conditions when attention or working memory are differentially deployed [61, 62]. Perception-based investigations tend to measure low-level perceptual cortex that has already been mapped in detail and measure the effects of attention or working memory within those regions [50, 55, 65]. Both approaches are important and should be fully integrated to garner a more complete and accurate localization of these attentional and working-memory systems.

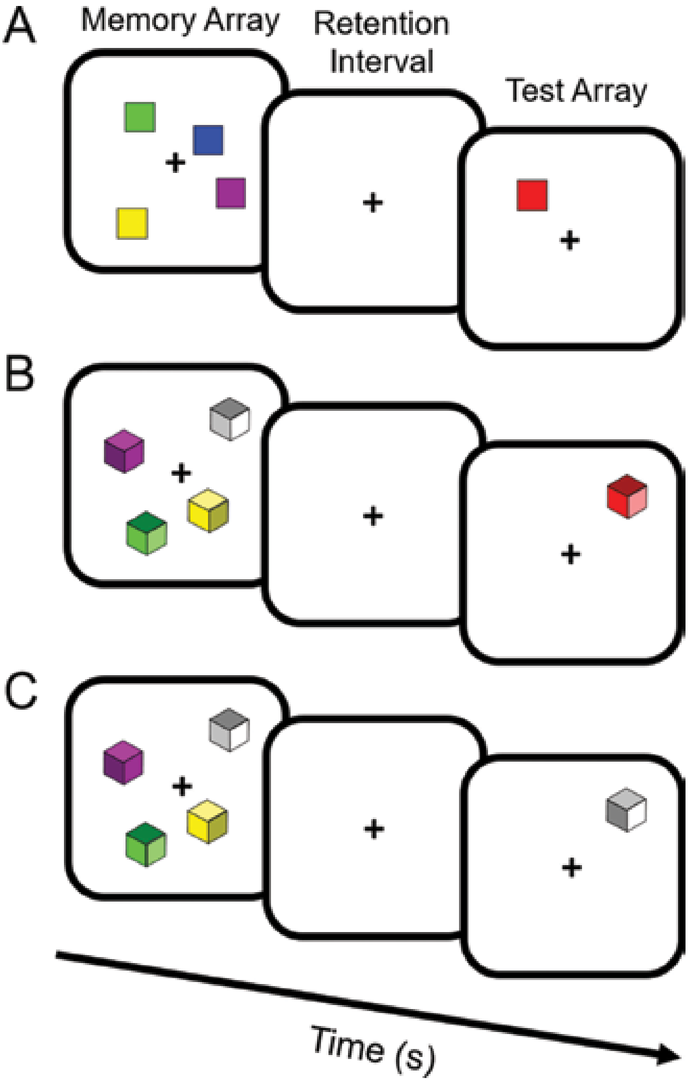

Figure 3.
Visual change-detection task. This task can be used to probe visual attention or working memory and is very similar to its auditory counterpart. Such tasks have three phases: first is encoding, when subjects are given ~100–500 ms to view the sample array; next is maintenance, which is short (~0–200 ms) for measuring attention and longer (~1000 ms) for working memory; last is the probe (lasting until the subject responds or with a time limit, often ~2000 ms). In this example, a set size of four is presented for the sample array and a probe array of one is used, though different set sizes are commonplace and often the probe array will be the same set size as the encoding array with a possibility of one object being changed. Typically there is an equal chance (50%) of the probe array containing a change or not. Generally subjects will be required to fixate centrally, particularly if fMRI, EEG, or PET recordings are being made. (A) Simple colored square stimuli are depicted here, often drawn from a small set of easily distinguished hues (in this case, 6). As a result, changes are always low in similarity, requiring low resolution to make accurate comparisons between encoding and test arrays, which is important at least for visual working-memory measurements. More complex stimuli can also be used as in (B) and (C). These stimuli are shaded cubes with the same hue set as in (A), but also have 6 possible shading patterns with the dark, medium, and light shaded sides on each cube. Changes between hues, as in (B), are equivalently low similarity to (A) and result in similar performance under visual working-memory conditions. Changes in shading patterns, as in (C), result in worse performance than (B) despite having the same number of possible pattern changes as hue changes in (A) or (B), because such changes require higher resolution representations in visual working memory. Adapted from Barton and Brewer [50].

2.2 Overview of auditory and visual attention research

Research into attention began in earnest in the auditory system after World War II with a very practical motivation. It had been noted that fighter pilots sometimes failed to perceive auditory messages presented to them over headphones despite the fact that the messages were completely audible. To solve this problem, Donald Broadbent began studying subjects with an auditory environment similar to the pilots, with multiple speech messages presented over headphones [34]. Based on his findings, he proposed a selective theory of attention, which was popular and persuasive, but ultimately required modification. Environments such as the one Broadbent studied are more commonly encountered at cocktail parties, in which multiple audible conversations are taking place, and people are able to attend to one or a small set of speech streams while attenuating the others. To study the 'cocktail party phenomenon,' the dichotic listening task was developed in the 1950s by Colin Cherry [66, 67]. Subjects were asked to shadow the speech stream presented to one ear of a set of headphones while another stream was presented to the other ear, and they demonstrated little knowledge of the nonshadowed (unattended) stream (**Figure 4**).

A host of studies followed up on the basic finding, revealing several attentional parameters within the context of that type of task (e.g., [30, 35, 40, 68–71]). Importantly, preferential processing of the attended stream relative to the unattended streams is not absolute; for example, particularly salient information, such as the name of the subject, could sometimes be recalled from an unattended stream, presumably by reorienting attention [39, 66, 67, 69]. The streams were typically differentiated spatially (e.g., to each ear through a headset), indicating a spatial aspect to attentional selection and therefore the attentional system. Similarly, the streams

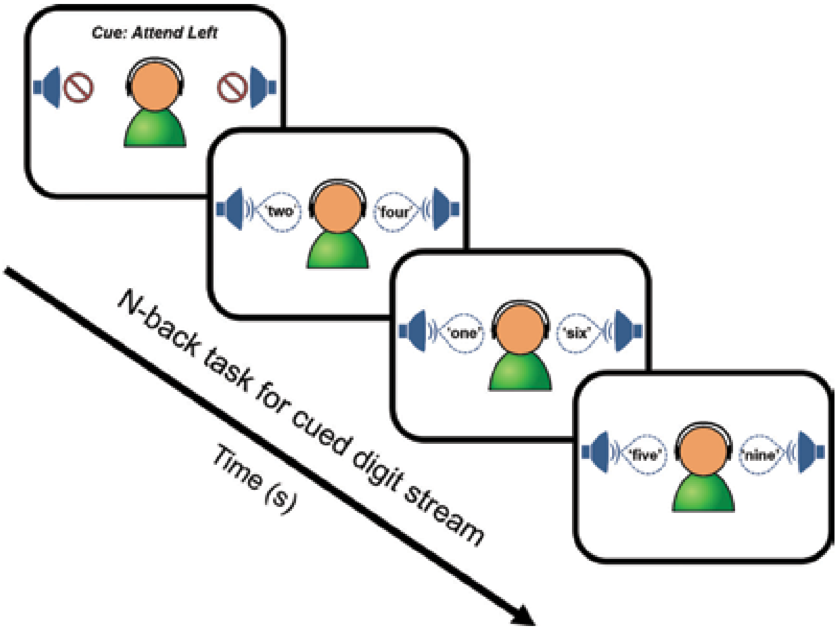

Figure 4.
Auditory spatial attention. Schematic of an example auditory spatial attention task (e.g., see [35, 40, 66, 67]). Each block typically starts with cue (auditory or visual) for the subject to attend left or right on the upcoming trial. Two simultaneous auditory streams of digits are presented as binaural, spatially lateralized signals. Behavioral studies in an anechoic chamber often use speaker physically located to the left and right of the subject; fMRI measurements do not have the option of such a set up, but instead can use differences in the interaural time difference (ITD) to produce a similarly effective lateralization for the two digit streams. The subjects attend to the cued digit stream and perform a 1-back task.

were also typically differentiated by the voice of the person speaking, indicating attentional selection based on the spectrotemporal characteristics of the speaker's voice such as the average and variance of pitch and speech rate (often reflecting additional information about the speaker, such as gender) [66–68, 72].

These findings are very similar to findings in the visual domain, indicating that attentional systems across senses are similarly organized. Visual attention can similarly be deployed to a small set of locations or to visual features with very little recall of nonattended visual stimuli [41]. Roughly analogous to speech shadowing are multiple-object-tracking tasks, which require subjects to visually track a small set of moving objects out of a group [47, 73]. Visual change-detection tasks are also very common, and they demonstrate very similar results as their auditory counterparts [50, 74, 75]. In sum, the evidence suggests that attentional systems are organized very similarly, perhaps identically, between at least vision and audition.

Despite these broad contributions, these types of tasks are of limited utility when tying behavior to cortical activity because the types of stimuli used are rather high-order (e.g., speech) with relatively uncontrolled low-level parameters. For example, the spectrotemporal profile of a stream of speech is complex, likely activating broad swaths of low-level sensory cortex in addition to higher-order regions dedicated to speech comprehension, including working and long-term memory [68, 72, 76, 77]. If one were to compare fMRI activity across auditory cortex in traditional dichotic listening tasks, the differences would have far too many variables for which to account before meaningful conclusions can be made about attentional systems. It may seem intuitive to compare cortical activity between conditions where identical speech stimuli have been presented and the subject either attended to the stimuli or did not. However, areas that have increased activity when the stimuli were attended could simply reflect higher-order processing that only occurs when attention is directed to the stimuli rather than directly revealing areas involved in attentional control. For example, recognition of particular words requires comparison of the speech stimulus to an internal representation, which requires activation of long-term memories of words [77]. Long-term memory retrieval does not happen if the subject never perceived the word due to attention being maintained on a separate speech stream, so such memory-retrieval activity would be confounded with attentional activity in the analysis [70].

Thus, simpler stimuli that are closer in nature to the initial spectrotemporal analyses performed by primary auditory cortex (PAC) are better suited for experiments intended to demonstrate attentional effects in cortex [24]. Reducing the speech comprehension element is a good first step, and research approached this by using a change-detection task and arrays of recognizable animal sounds (cow, owl, frog, etc.; **Figure 5**) [59]. These tests revealed what the researchers termed 'change deafness,' in which subjects often failed to identify changes in the sound arrays. Such inability to detect changes is entirely consistent with very limited attentional resources, and very similar to results of working-memory change-detection tasks [30, 53, 60, 78].

However, even these types of stimuli are not best suited to fMRI investigation at this stage of understanding due to their relative complexity compared to the basic spectrotemporal features of sounds initially processed in auditory cortex [12, 50]. As discussed in detail below, the auditory system represents sounds in spectral and temporal dimensions, and stimuli similar to those used to define those perceptual areas would be best suited now to evaluating the effects of attention in the auditory system (**Figure 6**) [5, 10].

2.3 Overview of auditory and visual working-memory research

Visual and auditory working memory were discovered in quick succession and discussed together in a very popular and influential model by Baddeley and Hitch

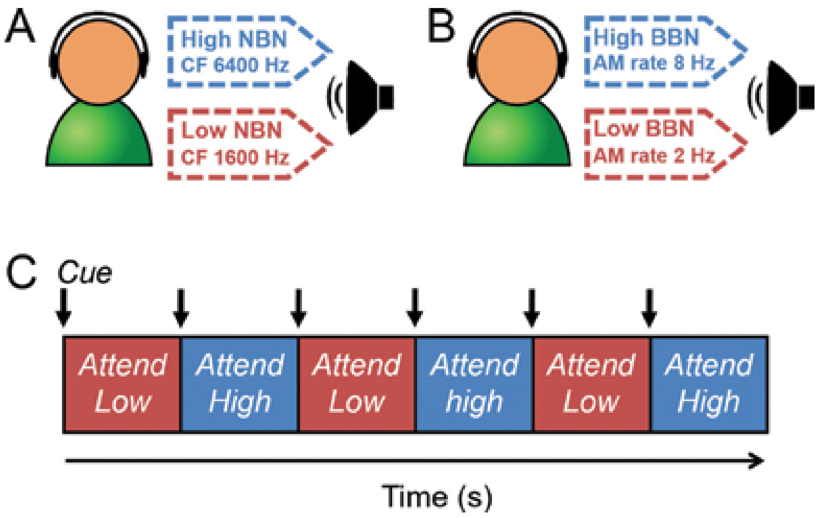

Figure 5.
Auditory feature attention. Schematic outlines a simple proposed attention task utilizing spectral (narrowband noise) and temporal (broadband noise) stimuli taken from the stimuli used by [10] to define auditory field maps. Subjects are asked to attend to one of two simultaneously presented stimuli, which are either (A) narrowband noise, in this case with central frequencies of 6400 and 1600 Hz and the same amplitude modulation (AM) rate of 8 Hz, or (B) broadband noise, in this case with AM rates of 2 and 8 Hz. (C) A proposed task that varies auditory feature attention, in which subjects are instructed to attend to each of the stimuli in an alternating pattern, cued by a short sound at the beginning of each block.

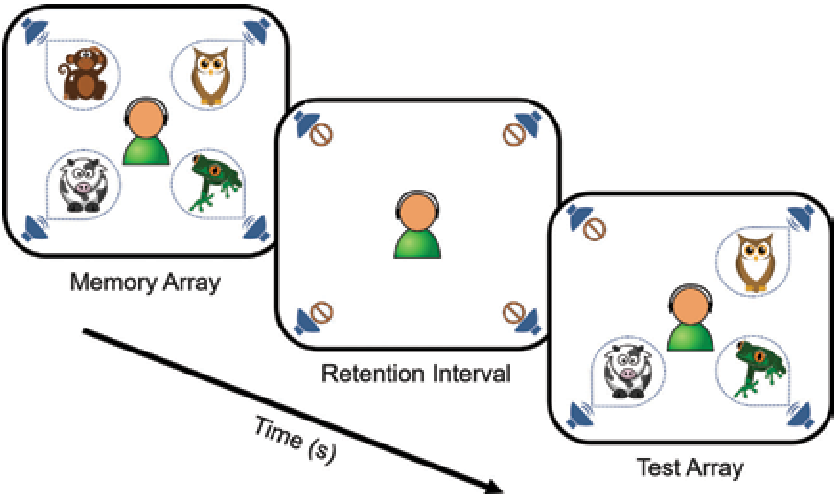

Figure 6.
Auditory object attention and working memory. Schematic of one trial in an auditory change-detection task (e.g., see change-deafness experiments in [59]). Subjects are first presented with an array of four distinct auditory objects (e.g., four different recordings of real animal sounds, randomized each trial from a larger set of iconic animal calls). In the initial memory array, the four animal sounds are initially presented binaurally and are temporally overlapped for a short time (e.g., 2 s). Within an anechoic chamber setup often used in psychoacoustic studies, these speakers may be physically positioned at the corners of a square; fMRI measurements do not have the option of such a set up, but instead can use differences in the interaural time difference (ITD) and interaural level difference (ILD) to produce a similarly effective virtual space. The subject's goal is typically to identify and remember all four animal sounds. The interstimulus interval is commonly filled with silence or white noise and can be varied in length to create shorter or longer retention intervals for attention or working-memory tasks, respectively. During the subsequent test array, subjects attempt to identify which one of the four auditory objects is now missing from the simultaneous animal sound presentations. In such auditory change-deafness paradigms, subjects fail to notice a large proportion of the changes introduced between the initial and test arrays.

linking sensory perception, working memory, and executive control [79–81]. The generally accepted modern model of working memory has changed somewhat from the original depiction, but the vast majority of research has been working within the framework (for reviews, see [30, 51, 53, 79, 81]). Each sense is equipped with its own perceptual system and three memory systems: sensory memory, working memory, and long-term memory. Direct sensory input, gated by attentional selection, is one of the two primary inputs into working memory. Sensory memory is a vivid trace of sensory information that persists after the information has vanished for a short time and is essentially equivalent to direct sensory input into working memory, again gated by attentional selection; one can reorient attention to aspects of the sensory trace as if it were direct sensation. Long-term memory is the second primary input into working memory, which is gated by an attention-like selection, generally referred to as selective memory retrieval. Working memory itself is a short-term memory workspace lasting a couple of seconds without rehearsal, in which sensory information is maintained and manipulated by a central executive [82]. The central executive is a deliberately vague term with nebulous properties; as a colleague often quips, "All we know of the central executive is that it's an oval," after its oval-shaped depiction in the Baddeley and Hitch model. There is ongoing debate as to the level of the hierarchy at which each system is integrated into that of the other senses, with no definitive solutions.

Visual working memory and visual sensory memory (i.e., 'iconic memory') were fundamentally measured by George Sperling in 1960 [52]. He presented arrays of simple visual stimuli for short periods of time and asked subjects to report what they had seen after a number of short delays. He discovered that subjects could only recall a small subset of stimuli in a large array, representing the limited capacity of visual working memory. Furthermore, they could recall a particular subset of the stimuli when cued after the presentation but before the sensory trace had faded (≤100 ms), indicating that visual sensory memory exists and that visual attention can be deployed to stimuli either during sensation or sensory memory. Over the next decade, George Sperling went on to perform similar measurements in the auditory system, delineating very similar properties for auditory perception, sensory memory, and working memory [83].

Without directly measuring brain activity, researchers concluded that sensory systems must be operating independently with dual-task paradigms in which subjects were asked to maintain visual, auditory, or both types of information in working memory. It was shown that subjects could recall ~3–4 'chunks' of information (which may not precisely reflect individual sensory locations or features) of each type, regardless of whether they were asked to maintain visual, auditory, or both types of information [49, 78]. If the systems were integrated, one would be able to allocate multisensory working-memory 'slots' to either sense, with a maximum number (e.g., 6–8) that could be divided between the senses as desired. Instead, subjects can maintain on average ~3–4 visual chunks and ~3–4 auditory chunks, without any ability to reallocate any 'slots' from one sense to the other.

While electroencephalogram (EEG) and positron emission topography (PET) recordings could broadly confirm the contralateral organization of the visual system and coarsely implicate the parietal and frontal lobes in attention and working memory, it was not until the advent of high-resolution fMRI that researchers could begin localizing attention and working memory in human cortex with any detail [6, 17, 37, 50, 84–90]. Model-based fMRI investigations have attempted to localize visual working memory by comparing BOLD activity in conditions where subjects are required to hold different numbers of objects in working memory [50, 62, 91, 92]. The logic goes that, because visual-working-memory models posit that a maximum of ~3–4 objects can be held in visual working memory on average, areas that increase their activity with arrays 1, 2, 3 objects and remaining constant with arrays of 4 or more objects should be areas controlling visual working memory. Such areas were found bilaterally in parietal cortex by multiple laboratories [57, 62, 91, 93], but activity related to visual working memory has also been measured in early visual cortex (e.g., V1 and hV4) [55, 65, 94], prefrontal cortex [95], and possibly in object-processing regions in lateral occipital cortex [62], indicating that working-memory tasks recruit areas throughout the visual-processing hierarchy. (We note that the report of object-processing regions is controversial, as the cortical coordinates reported in that study are more closely consistent with the human motion-processing complex, hMT+, than the lateral occipital complex [15, 17, 96, 97]). However, little has been done to measure visual-working-memory activity in visual field maps, and so these studies should be considered preliminary rather than definitive. Measurements within CFMs would, in fact, help to clear up such controversies.

Auditory-working-memory localization with fMRI has been quite limited compared to its visual counterpart, and largely concentrated on speech stimuli rather than fundamental auditory stimuli [30, 68]. As noted above with attention localization with fMRI, too many variables exist with highly complex stimuli, and as such, a different approach is necessary. Furthermore, even low-level auditory sensory areas have only very recently been properly identified [5, 10].

3. Auditory processing in human cortex

3.1 Inputs to auditory cortex

Auditory processing is essential for a wide range of our sensory experiences, including the identification of and attention to environmental sounds, verbal communication, and the enjoyment of music. The intricate sounds in our daily environments are encoded by our auditory system as the intensity of their individual component frequencies, comparable to a Fourier analysis [98]. This spectral sound information is thus one fundamental aspect of the auditory feature space (**Figure 7A, C**). The basilar membrane of the inner ear responds topographically to incoming sound waves with higher frequencies transduced to neural signals near the entrance to the cochlea and progressively lower frequencies transduced further along the membrane. This organized gradient of frequencies (i.e., tones) is referred to as tonotopy (i.e., a map of tones); this topography may also be termed cochleotopy, referring to a map of the cochlea. Tonotopic organization is maintained as auditory information is processed and passed on from the inner ear through the brainstem, to the thalamus, and into PAC along Heschl's gyrus (HG; **Figure 1**; for additional discussion, see [2, 5, 6, 12, 99, 100]). The preservation of such topographical organization from the basilar membrane of the inner ear to auditory cortex allows for a common reference frame across this hierarchically organized sensory system [6, 7, 12, 13, 22, 23].

A second fundamental aspect of the auditory feature space is temporal sound information, termed periodicity (**Figure 7B, D**) [10, 101, 102]. Human psychoacoustic studies indicate that there are separable filter banks (i.e., neurons with distinct receptive fields) for not only frequency spectra—as expected given tonotopy, but also temporal information [103–105]. The auditory nerve likely encodes such temporal information through activity time-locked to the periodicity of the amplitude modulation (i.e., the length of time from peak-to-peak of the temporal envelope) [101, 106]. Temporally varying aspects of sound are thought to preferentially active neurons selective for the onset and offset of sounds and for sounds of certain durations. Organized representations of periodicity in primates have been

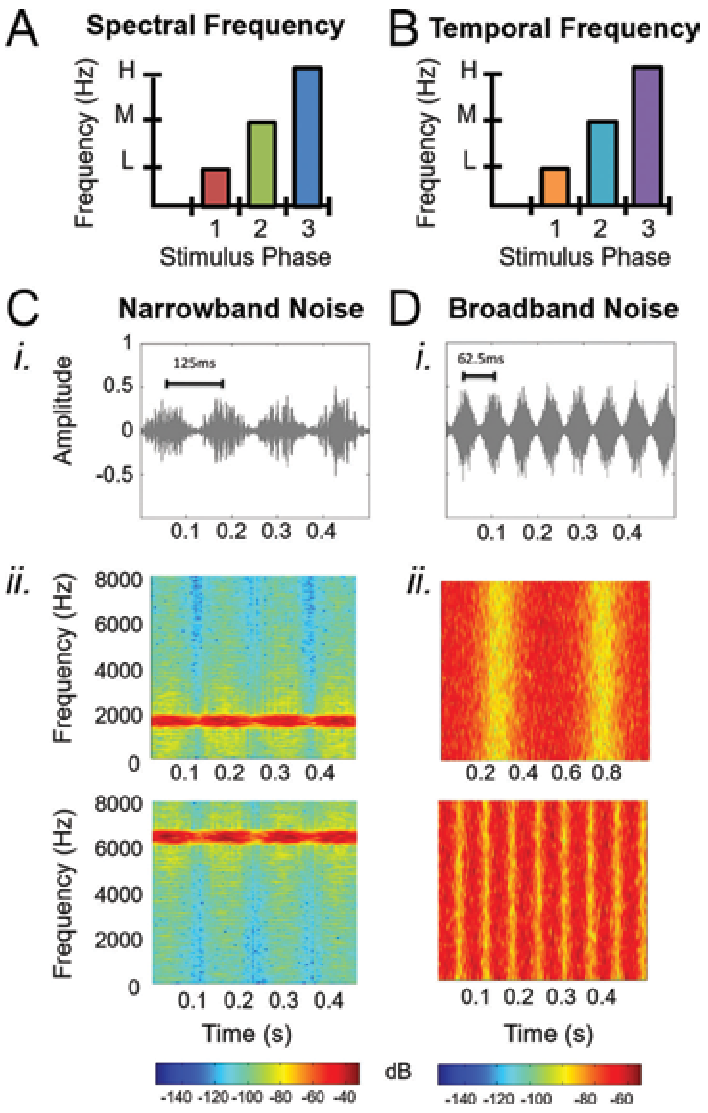

Figure 7.
Example tonotopic and periodotopic stimuli for auditory field mapping. (A) Three stimulus values for one dimension of auditory feature space (e.g., tonotopy) are depicted in the graph: 1—low (L, red); 2—medium (M, green); 3—high (H, blue). (B) Three stimulus values for a second dimension of auditory feature space (e.g., periodotopy) are depicted in the second graph: 1—low (L, orange); 2—medium (M, aqua); 3—high (H, purple). (C) Tonotopic representations can be measured using narrowband noise stimuli, which hold periodicity constant and vary frequency. (i) Sound amplitude (arbitrary units) for this stimulus set as a function of time in seconds. (ii) Sound spectrograms for two example narrowband noise stimuli with center frequencies (CF) of 1600 Hz (top) and 6400 Hz (bottom). Higher amplitudes in decibels (dB) are represented as 'warmer' colors (see dB legend below). (D) Periodotopic representations can be measured using broadband noise stimuli, which maintain constant frequency information and vary periodicity. (i) Sound amplitude (arbitrary units) for this stimulus set as a function of time in seconds. (ii) Sound spectrograms for two example broadband noise stimuli with amplitude modulation (AM) rates of 2 Hz (top) and 8 Hz (bottom). Higher amplitudes are again depicted as 'warmer' colors (see dB legend on bottom).

measured to date in the thalamus and PAC of macaque and human, respectively, and are termed periodotopy, a map of neurons that respond differentially to sounds of different temporal envelope modulation rates [5, 10, 107]. Repeating periodotopic gradients exist in the same cortical locations as, but are orthogonal to, tonotopic gradients, which allows researchers to use measurements of these two acoustic dimensions to identify complete AFMs.

3.2 fMRI measurements of auditory field maps

Measurements of the structure and function of human PAC and lower-level auditory cortex have been relatively few to date, with many studies hampered by methodological issues (for reviews, see [5, 23]. Precise measurements of AFMs across primary and lower-level auditory cortex are vital, however, for studying the neural underpinnings of such prominent auditory behaviors as attention and working memory. Recent research has now successfully applied fMRI methods commonly used to measure visual field maps to the study of AFMs in human auditory cortex.

3.2.1 Phase-encoded fMRI

The phase-encoded fMRI paradigm provides highly detailed *in vivo* measurements of CFMs in individual subjects [9, 10, 15, 108–111]. This technique measures topographical representations using stimuli that periodically repeat a set of values in an orderly sequence (**Figure 7**). The phase-encoded methods are specialized for AFM measurements by combining this periodic stimulus with a sparse-sampling paradigm (**Figure 8**) [10, 112–115]. Sparse-sampling separates the auditory stimulus presentation from the noise of the MR scanner during data acquisition to avoid contamination of the data by nonstimulus sounds [116–118].

The periodic stimulus allows for the use of a Fourier analysis to determine the value of the stimulus (e.g., 800 Hz frequency for tonotopy) that most effectively drives each cortical location [110]. The cortical response at a specific location is said to be ‘in phase’ throughout the scan with the stimulus value that most effectively activates it, hence the term ‘phase-encoded’ mapping. The alternate term ‘traveling-wave’ mapping arises from the consecutive activation of one neighboring cortical location after the other to create a wave-like pattern of activity across the CFM during the stimulus presentation. The phase-encoded paradigm only captures cortical activity that is at the stimulus frequency, thus excluding unrelated cortical activity and other sources of noise. Similarly, cortical regions that are not organized topographically will not be significantly activated by phase-encoded stimuli, as there would be no differential activation across the cortical representation [8, 15, 16]. The statistical threshold for phase-encoded cortical activity is commonly determined by coherence, which is a measure of the amplitude of the BOLD signal modulation at the frequency of the stimulus presentation (e.g., six stimulus cycles per scan), divided by the square root of the power over all other frequencies except the first and second harmonic (e.g., 12 and 18 cycles per scan) [15, 17, 110].

Measurement and analysis of phase-encoded CFM data must be performed within individual subjects rather than across group averages to avoid problematically blurring together discrete CFMs and their associated computations (for extended discussions, see [5, 15, 17]). CFMs may differ radically in size and anatomical position among individual subjects independent of brain size; this variation is reflected in associated shifts in cytoarchitectural and topographic boundaries [119–124]. In the visual system, for example, V1 can differ in size by at least a factor of three despite its location on the relatively stable calcarine sulcus [120]. Accordingly, when such data are group-averaged across subjects, especially through such approaches as aligning data from individual brains to an average brain with atlases such as Talairach space [125] or Montreal Neurological Institute (MNI) coordinates [126], the measurements will be blurred to such a degree that the measured topography of the CFMs is inaccurate or even lost. Blurring from such whole-brain anatomical co-alignment will thus cause different CFMs to be incorrectly averaged together into a single measurement, mixing data together from adjacent CFMs within each subject and preventing the analysis of the distinct computations of each CFM.

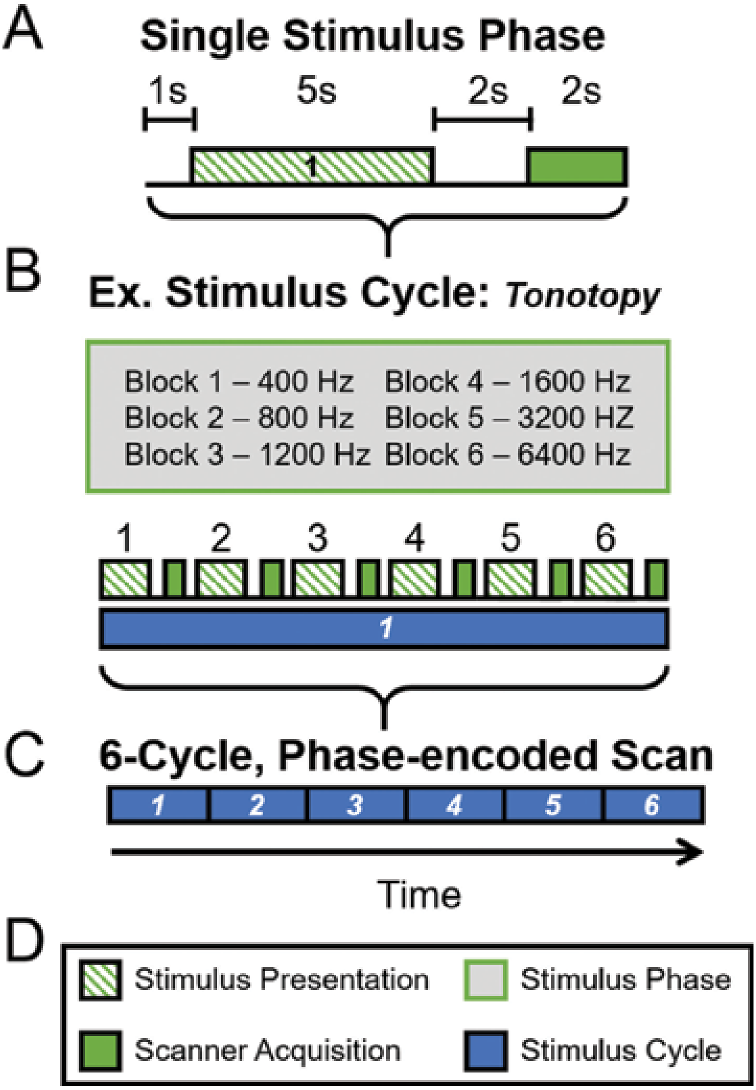

Figure 8.
Schematic of phase-encoded fMRI paradigm for auditory field mapping experiments. (A) Diagram of a single stimulus phase shows the components of a single block of one auditory stimulus presentation (striped green) followed by an fMRI data acquisition period (solid green). This sparse-sampling paradigm separates the auditory stimulus presentation from the noisy environment of the MR scanner acquisition. The timing of the acquisition (2 s delay) is set to collect the approximate peak response of auditory cortex to the stimulus, in accordance with the estimated hemodynamic delay. (B) Each phase (block) of an example tonotopic stimulus is displayed within the gray box above the colored blocks; one block thus represents one stimulus position in the 'phase-encoded' sequence. The diagram of an example stimulus cycle below this depicts six presentation blocks (striped green+ solid green) grouped together into one stimulus cycle (blue). Each block, or stimulus phase, in each cycle represents a specific frequency; e.g., for tonotopic measurements, the stimulus that is presented sequentially changes to each of the Hz listed in the gray box. The term 'traveling-wave' is also used to describe this type of phase-encoded stimulus presentation, as the stimuli produce a sequential activation of representations across a topographically organized cortical region. (C) Diagram shows a full, single scan comprising six cycles. (D) Legend denotes color-coding for diagrams above.

3.2.2 Criteria for auditory field map identification

In order to avoid the imprecise application of the term 'map' to topographical gradients or other similar patterns of cortical organization, the designation of an AFM—and CFMs in general—should be established according to several key criteria (**Figure 9**) (for reviews, see [5, 8, 15]). First, by definition, each AFM must contain at least the two orthogonal, nonrepeating topographical representations of fundamental acoustic feature space described above: tonotopy and periodotopy (**Figure 9A**) [10, 17, 21, 108, 110, 111]. When this criterion is ignored and the measurement of only one topographical representation is acquired (e.g., tonotopy), it is impossible to correctly identify boundaries among cortical regions. Measurements of the organization and function of specific regions of early auditory cortex in human long have mostly relied on tonotopic measurements alone, which has resulted in variable, conflicting, and ultimately unusable interpretations of the organization of human PAC and surrounding regions (for detailed reviews, see [5, 23]).

The representation of one dimension of sensory space—one topographical gradient along cortex like tonotopy—is not adequate to delineate an AFM, or CFMs in any sensory system. The measurement of a singular topographical dimension merely demonstrates that this particular aspect of sensory feature space is represented along that cortical region. The CFMs within that cortical region cannot be identified without measuring an orthogonal second dimension: a region of cortex with a large, confluent gradient for one dimension could denote a single CFM (**Figure 9Ai, ii**) or many CFMs (**Figure 9Ai, iii**), depending upon the organization of the overlapping second topography. Similarly, the two overlapping gradients must be approximately orthogonal, as they will otherwise not represent all the points in sensory space uniquely (**Figure 9B**) [15, 16, 127, 128]. As the complexity of adjacent gradients increases, the determination of the emergent CFM organization grows increasingly complicated.

Due to the relatively recent measurements of periodotopic representations in human auditory cortex and monkey midbrain, AFMs in core and belt regions can now be identified [10, 102]. The identification of periodotopy as the second key dimension of auditory feature space is strengthened by psychoacoustic studies, which show that separable filter banks occur not only for frequency spectra, but also temporal information, indicating the presence of neurons with receptive fields tuned to ranges of frequencies and periods [14, 103–105]. Additionally, representations of temporal acoustic information (i.e., periodicity) have been measured in the auditory system of other model organisms, including PAC in domestic cat and inferior colliculus in chinchilla [129, 130].

A second AFM criterion is that each of its topographical representations must be organized as a generally contiguous and orderly gradient [16, 128]. For such a gradient to develop, the representation must be organized such that it covers a full range of sensory space, in order from one boundary to the other (e.g., from lower to upper frequencies for tonotopy; **Figure 9C**). A topographical gradient is thus one of the most highly structured features of the cortical surface that can be measured using fMRI. The odds of two orderly, orthogonal gradients arising as a spurious pattern from noise in an overlapping section of cortex is extraordinarily low (for a calculation of the probability of spurious gradients arising from noise, see [19]).

Third, each CFM should contain representations of a considerable amount of sensory space. Differences in cortical magnification are likely among CFMs with different computational needs, but a large portion of sensory space is still expected to be represented (e.g., [15, 16, 19, 21, 97, 127, 131]). A high-quality fMRI measurement of the topography is necessary to adequately capture the sensory range and magnification. The quality of the measurement is dependent upon choosing an appropriate set of phase-encoded stimuli. The sampling density and range of values in the stimulus set both affect the accuracy and precision of the measurement. For example, the intensity (i.e., loudness) of the tonotopic stimulus alone can alter the width of the receptive fields of neurons in PAC and consequently increase the

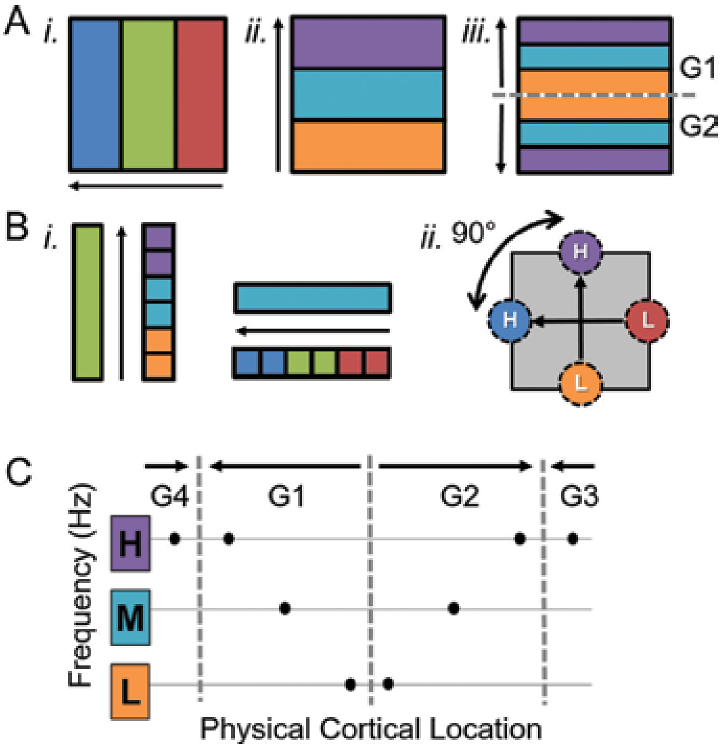

Figure 9.

Definition of auditory field maps (AFMs). (A) (i) Schematic of a single gradient of dimension 1 (e.g., tonotopy). Black arrow shows the low-to-high gradient for this tonotopic gradient. With only measurements of the single dimension of tonotopy, it cannot be determined whether the region within dimension 1 contains one or more cortical field maps without measuring a second, orthogonal gradient. (ii) Schematic of a single gradient of dimension 2 (e.g., periodotopy) overlapping the tonotopic gradient in (i) to form a single AFM like hA1. Black arrow shows the low-to-high gradient for this periodotopic gradient. Note the orthogonal orientation of the two gradients (i vs. ii) composing this AFM. (iii) schematic of an alternative gradient organization for periodotopy overlapping the same tonotopic gradient in (i). Black arrows now show two low-to-high gradients (G1: gradient 1, G2: gradient 2) of this second dimension within the same territory as the orthogonal low-to-high gradient in (i). The gray dotted line marks the boundary dividing this region into two AFMs. (B) (i) In a properly defined AFM, measurements along the cortical representation of a single value of tonotopy (e.g., green) span all values of periodotopy (e.g., orange to cyan to purple), and vice versa. (ii) Schematic of vectors drawn along a single CFM from centers of low-stimulus-value regions of interest (ROIs) to high-stimulus-value ROIs for dimensions 1 (e.g., red to blue) and 2 (e.g., orange to purple). The offset measured between the low-to-high vectors for each dimension should be approximately 90° to be considered orthogonal and thus allow for each voxel/portion of the map to represent a unique combination of dimension 1 and dimension 2 values. (C) The diagram demonstrates how gradient boundaries for one dimension of an AFM are determined. Black dots denote hypothetical measurement points along the cortical surface shown in (A, iii). Black arrows note gradient directions (low, L, to medium, M, to high, H). Dashed gray lines mark gradient reversals. Two gradients that span the full range of dimension 2 measurements can be divided into G1 and G2, with the representations of stimulus values increasing from low to high across the cortical surface in one gradient to the boundary where the representations in the next map then reverse back from high to low along the cortical surface in the next gradient. G3 and G4 (gradients 3 and 4, respectively) denote additional gradients continuing at reversal to regions outside the diagram. (for review, see [23]).

lateral spread of the BOLD signal measured in neuroimaging [132]. In addition, some degree of blurring in the measurements of the topography is expected due to such factors as the overlapping broad receptive fields, the inherent spatial spread of the fMRI signal, and measurement noise [64, 109, 133, 134]. The stimulus parameters and how they may affect the cortical responses should therefore be given careful consideration.

Fourth, the general features of the topographies composing the CFMs and the pattern of CFMs across cortex should both be consistent among individuals. It is essential to remember, nevertheless, that cytoarchitectural and topographic boundaries in PAC vary dramatically in size and anatomical location independent of overall brain size [119, 121–124, 135], as do CFMs across visual cortex [16, 17, 120, 136]. Regardless of these variations, the overall organization among specific CFMs and cloverleaf clusters will be maintained across individuals.

3.2.3 Definition of auditory field map boundaries

The measurement of AFMs is one of the few reliable *in vivo* methods to localize the distinct borders of the auditory core and belt regions in individual subjects [5, 10, 12, 23]. The boundaries of an AFM—and of CFMs in general—are determined by carefully defining the edges of overlapping sections of tonotopic and periodotopic gradients within a specific cortical region in an individual hemisphere (**Figure 9**). If a set of overlapping representations of the two dimensions is present in isolation, the boundary of the AFM can be estimated to be where the gradient responses end, although there will likely be some spatial blurring or spreading of the representation along these edges (**Figure 9Ai, ii**) [16, 17, 110, 137]. For multiple, adjacent representations that each span the full range of one dimension (e.g., low-to-high frequencies of tonotopy) can be divided into two sections at the point at which the gradients reverse (**Figure 9Ai, iii**). At the gradient reversals, the representations of stimulus values increase from low to high (or vice versa) across the cortical surface in one section to the boundary where the representations in the next AFM then reverse back from high to low (or vice versa) along the cortical surface in the next section (**Figure 9C**). Such phase-encoded fMRI measurements of the boundaries of the AFMs in human auditory cortex have been shown to be closely related to those determined by invasive human cytoarchitectural studies and nonhuman primate cytoarchitectural, connectivity, and tonotopic measurements [2, 5, 10, 121, 138–144].

At a scale of several centimeters, groups of adjacent CFMs are organized within both auditory and visual cortex into a macrostructural pattern called the cloverleaf cluster, named for the similarity of the organization of the individual CFMs composing a cluster to the leaves of a clover plant [8, 10, 15–20]. Within a cluster, one dimension of sensory topography is represented in concentric, circular bands from center to periphery of the cluster, and the second, orthogonal dimension separates this confluent representation into multiple CFMs with radial bands spanning the cluster center to periphery. In AFM clusters, a confluent, concentric tonotopic representation is divided into specific AFMs by reversal in the orthogonal periodotopic gradients. Neighboring cloverleaf clusters are then divided along the tonotopic reversals at the cluster boundaries.

While CFM clusters have consistent positions relative to one another across the cortical surface, CFMs within each cluster may be oriented differently among individuals as if rotating about a cluster's central representation. This inter-subject is consistent with the variability in molecular gradient expression that gives rise to the development of cortical topographical gradients [145–149]. This unpredictability of cluster anatomical location and rotation emphasizes the need for careful data analysis to be performed in individual subjects, in which common CFMs can be identified by analyzing the pattern of CFMs and cloverleaf clusters within that sensory system.

3.3 Organization of human auditory field maps

3.3.1 Auditory cortex organization in macaque monkey vs. human

Auditory processing in human cortex and in nonhuman primates occurs bilaterally along the temporal lobes near the lateral sulcus (**Figure 1**; e.g., [5, 10, 115, 121, 139–142, 144, 150–153]). In the macaque monkey model system upon which much of our understanding of human audition is based, converging evidence from cytoarchitectural, connectivity, electrophysiological, and neuroimaging studies have generally identified 13 auditory cortical areas grouped into core, medial and lateral belt, and parabelt regions that are associated with primary, secondary, and tertiary levels of processing, respectively (for extended discussions, see [2, 5, 154]). Auditory processing in macaque cortex begins along the superior temporal gyrus (STG) within three primary auditory areas: A1, R, and RT [140]. In contrast to early visual processing in which primary visual cortex is composed of V1 alone, primary auditory cortex is considered to be a core region composed of these three AFMs; all three areas contain the expanded layer IV arising from dense thalamic inputs and the high expression of cytochrome oxidase, acetylcholinesterase, and parvalbumin distinctive to primary sensory cortices [2, 142, 143, 150, 152, 154–157]. The eight belt regions are divided into four areas along both the lateral (CL, ML, AL, RTL) and medial (CM, RM, MM, RTM) sides of the core [158–160]. Along the lateral belt, two additional areas create the parabelt, which allocates auditory information to neighboring auditory cortex as well as to multimodal cortical regions [2, 161].

Based on cytoarchitectural, connectivity, and neuroimaging measurements, early auditory processing in human cortex has been shown to resemble the organization of lower-level macaque auditory processing [10, 23, 121, 144, 151–153, 162]. Over the ~25 million years of evolutionary separation between the species, the core, belt, and parabelt areas have rotated from the STG to Heschl's gyrus (HG), an anatomical feature unique to humans [11, 163]. The specific structure of HG differs across individuals, variably existing as a single or double gyrus. PAC is then either mostly centered on the single HG or overlapping both gyri in the case of two (**Figure 1B, C**) [122, 135, 136]. Core, belt, and parabelt areas have thus shifted in orientation from a strictly rostral-caudal axis for A1 to R to RT along macaque STG to a medial-lateral axis along human HG for hA1, hR, and hRT. The naming of the AFMs in human is based on the likely homology to macaque, but adds an 'h' to signify human [10].

3.3.2 Eleven human AFMs compose three cloverleaf clusters overlapping Heschl's gyrus

With our new understanding of periodotopic representations overlapping the previously identified tonotopic gradients, *in vivo* fMRI measurements can now identify the 11 AFMs that compose the core and belt regions of human auditory cortex (**Figure 10**) [5, 10, 12, 23]. Running from STG to the circular sulcus (CiS) along HG are three distinct, concentrically organized, tonotopic representations. The primary circular tonotopic gradient is one dimension of the HG cloverleaf cluster, with a confluent low-tone representation located centrally and expanding smoothly to high-tone representations at the outer edge (**Figure 10B, C**) [5]. The HG cluster is divided along the orthogonal periodotopic reversals into two AFMs each of core, medial belt, and lateral belt: hA1, hR, hMM, hRM, hML, and hAL (**Figure 10D, E**). Positioned at the tip of HG, hA1 is the largest of these core and belt AFMs, with the posterior/lateral region representing low tones and the anterior/medial region representing high ones. HA1 is involved in the most basic of cortical auditory computations, which is reflected in its representations of broad ranges of tonotopy and periodotopy [2].

A reversal in the tonotopic gradient along the anteromedial edge of the HG cluster divides it from the CM/CL cluster just past the tip of HG (**Figure 10B, C**). A high-periodicity gradient reversal splits this tonotopic gradient into hCM, and hCL, two regions associated with early language and speech processing as well as audiovisual integration (**Figure 10D, E**) [164]. Finally, the reversal in the tonotopic gradient along the posteriolateral edge of the HG cluster separates it from the RT cluster positioned where HG meets STG (**Figure 10B, C**). Two reversals in

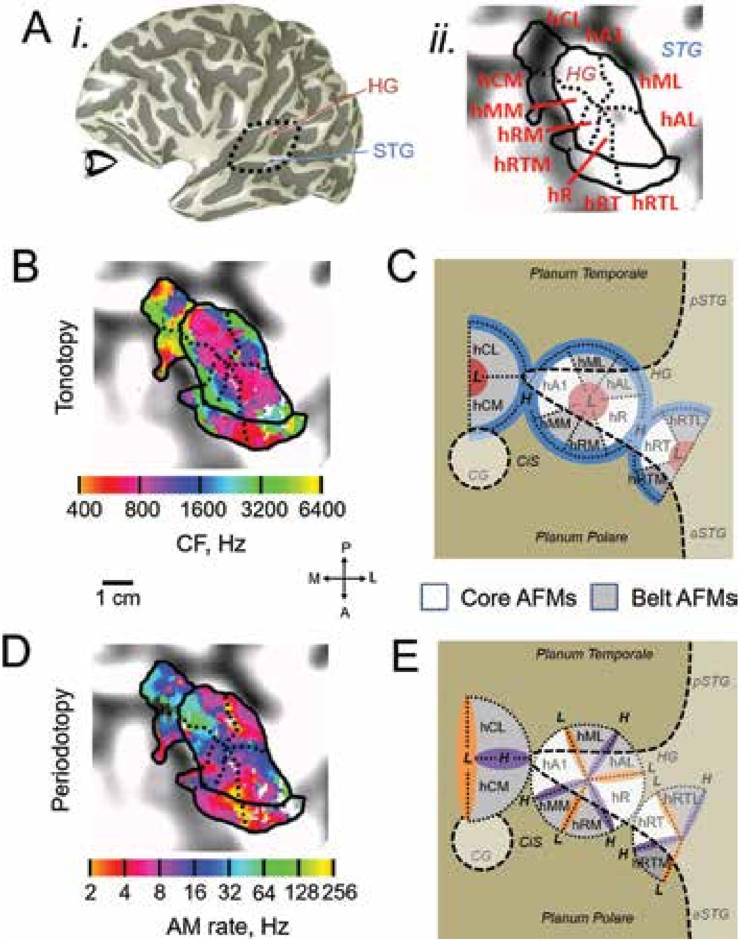

Figure 10.

Auditory field maps and cloverleaf clusters in human cortex. (A) Anatomical views of Heschl's gyrus (HG), superior temporal gyrus (STG) and surrounding auditory cortex in an individual subject's left hemisphere (S2). (i) Inflated 3-D rendering of the cortical surface. Light gray denotes gyri; dark gray denotes sulci. The approximate region presented in the other panels is indicated by the dotted black line. Note that this subject has a double peak along HG. (ii) flattened cortical surface of the region indicated by the dotted black line in (i). AFM boundaries between maps along tonotopic reversals are indicated by solid black lines. These tonotopic reversals constitute the separation of cloverleaf clusters from one another. AFM boundaries along periodotopic reversals are indicated by dotted black lines. These periodotopic reversals compose the separation between maps within a cloverleaf cluster. Red text indicates AFM names. (B) Tonotopic gradients measured using narrowband noise stimuli with a phase-encoded fMRI paradigm (example single-subject data from [10]). Color overlay indicates the preferred frequency range for each voxel. CF: center frequency in Hz. For clarity, only voxels within the core and belt AFMs are shown. Solid and dotted black lines are as in (A). Coherence ≥0.20. Inset scale bar designates 1 cm along the flattened cortical surfaces in (B, D). Inset legend indicates anatomical directions for (B-E). M: medial; L: lateral; A: anterior; P: posterior. (C) Diagram is based on individual-subject data measured by [10] in multiple phase-encoded fMRI experiments. Approximate positions of core AFMs (hA1, hR, hRT) are shown in white, and approximate positions of belt AFMs (hML, hAL, hRTL, hRTM, hRM, hMM, hCM, hCL) are shown in gray. Darker beige background indicates the plane of the lateral sulcus, while lighter beige overlay indicates gyri. Gyri are also marked with dashed black lines. HG: Heschl's gyrus. CG: circular gyrus; CiS: circular sulcus; a/p STG: anterior/posterior superior temporal gyrus. Diagram depicts the locations of tonotopic representations overlaid along the core and belt AFMs, with low (L) and high (H) tonotopic representations are marked in red and blue, respectively. Dotted black lines designate the boundaries between AFMs within three cloverleaf clusters: HG cluster with hA1; hCM/hCL cluster (partial cluster defined to date); hRTM/hRT/hRTL cluster (partial cluster defined to date). (D) Periodotopic representations measured using broadband noise stimuli with a phase-encoded fMRI paradigm. Data are from the same subject as shown for tonotopy in (B), with the color overlay now indicating the preferred period range for each voxel. AM rate: amplitude modulation rate in Hz. Other details are as in (B). (E) Diagram depicts periodotopic representations overlaid on the same example region of cortex as in (C). L and H now designate to the approximate locations of low (orange) or high (purple) periodotopic representations, respectively. Adapted from Barton et a.[10]. For a detailed review, see [5].

the periodotopic representations here divide the RT cluster into hRT, hRTM, and hRTL (**Figure 10D, E**). In macaque, these AFMs along STG are thought to subserve lower-level processing of auditory stimuli like temporally modulated environmental sounds [158, 159]. More research is needed to determine how what other AFMs form the CM/CL and RT clusters. Based on emerging data, it is likely that AFMs will also be a fundamental organization of auditory cortex adjacent to these cloverleaf clusters, such as planum temporale (PT), planum polare (PP) and STG.

3.4 Measuring attention and working memory in human AFMs

The characterization of AFMs and cloverleaf clusters will be crucial for the study of the structure and function of human auditory cortex, as these *in vivo* measurements allow for the systematic exploration of computations across a sensory system (for reviews, see [5, 17]). Such AFM organization provides a basic framework for the complex processing and analysis of input from the sensory receptors of the inner ear [5, 12, 17, 23]. The cloverleaf cluster organization of AFMs may also play a role in coordinating neural computations, with neurons within each cluster sharing computational resources such as common mechanisms to coordinate neural timing or short-term information storage [8, 12]. Similarly, vision studies suggest that functional specializations for perception are organized by cloverleaf clusters, as a particular cloverleaf cluster can be functionally differentiated from its neighbors by its pattern of BOLD responses, surface area, cortical magnification, processing specialization, and receptive field sizes [12, 16, 18, 19, 21, 165]. These distinctions indicate that CFMs within individual cloverleaf clusters are not only anatomically but also functionally related [15, 18, 20, 166].

The cluster organization is not necessarily thought to be driving common sensory functions, but rather reflects how multiple stages in a sensory processing pathway might arise during development across individuals and during evolution across species. It is likely that this cluster organization, like the topographic organization of CFMs, allows for efficient connectivity among neurons that represent neighboring aspects in sensory feature space [166–169]. Since the axons contained within one cubic millimeter of cortex can extend 3-4 km in length, efficient connectivity is vital for sustainable energetics in cortex [170].

The definitions of AFMs and the cloverleaf clusters they compose using phase-encoded fMRI will thus serve as reliable, independent localizers for investigations of attention and working memory in early auditory cortex across individuals. Measurements of individual AFMs along the cortical hierarchy will help reveal the distinct stages of top-down and bottom-up auditory processing. In addition, changes in AFMs can be tracked to study how auditory cortex changes under various attentional and working memory tasks and disorders (e.g., [145, 171–177]).

4. Conclusion

The human brain has sophisticated systems for perception, trace memory, attention, and working memory for audition and vision, and likely the other senses as well. These systems appear to be organized in a very similar manner for each sense, despite the inputs to each system and information content being quite different. Behavioral measures of the last several decades have led to the development of well-defined models of each system. These models form the basis for the investigation of their underlying architecture in the cortical structures of the human brain. EEG and PET have allowed for spatially coarse investigation of cortical activity, but with the advent of fMRI, it has become possible to make exceptionally detailed spatial measurements. The methods of investigation must be carefully crafted to best elicit activity reflecting the desired aspects of each system; not only must the tasks be appropriate for fMRI, the stimuli and task must be closely matched not just to the system being studied, but to the inputs into that system as well.

For both audition and vision, the sensory processing in cortex happens in cloverleaf clusters of CFMs. This organizational pattern has clearly been demonstrated in the lower tiers of the processing hierarchy and very likely is organized as such throughout. Because the CFMs across the entire hierarchy (or at least, most) of one sense can be measured in just one session in the fMRI scanner, they make incredibly efficient localizers. CFMs are be measured in individual subjects, and serve as functional localizers that can be used to average more accurately across subjects than anatomical localizers. As such, due to the pervasive and fundamental role CFMs play in sensory systems, they are also excellent candidates for measuring the effects of attention and working memory in cortex. To best accomplish this feat, it is proposed that stimuli that are similar to those used to measure CFMs are excellent candidates for use in traditional tasks used to define attentional and working-memory models.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant Number 1329255 and by startup funds from the Department of Cognitive Sciences at the University of California, Irvine.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author details

Brian Barton[1] and Alyssa A. Brewer[1,2]*

1 Department of Cognitive Sciences, University of California, Irvine, CA, USA

2 Department of Language Science, University of California, Irvine, CA, USA

*Address all correspondence to: aabrewer@uci.edu

References

[1] Van Essen DC, Felleman DJ, DeYoe EA, Olavarria J, Knierim J. Modular and hierarchical organization of extrastriate visual cortex in the macaque monkey. Cold Spring Harbor Symposia on Quantitative Biology. 1990;**55**:679-696

[2] Kaas JH, Hackett TA. Subdivisions of auditory cortex and processing streams in primates. Proceedings of the National Academy of Sciences of the United States of America. 2000;**97**(22):11793-11799. DOI: 10.1073/pnas.97.22.11793

[3] Krubitzer L. The magnificent compromise: Cortical field evolution in mammals. Neuron. 2007;**56**(2):201-208. DOI: 10.1016/j.neuron.2007.10.002

[4] Schreiner CE, Winer JA. Auditory cortex mapmaking: Principles, projections, and plasticity. Neuron. 2007;**56**(2):356-365. DOI: 10.1016/j.neuron.2007.10.013

[5] Brewer AA, Barton B. Maps of the auditory cortex. Annual Review of Neuroscience. 2016;**39**:385-407. DOI: 10.1146/annurev-neuro-070815-014045

[6] Wessinger CM, VanMeter J, Tian B, Van Lare J, Pekar J, Rauschecker JP. Hierarchical organization of the human auditory cortex revealed by functional magnetic resonance imaging. Journal of Cognitive Neuroscience. 2001;**13**(1):1-7

[7] Kaas JH. Topographic maps are fundamental to sensory processing. Brain Research Bulletin. 1997;**44**(2):107-112

[8] Wandell BA, Brewer AA, Dougherty RF. Visual field map clusters in human cortex. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 2005;**360**(1456):693-707. DOI: 10.1098/rstb.2005.1628

[9] Sanchez-Panchuelo RM, Francis S, Bowtell R, Schluppeck D. Mapping human somatosensory cortex in individual subjects with 7T functional MRI. Journal of Neurophysiology. 2010;**103**(5):2544-2556. DOI: 10.1152/jn.01017.2009

[10] Barton B, Venezia JH, Saberi K, Hickok G, Brewer AA. Orthogonal acoustic dimensions define auditory field maps in human cortex. Proceedings of the National Academy of Sciences of the United States of America. 2012;**109**(50):20738-20743. DOI: 10.1073/pnas.1213381109

[11] Krubitzer LA, Seelke AM. Cortical evolution in mammals: the bane and beauty of phenotypic variability. Proceedings of the National Academy of Sciences. 2012;**109**(Supplement 1): 10647-10654

[12] Brewer AA, Barton B. Cloverleaf clusters: A common macrostructural organization across human visual and auditory cortex. In: Heinbockel T, editor. Sensory Nervous System. London, UK: IntechOpen; 2018. pp. 127-160. DOI: 10.5772/intechopen.77964

[13] Felleman DJ, Van Essen DC. Distributed hierarchical processing in the primate cerebral cortex. Cerebral Cortex. 1991;**1**(1):1-47

[14] Santoro R, Moerel M, De Martino F, Valente G, Ugurbil K, Yacoub E, et al. Reconstructing the spectrotemporal modulations of real-life sounds from fMRI response patterns. Proceedings of the National Academy of Sciences of the United States of America. 2017;**114**(18):4799-4804. DOI: 10.1073/pnas.1617622114

[15] Brewer AA, Barton B. Visual field map organization in human visual cortex. In: Molotchnikoff S, Rouat J, editors. Visual Cortex—Current Status and Perspectives. Rijeka, Croatia: InTech; 2012. pp. 29-60. DOI: 10.5772/51914

[16] Brewer AA, Liu J, Wade AR, Wandell BA. Visual field maps and stimulus selectivity in human ventral occipital cortex. Nature Neuroscience. 2005;**8**(8):1102-1109. DOI: 10.1038/nn1507

[17] Wandell BA, Dumoulin SO, Brewer AA. Visual field maps in human cortex. Neuron. 2007;**56**(2):366-383. DOI: 10.1016/j.neuron.2007.10.012

[18] Kolster H, Peeters R, Orban GA. The retinotopic organization of the human middle temporal area MT/V5 and its cortical neighbors. The Journal of Neuroscience. 2010;**30**(29):9801-9820. DOI: 10.1523/JNEUROSCI.2069-10.2010

[19] Barton B, Brewer AA. Visual field map clusters in high-order visual processing: Organization of V3A/V3B and a new cloverleaf cluster in the posterior superior temporal sulcus. Frontiers in Integrative Neuroscience. 2017;**11**:4. DOI: 10.3389/fnint.2017.00004

[20] Kolster H, Mandeville JB, Arsenault JT, Ekstrom LB, Wald LL, Vanduffel W. Visual field map clusters in macaque extrastriate visual cortex. The Journal of Neuroscience. 2009;**29**(21):7031-7039. DOI: 10.1523/JNEUROSCI.0518-09.2009

[21] Press WA, Brewer AA, Dougherty RF, Wade AR, Wandell BA. Visual areas and spatial summation in human visual cortex. Vision Research. 2001;**41**(10-11):1321-1332

[22] Van Essen DC. Organization of Visual Areas in macaque and human cerebral cortex. In: Chalupa LM, Werner JS, editors. The Visual Neurosciences. Boston: Bradford Books; 2003. pp. 507-521

[23] Brewer AA, Barton B. Human auditory cortex. In: Hickok G, Small SL, editors. Neurobiology of Language. Cambridge: Academic Press, Elsevier; 2016. pp. 49-58. DOI: 10.1016/B978-0-12-407794-2.00005-5

[24] Kaya EM, Elhilali M. Modelling auditory attention. Philosophical Transactions of The Royal Society B: Biological Sciences. 2017;**372**(1714):1-10. DOI: 10.1098/rstb.2016.0101

[25] Fritz JB, Elhilali M, David SV, Shamma SA. Auditory attention—Focusing the searchlight on sound. Current Opinion in Neurobiology. 2007;**17**(4):437-455

[26] Driver J. A selective review of selective attention research from the past century. British Journal of Psychology. 2001;**92**(1):53-78

[27] Dalton P, Spence C. Selective attention in vision, audition, and touch. In: Byrne J, editor. Learning and Memory: A Comprehensive Reference. Oxford: Elsevier; 2008. pp. 243-257

[28] Corbetta M, Shulman GL. Control of goal-directed and stimulus-driven attention in the brain. Nature Reviews. Neuroscience. 2002;**3**(3):201-215. DOI: 10.1038/nrn755

[29] Corbetta M, Kincade JM, Shulman GL. Neural systems for visual orienting and their relationships to spatial working memory. Journal of Cognitive Neuroscience. 2002;**14**(3):508-523. DOI: 10.1162/089892902317362029

[30] Spence C, Santangelo V. Capturing spatial attention with multisensory cues: A review. Hearing Research. 2009;**258**(1-2):134-142. DOI: 10.1016/j.heares.2009.04.015

[31] Santangelo V, Olivetti Belardinelli M, Spence C, Macaluso E. Interactions between voluntary and stimulus-driven spatial attention mechanisms across sensory modalities. Journal of Cognitive Neuroscience. 2009;**21**(12):2384-2397. DOI: 10.1162/jocn.2008.21178

[32] Hurlstone MJ, Hitch GJ, Baddeley AD. Memory for serial order across domains: An overview of the literature

and directions for future research. Psychological Bulletin. 2014;**140**(2):339-373. DOI: 10.1037/a0034221

[33] Treisman AM. Strategies and models of selective attention. Psychological Review. 1969;**76**(3):282

[34] Broadbent DE. Perception and Communication. New York: Pergamon Press; 1958. p. 338

[35] Bharadwaj HM, Lee AK, Shinn-Cunningham BG. Measuring auditory selective attention using frequency tagging. Frontiers in Integrative Neuroscience. 2014;**8**:6. DOI: 10.3389/fnint.2014.00006

[36] Woods DL, Alain C, Diaz R, Rhodes D, Ogawa KH. Location and frequency cues in auditory selective attention. Journal of Experimental Psychology: Human Perception and Performance. 2001;**27**(1):65

[37] Zatorre RJ, Mondor TA, Evans AC. Auditory attention to space and frequency activates similar cerebral systems. NeuroImage. 1999;**10**(5):544-554

[38] Zatorre RJ, Bouffard M, Ahad P, Belin P. Where is' where'in the human auditory cortex? Nature Neuroscience. 2002;**5**(9):905

[39] Maddox RK, Atilgan H, Bizley JK, Lee AK. Auditory selective attention is enhanced by a task-irrelevant temporally coherent visual stimulus in human listeners. eLife. 2015;**4**:1-11. DOI: 10.7554/eLife.04995

[40] Kong L, Michalka SW, Rosen ML, Sheremata SL, Swisher JD, Shinn-Cunningham BG, et al. Auditory spatial attention representations in the human cerebral cortex. Cerebral Cortex. 2014;**24**(3):773-784. DOI: 10.1093/cercor/bhs359

[41] Awh E, Pashler H. Evidence for split attentional foci. Journal of Experimental Psychology: Human Perception and Performance. 2000;**26**(2):834

[42] Conway AR, Cowan N, Bunting MF. The cocktail party phenomenon revisited: The importance of working memory capacity. Psychonomic Bulletin & Review. 2001;**8**(2):331-335

[43] Alain C, Izenberg A. Effects of attentional load on auditory scene analysis. Journal of Cognitive Neuroscience. 2003;**15**(7):1063-1073

[44] Shinn-Cunningham BG. Object-based auditory and visual attention. Trends in Cognitive Sciences. 2008;**12**(5):182-186

[45] Saygin AP, Sereno MI. Retinotopy and attention in human occipital, temporal, parietal, and frontal cortex. Cerebral Cortex. 2008;**18**(9):2158-2168. DOI: 10.1093/cercor/bhm242

[46] Lauritzen TZ, D'Esposito M, Heeger DJ, Silver MA. Top-down flow of visual spatial attention signals from parietal to occipital cortex. Journal of Vision. 2009;**9**(13):18.1-18.14. DOI: 10.1167/9.13.18

[47] Cavanagh P, Alvarez GA. Tracking multiple targets with multifocal attention. Trends in Cognitive Sciences. 2005;**9**(7):349-354. DOI: 10.1016/j.tics.2005.05.009

[48] Baddeley A. Working memory: Looking back and looking forward. Nature Reviews. Neuroscience. 2003;**4**(10):829-839

[49] Cowan N. The magical number 4 in short-term memory: A reconsideration of mental storage capacity. Behavioral and Brain Sciences. 2001;**24**(01):87-114

[50] Barton B, Brewer AA. Visual working memory in human cortex. Psychology (Irvine). 2013;**4**(8):655-662. DOI: 10.4236/psych.2013.48093

[51] Postle BR. Working memory as an emergent property of the mind and brain. Neuroscience. 2006;**139**(1):23-38. DOI: 10.1016/j.neuroscience.2005.06.005

[52] Sperling G. The information available in brief visual presentations. Psychological Monographs: General and Applied. 1960;**74**(11):1-29. DOI: 10.1037/h0093759

[53] Awh E, Vogel EK, Oh S-H. Interactions between attention and working memory. Neuroscience. 2006;**139**(1):201-208

[54] Hitch GJ, Hu Y, Allen RJ, Baddeley AD. Competition for the focus of attention in visual working memory: Perceptual recency versus executive control. Annals of the New York Academy of Sciences. 2018;1424(1): 64-75. DOI: 10.1111/nyas.13631

[55] Serences JT, Ester EF, Vogel EK, Awh E. Stimulus-specific delay activity in human primary visual cortex. Psychological Science. 2009;**20**(2):207-214

[56] Postle BR, Zarahn E, D'Esposito M. Using event-related fMRI to assess delay-period activity during performance of spatial and nonspatial working memory tasks. Brain Research Protocols. 2000;**5**(1):57-66

[57] Sereno AB, Amador SC. Attention and memory-related responses of neurons in the lateral intraparietal area during spatial and shape-delayed match-to-sample tasks. Journal of Neurophysiology. 2006;**95**(2):1078-1098. DOI: 10.1152/jn.00431.2005

[58] Pashler H. Familiarity and visual change detection. Perception & Psychophysics. 1988;**44**(4):369-378

[59] Pavani F, Turatto M. Change perception in complex auditory scenes. Perception & Psychophysics. 2008;**70**(4):619-629

[60] Simons DJ, Ambinder MS. Change blindness. Current Directions in Psychological Science. 2005;**14**(1):44-48

[61] Todd JJ, Marois R. Capacity limit of visual short-term memory in human posterior parietal cortex. Nature. 2004;**428**(6984):751

[62] Xu Y, Chun MM. Dissociable neural mechanisms supporting visual short-term memory for objects. Nature. 2006;**440**(7080):91-95. DOI: 10.1038/nature04262

[63] Logothetis NK. The neural basis of the blood-oxygen-level-dependent functional magnetic resonance imaging signal. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 2002;**357**(1424):1003-1037

[64] Logothetis NK, Wandell BA. Interpreting the BOLD signal. Annual Review of Physiology. 2004;**66**:735-769. DOI: 10.1146/annurev.physiol.66.082602.092845

[65] Sligte IG, Scholte HS, Lamme VAF. V4 activity predicts the strength of visual short-term memory representations. Journal of Neuroscience. 2009;**29**(23):7432-7438. DOI: 10.1523/Jneurosci.0784-09.2009

[66] Cherry EC. Some experiments upon the recognition of speech with one and two ears. Journal of the Acoustical Society of America. 1953;**25**:975-979

[67] Cherry EC. Some further experiments upon the recognition of speech, with one and with two ears. Journal of the Acoustical Society of America. 1954;**26**:554-559

[68] Kristensen LB, Wang L, Petersson KM, Hagoort P. The interface between language and attention: Prosodic focus marking recruits a general attention network in spoken language comprehension. Cerebral Cortex.

2013;**23**(8):1836-1848. DOI: 10.1093/cercor/bhs164

[69] Wood N, Cowan N. The cocktail party phenomenon revisited: How frequent are attention shifts to one's name in an irrelevant auditory channel? Journal of Experimental Psychology. Learning, Memory, and Cognition. 1995;**21**(1):255-260

[70] Wood NL, Cowan N. The cocktail party phenomenon revisited: Attention and memory in the classic selective listening procedure of Cherry (1953). Journal of Experimental Psychology. General. 1995;**124**(3):243-262

[71] Kaya EM, Elhilali M. Correction to 'Modelling auditory attention' Philosophical Transactions of The Royal Society B: Biological Sciences. 2017;**372**(1727):1. DOI: 10.1098/rstb.2017.0194

[72] Alho J, Green BM, May PJ, Sams M, Tiitinen H, Rauschecker JP, et al. Early-latency categorical speech sound representations in the left inferior frontal gyrus. NeuroImage. 2016;**129**:214-223. DOI: 10.1016/j.neuroimage.2016.01.016

[73] Culham JC, Brandt SA, Cavanagh P, Kanwisher NG, Dale AM, Tootell RB. Cortical fMRI activation produced by attentive tracking of moving targets. Journal of Neurophysiology. 1998;**80**(5):2657-2670

[74] Barton B, Ester EF, Awh E. Discrete resource allocation in visual working memory. Journal of Experimental Psychology. Human Perception and Performance. 2009;**35**(5):1359

[75] Awh E, Barton B, Vogel EK. Visual working memory represents a fixed number of items regardless of complexity. Psychological Science. 2007;**18**(7):622-628. DOI: 10.1111/j.1467-9280.2007.01949.x

[76] Hickok G, Poeppel D. The cortical organization of speech processing. Nature Reviews. Neuroscience. 2007;**8**(5):393-402. DOI: 10.1038/nrn2113

[77] Chomsky N. Knowledge of Language: Its Nature, Origin, and Use. Westport, Connecticut: Praeger Publishers; 1986

[78] Luck SJ, Vogel EK. The capacity of visual working memory for features and conjunctions. Nature. 1997;**390**(6657):279-281. DOI: 10.1038/36846

[79] Baddeley AD, Hitch GJ. Working memory. In: Bower GA, editor. ThePsychology of Learning and Motivation. Advances in Research and Theory. New York: Academic Press; 1974. p. 47-89. DOI: 10.1016/S0079-7421(08)60452-1

[80] Baddeley AD. The episodic buffer: A new component of working memory? Trends in Cognitive Sciences. 2000;**4**(11):417-423

[81] Baddeley AD, Hitch GJ, Allen RJ. From short-term store to multicomponent working memory: The role of the modal model. Memory & Cognition. 2018:1-14. https://doi.org/10.3758/s13421-018-0878-5

[82] Baddeley A. Exploring the central executive. The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology. 1996;**49**(1):5-28

[83] Sperling G, Speelman RG. Acoustic similarity and auditory short-term memory: Experiments and a model. In: Norman DA, editor. Models of human memory. New York: Academic Press; 1970. p. 151-202

[84] Niemeier M, Goltz HC, Kuchinad A, Tweed DB, Vilis T. A contralateral preference in the lateral occipital area: Sensory and attentional mechanisms.

Cerebral Cortex. 2005;**15**(3):325-331. DOI: 10.1093/cercor/bhh134

[85] Luck SJ, Woodman GF, Vogel EK. Event-related potential studies of attention. Trends in Cognitive Sciences. 2000;**4**(11):432-440

[86] Postle BR, Berger JS, Taich AM, D'Esposito M. Activity in human frontal cortex associated with spatial working memory and saccadic behavior. Journal of Cognitive Neuroscience. 2000;**12**(supplement 2):2-14. DOI: 10.1162/089892900564028

[87] Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC. Visual topography of human intraparietal sulcus. The Journal of Neuroscience. 2007;**27**(20):5326-5337. DOI: 10.1523/JNEUROSCI.0991-07.2007

[88] Kastner S, DeSimone K, Konen CS, Szczepanski SM, Weiner KS, Schneider KA. Topographic maps in human frontal cortex revealed in memory-guided saccade and spatial working-memory tasks. Journal of Neurophysiology. 2007;**97**(5):3494-3507. DOI: 10.1152/jn.00010.2007

[89] Szczepanski SM, Konen CS, Kastner S. Mechanisms of spatial attention control in frontal and parietal cortex. The Journal of Neuroscience. 2010;**30**(1):148-160. DOI: 10.1523/JNEUROSCI.3862-09.2010

[90] Zeki S, Bartels A. Toward a theory of visual consciousness. Consciousness and Cognition. 1999;**8**(2):225-259. DOI: 10.1006/ccog.1999.0390

[91] Xu Y, Chun MM. Visual grouping in human parietal cortex. Proceedings of the National Academy of Sciences of the United States of America. 2007;**104**(47):18766-18771. DOI: 10.1073/pnas.0705618104

[92] Xu Y, Chun MM. Selecting and perceiving multiple visual objects. Trends in Cognitive Sciences. 2009;**13**(4):167-174. DOI: 10.1016/j.tics.2009.01.008

[93] Phan ML, Schendel KL, Recanzone GH, Robertson LC. Auditory and visual spatial localization deficits following bilateral parietal lobe lesions in a patient with Balint's syndrome. Journal of Cognitive Neuroscience. 2000;**12**(4):583-600

[94] Harrison SA, Tong F. Decoding reveals the contents of visual working memory in early visual areas. Nature. 2009;**458**(7238):632-635. DOI: 10.1038/nature07832

[95] Rypma B, Berger JS, D'Esposito M. The influence of working-memory demand and subject performance on prefrontal cortical activity. Journal of Cognitive Neuroscience. 2002;**14**(5):721-731. DOI: 10.1162/08989290260138627

[96] Larsson J, Heeger DJ. Two retinotopic visual areas in human lateral occipital cortex. The Journal of Neuroscience. 2006;**26**(51):13128-13142. DOI: 10.1523/JNEUROSCI.1657-06.2006

[97] Huk AC, Dougherty RF, Heeger DJ. Retinotopy and functional subdivision of human areas MT and MST. Journal of Neuroscience. 2002;**22**(16):7195-7205. DOI: 20026661

[98] Spoendlin H. Sensory neural organization of the cochlea. The Journal of Laryngology and Otology. 1979;**93**(9):853-877

[99] Ress D, Chandrasekaran B. Tonotopic organization in the depth of human inferior colliculus. Frontiers in Human Neuroscience. 2013;**7**:586. DOI: 10.3389/fnhum.2013.00586

[100] Chang KH, Thomas JM, Boynton GM, Fine I. Reconstructing tone sequences from functional magnetic resonance imaging blood-oxygen level dependent responses within human primary auditory cortex.

Frontiers in Psychology. 2017;**8**:1983. DOI: 10.3389/fpsyg.2017.01983

[101] Langner G, Sams M, Heil P, Schulze H. Frequency and periodicity are represented in orthogonal maps in the human auditory cortex: Evidence from magnetoencephalography. Journal of Comparative Physiology. A. 1997;**181**(6):665-676

[102] Baumann S, Griffiths TD, Sun L, Petkov CI, Thiele A, Rees A. Orthogonal representation of sound dimensions in the primate midbrain. Nature Neuroscience. 2011;**14**(4):423-425. DOI: 10.1038/nn.2771

[103] Dau T, Kollmeier B, Kohlrausch A. Modeling auditory processing of amplitude modulation. II. Spectral and temporal integration. The Journal of the Acoustical Society of America. 1997;**102**(5 Pt 1):2906-2919

[104] Ewert SD, Dau T. Characterizing frequency selectivity for envelope fluctuations. The Journal of the Acoustical Society of America. 2000;**108**(3 Pt 1):1181-1196

[105] Hsieh IH, Saberi K. Detection of sinusoidal amplitude modulation in logarithmic frequency sweeps across wide regions of the spectrum. Hearing Research. 2010;**262**(1-2):9-18. DOI: 10.1016/j.heares.2010.02.002

[106] Langner G, Schreiner CE. Periodicity coding in the inferior colliculus of the cat. I. Neuronal mechanisms. Journal of Neurophysiology. 1988;**60**(6):1799-1822

[107] Baumann S, Joly O, Rees A, Petkov CI, Sun L, Thiele A, et al. The topography of frequency and time representation in primate auditory cortices. eLife. 2015;**4**:1-15. DOI: 10.7554/eLife.03256

[108] DeYoe EA, Carman GJ, Bandettini P, Glickman S, Wieser J, Cox R, et al. Mapping striate and extrastriate visual areas in human cerebral cortex. Proceedings of the National Academy of Sciences of the United States of America. 1996;**93**(6):2382-2386

[109] Engel SA, Glover GH, Wandell BA. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. Cerebral Cortex. 1997;**7**(2):181-192

[110] Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, et al. fMRI of human visual cortex. Nature. 1994;**369**(6481):525. DOI: 10.1038/369525a0

[111] Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, et al. Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. Science. 1995;**268**(5212):889-893

[112] Humphries C, Liebenthal E, Binder JR. Tonotopic organization of human auditory cortex. NeuroImage. 2010;**50**(3):1202-1211. DOI: 10.1016/j.neuroimage.2010.01.046

[113] Talavage TM, Sereno MI, Melcher JR, Ledden PJ, Rosen BR, Dale AM. Tonotopic organization in human auditory cortex revealed by progressions of frequency sensitivity. Journal of Neurophysiology. 2004;**91**(3):1282-1296. DOI: 10.1152/jn.01125.2002

[114] Petkov CI, Kayser C, Augath M, Logothetis NK. Optimizing the imaging of the monkey auditory cortex: Sparse vs. continuous fMRI. Magnetic Resonance Imaging. 2009;**27**(8):1065-1073. DOI: 10.1016/j.mri.2009.01.018

[115] Joly O, Baumann S, Balezeau F, Thiele A, Griffiths TD. Merging functional and structural properties of the monkey auditory cortex. Frontiers in Neuroscience. 2014;**8**:198. DOI: 10.3389/fnins.2014.00198

[116] Bandettini PA, Jesmanowicz A, Van Kylen J, Birn RM, Hyde JS. Functional MRI of brain activation induced by scanner acoustic noise. Magnetic Resonance in Medicine. 1998;**39**(3):410-416

[117] Gaab N, Gabrieli JD, Glover GH. Assessing the influence of scanner background noise on auditory processing. II. An fMRI study comparing auditory processing in the absence and presence of recorded scanner noise using a sparse design. Human Brain Mapping. 2007;**28**(8):721-732. DOI: 10.1002/hbm.20299

[118] Scarff CJ, Dort JC, Eggermont JJ, Goodyear BG. The effect of MR scanner noise on auditory cortex activity using fMRI. Human Brain Mapping. 2004;**22**(4):341-349. DOI: 10.1002/hbm.20043

[119] Clarke S, Morosan P. Architecture, connectivity, and transmitter receptors of human auditory cortex. In: Poeppel D, Overath T, Popper A, Richard R, editors. The Human Auditory Cortex. New York: Springer; 2012. pp. 11-38

[120] Dougherty RF, Koch VM, Brewer AA, Fischer B, Modersitzki J, Wandell BA. Visual field representations and locations of visual areas V1/2/3 in human visual cortex. Journal of Vision. 2003;**3**(10):586-598. DOI: 10:1167/3.10.1

[121] Galaburda A, Sanides F. Cytoarchitectonic organization of the human auditory cortex. The Journal of Comparative Neurology. 1980;**190**(3):597-610. DOI: 10.1002/cne.901900312

[122] Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K. Human primary auditory cortex: Cytoarchitectonic subdivisions and mapping into a spatial reference system. NeuroImage. 2001;**13**(4):684-701. DOI: 10.1006/nimg.2000.0715

[123] Rademacher J, Caviness VS Jr, Steinmetz H, Galaburda AM. Topographical variation of the human primary cortices: Implications for neuroimaging, brain mapping, and neurobiology. Cerebral Cortex. 1993;**3**(4):313-329

[124] Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, et al. Probabilistic mapping and volume measurement of human primary auditory cortex. NeuroImage. 2001;**13**(4):669-683. DOI: 10.1006/nimg.2000.0714

[125] Talairach J, Tournoux P. Col-Planar Stereotax Atlas of the Human Brain. New York: Thieme Medical Publishers; 1988

[126] Collins DL, Neelin P, Peters TM, Evans AC. Automatic 3D intersubject registration of MR volumetric data in standardized Talairach space. Journal of Computer Assisted Tomography. 1994;**18**(2):192-205

[127] Wade AR, Brewer AA, Rieger JW, Wandell BA. Functional measurements of human ventral occipital cortex: Retinotopy and colour. Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences. 2002;**357**(1424):963-973. DOI: 10.1098/rstb.2002.1108

[128] Tyler CW, Wade AR. Extended concepts of occipital retinotopy. Current Medical Imaging Review. 2005;**1**:3190329

[129] Langner G, Dinse HR, Godde B. A map of periodicity orthogonal to frequency representation in the cat auditory cortex. Frontiers in Integrative Neuroscience. 2009;**3**:27. DOI: 10.3389/neuro.07.027.2009

[130] Langner G, Albert M, Briede T. Temporal and spatial coding of periodicity information in the inferior colliculus of awake chinchilla

(Chinchilla laniger). Hearing Research. 2002;**168**(1-2):110-130

[131] Arcaro MJ, McMains SA, Singer BD, Kastner S. Retinotopic organization of human ventral visual cortex. The Journal of Neuroscience. 2009;**29**(34):10638-10652. DOI: 10.1523/JNEUROSCI.2807-09.2009

[132] Tanji K, Leopold DA, Ye FQ, Zhu C, Malloy M, Saunders RC, et al. Effect of sound intensity on tonotopic fMRI maps in the unanesthetized monkey. NeuroImage. 2010;**49**(1):150-157. DOI: 10.1016/j.neuroimage.2009.07.029

[133] Hart HC, Hall DA, Palmer AR. The sound-level-dependent growth in the extent of fMRI activation in Heschl's gyrus is different for low- and high-frequency tones. Hearing Research. 2003;**179**(1-2):104-112

[134] Menon RS, Kim SG. Spatial and temporal limits in cognitive neuroimaging with fMRI. Trends in Cognitive Sciences. 1999;**3**(6):207-216

[135] Leonard CM, Puranik C, Kuldau JM, Lombardino LJ. Normal variation in the frequency and location of human auditory cortex landmarks. Heschl's gyrus: Where is it? Cerebral Cortex. 1998;**8**(5):397-406

[136] Amunts K, Schleicher A, Burgel U, Mohlberg H, Uylings HB, Zilles K. Broca's region revisited: Cytoarchitecture and intersubject variability. The Journal of Comparative Neurology. 1999;**412**(2):319-341

[137] Dumoulin SO, Wandell BA. Population receptive field estimates in human visual cortex. NeuroImage. 2008;**39**(2):647-660. DOI: 10.1016/j.neuroimage.2007.09.034

[138] de la Mothe LA, Blumell S, Kajikawa Y, Hackett TA. Cortical connections of the auditory cortex in marmoset monkeys: Core and medial belt regions. The Journal of Comparative Neurology. 2006;**496**(1):27-71. DOI: 10.1002/cne.20923

[139] Petkov CI, Kayser C, Augath M, Logothetis NK. Functional imaging reveals numerous fields in the monkey auditory cortex. PLoS Biology. 2006;**4**(7):e215. DOI: 10.1371/journal.pbio.0040215

[140] Kaas JH, Hackett TA. Subdivisions of auditory cortex and levels of processing in primates. Audiology & Neuro-Otology. 1998;**3**(2-3):73-85

[141] Pandya DN, Sanides F. Architectonic parcellation of the temporal operculum in rhesus monkey and its projection pattern. Zeitschrift für Anatomie und Entwicklungsgeschichte. 1973;**139**(2):127-161

[142] Galaburda AM, Pandya DN. The intrinsic architectonic and connectional organization of the superior temporal region of the rhesus monkey. The Journal of Comparative Neurology. 1983;**221**(2):169-184. DOI: 10.1002/cne.902210206

[143] Morel A, Garraghty PE, Kaas JH. Tonotopic organization, architectonic fields, and connections of auditory cortex in macaque monkeys. The Journal of Comparative Neurology. 1993;**335**(3):437-459. DOI: 10.1002/cne.903350312

[144] Fullerton BC, Pandya DN. Architectonic analysis of the auditory-related areas of the superior temporal region in human brain. The Journal of Comparative Neurology. 2007;**504**(5):470-498. DOI: 10.1002/cne.21432

[145] Muckli L, Naumer MJ, Singer W. Bilateral visual field maps in a patient with only one hemisphere. Proceedings of the National Academy of Sciences of the United States of America. 2009;**106**(31):13034-13039. DOI: 10.1073/pnas.0809688106

[146] Brewer AA, Barton B. Developmental plasticity: FMRI investigations into human visual cortex. In: Papageorgiou TD, Christopoulos G, Smirnakis S, editors. Advanced Brain Neuroimaging Topics in Health and Disease - Methods and Applications. Rijeka, Croatia: InTech; 2014. pp. 305-334. DOI: 10.5772/58256

[147] Brewer AA. Visual maps: To merge or not to merge. Current Biology. 2009;**19**(20):R945-R947. DOI: 10.1016/j.cub.2009.09.016

[148] Katz LC, Shatz CJ. Synaptic activity and the construction of cortical circuits. Science. 1996;**274**(5290):1133-1138

[149] Polleux F, Ince-Dunn G, Ghosh A. Transcriptional regulation of vertebrate axon guidance and synapse formation. Nature Reviews. Neuroscience. 2007;**8**(5):331-340. DOI: 10.1038/nrn2118

[150] Merzenich MM, Brugge JF. Representation of the cochlear partition of the superior temporal plane of the macaque monkey. Brain Research. 1973;**50**(2):275-296

[151] Sweet RA, Dorph-Petersen KA, Lewis DA. Mapping auditory core, lateral belt, and parabelt cortices in the human superior temporal gyrus. The Journal of Comparative Neurology. 2005;**491**(3):270-289. DOI: 10.1002/cne.20702

[152] Dick F, Tierney AT, Lutti A, Josephs O, Sereno MI, Weiskopf N. In vivo functional and myeloarchitectonic mapping of human primary auditory areas. The Journal of Neuroscience. 2012;**32**(46):16095-16105

[153] Rivier F, Clarke S. Cytochrome oxidase, acetylcholinesterase, and NADPH-diaphorase staining in human supratemporal and insular cortex: Evidence for multiple auditory areas. NeuroImage. 1997;**6**(4):288-304. DOI: 10.1006/nimg.1997.0304

[154] Hackett TA. Information flow in the auditory cortical network. Hearing Research. 2011;**271**(1-2):133-146. DOI: 10.1016/j.heares.2010.01.011

[155] Hackett TA, Stepniewska I, Kaas JH. Subdivisions of auditory cortex and ipsilateral cortical connections of the parabelt auditory cortex in macaque monkeys. The Journal of Comparative Neurology. 1998;**394**(4):475-495

[156] Jones EG, Dell'Anna ME, Molinari M, Rausell E, Hashikawa T. Subdivisions of macaque monkey auditory cortex revealed by calcium-binding protein immunoreactivity. The Journal of Comparative Neurology. 1995;**362**(2):153-170. DOI: 10.1002/cne.903620202

[157] Molinari M, Dell'Anna ME, Rausell E, Leggio MG, Hashikawa T, Jones EG. Auditory thalamocortical pathways defined in monkeys by calcium-binding protein immunoreactivity. The Journal of Comparative Neurology. 1995;**362**(2):171-194. DOI: 10.1002/cne.903620203

[158] Kusmierek P, Rauschecker JP. Functional specialization of medial auditory belt cortex in the alert rhesus monkey. Journal of Neurophysiology. 2009;**102**(3):1606-1622. DOI: 10.1152/jn.00167.2009

[159] Rauschecker JP, Tian B, Hauser M. Processing of complex sounds in the macaque nonprimary auditory cortex. Science. 1995;**268**(5207):111-114

[160] Tian B, Rauschecker JP. Processing of frequency-modulated sounds in the lateral auditory belt cortex of the rhesus monkey. Journal of Neurophysiology. 2004;**92**(5):2993-3013. DOI: 10.1152/jn.00472.2003

[161] Kajikawa Y, Frey S, Ross D, Falchier A, Hackett TA, Schroeder CE. Auditory properties in the parabelt regions of the superior temporal gyrus in the awake macaque monkey:

An initial survey. The Journal of Neuroscience. 2015;**35**(10):4140-4150. DOI: 10.1523/JNEUROSCI.3556-14.2015

[162] von Economo C, Koskinas G. Die Cytoarchitectonik der Hirnrinde des erwachsenen menschen. Berlin: Julius-Springer; 1925

[163] Hedges S, Kumar S. Genomic clocks and evolutionary timescales. Trends in Genetics. 2003;**19**:200-206

[164] Kayser C, Petkov CI, Augath M, Logothetis NK. Functional imaging reveals visual modulation of specific fields in auditory cortex. The Journal of Neuroscience. 2007;**27**(8):1824-1835. DOI: 10.1523/JNEUROSCI.4737-06.2007

[165] Bartels A, Zeki S. The architecture of the colour Centre in the human visual brain: New results and a review. The European Journal of Neuroscience. 2000;**12**(1):172-193

[166] Moradi F, Heeger DJ. Inter-ocular contrast normalization in human visual cortex. Journal of Vision. 2009;**9**(3):13.1-13.22. DOI: 10.1167/9.3.13

[167] Mitchison G. Neuronal branching patterns and the economy of cortical wiring. Proceedings of the Biological Sciences. 1991;**245**(1313):151-158. DOI: 10.1098/rspb.1991.0102

[168] Chklovskii DB, Koulakov AA. Maps in the brain: What can we learn from them? Annual Review of Neuroscience. 2004;**27**:369-392. DOI: 10.1146/annurev.neuro.27.070203.144226

[169] Shapley R, Hawken M, Xing D. The dynamics of visual responses in the primary visual cortex. Progress in Brain Research. 2007;**165**:21-32. DOI: 10.1016/S0079-6123(06)65003-6

[170] Braitenberg V, Schüz A. Cortex: Statistics and Geometry of Neuronal Connectivity. 2nd ed. Berlin: Springer; 1998

[171] Baseler HA, Brewer AA, Sharpe LT, Morland AB, Jagle H, Wandell BA. Reorganization of human cortical maps caused by inherited photoreceptor abnormalities. Nature Neuroscience. 2002;**5**(4):364-370. DOI: 10.1038/nn817

[172] Hoffmann MB, Tolhurst DJ, Moore AT, Morland AB. Organization of the visual cortex in human albinism. The Journal of Neuroscience. 2003;**23**(26):8921-8930

[173] Hoffmann MB, Kaule FR, Levin N, Masuda Y, Kumar A, Gottlob I, et al. Plasticity and stability of the visual system in human achiasma. Neuron. 2012;**75**(3):393-401. DOI: 10.1016/j.neuron.2012.05.026

[174] Baseler HA, Gouws A, Haak KV, Racey C, Crossland MD, Tufail A, et al. Large-scale remapping of visual cortex is absent in adult humans with macular degeneration. Nature Neuroscience. 2011;**14**(5):649-655. DOI: 10.1038/nn.2793

[175] Fine I, Wade AR, Brewer AA, May MG, Goodman DF, Boynton GM, et al. Long-term deprivation affects visual perception and cortex. Nature Neuroscience. 2003;**6**(9):915-916. DOI: 10.1038/nn1102

[176] Brewer AA, Barton B. Visual cortex in aging and Alzheimer's disease: Changes in visual field maps and population receptive fields. Frontiers in Psychology. 2014;**5**:74. DOI: 10.3389/fpsyg.2014.00074

[177] Brewer AA, Barton B. Changes in visual cortex in healthy aging and dementia. In: Moretti DV, editor. Update on Dementia. Rijeka, Croatia: InTech; 2016. pp. 273-310. DOI: 10.5772/61983

Chapter 2

The Influence on Cortical Brainwaves in Relation to Word Intelligibility and ASW in Room

Chiung Yao Chen

Abstract

The influence of indoor speech intelligibility and apparent source width (ASW) on the response of cortical brainwaves was studied using two variables, the time gap between direct and the first reflection (Δt_1, ms) and the initial (<80 ms) interaural cross-correlation function (IACC_{E3}). Comparisons were performed based on autocorrelation function (ACF) of continuous brainwave (CBW) and slow vertex response (SVR). The results are: (1) the effective delay time of ACF (τ_e) of β-waves (13–30 Hz) in the left hemisphere under changes in Δt_1 was significantly and positively correlated with speech intelligibility ($p < 0.001$). (2) As ASW increased, the relative amplitude of left hemisphere A (P2-N2) tended to decrease ($p < 0.05$) in SVRs, while N2 latency tended to increase ($p < 0.05$); the lateral lemniscus in the auditory nerve was suggested to be the reactive site. (3) With regard to hemispheric specialization in brain, speech intelligibility, the main temporal factor, was found to be controlled by the left hemisphere. A subjective spatial factor, ASW, the relative amplitude of SVR was also found to decrease in the left hemisphere; nevertheless, they are coherent while the N2 latency of SVR significantly prolonged in both left and right hemisphere under changes in IACC_{E3}.

Keywords: brainwaves, ACF, apparent sound width, speech intelligibility, subjective diffuseness, hemispheric specialization

1. Introduction

In human speech cognition, speech intelligibility integrates short-term memory and cerebral feedback [1]. However, important factors constituting the spatial impressions of sound also include certain related evaluation indicators, such as the listener's judgment of sound source direction (sense of direction) and distance (sense of proximity), apparent source width (ASW), and lateral envelopment (LEV). As suggested by Ando [2] and Beranek [3], the composition of such spatial impressions mainly depends on fluctuations of the magnitude of the interaural cross-correlation (IACC) and is especially affected by the degree of subjective diffusion of the sound field. However, listeners differ in their needs and perceptions regarding subjective diffusion and ASW.

With regard to neuron-psychology, Sperry [4] discovered the phenomenon of hemispheric disconnect. The cerebral specialization theory distinguishes between "speech functions" and "non-speech functions." Certain symbols in architectural design belong to non-speech functions. For instance, the range of non-speech functions includes aesthetic perception and the feeling of balance. In particular, many non-speech symbols can be observed in environmental design. Earlier research on audio and cerebral correlations found that such common medical problems as aphasia and disturbances in tone judgment originate in the left cerebral hemisphere. Therefore, this study suggested that cerebral responses to speech and non-speech symbol in the physical environment effectively substitute for the semantic differences (SD) caused by age-related and cultural differences. Cerebral responses to communication stimuli are a direct cross-cultural and cross-age reference indicator, which is similar to the principle behind polygraph tests performed by police to examine physiological responses.

This study suggested that cerebral responses can be used to clearly and consistently examine responses to change in "speech functions" of the physical environment, or speech intelligibility, when designing a sound field. Ando [2] considered "speech functions" to be an important temporal factor and the result of autocorrelation function (ACF) evaluations in the brain. Therefore, the environmental effects of temporal factors were examined in this study based on the influence of speech intelligibility on the correlation between "subjective perceptions" and cerebral responses, which served as the basis for the objective design of an acoustic environment. Akita et al. [5] indicated that when the sensory information received by listeners is analyzed by brainwaves, this does not represent their direct experience of changes in the environment, but rather the interaction between physiology and the environment. This phenomenon is common in daily life. The intensity of cerebral evoked responses is the optimal evaluation tool [6]. Soeta et al. [7] studied the effects of sound source features on subjective psychological responses and cerebral responses measured by magnetoencephalography (MEG) and reported that at different delay times of reflection sounds (Δt_1 = 0, 5, 20, 60, and 100 ms) and 50 alternations, the ACF effective delay time of α-waves recorded by MEG indicated subjective preferences regarding sound fields. The methods used in this study can be summarized as follows:

1. The first reflection delay (Δt_1) was changed to change speech intelligibility. The degrees (or process) of subjective recognition of Chinese monosyllables were determined by comparing ACF calculation results related to α-waves and β-waves among cerebral continuous brainwaves (CBW).

2. The IACC_{E3} was changed to change subjective ASW. Changes in the waveforms of auditory evoked potentials (AEPs) during listeners' perceptions of spatial ASW were analyzed.

2. Empirical methods

2.1 Psychological test of intelligibility

This study used monosyllabic speech sound articulation and IACC_{E3} to quantify changes in two subjective experiences, namely, speech intelligibility and ASW. With regard to speech intelligibility, the fifth group of common Chinese monosyllabic speech sounds used in Taiwan [8] (female voice, **Table 1**) was used. Test results related to this group of monosyllabic sounds are characterized by the largest disparity in error rates because most related sounds belong to "fricative sounds" (i.e., apical vowels, such as "zh," "ch," "sh," "r," "z," "ci," and "si" in Bopomofo system). The amounts of fricative and non- fricative rhymes are balance (eight versus ten, respectively).

Chinese monosyllable					
1	shy0	7	yu2	13	ching3
2	iur1	8	leau3	14	tzuen1
3	li0	9	shoou3	15	cha2
4	meei3	10	ian1	16	shuo4
5	tsae3	11	tsong1	17	he4
6	ru2	12	guang1	18	chye2

Note: The pronunciation of each syllable depends on the tone (one of five pitch contours) used, which is indicated by a number attached to the end of the syllable. For example, 0 denotes monosyllables pronounced with a soft puff of air.

Table 1.
List of the term of monosyllables [8].

The sound structure of Mandarin differs from that of other languages. In Mandarin, each character is pronounced as a monosyllable with one of five tones (i.e., types of pitch contour). Each of these tones (0–4), when used with a given monosyllable, causes the monosyllable to convey a meaning distinct from those conveyed when the monosyllable is used with the other four tones. Utterance lengths in the experiment were set to 400–500 ms. Monosyllabic presents were separated by 2.5 s. The experiment was arranged according to the arrangement used in the study by Chen et al. [9].

The experiment was conducted in front of two overlapping loudspeakers in a semi-anechoic room (4 × 3 and 4 m in height) at Chaoyang University of Technology. The loudspeakers (Fostex NF-1A) were located at 1.5 m right front of the center of a listener's head. The first reflected sound was given off by the upper loudspeaker ($\eta = 15$) while another gave off the direct sound ($\eta = 0$). To vary speech intelligibility, the speech signal was assumed that emitted from the stage with a direct and a reflection sound reflected through the ceiling of the stage. The listening level was adjusted to a usual communicative sound volume of 62 dB(A) at the center of the room. The level of background noise in the semi-anechoic room was 32–42 dB(A), then the S/N ratio are approximate to 30–40 dB. The setup of the instrumental diagram (EEG recordings) could be referred to **Figure 1**, since they were same as that in the spatial ASW experiment stated below. The settings of the physical parameters used in the experiment are shown in **Table 2**. **Figure 2** shows the experimental results that indicate 62 listeners who were significantly able to distinguish sounds using percentage syllable articulation (PSA) tests [10]. To determine PSA, those written syllables are compared with the original syllables to find the percentage of syllables written correctly.

2.2 Psychological quantification test of ASW

The paired-comparison method [11] was used in the psychological quantification test of subjective ASW. The experiment was conducted in the same venue as the first experiment. Three loudspeakers (one for direct sounds and two for reflected sounds) were located at 1.5 m from the center of a listener's head; the incidence combinations ($\boldsymbol{\xi}, \boldsymbol{\eta}$) are: (0°, ±15°), (0°, ±55°), (0°, ±90°) and (0°, +15°, −55°) on the horizontal plane. 2 kHz pure-tone (1 ms) sounds were produced. The $\mathrm{IACC_{E3}}$ (0.35, 0.57, 0.68, and 0.81) [12, 13] of the sound field was changed by changing the angle of incidence stated above and the sound pressure level. As a result, different subjective ASWs were generated (**Table 3**). The instrumental setup of testing spatial ASW and the process of AEPs recordings are interpreted in **Figure 1**.
The participants (80 students) determined ASWs using paired comparisons. The interval between sound prompts within one group was 2 s and the interval between groups was 10 s; in total, six groups were used. The participants were asked to

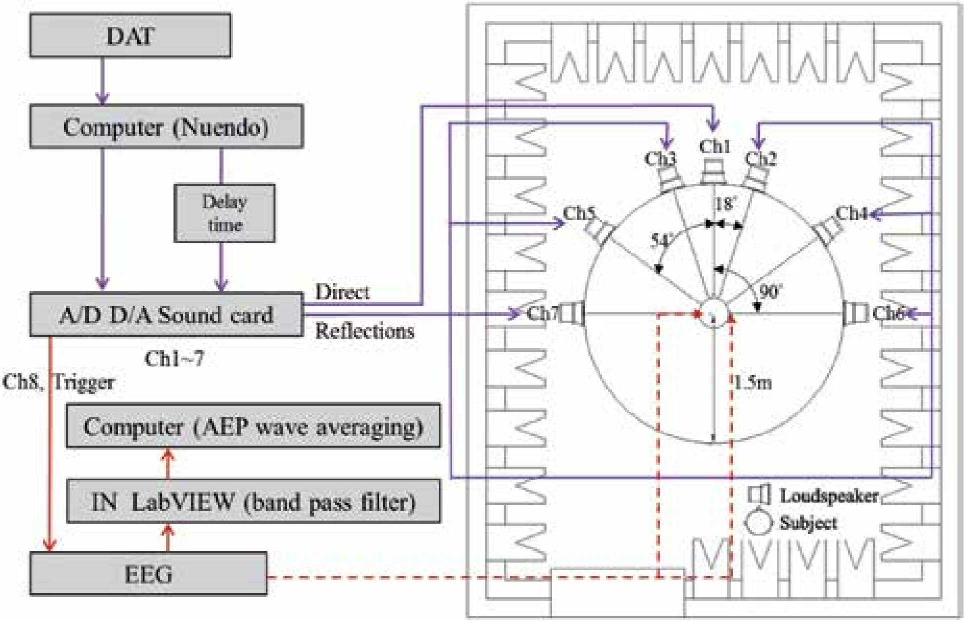

Figure 1.
The setup of the instrumental diagram (audio arrangement and EEG recordings).

Item	Conditions of experiments
Δt_1 (ms)	Delay gap: 0 ms, 35 ms, 100 ms, 150 ms, 200 ms
SPL of individual loudspeakers	Direct sound: 60 dB(A); first reflection, Δt_1: 55 dB(A)
Reverberation times	$RT \doteqdot 0.1$ s

Table 2.
The setting of the physical parameters in subjective articulation test of monosyllables.

immediately determine and record the relative probability of ASWs. Each questionnaire was conducted for 1 min. The psychological scale values of ASWs are shown in **Figure 3** calculated using Thurstone's Case V [11]. Non-linear correlation was observed in the IACC_{E3} result [14].

2.3 Brainwave physiological experiment methods

2.3.1 Brainwave analysis method

After the fast Fourier transform (FFT) was applied to the brainwaves, ACF of CBW calculations were performed for the α-waves (8–13 Hz) and β-waves (13–30 Hz) of the left and right hemispheres. In the earlier study by Chen and Ando [15], 100 Hz α-waves and 500 Hz β-waves were sampled according to the sampling frequency laws and, after A/D conversion (16 bits), input into a computer to calculate the effective duration (τ_e) of CBWs' ACF (**Figure 4**). In ACF calculations of τ_e values in the study by Chen and Chan [16], the 0.3 s integration time (2 T) of monosyllabic speech sounds was suggested to be the most effective. Eventually, the monosyllabic signals were played in this study included simulation of the first delay time [17]. Therefore, the integration time (2 T) of ACF of continuous brainwaves (CBW) used in calculation was adjusted to 0.5 s. As shown in **Figure 4**, substantial differences were observed in the ACF waveforms of α-waves and β-waves under the same first delay time settings.

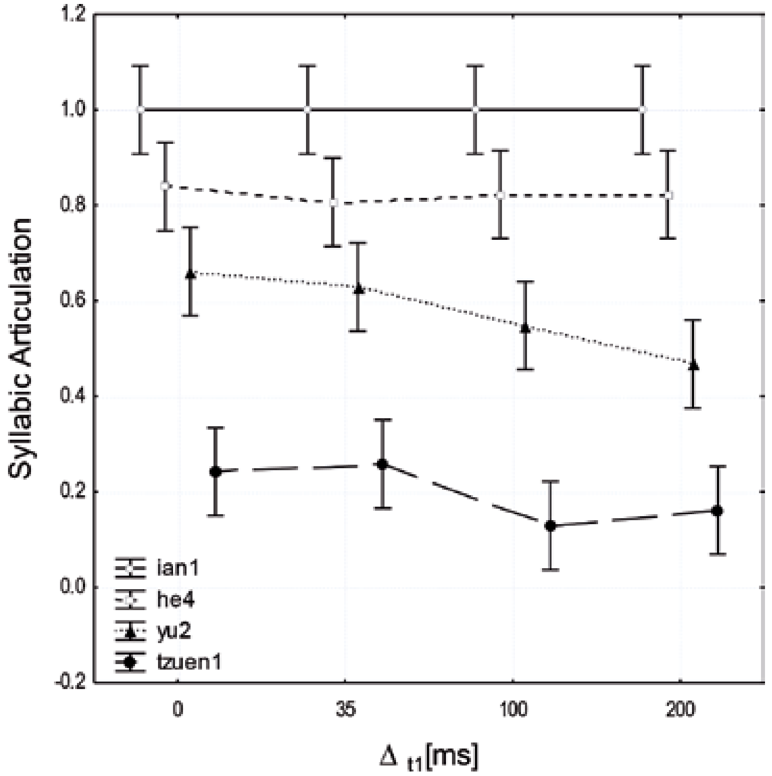

Figure 2.
The percentage syllabic articulation of monosyllable functioning initial time delay of a sound field.

IACC$_{E3}$ (setup values)	Amplitude of direct sound (A$_0$)	I-1, SPL/dB(A)	Amplitude of first reflection (A$_1$)
0.35	1	62.6	0.8
0.57	1	62.6	0.8
0.68	1	55.4	0.4
0.81	1	64.0	0.2
I-2, SPL/dB(A)	**Amplitude of second reflection (A2)**	**I-3, SPL/dB(A)**	**Σ*L* Total SPL dB(A)**
59.4	0.8	59.4	65
59.4	0.8	59.4	65
53.4	0.4	55.4	65
53.4	0.2	53.4	65

Note: I-1, I-2 and I-3 denote the sound intensity of direct sound and 1st and 2nd reflections sound measured at the location of the head top of the participants. a denotes the amplitude of the direct sound, 1st reflective and 2nd reflective sound by A_0, A_1 and A_2.

Table 3.
The parameters of subjective source apparent width (ASW) test arranged by 2 kHz pure tone burst.

To explore the changes in subjective perceptions of ASW, AEPs of nine participants were induced, recorded and analyzed as in the psychological intelligibility experiment. However, a spatial impression of a sound signal is a short-term memory phenomenon. Therefore, waveforms induced by the brain AEPs are normally used to observe changes in responses to weak brainwave signals (about 10–100 μV in

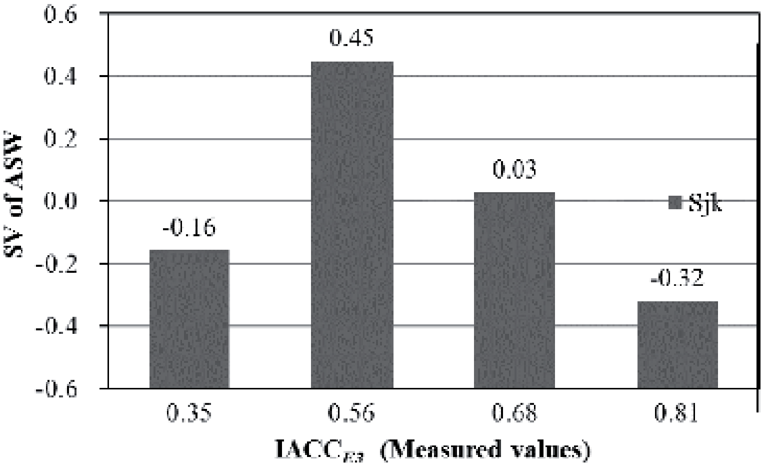

Figure 3.
The scale values of subjective ASW test functioning IACC$_{E3}$.

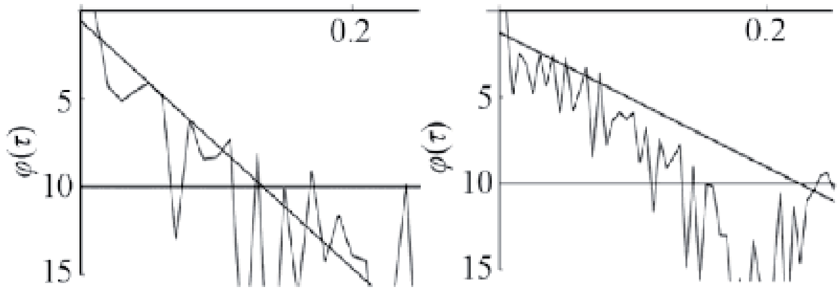

Figure 4.
The ACF curve of α-wave (left) and β-wave (right) recorded in relation to the monosyllable "tzuen1" were announcing.

amplitude when measured from the scalp). Clear consistent brain waveforms are usually obtained by applying the signal averaging method [18] to responses that occur within 500 ms after auditory stimulation (**Figure 5**). In this study, 180 times of averaging process was applied here since the wave form of slow vertex responses (SVR) were clearly obtained. The movements (latency) of waveform peaks and troughs in the wave relative amplitude can reflect the activation of different parts of auditory nerves [19, 20]. As shown in **Figure 5**, this study changed ASW perceptions by changing the sound arrival orientation and energy while fixed reflection delay and echo times ($\Delta t_1 < 15.5$ ms, $RT \doteqdot 0.1$ s) (**Table 3**). This study suggested that brain waveforms could be observed using SVR because the preset reaction time exceeded 10 ms. In such way, AEPs were obtained. Consequently, when potentials P1, P2, N1, and N2 of posterior waveforms (30–200 ms) among different AEPs were generated by different sound stimulus, the relative amplitudes of P1-N1 and P2-N2, and P1, P2, N1, and N2 latency were observed [18].

2.3.2 Brainwave recording method

With regard to brainwave recording, eight participants (and other nine in the AEPs experiments) sat on comfortable office chairs in the semi-anechoic room at Chaoyang University of Technology and their brainwaves were induced and

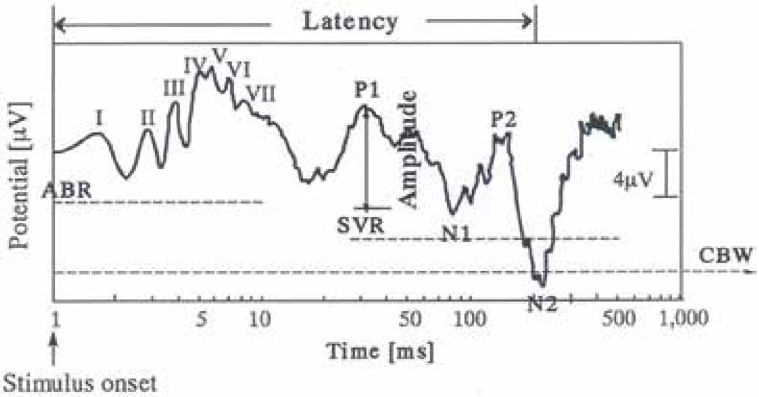

Figure 5.
The diagram illustrates brain waves in different time domains and their index at the peak or trough by Ichikawa [17].

recorded. The room temperature was maintained at 22 ± 2°C. All subjects were prohibited from drinking any alcohol for a period of 3 days before the brainwave recordings were conducted, and they refrained from smoking for 1 h before both experiments. They were instructed to concentrate on listening to the signals during the presentation. The participating subjects were eight male students (plus another nine male students in the AEPs experiments) aged 22–24 years old with normal hearing ability, as confirmed by an audiometry test and right-handed test (self-administered). The audiometry test detects sensorineural hearing loss (damage to the nerve or cochlea) and conductive hearing loss (damage to the eardrum or the tiny ossicle bones). Pure-tone subjective audiometry, in which air conduction hearing thresholds in decibels (dB) for a frequency range of 250–8000 Hz are plotted on an audiogram for each ear independently, was applied. All of the subjects had to be qualified as normal with a pure-tone audiogram (less than 25 dB) for both ears prior to the brainwave experiments and questionnaires.

This procedure has been applied in many studies, such as those by Chen et al. [9], Ando et al. [19], and Ando et al. [20], among others.

Electrodes used to explore brainwaves were positioned at the participants' T3 and T4 head points according to the international 10–20 system [21]. Electric potentials were examined using eardrops on the left and right sides. Unipolar induction of continuous brainwaves in the left and right hemispheres was performed. The G2 electrode was attached between the eyebrows for eye movement reference. The electrode system was grounded each time the brainwaves were recorded in order to avoid external electric interference. The settings of the simulated sound field were similar to that in the aforementioned psychological experiment [16]. The collected brainwave data was analyzed and processed by NI LabVIEW software. The setup of the instrumental diagram is shown in **Figure 1**. During the brainwave experiments, the subjects had to be relaxed while paying close attention to the sound stimuli. Brainwaves are extremely sensitive to any incoming stimuli or stress. For the purpose of this study, a relaxed state but one also focused on environmental variations was considered the best condition for the subjects during the brainwave recording process. For the recordings, periods of blinking had to be disregarded. Thus, a monitor was set up in the anechoic chamber to identify these periods, and these sections were later removed from the recordings.

3. Empirical results and discussion

3.1 Influence of speech intelligibility on brainwaves

Monosyllabic speech sounds had a major effect on both α-waves ($F = 12.96$ (9, 2488), $p < 0.001$) and β-waves ($F(9, 2488) = 5.21$, $p < 0.001$) at different first reflection delay times. As shown in **Figures 6** and **7**, the ACF of β-waves recorded in the left hemisphere was positively correlated with subjective perceptions of speech intelligibility. With regard to α-waves in the left hemisphere, brainwave

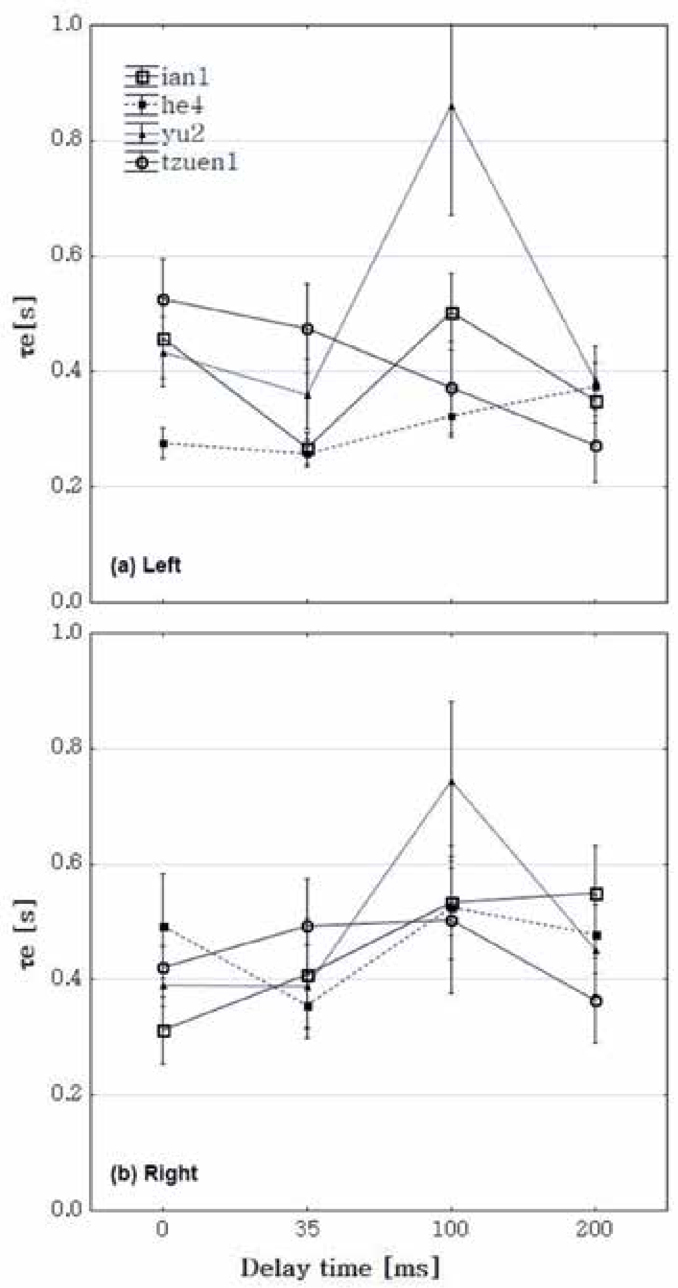

Figure 6.
Relationship between τ_e of ACF, α-wave and Δt_1 of sound field.

responses tended to increase at 100 ms for all sounds apart from "tzuen1." It is not clear whether these results were related to the nasal sounds "uen." The psychological experiment results (**Figure 2**) showed that the lowest articulation rates were observed for four sounds at 100 ms reflection delays. An opposite tendency was

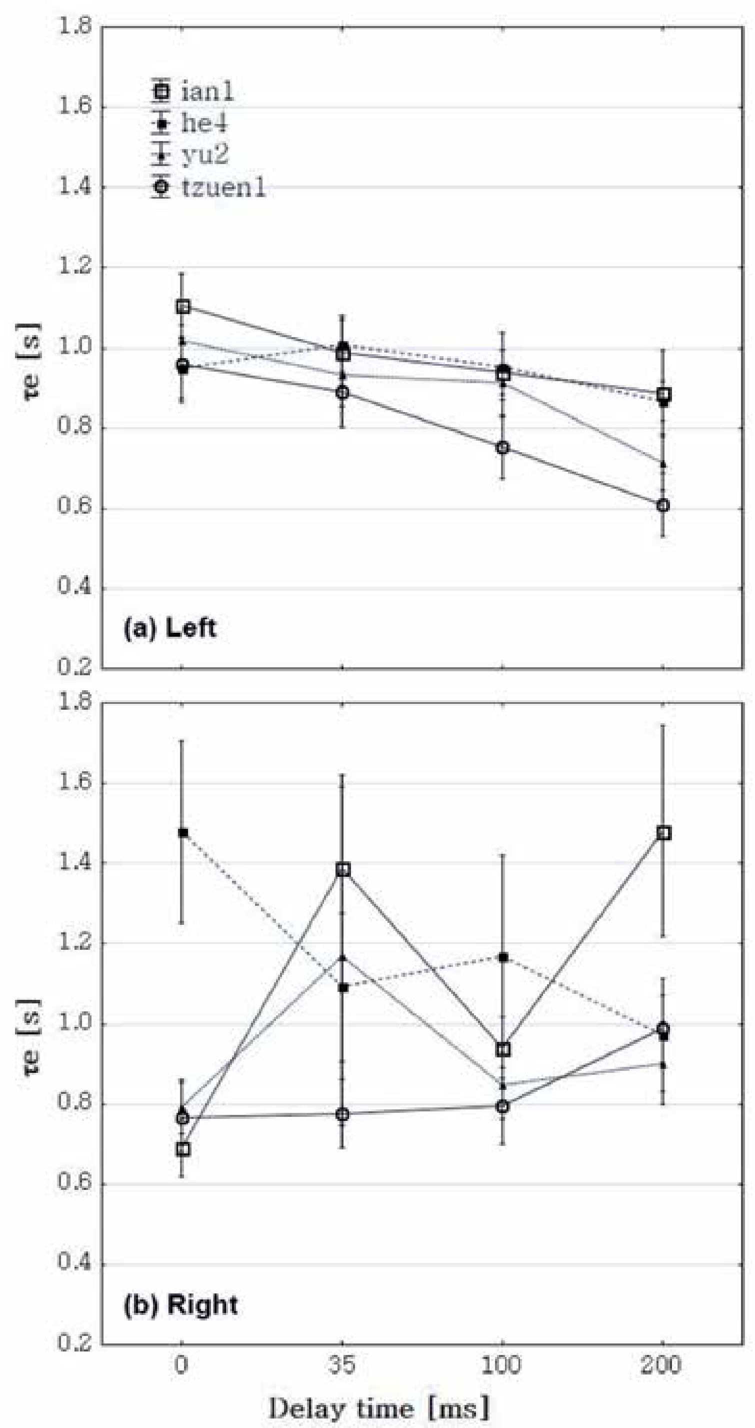

Figure 7.
Relationship between τ_e of ACF, β-wave and Δt_1 of sound field.

detected in the articulation rate results related to "yu2," "he4," and "ian1." The 100 ms delay is close to the 135 ms slow response delay proposed in Ando's [22] study on sound field preferences (echo disturbance). The displeased response of α-waves to the delay time of reflection [14] requires further investigation.

Figure 7 shows changes in β-waves. Consistent results were obtained with regard to the influence of the delay time of reflection on the left hemisphere ($F(9, 2488) = 5.21, p < 0.001$). However, no significant differences were observed in psychological reactions to speech intelligibility. A significant relation between articulation rates and the order of reactions was detected in the mean values related to the right hemisphere. Thus, left hemisphere showed changes in β-waves in relation to the order of delay time on reflection but speech intelligibility reactions.

3.2 Changes in subjective perception of ASW and brainwaves

The findings related to SVR to evoked potentials of nine participants are shown in **Figure 8**. With regard to the left hemisphere, SVR relative amplitude were consistently and inversely related to quantified psychological scale values ($F(1, 16) = 4.90$, $p < 0.05$). However, clear results were difficult to obtain due to the small difference between ASW (−0.16) and ASW (0.03).

Latency changes in the left and right hemispheres indicated the presence of a significant difference between ASW (0.03) and ASW (0.45) only at N2 in the left hemisphere ($F(1, 16) = 11.09, p < 0.05$). The tendency of ASW (0.45) latency being smaller than ASW (0.03) latency in the left hemisphere can be seen from **Figure 9**, whereas in the right hemisphere, ASW (0.45) latency was consistently

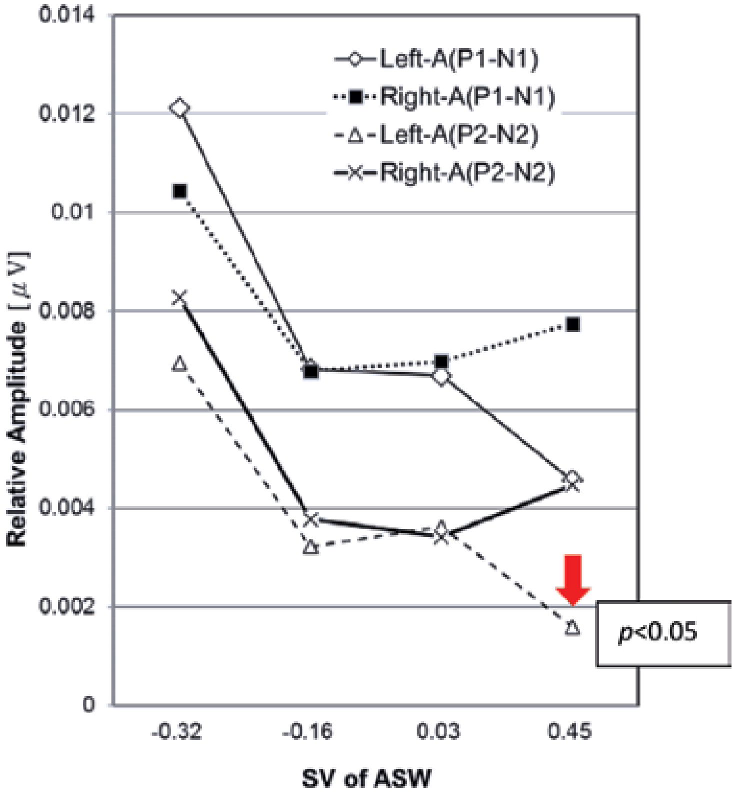

Figure 8.
Relationship between potential, SVR and ASW of sound field.

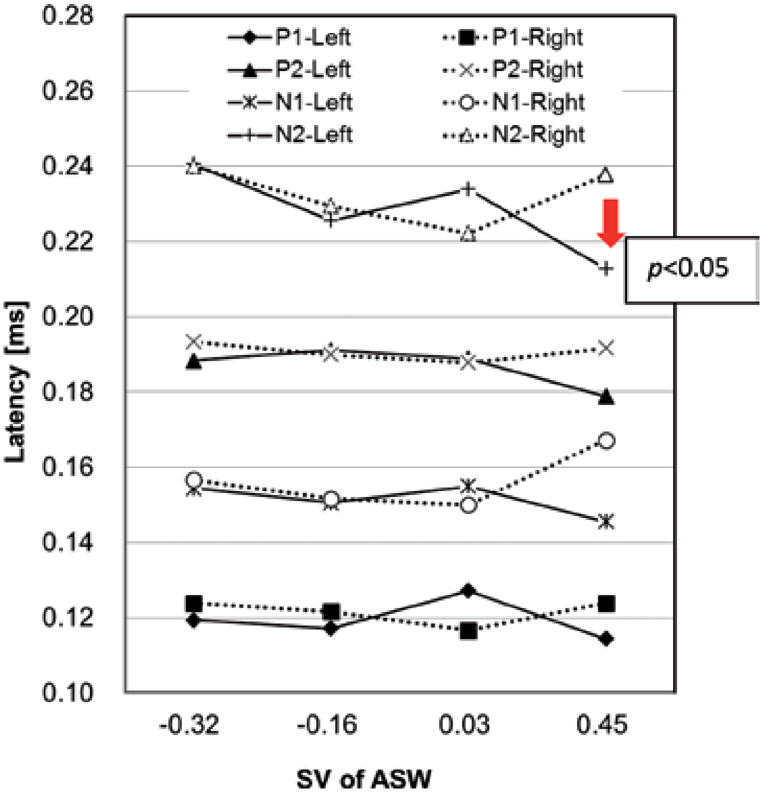

Figure 9.
Relationship between latency, SVR and ASW of sound field.

larger than ASW (0.03) latency. The results showed that relative amplitude in the left hemisphere were caused by subjective perceptions of ASW, which influenced the participants' preference toward a sound field. The consistency of latency at N2 was due to the activation of neural sites, which was clearly observed between ASW (0.03) and ASW (0.45), as well as at $IACC_{E3}$ of 0.56–0.68. Thus, the brain did not have a major effect on the corresponding changes at the extreme $IACC_{E3}$ values of 0.35 and 0.81, which corresponded to the psychological reaction results presented in **Figure 3**.

4. Conclusions

The arrangements and results of the aforementioned brainwave experiments indicated that when simple physical changes in a sound field and complex psychological feedbacks affect cerebral brainwave reactions, the correspondence of the cerebral specialization theory with the results becomes very complicated. In general, in this study, the left hemisphere tended to be activated in both temporal and spatial aspects based on the sound field. When the participants' brainwaves were recorded during the judgment task, the brain activation in the right hemisphere tended to reflect the discriminated object more closely. When CBW were observed during research on speech intelligibility, the left hemisphere showed clear reactions to the first reflection delay time of sound field (**Figure 7**). However, the degree of speech intelligibility is a reflection of the complex thinking process that occurs in the right hemisphere (cerebral feedback). This phenomenon was supported by the subjective ASW experiment. With regard to changes in spatial factors, the left hemisphere received information about sound field changes when the $IACC_{E3}$ value changed. ASW changes between (ASW (0.03) and ASW (0.45)), which were more evident in the right hemisphere, affected both right and left hemispheres. They are coherent while the N2 latency of SVR significantly prolonged in both left and right hemispheres under changes of subjective diffuseness in $IACC_{E3}$ found by Ando et al. [20]. Different sites are activated by brainwaves during focused and ambient use of the brain.

Cerebral specialization has been reported to be determined by focused conscious decisions. For instance, Floel et al. [23] conducted a spatial—visual focus experiment and used a Doppler ultrasound system and magnetic resonance imaging (MRI) equipment to observe the brain reactions of right-handed participants; the researchers found that both spatial recognition and speech functions were activated in the right hemisphere, which corresponded to clinical experiment results.

Nevertheless, for CBW researches, we conclude that α-waves (8–12 Hz) mainly responds to the emotional reactions; β-waves (13–30 Hz) reacts to the auditory matter drift (**Figures 6** and **7**). But the left hemisphere leads focus or attention on the varying of situational conditions (**Figures 8** and **9**), and the right one blends with imaginable feeling and experience. Hemispheric specialization has to pay attention to the conditioned response, conscientious and careful detail to setup each brainwaves' experiment.

Acknowledgements

I would like to express my thanks to the graduated students, Yong-Shang Chen and Qi-Wen Lin who have, in brainwaves' experiments, helped me in the course of preparing this study. In particular, I wish to thank Professor Em. Yoichi Ando, who kindly gave me directions on my brain research and data analysis methodology. I am also indebted to the Ministry of Science and Technology Taiwan, for their 2 years period (2007 and 2012) of financial support to complete this research. Meanwhile, the application conformed to academic ethics for the conduct of research. Moreover, during the brainwave experiments, the subjects were assured of their safety and told the procedure was non-invasive. Special thanks are due to my many colleagues for their participation in the experiments involving the subjective judgments and the brainwaves' recordings.

Glossary of symbols

ASW	apparent source width, a sound perception of the subjective diffuseness occurred from beginning to 80 ms of stimulus
Δt_1	delay gap between direct and first reflection in a defuse sound field
$IACC_{E3}$	binaural initial (<80 ms) interaural cross-correlation function
ACF	autocorrelation function
τ_e	effective delay of autocorrelation function (ACF)
CBW	continuous brainwave, a term to distinguish from an evoked potential (EP) or evoked response within EEG
SVR	slow vertex response, an evoked potential is a direct result after a specific sensory stimulus in the period of 10–500 ms
LEV	listener envelopment, a sound perception of the subjective diffuseness occurred after 80 ms of stimulus
SD	semantic differences, a method of questionnaire employed the scale of responses caused by a psychological affection

/yu2/	example of a monosyllable in Taiwanese's life speech
η	vertical angles at a median plane, 0° started from the front of head at ear height
ξ	angles at clockwise horizontal plane, 0° started from the front of head at ear height
SPL	sound pressure level measured by a sound level meter in a fast time-weighting mode
AEP	auditory evoked potential
PSA	percentage syllable articulation

Author details

Chiung Yao Chen
Department of Architecture, Chaoyang University of Technology, Taiwan

*Address all correspondence to: chychen@cyut.edu.tw

References

[1] Schmidt RA, Wrisberg CA. Motor Learning and Performance: A Problem-Based Learning Approach. 2nd ed. Champaign, IL: Human Kinetics, California - Los Angeles; 2000. pp. 255-287. Chapter 10

[2] Ando Y. Concert Hall Acoustics. Berlin Heidelberg, New York: Springer; 1985. pp. 41-44. DOI: 10.1007/978-3-642-69810-1

[3] Beranek LL. Concert hall acoustics. The Journal of the Acoustical Society of America. 1992;**92**(1):1-39. DOI: 10.1121/1.404283

[4] Sperry RW. Lateral preference in the identification of patterned stimuli. The Journal of the Acoustical Society of America. 1970;**47**:574-578. DOI: 10.1121/1.1911932

[5] Akita T, Fujii T, Hirate K, Yasuoka M. A research on auditory information processing in the brain when a person at his task by eans of measurement and analysis of auditory evoked potential. Journal of Architecture, Planning and Environmental Engineering. 1998;**507**:61-69

[6] Jian YH, Chen JS, Huang SH, et al. Measurements on emotional response using EEG. Journal of Advance Engineering Taiwan. 2006;**33**(1): 123-131 (in Chinese)

[7] Soeta Y, Nakagawa S, Yano T, et al. Magnetoencephalographic responses in relation to the factors extracted from autocorrelation and inter-aural cross-correlation functions. Journal of South China University of Technology (Natural Science Edition). 2007;**35**(s):80-83

[8] Chen CY, Chen LS, Lin W. A study on evaluation method of Chinese articulation standard of speech intelligibility for sound field in Taiwan. Journal of Architecture, Architectural Institute of the Republic of China. 2002;**43**:27-36 (in Chinese)

[9] Chen CY, Soeta H, Ando Y. Analysis of time-varying continuous brainwaves: Cross correlation function of cortical alpha waves. In: 20th Symposium of Taiwan Architecture Institute, Taipei. 1999 (in Chinese)

[10] Benesty J, Sondhi MM, Huang Y. Springer Handbook of Speech Processing. Berlin, Germany: Springer-Verlag; 2007. ISBN: 978-3-540-49125-5. DOI: 10.1007/978-3-540-49127-9_1

[11] Thurstone LL. A law of comparative judgment. Psychological Review. 1927;**34**:273-289

[12] Chen CY, Chang CH. Sound localization in respect of magnitude of inter-aural cross-correlation function: Two reflections in simulated sound field as example. Chaoyang Journal of Design. 2000;**1**:101-114 (in Chinese)

[13] Chen CY, You CR. A comparative study of subjectively apparent source width between traditional Chinese and Western classic musical instruments recital. Chaoyang Journal of Design. 2011;**12**:96-116 (in Chinese)

[14] Ando Y, Kurihara Y. Nonlinear response in evaluating the subjective diffuseness of sound fields. The Journal of the Acoustical Society of America. 1986;**80**(3):833-836. DOI: 10.1121/1.393906

[15] Chen CY, Ando Y. On the relationship between the autocorrelation function of the α-waves on the left and right hemispheres and subjective preference for the reverberation time of music sound field. Journal

of Architecture, Planning and Environmental Engineering. 1996;**489**:73-80

[16] Chen CY, Chan MH. A study of the Chinese speech intelligibility in halls in relation to the autocorrelation function model: The case of Chinese in Taiwan. Journal of Architecture, Architectural Institute of the Republic of China. 2006;**57**:55-68 (in Chinese)

[17] Chen YS. Analysis of cortical continuous brainwaves in relation to speech intelligibility in room [thesis of Master Degree]. 2010. pp. 33-38

[18] Ichikawa G. The logarithmical presentation for auditory evoked potential. Audiology Japan. 1983;**26**:735-739

[19] Ando Y, Kang SH, Morita K. On the relationship between auditory-evoked potential and subjective preference for sound field. Journal of the Acoustical Society of Japan (E). 1987;**8**:197-204. DOI: 10.1250/ast.8.197

[20] Ando Y, Kang SH, Nagamatsu H. On the auditory-evoked potential in relation to the IACC of sound field. Journal of the Acoustical Society of Japan (E). 1987;**8**:183-190. DOI: 10.1250/ast.8.183

[21] Jasper HH. The 10/20 electrode system of the international federation. Electroencephalography and Clinical Neurophysiology. 1958;**10**:371-375

[22] Ando Y. Architectural Acoustics. New York, Inc.: Springer-Verlag; 1998. pp. 89-95. DOI: 10.1121/1.423953

[23] Floel A, Jansen A, Deppe M, et al. A typical hemispheric dominance for attention: Functional MRI topography. Journal of Cerebral Blood Flow and Metabolism. 2005;**25**:1197-1208

Chapter 3

Neuroplasticity and the Auditory System

Caroline Donadon, Milaine Dominici Sanfins,
Piotr Henry Skarzynski, Maria Francisca Colella dos Santos
and Stavros Hatzopoulos

Abstract

This chapter will present information on the central auditory nervous system with a special focus in the auditory pathways. The intrinsic and extrinsic aspects of neuroplasticity will be described, and the neuroplasticity of the auditory system will be presented in detail. These topics are the basis of the auditory training (AT) program for central auditory processing disorders.

Keywords: auditory pathways, neuroplasticity, auditory training

1. Introduction

1.1 The central auditory nervous system

The central nervous system is bilaterally symmetrical, and it is composed of seven main regions: the spinal cord, the bulb, the bridge, the cerebellum, the midbrain, the diencephalon, and the cerebral hemispheres. Each of these neural regions performs a number of specific functions. Additionally, each function, whether it is a sensory, a motor, or another integrative task, is performed by more than one neural pathway [1, 2].

The linguistical cerebral functions are mainly located in the auditory cortex, which is divided into four anatomically distinct lobes: the frontal, the parietal, the occipital, and the temporal lobe. The latter is responsible for the function of hearing as well as for various aspects of learning, of memory, and of emotions (**Figure 1**).

The cerebral hemispheres are characterized by two important organizational features:

i. Each hemisphere is primarily related to specific sensory and motor processes on the opposite (contralateral) side of the body. The structure connecting the two hemispheres is the corpus callosum [2].

ii. Although both hemispheres appear to be similar (in humans), they are not structurally fully symmetrical nor they have equivalent functions.

Broca, Wernicke, and Penfield were pioneers in unraveling the functions of the temporal lobe. Penfield found that a stimulation of primary auditory areas produced gross auditory sensations, whereas stimulation in the superior temporal gyrus produced altered perception of auditory sounds, illusions, and hallucinations.

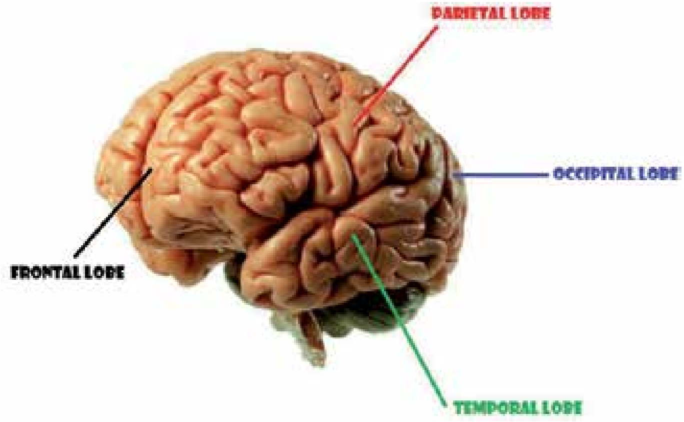

Figure 1.
The four cerebral lobes of the auditory cortex.

Studies on epileptic patients, whose hemispheres were separated by a section of the corpus callosum, have allowed us to understand numerous details related to the concept of hemispheric specialization [3].

An important contribution for the understanding of hemispheric function was achieved through the development of the Wada test [4]. The latter was developed in order to determine the dominant hemisphere for speech, so that inadvertent lesions of the speech centers, during neurosurgical procedures, could be avoided. During the test, the patient is instructed to count aloud, while sodium nitrite, a fast-acting barbiturate, is injected into the left or right carotid artery. The drug preferentially accesses the hemisphere on the same side as the injection, causing a brief speech dysfunction. If the dominant hemisphere's speech center is affected, the patient usually stops counting. With this test, the relationship between hemispheric specialization and laterality can be assessed, especially in left-handed subjects. Data in the literature [4] suggest that 96% of the right-handed people have dominant speech centers in the left hemisphere, while 15% of left-handed people have dominant speech centers in the right hemisphere. In some left-handed people, speech is controlled by both hemispheres. In such cases the administration of sodium nitrite does not suppress speech. Similar results were observed in trials involving hearing. When sounds were presented at the same time in both ears, it was found that in right-handed people the left ear performed better with nonverbal sounds.

The auditory system can be considered a high-performance signal-processing region, presenting a complex built-in hierarchy. Within the system each structure has a specific function, and progressing upwards along the pathways, these functions become more specific and dependent on the functional and physiological integrity of the previous structures. The auditory system plays an essential role for the communication among the members of the same species. Additionally, humans use the sensory inputs of the auditory system to identify different sounds leading to immediate actions, such as the status of alertness caused by the perception of siren sounds from police cars and ambulances.

From a simplified point of view, the auditory system consists of two areas, namely, the auditory periphery and the central auditory system. In each there are several structures which are stimulated when a sound stimulus is presented. Some researchers [5], however, have disputed this simple division of the auditory system suggesting instead a three-stage depiction:

i. The periphery, which captures and converts the acoustic sound stimuli into electrical neural pulses.

ii. The brainstem, which performs the initial processing of the information through modulation and interaction of the signals.

iii. The thalamocortical region, which is responsible for more advanced functions and produces emotional, cognitive, and linguistic responses from the acoustical stimuli

The efficiency of this set of interconnected structures depends primarily on the auditory experience. The simplest auditory task is influenced by high-level functions that include motivation, memory, and decision-making [6].

From a functional point of view, the following functions (among others) are assigned to CANS: the ability to detect and discriminate sound sources, the separation of acoustic stimuli from the background noise, the process of understanding the incoming stimuli, and the process of recognizing the stimuli as something familiar, though memory connections [7].

There are two pathways in the CANS: the ascending (afferent) and the descending (efferent) pathway. The afferent pathway is the path that an impulse, generated in the hair cells, travels along to the auditory cortex), whereas the efferent auditory pathway is a similar path, but in the opposite direction, conducting impulses from the auditory cortex to the hair cells [8].

The afferent and efferent pathways act in an integrated way. The afferent auditory pathway has a bilateral and predominantly contralateral auditory representation. The propagation of auditory information occurs via the cochlear nuclei, superior olivary complex, lateral lemniscus, inferior colliculus, and medial geniculate body up to the auditory area of the temporal lobe in the cerebral cortex. The efferent auditory pathway is composed by the medial and lateral olivocochlear bundles which have anatomical and physiological differences, coordinating the independent function of the two ears [9]. The function of the efferent auditory feedback includes the electrical modulation of the outer hair cells in the cochlea, the reduction of the cochlear nerve action potentials, the protection against noise, the localization of a sound source, the improvement in the detection of sound sources in noisy environments, and the focusing of attention to the incoming acoustic stimuli, which is less effective in patients with tinnitus [9, 10] (**Figure 2**).

It should be emphasized that information from sound stimuli is sent to the brain through both ipsilateral and contralateral pathways, the result of which provides data on signal timing and stimulus intensity. These are passed on to associative areas which process the data in a differentiated way, with the left hemispheric dominance for the processing of language and the right hemisphere for the process of melody. The data on the hemispheric dominance are derived from studies in patients with cortical lesions revealing a loss of recognition of family songs and prosody in patients with right-sided injury and poor recognition of verbal language and symbolic noise in patients with a left lesion [11].

Rees [12] identified auditory perception deficits as one of the causes of language disorders and concluded that auditory abilities seem to play a major role in language and learning. Lubert [13] reported that a deficiency in the ability to detect acoustic characteristics of an auditory signal was overwhelmingly important that affected children never achieved good performance in language tasks.

There is a strong interest in the literature on the impact of auditory processing deficits on language skills and reading [14]. However, the nature of the relationship between auditory and speech processing continues to be debated [15]. The starting point for differentiating auditory processing from that of language consists of a knowledge of acoustic, phonemic, and linguistic characteristics in a behavioral and neurological way. Individuals who have primary deficits in their auditory perceptual

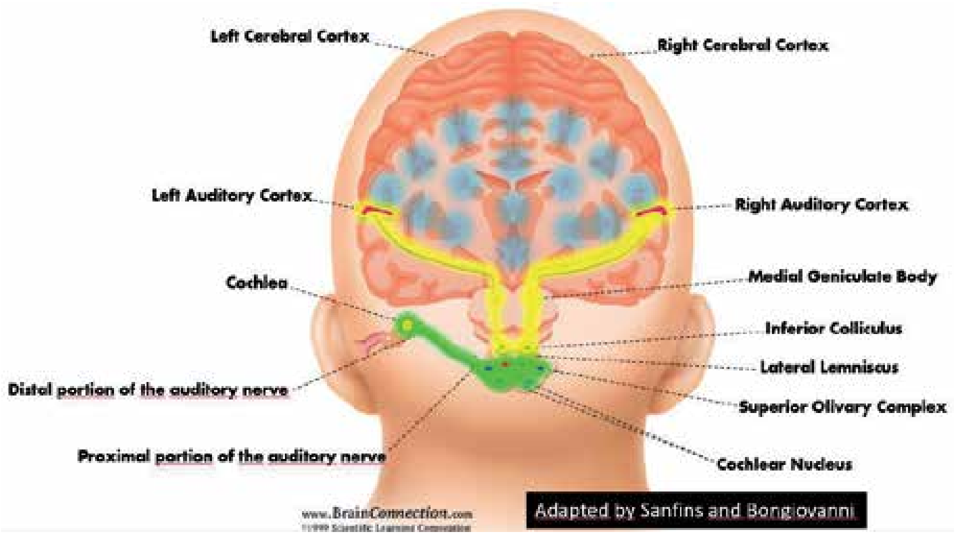

Figure 2.
Schematic representation of the auditory pathways.

abilities therefore have similar symptoms to those who have other pathologies such as dyslexia and attention deficit hyperactivity disorder, and as a consequence they may have attention and executive deficits as well [16].

1.2 Neuroplasticity

The human brain has certain time periods—called critical periods—during which it is conducive to neuroplasticity. In these critical periods, capacities are shaped, perfected, or altered as a result of experience.

In humans, cortical neuroplasticity is most pronounced in the first years of life. During this developmental period, cortical neurons are extensively stimulated, and in this way, synapses mature and developed. In addition, various sensory and cognitive systems interact and adjust their functional properties based on prior experience and learning. Younger brains seem to be more able to change as a result of persisting stimuli [17]. These are usually related to changes (i) in behavior, (ii) in the environment, and (iii) in the neural processes.

Over the course of a lifetime, a lack of experience during critical developmental periods can hinder learning [19]. A critical period can be described by the simultaneous presence of these three conditions:

i. The information must be reliable and extremely precise.

ii. The neural connections need to be intact to be able to process information, either through inhibitory and/or excitatory connections [18].

iii. Mechanisms must be in place to sustain the plasticity process, such as modifications in the morphologies of axons and dendrites and modification to synaptic connections.

It should be emphasized that simple skills require the use of less specialized neural circuits, while more complex abilities depend on the use of more specialized ones. The simplest neural circuits need first to be activated and to be efficient, before new neural circuits can be made reliable [20, 21].

Stimulation of any skill during the critical period of development is an extremely important factor for the success of any intervention process. However, it is important to note that adult brains also have a proven ability to change. Thus, different neuronal systems can be activated regardless of the age of the individual [22, 23].

Activities which depend on the integration of different neuronal systems engage in a multimodal cerebral activation, and thus they can enhance neuroplasticity. A good example of such multisensory stimulation is the process of learning to play music [24–26]. Paraskevopoulos et al. [25] demonstrated that musicians who started their training as young adults had a greater activation of the prefrontal cortex than musicians with only short-term training. Data in the literature suggest that a wide range of beneficial effects can be manifested by elderly musical students, including improvements in attention, memory, motor function, executive function, creativity, anxiety reduction, and visual scanning [27–30].

Intrinsic and extrinsic factors can cause changes in brain cells. Data from the literature suggest that new neurons are present after 6–8 weeks from the time an adult undertakes a new skill [31, 32]. It is therefore suggested that learning and maintaining a new activity should be encouraged in order to activate neural circuits and create new synapses.

Neuroplasticity has been associated with a delayed onset of dementia. Broolmeyer et al. [33] state that brain plasticity should be made a priority in dealing with individuals who have dementia. Concomitantly, age-related cognitive decline can be delayed, interrupted, or even reversed by introducing tasks that involve multimodal neuronal stimulation.

By recognizing the importance of neuroplasticity, professionals involved in rehabilitation are encouraged to turn their efforts toward stimulating, motivating, creating, and developing new strategies for the treatment of their patients.

1.3 Neuroplasticity of the auditory system

Like other systems, the development of the central auditory nervous system depends on a critical period during the first years of life when responses to different stimuli and sound environments are gradually established. In the auditory system the capacity for anatomical and functional modification is called auditory neuroplasticity [34].

The cortical areas that encompass the auditory system develop rapidly in the first years of life, due to an abundance of neuronal connections [35, 36]. At approximately 4 years of age, the neurons responsible for hearing go through a process which is called pruning, where neurons and synapses which are not activated are eliminated from the system [37].

Although the plasticity due to experience is far greater in the first years of life, it is known that the auditory system has some malleability throughout life [37]. Sharma et al. [38] established that there was a difference between what is known as a critical period and a sensitive period. According to Sharma, the critical period ends suddenly, and the neural system is unable to adapt to stimuli; in contrast, the sensitive period is an ideal neuroplastic period during which sound can be introduced into the auditory cortex and promote normal age-appropriate development.

Preterm infants who remain long periods in a neonatal intensive care unit (NICU) are often exposed to high ambient noise levels, generated by the hospital equipment. The high-frequency sounds can cause acoustic trauma and hamper the proper development of the central auditory nervous system [39]. According to Zhang et al., excessive noise at critical periods of development can lead to impaired cortical tonotopic maps, resulting in a reduction in neural synchrony and a decreased sensitivity to particular frequencies [40]. In addition, the extra noise can mask speech sounds, thereby impoverishing the auditory experience. As a result, infants can become more sensitive to noise and focus their attention on this type of sound stimulus instead of ignoring it and focusing on speech [41]. Among preterm infants there is a high rate of impairment of hearing, language, and attention; on the other hand, a home environment rich in post-NICU auditory and linguistic stimuli favors auditory neuroplasticity, meaning that premature infants then have a good chance of developing normal speech, language, and learning [42].

One way to observe plasticity in the auditory system is by monitoring patients undergoing cochlear implantation. Even after a period of auditory deprivation due to hearing loss, it is possible for the brain's auditory system to reorganize and develop better hearing abilities. Research on children implanted at the age of 3, 5, and 7 years has demonstrated that cortical auditory development can be mixed, with some children presenting cortical auditory evoked potential responses (notably P1) within normal limits, while others do not seem to achieve normal central auditory maturity. These findings are consistent with positron emission tomography (PET) imaging tests performed before and after cochlear implantation. It appears that 3.5 years of age is the end of the sensitive period for cochlear implantation in children with congenital deafness; this age is approximately when the observed exponential increase in synaptic density ends and begins to decrease [35]. Beyond 7 years of age, neuroplasticity in the central auditory system is significantly reduced; if new sounds are introduced after this time, the auditory cortex is unable to process auditory information normally [38, 43]. Research on the development of speech and language skills in children has indicated significantly better outcomes in those who received cochlear implants at younger ages [44, 45].

Neuroplasticity can be observed in individuals with central auditory processing disorder (CAPD) who have undergone auditory training. Training is a therapeutic procedure involving auditory stimulation that leads to reorganization (remapping) of the cortex and brainstem, improving synaptic efficiency and increasing neural density. These neurophysiological changes, reflected on behavioral changes, have encouraged the use of this rehabilitation strategy [46–48].

1.4 Auditory training in central auditory processing disorder

It is well established that listeners with CAPD exhibit diverse behaviors such as poor listening skills, difficulty learning through the auditory modality, difficulty following auditory instructions, difficulty in understanding when there is background noise, requesting information to be repeated, poor auditory attention, easily distracted, deficits with phonological awareness and phonic skills, weak auditory memory, delayed response to verbal stimuli, and difficulty with spelling, reading, and learning [49].

A diagnosis of impaired central auditory processing is done by applying a battery of behavioral and electrophysiological procedures. The results provide information about the physiological mechanisms in the auditory system and a profile of abilities that are altered and those that are preserved. Based on this diagnostic information, rehabilitation should start as soon as possible in order to minimize the effects of CAPD on language development. One strategy is the use of auditory training (AT), defined as the set of (acoustic) conditions and/or tasks designed to activate auditory and related systems such that neural connections and the associated auditory behavior is improved [50].

The general aim of AT when applied to individuals with CAPD is to improve auditory skills such as sound localization and lateralization, auditory discrimination, auditory pattern recognition, temporal aspects of audition, and auditory discrimination against competing acoustic signals [51]. Formal and informal AT procedures are conducted by audiologists in the clinical environment. The difference between them is that formal training needs to be acoustically controlled, with a strict control over the stimulus generation and presentation. The combination of formal and informal AT procedures offers a flexible approach which presents positively effective outcomes [50].

The management of CAPD requires a multidisciplinary team, since the pathology commonly appears with other disorders (attention deficit/hyperactivity disorder), learning and language disabilities, or dyslexia. The multidisciplinary team members are often speech-language pathologists, psychologists, neuropsychologists, neuropediatric specialists, teachers and parents, or other specialists involved in the child's overall care [52] (**Figure 3**).

The therapy to enhance auditory skills should be evidence-based, individualized, and segmented into bottom-up and top-down treatments. A bottom-up approach is based on the premise that difficulties in central auditory processing (CAP) lead to impaired auditory perception, language, reading, and communication. The objective of the bottom-up therapy is to improve speech perception. The top-down approach includes auditory cohesion, auditory attention, and metacognitive and metalinguistic activities [52, 53].

The AT program should follow some important principles:

- It should be frequent, challenging, and motivating, using age and language appropriate for the patient.
- It should include diverse tasks to maintain motivation.
- It should be gradual in difficulty over time.
- It should employ a follow-up on acquired responses (achieving response rates >70% is an indication that the task needs to be more demanding).
- It should use monitoring and feedback based on psychophysical, electrophysiological, and questionnaire-based information [50, 52].

The results obtained in the diagnostic battery will guide the therapeutic planning, which should include tasks aimed at discriminating sound intensity, frequency, and duration; phoneme discrimination; time perception discrimination; temporal ordering and sequencing; pattern recognition, location, and lateralization; and recognition of auditory information in the presence of competitive signals. Other aspects may include study of interhemispheric information transfer and binaural listening [51, 54].

In addition, modifications are important depending on the environment. To improve access to auditory information outside the therapy room, teachers and parents also need to help with CAPD treatment strategies. Simple changes may bring many benefits to learning. Options may include:

i. Preferential seating

ii. Addition of visual cues

iii. Clear language

iv. Making frequent checks for understanding

v. Repetition or rephrasing

vi. Multimodality cues and hands-on demonstrations

vii. Pre-teaching of new information and new vocabulary

viii. Provision of a notetaker

ix. Gaining attention prior to speaking

x. Positive reinforcement

xi. Reduce background noise

xii. FM systems

Monitoring progress of the patient is important since it allows the therapist to measure the appropriateness of the AT program and provides a basis for feedback to the patient and parents [50]. Ideally, three types of monitoring should be employed to measure auditory changes: psychophysical, electrophysiological, and questionnaires. These measures should be obtained before and after hearing training. Several questionnaires are available and can be answered by the patient and/or individuals interacting with him or her, such as parents, teachers, and other professionals.

Several questionnaires are described in the literature, such as the Children's Auditory Performance Scale (CHAPS) [55, 56], Screening Instrument for Targeting Educational Risk (SIFTER) [56, 57], Children's Home Inventory of Listening Difficulties (CHILD) [58], and the Scale of Auditory Behaviors (SAB) [58].

A large number of studies provide definitive evidence for the plasticity of the auditory system evidenced by behavioral changes in both animals [59–61] and in humans [62–68]. A recent study by Donadon et al. [69], whose objective was to investigate auditory training in children and adolescents suffering from otitis media

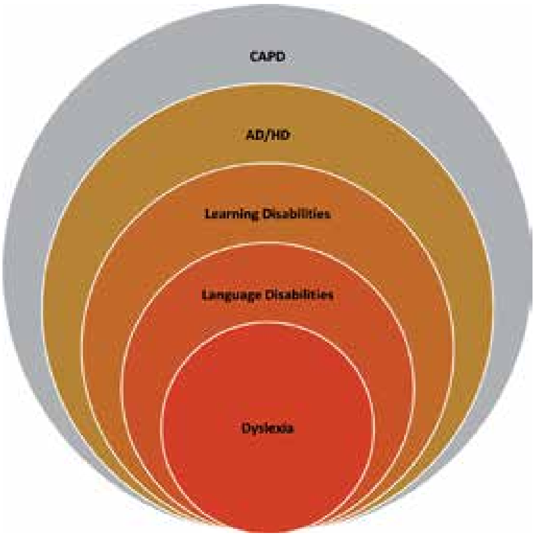

Figure 3.
CAPD and associated pathologies.

with a documented history of bilateral ventilation tube insertion, highlighted some aspects of auditory neuroplasticity. According to the data from the study, the participants were randomly divided into two groups: (i) auditory training and (ii) visual training. In the behavioral tests during the pre-intervention evaluation, no statistical differences were detected. However, after the auditory training program, there was an improvement in the subjects' performance for auditory abilities. In addition, comparing the two types of intervention (visual vs. auditory), the behavioral tests revealed better responses to the post-intervention auditory training. The results, assessed through behavioral tests on subjects with a history of bilateral otitis media, suggest that auditory training provided beneficial gains for all auditory abilities.

2. Conclusion

The central auditory nervous system is responsible for the processing of auditory information. It is highly complex and plastic, being able to reorganize itself in response to auditory stimulation. Auditory training promotes behavioral and electrophysiological changes due to the neurophysiology of the brain's plasticity. The latter enables the positive performance of the auditory training, which is an important rehabilitation strategy for individuals with central auditory processing disorders.

Abbreviations

AD/HD	attention deficit/hyperactivity disorder
AT	auditory training
CANS	central auditory nervous system
CAP	central auditory processing
CAPD	central auditory processing disorder
CNS	central nervous system
critical period	the time during which the neural system is unable to adapt
sensitive period	the ideal period for neuroplasticity to occur
Wada test	a test for determining the dominant hemisphere for speech

Author details

Caroline Donadon[1], Milaine Dominici Sanfins[2]*, Piotr Henry Skarzynski[3,4,5],
Maria Francisca Colella dos Santos[6] and Stavros Hatzopoulos[7]

1 Program in Child and Adolescent Health, Faculty of Medical Sciences,
University of Campinas, Campinas, Brazil

2 Centro de Eletrofisiologia e Neuroaudiologia Avançada (CENA), São Paulo, Brazil

3 Department of Teleaudiology and Hearing Screening, World Hearing Center,
Institute of Physiology and Pathology of Hearing, Warsaw/Kajetany, Poland

4 Department of Heart Failure and Cardiac Rehabilitation, Warsaw, Poland

5 Institute of Sensory Organs, Warsaw, Poland

6 Department of Human Development and Rehabilitation, State University of
Campinas, Campinas, Brazil

7 Clinic of Audiology and ENT, University of Ferrara, Ferrara, Italy

*Address all correspondence to: msanfins@uol.com.br

References

[1] Donadon C, Dominici Sanfins M, Skarzynski PH, Colella dos Santos MF, Hatzopoulos S. Neuroplasticity and the auditory system. In: Hatzopoulos S, Ciorba A, Skarzynski PH, editors. The Human Auditory System – Basic Features and Updates on Audiological Diagnosis and Therapy. London: IntechOpen; 2019. DOI: 10.5772/intechopen. 90085

[2] Kandel E, Schwartz J, Jessell T. Fundamentos da Neurociência e do Comportamento. Rio de Janeiro: Guanabara Koogan S.A; 2000

[3] Sperry R. Cerebral organization and behavior. The American Association for the Advancement of Science. Science. 1960;**133**:1749-1757

[4] Rasmussen T, Milner B. The role of early left-brain injury in determining lateralization of cerebral speech functions. Annals of the New York Academy of Sciences. 1977;**30**:355-369

[5] Finger S. Origins of Neuroscience: A History of Explorations into Brain Function. New York: Oxford University Press; 1994

[6] Bellis T. Differential diagnosis of (central) auditory processing disorder in older listeners. In: Musiek F, Chermak G, editors. Handbook of (Central) Auditory Processing Disorder: Auditory Neuroscience and Diagnosis. San Diego: Plural Publishing; 2007. pp. 319-346

[7] Pickles J. An Introduction to the Physiology of Hearing. Berkeley: Academic Press; 1988

[8] Berne A, Delrieu G, Creutin J, Obled C. Temporal and spatial resolution of rainfall measurements required for urban hydrology. Journal of Hydrology. 2004;**299**:166-179

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

[10] Breuel M, Sanchez T, Bento R. Vias auditivas eferentes e seu papel no sistema auditivo. Arquivos de Otorrinolaringologia. 2001;**5**(2):62-67

[11] Musiek F, Reeves A. Asymmetries of the auditory areas of the cerebrum. Journal of the American Academic of Audiology. 1990:240-245

[12] Rees N. Auditory processing factors in language disorders: The view from procrustes'bed. Journal of Speech and Hearing Disorders. 1973;**38**:304-315

[13] Lubert N. Auditory perceptual impairments in children with specific language disorders: A review of the literature. The Journal of Speech and Hearing Disorders. 1981;**46**(1):1-9

[14] Sharma M, Purdy S, Kelly A. Comorbidity of auditory processing, language, and reading disorders. Journal of Speech, Language, and Hearing Research. 2009:706-722

[15] Richard G. The role of the speech-language pathologist in identifying and treating children with auditory processing disorder. Language, Speech, and Hearing Services in Schools. 2011;**42**(3):241-245

[16] Richard G. Language processing versus auditory processing. In: Geffner D, Ross-Swain D, editors. Auditory Processing Disorders: Assessment, Management and Treatment. 2a ed. San Diego: Plural Publishing; 2013

[17] Kennard M. Reorganization of motor function in the cerebral cortex of monkeys deprived of motor and premotor areas in infancy. Journal of Neurophysiology. 1938;**1**:477-496

[18] Fagiolini M, Hensch T. Inhibitory threshold for critical-period activation in primary visual cortex. Nature. 2000;**404**:183-186

[19] Immelmann K. Sexual imprinting in birds. Advances in the Study of Behavior. 1972;**4**:147-174

[20] Jones E. Cortical and subcortical contributions to activity-dependent plasticity in primate somatosensory cortex. Annual Review of Neuroscience. 2000;**23**:1-37

[21] Daw N. Critical periods and strabismus: What questions remain? Optometry and Vision Science. 1997;**74**:690-694

[22] Irvine D, Fallon J, Kamke M. Plasticity in the adult central auditory system. Acoustics Australia. 2006;**34**:13-17

[23] Mecklenburg D, Babighian G. Cochlear performance as an indicator of auditory plasticity in humans. In: Salvi R, Henderson D, Fiorino F, Colletti V, editors. Auditory System Plasticity and Regeneration. New York: Tieme Medical; 1996. pp. 395-404

[24] Kuchenbuch A, Paraskevopoulos E, Herholz SC, Pantev C. Audio-tactile integration and the influence of musical training. PLoS One. 2014;**9**(1):e85743

[25] Paraskevopoulos E, Kuchenbuch A, Herholz SC, Pantev C. Multisensory integration during short-term music reading training enhances both uni- and multisensory cortical processing. Journal of Cognitive Neuroscience. 2014;**26**(10):2224-2238

[26] Pantev C, Paraskevopoulos E, Kuchenbuch A, Lu Y, Herholz SC. Musical expertise is related to neuroplastic changes of multisensory nature within the auditory cortex. The European Journal of Neuroscience. 2015;**41**(5):709-717

[27] Benz S, Sellaro R, Hommel B, Colzato LS. Music makes the world go round: The impact of musical training on non-musical cognitive functions—A review. Frontiers in Psychology. 2015;**6**:2023

[28] Hars M, Herrmann F, Gold G, Rizzoli R, Trombetti A. Effect of music-based multitask training on cognition and mood in older adults. Age and Ageing. 2014;**43**:196-200

[29] Bottiroli S, Rossi A, Russo R, T V, Cavallini E. The cognitive effects of listening to background music on older adults: Processing speed improves with upbeat music, while memory seems to benefit from both upbeat and downbeat music. Frontiers in Aging Neuroscience. 2014;**6**:284-291

[30] Seinfeld S, Figueroa H, Ortiz-Gil J, Sanchez-Vives MV. Effects of music learning and piano practice on cognitive function, mood and quality of life in older adults. Frontiers in Psychology. 2013;**4**:810

[31] Ge S, Yang C, Hsu K, Ming G, Song H. A critical period for enhanced synaptic plasticity in newly generated neurons of the adult brain. Neuron. 2007;**54**:559-566

[32] Dranovsky A, Picchini A, Moadel T, Sisti A, Yamada A, Kimura S, et al. Experience dictates stem cell fate in the adult hippocampus. Neuron. 2011;**70**:908-923

[33] Brookmeyer R, Johnson E, Ziegler-Graham K, HM A. Forecasting the global burden of Alzheimer's disease. Alzheimer's & Dementia. 2007;**3**:186-191

[34] Pascual-Leone A, Amedi A, Fregni F, Merabet L. The plastic human brain cortex. Annual Review of Neuroscience. 2005;**28**:377-401

[35] Huttenlocher P, Dabholkar A. Regional differences in synaptogenesis in human cerebral cortex. Journal of Comparative Neurology. 1997;**387**:167-178

[36] Huttenlocher P, de Courten C. The development of synapses in striate cortex of man. Human Neurobiology. 1987;**6**:1-9

[37] Sanes D, Woolley S. A behavioral framework to guide research on central auditory development and plasticity. Neuron. 2011;**72**:912-929

[38] Sharma A, Glick H, Campbell J, Torres J, Dorman M, Zeitler DM. Cortical plasticity and reorganization in pediatric single-sided deafness pre- and postcochlear implantation: A case study. Otology & Neurotology. 2016;**37**(2):e26-e34

[39] Lahav A, Skoe E. An acoustic gap between the NICU and womb: A potential risk for compromised neuroplasticity of the auditory system in preterm infants. Frontiers in Neuroscience. 2014;**8**:1-8

[40] Zhang L, Bao S, Merzenich M. Persistent and specific influences of early acoustic environments on primary auditory cortex. Nature Neuroscience. 2001;**4**:1123-1130

[41] Skoe E, Chandrasekaran B. The layering of auditory experiences in driving experience-dependent subcortical plasticity. Hearing Research. 2014;**311**:36-48

[42] Skoe E, Krizman J, Kraus N. The impoverished brain: Disparities in maternal education affect the neural response to sound. The Journal of Neuroscience. 2013;**33**(44):17221-17231

[43] Sharma A, Glick H. Cross-modal re-organization in clinical populations with hearing loss. Brain Sciences. 2016;**6**(1)

[44] Geers AE, Nicholas J, Tobey E, Davidson L. Persistent language delay versus late language emergence in children with early cochlear implantation. Journal of Speech, Language, and Hearing Research. 2016;**59**(1):155-170

[45] Holt R, Svirsky M. An exploratory look at pediatric cochlear implantation: Is earliest always best? Ear and Hearing. 2008;**29**:492-511

[46] Knudsen E. Capacity of plasticity in the adult owl auditory system expanded by juvenile experience. Science. 1998;**279**:1531-1533

[47] de Boer J, Thornton ARD. Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. The Journal of Neuroscience. 2008;**28**:4929-4937

[48] Johnson KL, Nicol T, Zecker SG, Kraus N. Developmental plasticity in the human auditory brainstem. The Journal of Neuroscience. 2008;**28**:4000-4007

[49] Keith RW. Random Gap Detection Test. Missouri: Auditec of Saint Louis; 2000

[50] Musiek FE, Chermak GD, Weihing J. Auditory training. In: Chermak GD, Musiek FE, editors. Handbook of Central Auditory Processing Disorder—Comprehensive Intervention. San Diego: Plural Publishing; 2014

[51] American Speech-Language-Hearing Association. (Central) Auditory Processing Disorders [Technical Report]. Available at: www.asha.org/policy.2005

[52] Bellis TJ. Assessment and Management of Central Auditory Processing Disorders in the Educational Setting from Science to Practice. Plural Publishing; 2011

[53] Chermak GD, Musiek FE. Central Auditory Processing Disorders: New Perspectives. San Diego: Singular Publishing Group; 1997. 374p

[54] AAA (American Academy of Audiology). Guidelines for the

diagnosis, treatment and management of children and adults with central auditory processing disorder. Retrieved from: http://www.audiology.or/resources/documentlibrary/documents/CAPD%20guidelines%208-2010.pdf

[55] Lam E, Sanchez L. Evaluation of screening instruments for auditory processing disorder (APD) in a sample of referred children. The Australian and New Zealand Journal of Audiology. 2007;**29**:26-39

[56] Wilson W, Jackson A, Pener A, Rose C, Wilson J, Heine C, et al. The CHAPS, SIFTER and TAPS-R as predictors of (C)AP skills and (C)APD. Journal of Speech, Language and Hearing Research. 2011;**54**(2):278-291

[57] Emanuel DC. The auditory processing battery: Survey of common practices. Journal of the American Academy of Audiology. 2002;**13**(2):93-117

[58] Schow RL, Seikel JA. Screening for (central) auditory processing disorder. In: Chermak G, Musiek F, editors. Handbook of (Central) Auditory Processing Disorder: Auditory Neuroscience and Diagnosis. San Diego, CA: Plural Pub; 2006. pp. 137-161

[59] Likenhoker B, Knudsen E. Incremental training increases the plasticity of the auditory space map in adult barn owls. Nature. 2002;**419**:293-296

[60] von der Likenhoker E, Ohe C, Knudsen E. Anatomical traces of juvenile learning in the auditory system of adult barn owls. Nature Neuroscience. 2004;**8**:93-98

[61] Recanzone G, Schreiner C, Merzenich M. Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys. Journal of Neuroscience. 1993;**13**(97-103)

[62] Hayes E, Warrier C, Nicol T, Zecker S, Kraus N. Neural plasticity following auditory training in children with learning problems. Clinical Neurophysiology. 2002;**114**:673-684

[63] Bao S, Chang E, Woods J, Merzenich M. Temporal plasticity in the primary cortex induced by operant perceptual learning. Nature Neuroscience. 2004;**7**:974-981

[64] Musiek FE. Temporal (auditory) training for (C)APD. Hearing Journal. 2005;**58**:46

[65] Alonso R, Schochat E. The efficacy of formal auditory training in children with (central) auditory processing disorder: Behavioral and electrophysiological evaluation. Brazilian Journal of Otorhinolaryngology. 2009;**75**:726-732

[66] Kraus N, Chandrasekaran B. Music training for the development of auditory skills. Nature Review Neuroscience. 2010;**11**:599-605

[67] Barcroft J, Sommers M, Tye-Murray N, Mauze E, Schroy C, Spehar B. Tailoring auditory training to patient needs with single and multiple talkers: Transfer appropriate gains on a four-choice discrimination test. International Journal of Audiology. 2011;**50**:802-808

[68] Sharma M, Purdy SC, Kelly AS. A randomized control trial of interventions in school-aged children with auditory processing disorders. International Journal of Audiology. 2012;**51**:506-518

[69] Donadon C, Sanfins MD, Borges LR, Colella-Santos MF. Auditory training: Effects on auditory abilities in children with history of otitis media. International Journal of Pediatric Otorhinolaryngology. 2019;**118**:177-180

Chapter 4

Otitis Media, Behavioral and Electrophysiological Tests, and Auditory Rehabilitation

Milaine Dominici Sanfins, Piotr Henryk Skarzynski and Maria Francisca Colella-Santos

Abstract

For speech and language to develop, an intact and active auditory system is of fundamental importance. The central auditory nervous system (CANS) can be hampered by several occurrences, including otitis media (OM) originating from inflammation in the middle ear and which is often associated with the accumulation of infected (or sometimes noninfected) fluid. OM can have a diffuse effect on cognitive and linguistic abilities, affecting both speech and phoneme perception through a failure to discriminate, store, and reproduce the acoustic contrasts necessary for comprehension. It is especially common in the first years of school. In addition, OM can generate internal noise from the presence of middle ear fluid near the cochlea, which can lead to changes in speech perception, distortion in acoustic images, and a reduction in the speed and accuracy of decoding speech. Evaluating the effectiveness of the CANS is recommended in cases where there have been repeated episodes of OM. Very useful information can be gained from behavioral and electrophysiological tests. The tests allow functional diagnoses to be made and can also reveal clinical and subclinical changes. In this way, they allow information to be collected, which can help in making a prognosis and planning intervention strategies.

Keywords: otitis media, auditory processing, electrophysiology, frequency following response, evoked auditory potential; long latency auditory evoked potential

1. Introduction

Otitis media (OM) is a common childhood disease. Research has shown that recurrent episodes can induce changes or delay the development of the central auditory nervous system, leading to central auditory processing disorder (CAPD). In this chapter, we present results obtained in the behavioral and electrophysiological evaluation of the auditory processing of children and adolescents with OM over the first few years of life. In addition, we discuss aspects of the auditory rehabilitation process itself.

2. Auditory system and otitis media

Language plays an essential role in perceptual organization, including the reception and structuring of information, learning, and social interactions. Language enables us to communicate with each other and acquire and transmit experience and knowledge. The development of speech and language requires a functional auditory system capable of detecting sound, paying attention, remembering, discriminating, and perceiving location. Any interruption to development will lead to significant functional impairments, not only in language but also in cognitive, intellectual, cultural, and social development [1, 2].

Central auditory processing (CAP) is defined as the efficiency and effectiveness with which the central auditory nervous system uses auditory information. It refers to the perceptual processing of auditory information and to the neurobiological activity underlying this processing that gives rise to electrophysiological auditory potentials [3, 4]. The efficient analysis and interpretation of normal auditory information involves several subprocesses and skills, and includes neural mechanisms underlying a range of auditory behaviors such as sound localization and lateralization; auditory discrimination; recognition of auditory patterns; temporal aspects of the hearing (integration, discrimination, resolution, temporal masking); auditory performance in the presence of competing acoustic signals (which includes dichotic listening); and decoding degraded acoustic signals [5, 6].

This whole process involves a complex system of neurons located in several stations of the auditory system. The initial analysis of the stimulus occurs in the peripheral auditory system, constituted by the external and middle ear, responsible for the capture, transduction, and processing of the sound stimulus. The stimulus arrives first at the cochlear nucleus and encephalic trunk, followed by the upper olivary complex, lateral lemniscus, inferior colliculus, and medial geniculate body, and finally reaches the primary area of auditory reception in the temporal lobe of each hemisphere. From the primary auditory cortex of each hemisphere, the signals travel to other regions of the brain-the association areas-both in the same hemisphere and in the opposite hemisphere. As the auditory information travels by ipsi- and contralateral routes, it undergoes increasingly complex levels of processing. This processing occurs both hierarchically and serially, as well as in parallel or overlapping. The result of combining serial and parallel processing makes the system highly efficient and redundant. In addition to ascending pathways, there are also descending pathways that can moderate the response to a received acoustic stimulus [7–8].

Central auditory processing disorder (CAPD) is a dysfunction of the central auditory nervous system that leads to hearing difficulties. It can lead to, or be associated with, changes in language, learning, cognition, or other communicative functions [3–5, 9]. In the pediatric population, there are several possible causes of the disorder, among them otitis media [10, 11].

Otitis media with effusion (OME) is a clinical entity characterized by the presence of effusion in the middle ear, without perforation of the tympanic membrane, but with an acute infection that lasts for a period of at least 3 months. The condition is common enough to be called an "occupational hazard of early childhood" [12] because about 90% of children have OM before school age and they develop, on average, four episodes of OM per year. OM may occur during an upper respiratory infection or occur spontaneously because of poor Eustachian tube function or an inflammatory response following a previous OM, most often between the ages of 6 months and 4 years [13, 14]. In the first year of life, 50% of children will experience OM, increasing to 60% by age 2. When primary school children aged 5–6 years were screened for OM, about 1 in 8 was found to have fluid in one or both ears [15] **Figure 1a–d**.

Most episodes of OM resolve spontaneously within 3 months, but about 30–40% of children have repeated OM episodes and 5–10% of episodes last 1 year [13]. At least 25% of OM episodes persist for 3 months and may be associated with hearing loss which is usually noticed by parents or teachers as inattention, needing to ask several times, disinterest, and poor school achievement.

OM impairs sound transmission to the inner ear by reducing mobility of the tympanic membrane and ossicles, thereby reflecting acoustic energy back into the ear canal instead of allowing it to pass freely to the cochlea.

Diagnosis is performed by otoscopy and confirmed by a basic audiological evaluation. Under otoscopy, a retracted, opaque tympanic membrane with reduced mobility is seen. In the vast majority of cases, a yellowish liquid line, sometimes with air bubbles, is visible through the tympanic membrane. In the audiological evaluation, the result can range from normal hearing to moderate conductive hearing loss (HL of 0–55 dB) [16]. The mean hearing loss associated with OM in children is 28 dB, while a lesser proportion (~20%) exceeds 35 dB, with a type B tympanometric curve characteristic of effusion. Auditory losses are characterized by being fluctuating, temporary, and asymmetrical [17]. The mild degree of loss is sufficient to impair certain auditory functions, and the fluctuating nature (which may change to periods of normal hearing) leads to variable stimulation of the central auditory nervous system. The effect is to make it difficult to perceive sounds, and leads to diffuse cognitive and linguistic abilities affecting both speech and the perception of phonemes; school performance also suffers [18]. In addition, the fluid in the middle ear can cause noise near the cochlea, producing a distorted perception of sounds.

Depending on the clinical history and functional conditions of the child's middle ear, treatment involves either clinical or surgical management. In small children with OME, the most common surgical procedure is tympanotomy with ventilation tube placement, which drains fluid from the middle ear and thus restores hearing. Diagnosis and treatment is essential, since in an acute episode of OM fluids can remain in the middle ear for 3–12 months; in 10–30% of children, the fluid remains for 2–3 months. Thus, a child who has had three to four OME episodes before the

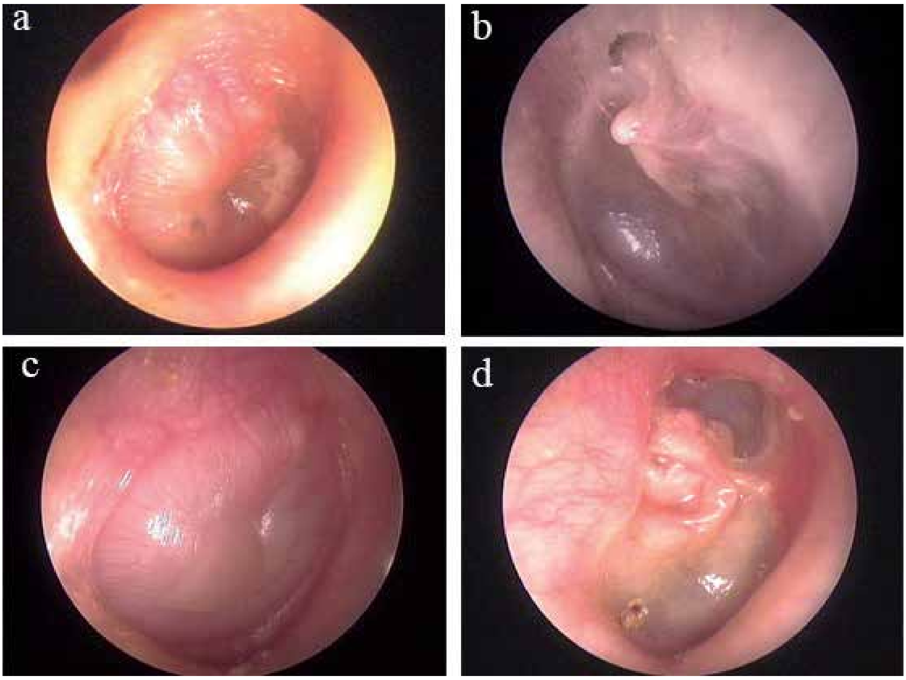

Figure 1
(a–d) Otitis media with effusion (OME). Personal collection.

age of three can have had 12 months of conductive hearing loss, which is a third of the period considered critical for development and learning [19]. The periods of auditory deprivation during the active periods of OME over the first years of life can delay the maturation of the structures in the CANS and consequently impair auditory abilities associated with central auditory processing.

Therefore, evaluation of auditory processing is fundamental in children with a history of otitis media in order to allow diagnosis, intervention, and guidance.

3. Testing the central auditory processing of children with a history of otitis media

To evaluate central auditory processing in children with a history of OM, it is recommended that a battery of test procedures be used by which the mechanisms and auditory abilities involved in the analysis and interpretation of sounds can be investigated. Due to the complexity of CANS, no single test is sufficient to explore its nature [3, 4]. Since the 1950s, numerous tests have been developed to evaluate central hearing function. These tests differ in that each presents different types of stimuli (verbal or nonverbal) and involves presentation to one or both ears (monaural or binaural). Each test is designed to evaluate a particular auditory mechanism or auditory ability and consequently probes different areas and functions of the CANS. Below the tests are divided into categories according to the way in which the stimuli are presented to the ears, the nature of the auditory tasks involved, and the method or approach used. Other currently accepted classifications involve categorizing them as binaural interaction tests, dichotic tests with verbal and non-verbal sounds (binaural integration and separation), monaural tests using low redundancy stimuli, time processing tests, and electroacoustic and electrophysiological procedures [20].

A comprehensive assessment allows for correct quantification and qualification of the various CANS mechanisms and dysfunctions and provides important information for planning and managing treatment.

3.1 Behavioral evaluation

Research by Colella-Santos et al. [11] involved 50 children (28 boys, 22 girls, mean age 11.2 years) with a documented history of bilateral SOM in the first 6 years of life and who had bilateral tympanostomy tube insertion (experimental group, EG); a control group (CG) consisted of 40 children (17 boys, 23 girls, mean age 10.7 years) with no history of otitis media. All children had auditory thresholds within normal limits on the day of evaluation and had a type A tympanometric curve. They were all evaluated with the tests described below [21–23]. The tests were the dichotic digits test, synthetic sentence identification test, gaps-in-noise test, and frequency pattern test. Details are as follows **Figure 2a** and **b**.

3.1.1 Dichotic digits (DD)

The DD test as developed in Brazil consists of four presentations of a list of two-syllable digits in Brazilian Portuguese, in which four different digits are presented simultaneously, two in each ear. The list contains 40 randomly arranged pairs of digits presented at 50 dB HL. The digits used to form the numbers are four, five, seven, eight, and nine. The participants are instructed to listen to two numbers in each ear and repeat all the numbers they hear. The order does not matter. The dichotic digits test verifies binaural integration ability [21].

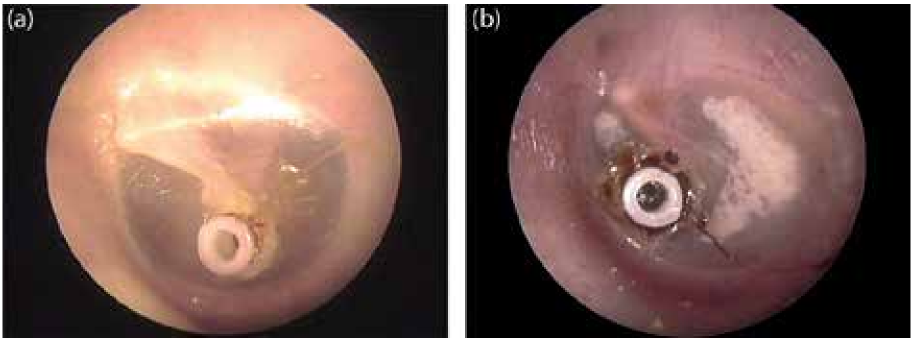

Figure 2.
(a–b) Tympanostomy tube insertion. Personal collection.

3.1.2 Synthetic sentence identification (SSI)

The SSI test consists of the presentation of 10 Brazilian Portuguese sentences at 40 dB HL, in the presence of a competing children's story in the same ear at a signal-to-noise ratio of 0, −10, or −15 dB. The task of the subject is to listen to the sentence and point to it in a frame. The ability analyzed in this test is figure-ground discrimination [21].

3.1.3 Frequency pattern test (FPT)

The FPT test is composed of three 150 ms tones presented at 50 dB HL and separated by 200 ms. The tones in each triplet are combinations of two sinusoids, 880 and 1122 Hz, which are designated as low frequency (L) and high frequency (H), respectively. Thus, there are six possible combinations of the three-tone sequence (LLH, LHL, LHH, HLH, HLL, and HHL). The subjects are instructed that they will hear sets of three consecutive tones that vary in pitch. Their task is to repeat the pattern by humming and verbalizing the frequency pattern (e.g., high–low–high). The FPT test checks temporal ordering ability [22].

3.1.4 Gaps-in-noise (GIN)

The GIN test consists of a series of 6-second segments of broad-band noise presented at 50 dB HL with 0–3 gaps embedded within each segment. The gaps vary in duration from 2 to 20 ms. The gap-detection threshold is defined as the shortest duration that is correctly identified at least four out of six times. The participants are instructed to indicate each time they perceive a gap. The GIN test measures temporal resolution ability [23].

To establish a difference between the right and left ears of subjects in the EG, it was necessary that there was a statistically significant difference in both the Dichotic Digits ($p = 0.001$) and GIN ($p = 0.004$) tests. No significant difference was found for gender in the behavioral tests. It was observed that the EG had lower mean responses than the CG for the DD test of approximately 5% in both ears; for the FPT 9.6% (humming) and 30% (naming); and 8% for the SSI test. For the GIN test, there was a statistically significant difference in the gap-detection threshold between the groups, with the highest threshold obtained in the EG compared to the CG (the higher the threshold, the worse the performance).

In summary, there was a negative effect of OM on the auditory skills of figure-background discrimination, resolution, and temporal ordering. The poorer results in CAP behavioral tests in the EG participants can be explained by the fact that OM, by generating a fluctuating auditory threshold and causing temporary auditory deprivation, hampers the maturation of auditory abilities (such as binaural integration, resolution, temporal ordering, and discrimination) which are fundamental for understanding speech. During this period of auditory deprivation due to episodic OM, the CANS received inconsistent and incomplete auditory information. That is, the period between clinical assessment and the decision to perform surgery may have been too long **Table 1**.

Recent research has demonstrated associations similar to those found in the present study. Borges et al. [11] studied the effect of OM in 69 children of different socioeconomic levels who underwent surgical intervention (insertion of ventilation tubes) and observed worse performance in both the DD and GIN tests. The authors concluded that a history of OM can lead to changes in central auditory functioning, regardless of socioeconomic status.

Khavarghazalani et al. [24] evaluated 12 children with a history of OM who had undergone surgical intervention for insertion of ventilation tubes and found worse performance in the DD and GIN responses than in normals.

Gravel and Wallace [25] also found a significant increase in signal-to-noise ratio in a prospective study of children with a history of OM. There was worse performance on the SSI test (responsible for the figure-ground ability) in the OM group.

Tomlin and Rance [26] recommend that children with a history of OM undergo an evaluation of spatial processing upon entering school. They studied 35 children with a history of chronic OM and found a statistically worse performance compared to the control group in the listening in spatialized noise-sentences test (LISN-S). They concluded that these children have altered spatial processing, difficulty in focusing attention on the relevant stimulus, and difficulty in simultaneously suppressing competing stimuli coming from other directions. It is hypothesized that fluctuating access to binaural cues, caused by OM, may negatively affect the development of spatial processing in the CANS.

3.2 Electrophysiological evaluation

Auditory evoked potentials are an extremely useful instrument for the study of auditory perception and its disorders, especially when a range of stimuli are used [27].

Procedure	**Ear**	**Control group**			**Experimental group**			
		N	**∑ (%)**	**SD**	**N**	**∑ (%)**	**SD**	***p*-value**
DD	R	40	98.93	1.86	50	95.40	5.16	<0.001
	L	40	97.93	4.15	50	92.55	7.95	<0.001
FPT								
Humming	B	80*	73.50	21.2	100*	42.7	22.2	<0.001
Verbalizing	B	80*	73.50	21.2	100*	42.7	22.2	<0.001
SSI	B	80*	67.5	13.9	100*	59.8	16.9	0.020
GIN	R	40	4.65	1.00	50	6.22	1.40	<0.001
	L	40	4.72	1.06	50	6.56	1.52	<0.001

*n = number, * = number of ears, B = both, R = right, L = left; ∑ = mean, SD = standard deviation, DD = dichotic digits, SSI = synthetic sentence identification, FPT = frequency pattern test, GIN = gaps-in-noise.*

Table 1.
Behavioral evaluation values of central auditory processing between control and experimental groups.

3.2.1 Click ABR

In the literature, there are contradictory results in Click ABR responses in individuals with a history of OM. Chambers et al. [28] and Folsom et al. [29] identified an increase in the latency of waves III and V in a group of children with a history of OM, whereas Shaffer [30] did not find a statistically significant difference in Click ABR responses in individuals with and without a history of OM. The majority of studies relating Click ABR results with OM history have investigated latency values; however, Maruthy and Mannarukrishnaiah [31] found a reduction in the amplitude of waves I and III. Sanfins et al. [32] observed statistically significant differences in the absolute latencies of waves I and V as well as in the amplitude of waves III and V from children with a history of bilateral OME compared to their healthy peers. Colella-Santos et al. [11] reported a significant increase in the absolute latency of wave III associated with a decrease in amplitude in children with bilateral OME. Finally, Sanfins [33] reported alterations in the values of waves III and V for both groups of children with a history of OME, seeing both bilateral and unilateral alterations **Figure 3**.

In animals, the effect of conductive hearing loss on CANS was studied by unilaterally removing the malleus and applying a fluid to simulate OM [34], finding a decrease in neuronal activity due to changes in various structures (wave III), upper olivary complex (wave IV), and lateral lemniscus (wave V). At the same time, based on the results of Maruthy and Mannarukrishnaiah [31], it has been suggested that the auditory nerve and cochlear nuclei are more susceptible to modifications after OM infection.

Sanfins et al. [32] suggest that different modifications may occur in CANS structures depending on the unilaterality or bilaterality of the infection. In episodes of bilateral OME, the latency values indicated that the auditory nerve (wave I, wave III) and the lateral lemniscus (wave V) were affected, whereas in unilateral OME, the cochlear nuclei (wave III) was affected. However, when the amplitudes were analyzed, the structures involved were the cochlear nuclei (wave III) and the lateral lemniscus (wave V), both for children with unilateral and bilateral involvement. It should be noted that when evaluating click ABR, the amplitude values show greater variability than the latencies. It is important to emphasize that a unilateral OM may not provide a better performance in the processing of auditory information than bilateral OM. The use of only one ear can lead to damage to the functionality of

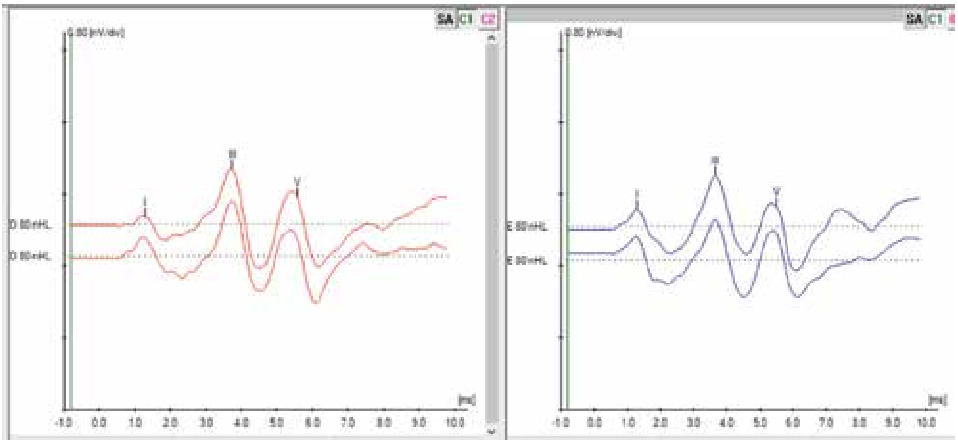

Figure 3.
Click ABR. Personal collection.

the CANS and, over time, activities that depend on binaural auditory processing (binaural interaction and binaural integration, among others) can be compromised due to the auditory imbalance arising from OM.

3.2.2 Frequency following response

Few studies have investigated the frequency following response (FFR) in cases of otitis media. A study of two groups of children with a history of bilateral OM (recent onset and long-term) showed that FFR responses were affected in a statistically significant way in the onset portions (waves V and A) and offset portion (wave O), along with reduced values of the VA complex (more specifically VA slope) when responses between the groups were compared. The findings suggest that long-term OM in children is associated with a reduced neural conduction velocity relative to the processing time of speech stimuli, either at the beginning (onset) or final portion (offset), resulting in a decrease in the coding of speech in the brainstem [35] **Figure 4**.

Sanfins et al. [32] reported that children with a history of SOM present an increase in the absolute latency of all FFR waves compared to children with no history of otological problems. In addition, children without hearing loss have more coherent responses in both ears, whereas the group of children with a history of OME has a greater dispersion of latencies in all FFR components (**Figure 5**). Colella-Santos et al. [11] also reported a decrease in VA slope in girls with OME.

3.2.3 Long latency auditory evoked potential

3.2.3.1 Tone burst

The literature reports alterations in the components of the LLAEP in children with language disorders and also in those with phonological disorders [36] changes that are frequently associated with problems arising from OM. Researchers note that OM can lead to changes in central auditory pathways [30, 37, 38]. However, there are few studies that have associated the LLAEP responses in children with a history of OM **Figure 6**.

Maruthy and Mannarukrishnaiah [31], Shaffer [30], Sanfins [33], and Colella-Santos [11] reported similar results, i.e., the presence of LLAEP changes in children with a history of OME. In the studies by Maruthy and Mannarukrishnaiah [32], all components of the LLAEP (P1, N1, P2, and N2) were significantly longer in

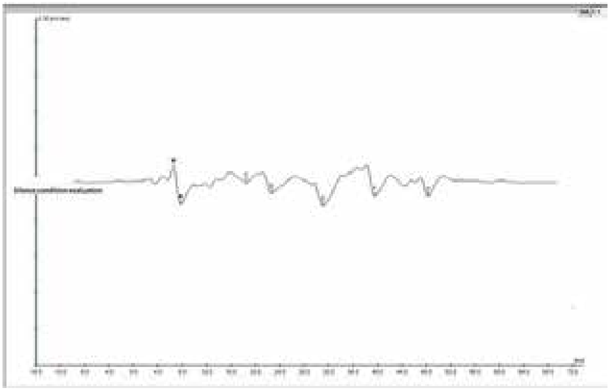

Figure 4.
FFR. Personal collection.

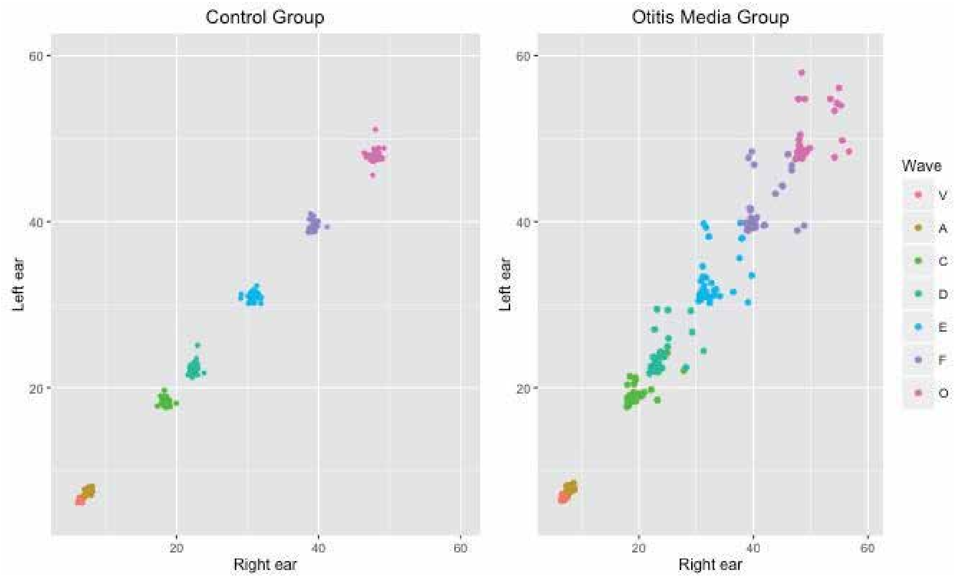

Figure 5.
Comparison (left vs. right ear) of absolute latency values of FFR components in children with a history of otitis media (right panel) and children with no history of otitis media (left panel), from Sanfins et al. [32].

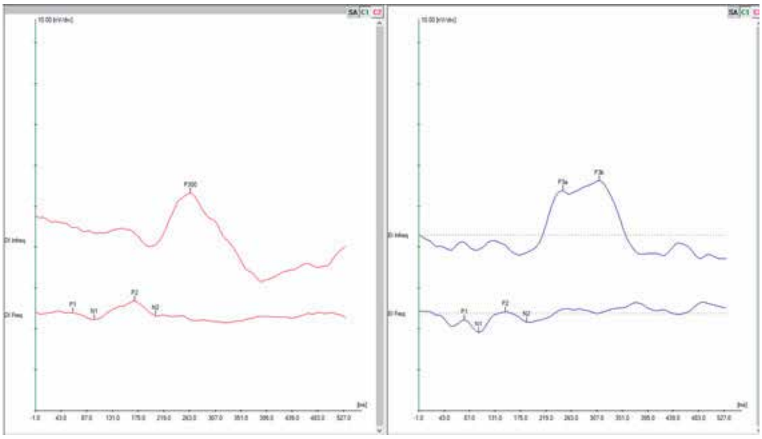

Figure 6.
LLAEP. Personal collection.

children with an SOM history. Shaffer [31] showed an increase in the latencies of N1 and P2, associated with the absence of the P300, in the majority of children evaluated. Sanfins [34] found prolongation of latencies only for P2 and N2 (for females), in comparison with the responses of children without otological alterations. Colella-Santos [11] observed an increase in P2, N2, and P300 latencies in children with a history of OME.

3.2.3.2 Speech

The LLAEP with verbal stimuli provides additional information about the biological processes involved in speech processing, enabling the collection of information complementary to those obtained by standard behavioral evaluations [30, 39, 40].

In the studies of Sanfins [33], children with bilateral OME presented prolonged latencies for N1, P2, N2 (female), and P300, in comparison with responses of children without auditory changes. Children with unilateral OME had prolonged latencies for P2 and P300 in comparison to the responses from healthy children.

The evaluation of the LLAEP using both nonverbal and verbal stimuli seems to be able to identify neurophysiological changes resulting from OM. However, it is important to note that, in unilateral OM episodes, only verbal sound stimuli (speech LLAEP) seem to be able to differentiate groups on the basis of latency. OM impairs speech perception as a result of a failure to recognize sound signals (discrimination, storage, memory). Therefore, the more accurate identification of LLAEP changes with verbal and non-verbal stimuli may relate to underlying OM.

4. Auditory rehabilitation

It is known that hearing loss due to OM during childhood development may result in long-term changes in neural function, structure, and connectivity. The changes are associated with a series of sensory, cognitive, and social difficulties suggestive of impaired brain function [41, 42] which may culminate in central auditory processing disorder (CAPD) [11].

Intervention for CAPD should be initiated as soon as the diagnosis, made through a series of behavioral and electrophysiological procedures, demonstrates the involvement of the CANS. Early identification, followed by intensive intervention, makes best use of the brain's inherent plasticity. Successful treatment outcomes depend on stimulation and repeated practice that induce cortical reorganization (and possibly reorganization of the brainstem), which is reflected in behavioral change [43–45].

Neuroplasticity is the key to the effectiveness of repeated auditory stimulation. Through experience and stimulation it induces reorganization of the cortex and brainstem, improving synaptic efficiency and neural density, giving rise to associated cognitive and behavioral changes [46–48]. The ability of the CANS to adapt to internal and external changes has important implications for learning [49].

Auditory training (AT) is defined as a set of (acoustic) conditions and/or tasks designed to activate the auditory system and related structures in such a way that their underlying neural processes and associated auditory behavior is altered in a positive way [8]. Both formal and informal AT procedures are conducted by audiologists in clinics; the difference between them is that formal training is acoustically controlled, meaning control over stimulus generation and presentation. Combined formal and informal AT offers an approach that provides more intensive practice and leads to better treatment efficacy [8]. AT performed in an individual with CAPD should include activities that aim to improve auditory skills such as sound localization and lateralization tasks, auditory discrimination, auditory pattern recognition, temporal aspects of audition, and auditory discrimination among competing acoustic signals [4].

Donadon and colleagues [50] have studied the efficacy of AT through behavioral CAP tests in children with a history of OM who had undergone bilateral tympanotomy for insertion of ventilation tubes. The sample consisted of 34 subjects who were divided into two groups: an auditory training group (ATG) formed by 20 children and adolescents, aged 8–13 years, diagnosed with CAPD, who were given an auditory training program; and a visual training group (VTG) formed of 14 children and adolescents, aged 9–13 years, diagnosed with CAPD who were given a visual training program. All subjects underwent peripheral auditory evaluation and behavioral evaluation of their CAP (using the dichotic digit test, sentence identification test with ipsilateral competing message, gaps-in-noise test, frequency pattern test, and dichotic vowel test). Auditory training was given through repeated verbal and non-verbal stimuli and associated tasks (available at the website www.afinandoocerebro.com.br) via headphones in an acoustic booth (the intensity was set at 50 dB HL). Each session lasted between 40 and 45 minutes and was performed once a week. The stimulation protocol was developed with the purpose of developing the auditory abilities of:

i. binaural integration-through dichotic listening exercises;

ii. temporal resolution-by means of minimum time interval perception exercises;

iii. temporal ordering-using nonverbal tasks related to frequency, intensity, and duration; and

iv. figure-background exercises with competing noise.

The visual stimulation protocol was elaborated using varied stimuli and tasks from the website via a 15″ notebook positioned in front of the subject on a table arranged in a sound booth. The stimulation protocol was designed with the purpose of stimulating the visual abilities of:

i. visual background;

ii. visual closure;

iii. perception and discrimination of sizes and formats; and

iv. visual memory.

All subjects were reevaluated after 8 weeks with the same battery of behavioral tests as performed at the initial evaluation. In the ATG the results showed a statistically significant difference in the abilities of binaural integration (p = 0.001), temporal ordering (p < 0.0001), temporal resolution (p < 0.0001), and bottom figure (<0.0001) in a comparison of before and after AT. These results suggest that the auditory stimulation performed during AT induced changes in the central auditory nervous system, as demonstrated by the better values recorded in the behavioral tests after intervention. Behavioral changes observed after AT in this population with a history of OM point to evidence of neuroplasticity, since auditory stimulation brought about improvements to the identified impaired hearing abilities.

For the visual training group, however, there was no significant difference in performance for any CAP behavioral tests when comparing pre and post interventions. Thus, auditory training appears to be effective as an intervention strategy for re-adjusting the auditory skills in subjects with a history of OM. Auditory stimulation brought about improvements in impaired hearing skills. AT was able to reorganize the neural substrate, providing appropriate experiences, shaping existing circuits in the CANS, and increasing neural density, reflected by an improvement in the behavioral evaluation **Figure 7**.

Modifications to a child's environment are also important aspects for teachers and parents to address in order to help individuals with CAPD improve access to

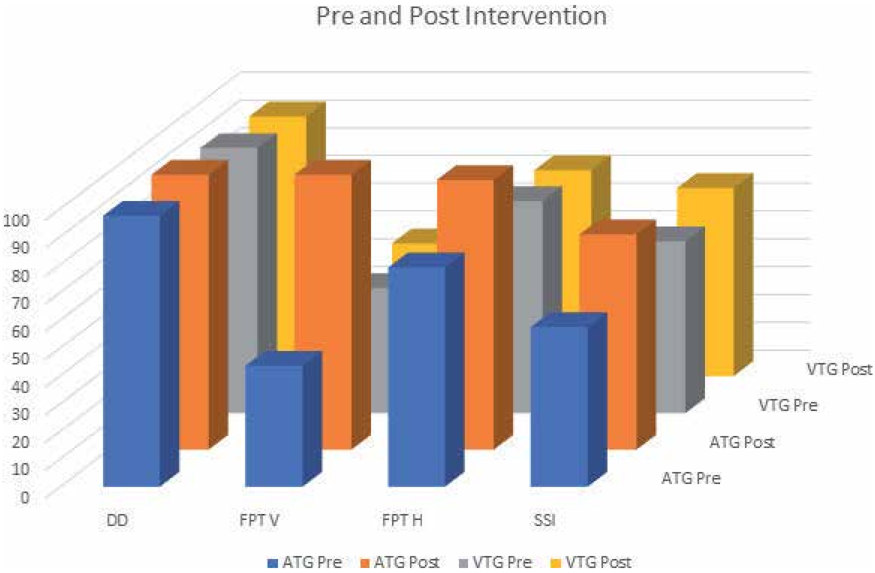

Figure 7.
Comparison of performance in behavioral evaluation pre and post intervention by groups. ATG Pre = auditory training pre intervention; ATG Post = auditory training post intervention; VTG Pre = visual training pre intervention; VTG Post = visual training post intervention; DD = dichotic digits; FPT V = frequency pattern test verbalizing; FPT H = frequency pattern test humming; SSI = synthetic sentence identification.

auditory information outside the therapy room. Some simple changes may bring many benefits to learning. Common recommendations for individuals with auditory disorders include the following:

- Preferred seating arrangements
- Addition of visual cues
- Clear language
- Making frequent checks for understanding
- Repetition or rephrasing
- Multimodality cues and hands-on demonstrations
- Preteaching of new information and new vocabulary
- Provision of a notetaker
- Recording information pictorially
- Gaining attention prior to speaking
- Positive reinforcement
- Reducing environmental noise
- FM systems

5. Summarize

- The negative effects of otitis media on the development of auditory abilities in children and the maturation of their central auditory pathways is undeniable;

- Early medical intervention in OM and family counseling is extremely important;

- The aim should be to avoid prolonged auditory fluctuation caused by OM, thereby minimizing the effects generated by fluid in the middle ear in the development of auditory abilities;

- The overall recommendation is that audiological diagnosis should include both behavioral evaluations and electrophysiological testing of auditory processing;

- In cases of auditory processing disorder, research shows that auditory training is the most effective procedure to re-adjust auditory skills.

Author details

Milaine Dominici Sanfins[1]*, Piotr Henryk Skarzynski[2,3,4]
and Maria Francisca Colella-Santos[5]

1 CENA – Centro de Eletrofisiologia e Neuroaudiologia Avançada, São Paulo, Brazil

2 Department of Teleaudiology and Hearing Screening, World Hearing Center, Institute of Physiology and Pathology of Hearing, Warsaw, Poland

3 Department of Heart Failure and Cardiac Rehabilitation, Medical University of Warsaw, Poland

4 Institute of Sensory Organs, Poland

5 Faculty of Medical Science, State University of Campinas, Campinas, Brazil

*Address all correspondence to: msanfins@uol.com.br

References

[1] Baroch K. Universal newborn hearing screening: Fine-tuning process. Current Opinion in Otolaryngology & Head and Neck Surgery. 2003;**11**:003424-003427

[2] Yoshinaga-Itano C, Sedey A. Language, speech, and social emotional development of children who are deaf or hard of hearing: The early years. Volta Review. 2000;**100**(5):298

[3] AAA (American Academy of Audiology). Clinical practice guidelines for the diagnosis, treatment and management of children and adults with central auditory processing disorder. 2010. Available from: http://audiology.org/resources/documentlibrary/Documents/CAPD%20Guidelines%208-2010.pdf

[4] ASHA (American Speech-Language-Hearing Association). Central Auditory Processing Disorders [technical report]. Rockville, MD; 2005. Available from: http://www.asha.org/docs/html/TR2005-00043.html

[5] Bellis TJ. Assessment and Management of Central Auditory Processing Disorders in the Educational Setting: From Science to Practice. Clifton Park, NY: Thomson Learning; 2003. pp. 103-139

[6] Chermak GD, Musiek FE. Central Auditory Processing Disorders: New Perspectives. San Diego, CA: Singular; 1997

[7] Musiek FE, Chermak GD. Handbook of Central Auditory Processing Disorder: Auditory Neuroscience and Diagnosis. San Diego, CA: Plural Publishing; 2014

[8] Geffner D, Ross-Swain D. Auditory Processing Disorders: Assessment, Management and Treatment. San Diego, CA: Plural Publishing; 2013

[9] Billiet CR, Bellis TJ. The relationship between brainstem temporal processing and performance on tests of central auditory function in children with reading disorders. Journal of Speech, Language, and Hearing Research. 2011;**54**:228-242

[10] Borges L, Paschoal J, Colella-Santos MF. (Central) auditory processing: The impact of otitis media. Clinics. 2013;**68**(7):954-959

[11] Colella-Santos MF, Sanfins MD, Donadon C, Borges LR. Otitis media: Long-term effect on central auditory nervous system. BioMed Research International. 2019;**2019**:10. Article ID: 8930904. DOI: https://doi.org/10.1155/2019/8930904

[12] Rosenfeld RM. A Parent's Guide to Ear Tubes. Hamilton, Canada: BC Decker Inc; 2005

[13] Tos M. Epidemiology and natural history of secretory otitis. The American Journal of Otology. 1984;**5**:459-462

[14] Mandel EM, Doyle WJ, Winther B, Alper CM. The incidence, prevalence and burden of OM in unselected children aged 1-8 years followed by weekly otoscopy through the "common cold" season. International Journal of Pediatric Otorhinolaryngology. 2008;**72**:491-499

[15] Martines F, Bentivegna D, Di Piazza F, Martinciglio G, Sciacca V, Martines E. The point prevalence of otitis media with effusion among primary school children in Western Sicily. European Archives of Oto-Rhino-Laryngology. 2010;**267**:709-714

[16] Gravel JS. Hearing and auditory function. In: Rosenfeld RM, Bluestone CD, editors. Evidence-Based Otitis Media. 2nd ed. Hamilton, Canada: BC Decker Inc; 2003. pp. 342-359

[17] Sabo DL, Paradise JL, Kurs-Lasky M, Smith CG. Hearing levels in infants and young children in relation to testing technique, age group, and the presence or absence of middle-ear effusion. Ear and Hearing. 2003;**24**:38-47

[18] Katz J, Zalewski TR, Brenner MJ. Otitis media and central auditory processing disorder (CAPD). In: Geffner D, Ross-Swain D, editors. Auditory Processing Disorders. 3rd ed. San Diego, CA: Plural Publishing; 2019. pp. 307-326

[19] Rosenfeld RM, Shin JJ, Schwartz SR, Coggins R, Gagnon L, Hackell JM, et al. Clinical practice guideline: Otitis media with effusion (update). Otolaryngology and Head and Neck Surgery. 2016;**154**(1S):S1-S41

[20] Baran JA. Test battery principles and considerations. In: Musiek FE, Chermak GD, editors. Handbook of Central Auditory Processing Disorder: Auditory Neuroscience and Diagnosis. San Diego, CA: Plural Publishing; 2014. pp. 291-323

[21] Pereira LD, Schochat E. Testes Auditivos Comportamentais Para Avalia¸ ca˜o Do Processamento Auditivo Central. Barueri, Brazil: Pro´ Fono; 2011

[22] Musiek FE, Baran JA, Pinheiro ML. Duration pattern recognition in normal subjects and in patients with cerebral and cochlear lesions. Audiology. 1990;**29**:304-313

[23] Musiek FE, Zaidan EP, Baran JA, Shinn JB, Jirsa RE. Assessing temporal processes in adults with LD: The GIN test. In: Convention of American Academy of Audiology. Vol. 203. Salt Lake City, Utah, USA: AAA; 2004

[24] Khavarghazalani B, Farahani F, Emadi M, Hosseni Dastgerdi Z. Auditory processing abilities in children with chronic otitis media with effusion. Acta Oto-Laryngologica. 2016;**136**(5):456-459

[25] Gravel JS, Wallace IF, Ruben RJ. Auditory consequences of early mild hearing loss associated with otitis media. Acta Oto-Laryngologica. 1996;**116**(2):219-221

[26] Tomlin D, Rance G. Long-term hearing deficits after childhood middle ear disease. Ear and Hearing. 2014;**35**(6):e233-e242

[27] Burkard R, Don M, Eggermont J. Auditory Evoked Potentials: Basic Principles and Clinical Application. Philadelphia: Lippincott Williams & Wilkins; 2007

[28] Chambers R, Rowan L, Mathics M, Novak L. Auditory brain-stem responses in children with previous otitis media. Archives of Otolaryngology – Head & Neck Surgery. 1989;**115**:452-456

[29] Folsom R, Weber B, Thompson G. Auditory brainstem responses and children with early recurrent middle ear disease. The Annals of Otology, Rhinology, and Laryngology. 1983;**92**:249-253

[30] Shaffer EK. Auditory evoked potentials in children with and without otitis medua. Tejas Journal of Audiology and speech pathology. 1999;**XXIII**:10-20

[31] Maruthy S, Mannarukrishnaiah J. Effect of early onset otitis media on brainstem and cortical auditory processing. Behavioral and Brain Functions. 2008;**4**:17

[32] Sanfins M, Borges L, Donadon C, Hatzopoulos S, Skarzynski P, Colella-Santos M. Electrophysiological responses to speech stimuli in children with otitis media. The Journal of Health Science. 2017;**7**(4):9-19

[33] Sanfins M. Electrophysiological evaluation with verbal and non-verbal sounds in children with a history of otitis media. State University of Campinas; 2017. Available from: http://www.repositorio.unicamp.br/handle/REPOSIP/330747

[34] Tucci DL, Cant NB, Durham D. Conductive hearing loss results in changes in cytochrome oxidase activity in gerbil central auditory system. JARO. 2001;**3**:89-106

[35] El-Kabarity RH, Abdel Rahman TT, Abdel Kader HA, Sanyelbhaa H. Effect of otitis media with effusion on brainstem timing in children. Hearing, Balance and Communication. 2016;**14**(1):20-24

[36] Tonnquist-Uhlén I. Topography of auditory evoked cortical potentials in children with severe language impairment. Scandinavian Audiology. Supplementum. 1996;**25**(44):1-40

[37] Hall JW, Grose JH, Buss E, Dev MB, Drake AF, Pillsbury HC. The effect of otitis media with effusion on perceptual masking. Archives of Otolaryngology – Head & Neck Surgery. 2003;**129**(10):1056-1062

[38] Gravel J, Roberts J, Roush J, Grose J, Besing J, Burchinal M, et al. Early otitis media with effusion, hearing loss, and auditory processes at school age. Ear and Hearing. 2006;**27**(4):353-368

[39] Oates P, Kurtzberg D, Stapells D. Effects of sensorineural hearing loss on cortical event-related potential and behavioral measures of speech-sound processing. Ear and Hearing. 2002;**23**(5):399-415

[40] Martin B, Tremblay K, Korczack P. Speech evoked potential: From the laboratory to the clinic. Ear and Hearing. 2002;**23**(5):399-415

[41] King AJ, Parsons CH, Moore DR. Plasticity in the neural coding of auditory space in the mammalian brain. Proceedings of the National Academy of Sciences of the United States of America. 2000;**97**(22):11821-11828

[42] Hogan SC, Moore DR. Impaired binaural hearing in children produced by a threshold level of middle ear disease. Journal of the Association for Research in Otolaryngology. 2003;**4**(2):123-129

[43] Kolb B. Brain Plasticity and Behavior. Mahwah, NJ: Lawrence Erlbaum; 1995

[44] Merzenich M, Jenkins W. Cortical plasticity, learning and learning dysfunction. In: Julesz B, Kovacs I, editors. Maturational Windows and Adult Cortical Plasticity: SFI Studies in the Sciences of Complexity. Vol. XXIII. Reading, PA: Addison-Wesley; 1995. pp. 247-272

[45] Russo NM, Nicol TG, Zecker SG, Hayes EA, Kraus N. Auditory training improves neural timing in the human brainstem. Behavioural Brain Research. 2005;**156**(1):95-103

[46] de Boer J, Thornton ARD. Neural correlates of perceptual learning in the auditory brainstem: Efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. The Journal of Neuroscience. 2008;**28**:4929-4937

[47] Johnson KL, Nicol T, Zecker SG, Kraus N. Developmental plasticity in the human auditory brainstem. The Journal of Neuroscience. 2008;**28**:4000-4007

[48] Song JH, Skoe E, Wong PC, Kraus N. Plasticity in the adult human auditory brainstem following short-term linguistic training. Journal of Cognitive Neuroscience. 2008;**20**(10):1892-1902

[49] Bellis TJ. Assessment and Management of Central Auditory

Processing Disorders in the Educational Setting: From Science to Practice. USA: Plural Publishing; 2011

[50] Donadon C, Sanfins MD, Borges L, Colella-Santos M. Auditory training: Effects on auditory abilities in children with history of otitis media. International Journal of Pediatric Otorhinolaryngology. 2019;**118**:177-180

Section 2

Updates on Audiological Diagnosis and Therapy

Chapter 5

Electrical Stimulation of the Auditory System

Patrick J. Boyle

Abstract

In many healthcare systems electrical stimulation of the human auditory system, using cochlear implants, is a common treatment for severe to profound deafness. This chapter will describe how electrical stimulation manages to compensate for sensory-neural hearing loss by bypassing the damaged cochlea. The challenges involved in the design and application of cochlear implants will be outlined, including the programming of clinical systems to suit the needs of implanted patients. Today's variety of patient will be reviewed: unilaterally and bilaterally implanted, bimodal users of a cochlear implant as well as a contralateral hearing aid, CROS device users having either asymmetrical hearing loss or single-sided deafness. Alternative devices such as auditory brainstem implants will be described, and additionally the more experimental auditory mid-brain implants and intraneural stimulation approaches. Research that is likely to bring medium term benefits to the clinical application of cochlear implants will also be described.

Keywords: cochlear implant, electrical, stimulation, prosthesis

1. From the beginning to current practice

Electrical stimulation of the human auditory system is generally traced back to the pioneering experiments of Alessandro Volta, inventor of the battery. When Volta applied 50 volts to his own head, he reported hearing an unpleasant boiling sound [1]. However, the forerunner of a modern CI system is just over 60 years old: opportunistic stimulation of the auditory nerve [2] of a bilaterally deaf patient receiving a facial nerve graft. During the two decades following this work, various clinical studies [3–10] saw the implantation of single and then multi-channel cochlear implant (CI) systems in people suffering profound deafness. While many of these pioneers suffered ridicule at the hands of the mainstream scientific community, clinical considerations prevailed. The early devices that were produced in academic institutions were transferred to commercial organizations, these often building on prior medical device experience, for example experience gained in the pace maker field.

Today over half a million people, from babies under 6 months of age to adults in their late 90s, have been implanted with a CI. While it can be argued that the CI is the most successful medical device ever created, the outcomes are still highly variable (**Figure 1**). In the best of cases, CI users can make fluent use of a telephone, understand speech in adverse listening conditions where there is considerable competing noise and reverberation, hence enjoying independence spanning social lives and careers that would have been unimaginable without their CI device. Even where speech understanding is limited, a release from the isolation of deafness through access to environmental

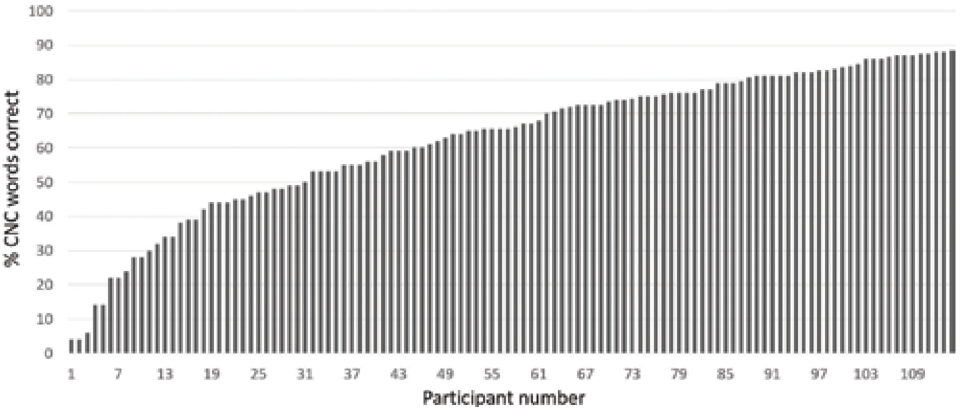

Figure 1.
Percent correct scores on the CNC word test ranked from poorest to best for 113 cochlear implant users showing a large variation in outcome reproduced from Holden et al. [79].

sounds, a reduction in the level of tinnitus and support of lip-reading with a reduction in the effort required for oral communication are all worthwhile benefits from use of a CI. It should also be noted that in many cases those most satisfied with their implant are not those who receive the highest scores on standardized tests of speech understanding.

The following sections will describe how electrical stimulation of the auditory system is achieved, with the main focus being on CI systems. The factors that influence outcome, so far as they are known, will be described, along with the challenges in delivering clinical service, both today and into the future. With the future in mind the major research topics that are currently being addressed will be outlined.

2. Overview of a cochlear implant system

Figure 2 shows the various components common to all of today's clinically applied cochlear implant systems. Sound is typically collected from microphones housed on a behind the ear (BTE) sound processor. The sound is first "cleaned" to remove noise and then processed to create the stimulation patterns destined for the implanted electrode array. Except in the case of one-piece processors, a lead connects the sound processor output to a radio frequency (RF) transmitter coil located above and behind the ear. The external coil is held in place over the implant's receiver coil through a pair of magnets: one external and one within the surgically implanted device under the skin. This arrangement supports reliable communication across the skin through the use of RF based telemetry. The RF signal provides both power for the implant's electronics and the information needed to produce electrical stimulation. Hence the implant consists of: its receiver coil, a hermetic package containing electronic circuits and an electrode lead assembly connecting to the electrode array that is placed inside the cochlea (**Figure 3**). In some of today's CI systems the sound processor and RF coil are a single component held in place by the magnet but having no wire or BTE part. This provides some esthetic advantage but may fall off more easily and compromises sound collection.

Additionally today's implants have the ability to make both physical and physiological measurements, using back-telemetry, to transmit these data to the sound processor. Through the use of wireless technology, information can be relayed to and from a host of devices: smartphones, tablets, laptops, remote microphones or other listening aids. Such connectivity leaves a CI user well placed to use many consumer devices to enhance their communication and support maintenance of their implant system.

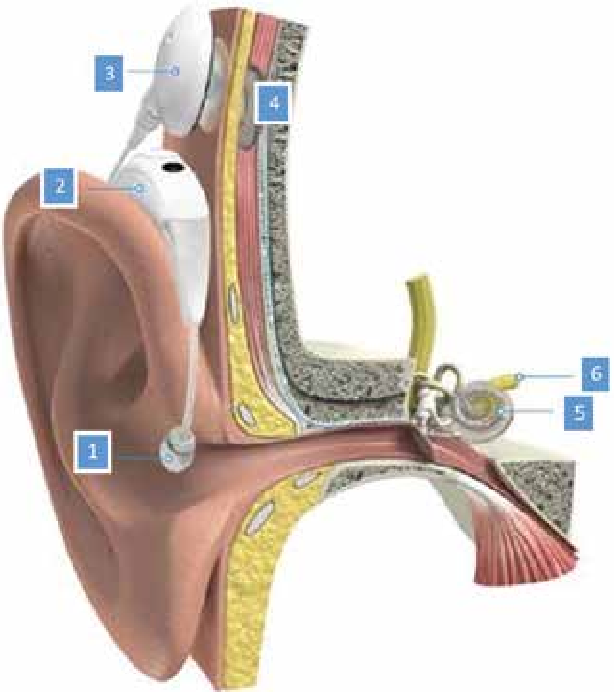

Figure 2.
The components of a behind the ear (BTE) model of cochlear implant showing (1) the T-mic placed in the external ear canal, (2) BTE sound processor, (3) radio frequency transmitting headpiece, (4) the implant body, (5) intra-cochlear electrode array and (6) the auditory nerve.

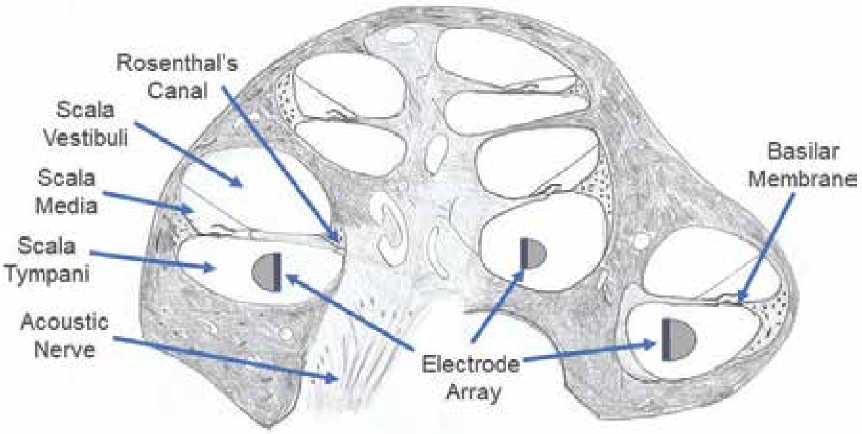

Figure 3.
A cross-section through the spirally-shaped cochlea showing the various compartments, including the scala tympani with a mid-scala located electrode array.

3. Electrical stimulation principles

In the earlier chapters of this book the auditory system has been described in some detail, including pathology that can result in the most debilitating degrees of hearing loss: severe to profound deafness. Fortunately, electrical stimulation can be delivered without external, middle or indeed even an inner ear. However, in the large majority of clinical cases the auditory nerve is intact and can be accessed via a very poorly or non-functioning cochlea: it being this damage that the CI bypasses. The operating principle of a CI is that small electrical currents are able to initiate activity on the auditory nerve that crudely mimics the activity produced by a normally functioning cochlea. Taking advantage of the cochlea's tonotopic organization, currents representing high frequency sounds are delivered to the base, while currents representing lower frequency sounds are delivered to more apical cochlear locations. This is achieved through the use of an electrode array containing multiple separate electrode contacts placed along the scala tympani (**Figure 4**). The number of intra-cochlear electrode contacts in clinical service varies by manufacturer between 12 and 24. In addition, either the shorting of adjacent contacts, or simultaneous delivery of synchronized stimulation patterns on multiple electrode contacts, seeks to increase the number of stimulation sites available [11–13].

In **Figure 4** an electrode array is shown placed in the scala tympani. Here the exposed electrode surface, from which stimulation current is delivered, faces towards the modiolar wall, behind which the auditory nerve's spiral ganglion cell bodies are located in Rosenthal's canal. The remainder of the array is composed of a soft silicone that supports the contacts and the insulated wires connecting the electrode contacts to the implant's electronics. A modiolar facing contact surface orients the electrical stimulation towards the primary stimulation targets, the spiral ganglion cell bodies. Since the perilymph fluid in the scala tympani is electrically conductive, it allows current to flow through the cochlea and achieve stimulation of neural elements. A downside is that current also tends to spread along the scala, rather than addressing only the area local to the electrode contact where we would like it to act. If peripheral processes still remain in the cochlea, extending from the cell bodies to the organ of Corti, then these may also be targets for electrical stimulation. Unfortunately it is not currently possible to accurately know the status of any individual's cochlea. It appears reasonable to assume that a great deal of the variability illustrated in **Figure 1** comes from variations in the health of those individuals' cochleae, this variation being part dependent on environment and disease and the individual's particular physiology, as well as how any one individual's immune system reacts to the insertion and presence of the electrode array itself.

The most common type of stimulation paradigm used in CIs today is monopolar stimulation (**Figure 5a**). Here the implant introduces current into the cochlea via a relatively small electrode contact. A typical surface area might be 0.2 mm^2. The density of current close to the electrode contact introduces a higher probability of activating elements close to the contact, the probability decreasing for elements

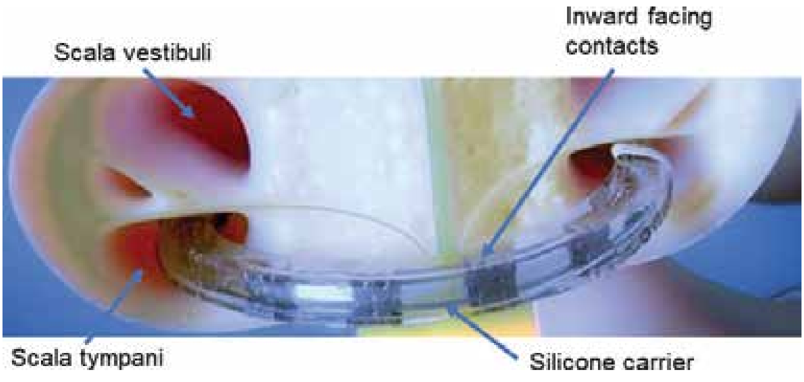

Figure 4.
View of how an electrode array will be positioned within the scala tympani of the cochlea, here with the electrode contacts facing towards the modiolar wall behind with the spiral ganglion cell bodies are located.

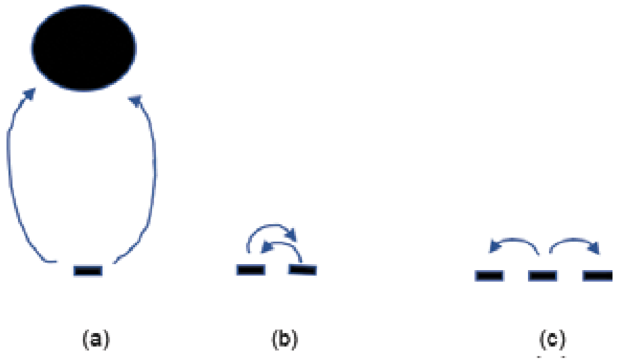

Figure 5.
Electrical stimulation configurations: (a) monopolar where current is returned to a large distant extra-cochlear return electrode, (b) bipolar where current flows between adjacent intra-cochlear electrodes and (c) tripolar where current is returned by two adjacent contacts.

at increasing distance. Monopolar stimulation current is returned to the implant using a distant extra-cochlear electrode that has at least 10 times the intra-cochlear electrode contact's surface area. This keeps the returning current density low, avoiding stimulation at this remote site. Typically there are two return electrodes on a cochlear implant, in case anything goes wrong with one of them. One is placed on the body of the implant while the other is on a separate flying lead, or on the electrode lead assembly but located outside of the cochlea. In **Figure 6** an implant can be seen with its various component parts indicated.

Alternative stimulation paradigms are sometimes used, but mainly for research. **Figure 5b** shows a bipolar stimulation paradigm. Here two intra-cochlear electrode contacts, separated by around 1 mm, operate as a pair. Stimulation is introduced by one contact and returned by the other. This in theory restricts stimulation to a small part of the cochlea, so should help with the spread of current mentioned above. However, in practice current introduced by one contact is returned to the other without activating enough neural tissue to create a loud enough hearing sensation. Hence, it is often necessary to form a bipolar pair using contacts that are not adjacent but for example are spaced 2 mm or more apart. In addition, the arrangement of contacts along the cochlea means that there will be a plane of zero field between the contacts, leading to a need to use higher currents and thus produce a wider spread in stimulation. In **Figure 5c** tripolar stimulation is shown. Now a group of three electrode contacts are used together. Stimulation is introduced via the middle contact and ideally half returned via the adjacent contact on either side. This avoids

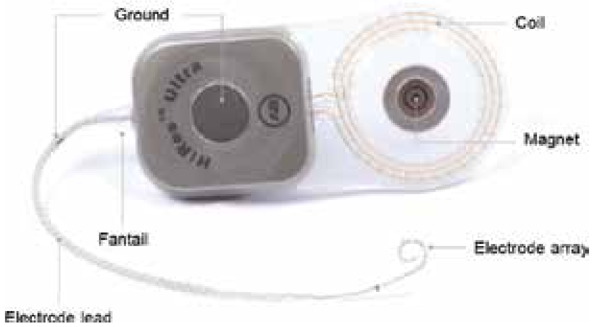

Figure 6.
A cochlear implant showing its various components. Note the two ground or return contacts, one on the case body and one on the electrode lead assembly.

the zero potential plane problem of bipolar stimulation and theoretically provides a tighter containment of stimulation. Again, in practice there is a need to recruit a given number of neurones to signal sufficient loudness and this means increasing the tripolar stimulation current on the middle contact. Eventually the current on the return electrodes will become high enough to cause stimulation, resulting in a wider spread of current than intended. In many cases it is not even possible to increase the current far enough to achieve sufficient loudness for a tripolar configuration. In such cases some of the current has to be returned to the remote extra-cochlear return electrode, a configuration referred to a partial tripolar. This even further increases the current spread. So, with these practical limitations in mind it is not difficult to see why monopolar is universally used as the default stimulation paradigm, despite the apparent large current spread that this entails.

3.1 Technical considerations

In the interests of simplicity some of the more technical aspects of electrical stimulation have been avoided in the text above, allowing focus on the broader application. In this section some of these more technical issues will be discussed. Where a reader is not interested in technical detail this section can be ignored.

Ohm's Law states that the electrical voltage difference required to drive a given current is directly proportional to the resistance through which the current has to flow. With changes in the resistance, or more generally the impedance if we consider frequency effects, a voltage source would lead to uncontrolled changes in the current being output. As described below this could lead to uncontrolled changes in loudness over time. Most of today's CI systems deliver electrical stimulation through one or more current sources. As the name implies, this circuit attempts to deliver the current requested of it regardless of how much electrical resistance is offered by the body. A current source is said to be in compliance when it is delivering the current requested of it. Given a finite amount of voltage being available within an implant, for example 8 volts, there will be a maximum impedance into which the implant's maximum output current can be delivered. With a typical stated maximum output current of 2000 μA, the maximum impedance for which this could be delivered will be 4000 Ohms (8 volts divided by 2000 μA, or $8/2 \times 10^{-3} = 4000$). For higher impedances the maximum stated output current will not be available. For lower impedances the implant will be limited to its maximum output current value, ensuring safe operation.

Since CI systems are designed to provide stimulation for essentially all waking hours, day after day over decades, their output must not damage the neural tissue that they are intended to stimulate. One obvious source of damage is the delivery of direct current (DC). If a DC current is applied to the body it will result a process called electrolysis. Here there will be a continuous transport of ions, charged atomic particles, leading to dissolution of the platinum electrode contact and destruction of the cochlear tissue: obviously a catastrophic situation. Research in animals indicates that even very small amounts of DC, 400 nA, is enough to cause tissue damage [14]. Several mechanisms are used in a CI to prevent the delivery of DC. Firstly, a balanced stimulus waveform is used, almost always a symmetrical pulse having two complementary phases (**Figure 7a**), although so long as the two phases contain an equal and opposite area they need not be symmetrical (**Figure 7b**). At the simplest interpretation, the first phase, referred to as cathodic, will depolarize neurones hence producing the electrical stimulation that we seek to achieve. The second, anodic, phase balances the stimulation resulting in no nett current being delivered to the body, thereby avoiding DC. Even with very careful design, there is likely to be some small imbalance between the two phases. To account for any in-balance,

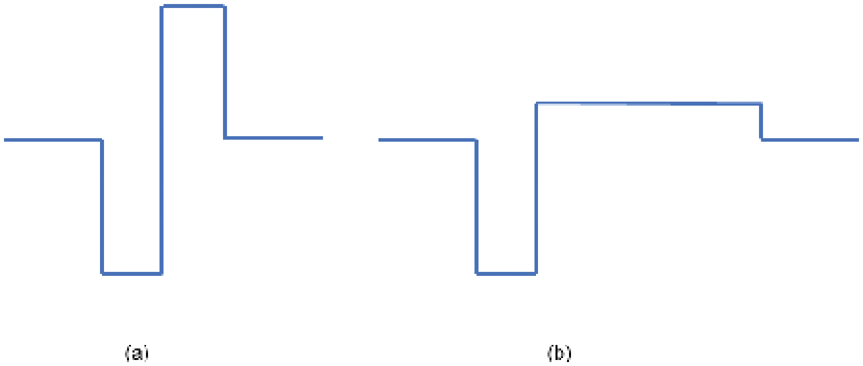

Figure 7.
Stimulation waveforms with balanced cathodic and anodic phases may have either symmetrical phases (a), or asymmetrical phases (b) where the area of each phases is identical.

following delivery of a stimulation pulse the electrode contact is connected to ground, ideally removing any residual DC. Finally, an electronic component called a capacitor is placed in series with the stimulating electrode contact. A capacitor does not allow DC to pass, so offers yet more protection in case some fault in the electronics interferes with either of the previous two protection mechanisms. Together these mechanisms appear successful in avoiding the delivery of DC. Devices do fail, particularly in the pediatric population, with reimplantation rates over tens of years being reported at 8% from the well-established Sydney clinic [15]. Typically half of these failures are medical issues and half device failures. However, in virtually all cases it is possible to re-implant the patient, with outcomes almost always being as good as those obtained when the original device was working well [16].

While we speak about stimulation current it is really the electrical charge that is at issue. Charge is simply the product of current times time and has units of Coulombs, C. Electro-chemical considerations mean that an electrode has a maximum charge injection capacity, such that a given size and material will only be able to handle a given charge limit in a reversible way, so that all of the charge injected in one phase can be removed in the second phase [17]. This is necessary to avoid the DC as discussed above. A conservative value for the maximum charge density, typically 30 $\mu C/cm^2$ [18], taking account of the electrode dimensions, is programmed into the implant's controlling software, ensuring that this limit is never exceeded. Animal experiments confirm that chronic stimulation with higher charge densities, for example 400 $\mu C/cm^2$, results in the dissolution of platinum but interestingly not to the loss of auditory neurones [19]. The loudness sensation produced by electrical stimulation is related to the amount of charge delivered in one phase of the stimulation waveform. There is also an effect of the rate at which stimulation pulses are delivered. However, for the stimulation rates used in clinical practice, typically 500–2000 pulses per second per channel (ppps/ch) the effect of rate is quite small and largely ignored.

Stimulation current flows through cochlear tissue as a result of voltage differences developed along the current's path. How these voltage differences are arranged along the length of a neurone's peripheral or medial process, or indeed across the cell body, determines which neurones depolarize, leading to action potentials being generated. The action potential may propagate to the medial synapse with the cochlear nucleus and hence initiate activity on the auditory pathway, leading to a sense of hearing being detected in the brain. Rather than stimulating a single neuron, typically hundreds, or even thousands, of neurones in a region of the cochlea will be addressed by a single electrode contact. These patterns of stimulation are interpreted as sound input by the higher levels of the auditory system, leading to the sense of electrical hearing. The next section will describe how sounds detected by the CI system's microphone will result in the generation of electrical stimulation patterns.

4. Operation of the cochlear implant system

The main cochlear implant system functions are shown schematically in **Figure 8**. Sound from the microphone is compressed by a single-channel automatic gain control (AGC) system. Compression ratios in CI systems tend to be substantially higher than those in acoustic hearing instruments: six to infinity, compared to two to three respectively. This reflects both the small electrical dynamic range of typically 10 dB [20] and the exponential like increase in loudness found for electrical hearing [21]. Both considerations require tight control of the stimulation current's amplitude to avoid discomfort. Research with different implant types shows a consistent advantage for slow-acting AGC, the benefit being a reduced compression of the information rich temporal modulations of speech [22–24], as well as a reduced co-modulation effect [25] associated with the single channel AGC.

Following AGC, the sound is broken into a number of frequency channels, this number varying between 12 and 24 channels, reflecting the number of intra-cochlear electrode contacts available in the implant model. In **Figure 8** only four channels are used to illustrate the principle. Today a Fast Fourier Transform (FFT) algorithm is often used to separate the incoming sound into discrete frequency channels. The amount of energy in each channel is then estimated by a rectification and low-pass filtering process. While the average energy is calculated over a period of perhaps 10 ms (milliseconds) or more, stimulation pulses will be delivered much more rapidly, typically once every millisecond. Hence calculations will be made that overlap in time in an attempt to follow the changes in speech energy over time. Next the acoustic energy in each channel is mapped to an electrical current amplitude that takes account of the CI user's sensitivity to electrical stimulation. The goal is to use smaller currents that barely produce a perception of electrical stimulation to represent low-intensity acoustic activity and larger currents that are perceived as loud to represent very intense acoustic events. This

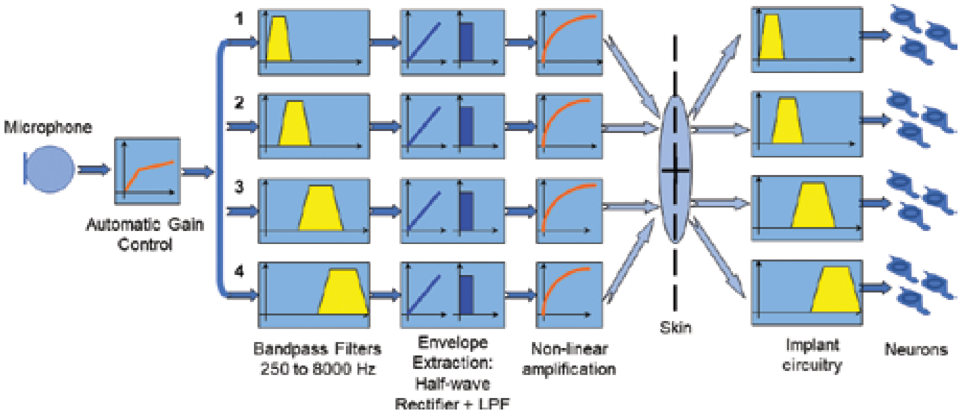

Figure 8.
A schematic of the sound processor system where sound is collected by the microphone, compressed by an automatic gain control, broken into discrete frequency channels, which have their energy assessed and mapped to the user's requirements. This information is then combined into a digital stream, transmitted by radio frequency to the implant where stimulation currents are generated.

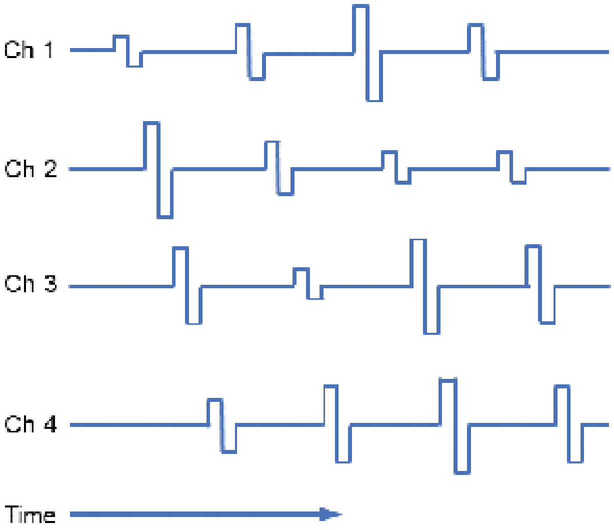

Figure 9.
An illustration of now non-simultaneous waveforms delivers information for each channel. Once a channel has been stimulated no more information may be delivered until the other channels have been updated.

needs to be managed separately for each channel, resulting in the continuous output of a stream of stimulation amplitudes for each channel. As shown schematically, these amplitudes are then combined together for transmission by the RF signal across the skin to the implant. Electronics inside the implant extract the digitally transmitted amplitudes, convert them to analogue values and then drive the implant's current source(s), resulting in stimulating currents being delivered by the intra-cochlear electrode contacts. For virtually all of today's clinical systems only one channel will be stimulated at a time. This approach avoids the channel interactions that would occur were channels presented simultaneously within the conductive scala tympani [26]. The disadvantage of this approach can be seen in **Figure 9** where a channel is only updated during its own time period, therefore, must wait until all the other channels have been updated until new information can be transmitted. Deliberately, very brief current pulses each of around 40–50 μs duration (20–25 μs/phase) are used, so that it is still possible to update each channel rapidly enough to keep up with the changes in acoustic energy over time. This often means stimulation at more than 1000 pps/ch. Such an approach generally leads to higher levels of speech understanding than where simultaneous stimulation is delivered [27, 28].

4.1 Sound coding strategies

Over the years the sound coding strategy, a software algorithm that relates audio from the sound processor microphones to the electrical patterns appearing at the electrode contacts, has changed. Initially it was believed that the damaged auditory system was not capable of transmitting much information, hence the most useful information was extracted from the speech and directly coded on sets of electrodes. For example an early feature extraction strategy F0F1F2 [29] extracted the first two formants of speech, F1 and F2, from which it is possible to estimate the vowel being articulated. Each formant range had a set of electrode contacts allocated, such that higher or lower frequencies for each formant lead to stimulation on more basal or more apical electrode contacts in that formant's electrode set. The rate at which pulses were delivered was related to the fundamental frequency (Fo) driving the vocal tract, leading to the strategies name. Such a strategy supported only very modest levels of speech understanding, around 8% correct for monosyllabic words presented in quiet [30]. The information extracted was limited to begin with and further reduced through errors generated in real life listening situations where background noise, reverberation and intensity and frequency response variations led to the algorithm making mistakes in both the extraction of formant frequencies and in the estimation of Fo.

It was eventually recognized that the brain was better at extracting information than the feature extraction algorithms and hence "whole-speech strategies" replaced feature extraction. Today's sound coding strategies simply average the energy in each channel's frequency range and generate levels of stimulation that represent this. In some cases a so-called n-or-m strategy will work out which subset (n) channels from the total (m) number available have the highest energy and then only stimulate this reduced set. Refinements to this may neglect adjacent channels on the basis that stimulating both will not add anything, so select a more distant lower amplitude electrode to transmit more information [31].

4.2 Neural population

Stimulation delivered by a CI system will result in the depolarization of neural elements, resulting in action potentials being generated that propagate to the next stage of the auditory system: the cochlear nucleus. With reference to the schematic of **Figure 8**, there is a population of spiral ganglion neurones associated with each electrode and hence each frequency channel of the CI system. As mentioned above, a channel's stimulation current will need to recruit a certain population of neurones whose firing indicates to the brain the amount of activity in a particular frequency range. Ideally, there will be a sufficient local neural population such that progressively increasing stimulation current initiates an appropriate number of action potentials, so that the brain correctly perceives the amount of acoustic activity in the channel's frequency range.

Unfortunately a discrete neural population for each channel as shown in **Figure 10a** is not always available. In **Figure 10b** only a reduced neural population is available for each channel. Hence, when there is a lot of activity in one channel, requiring recruitment of a full population of neurones, these are not available locally. It is still possible to increase the stimulation current, spread the electrical field further away from the electrode and depolarize neurones that should really be associated with another channel. While this will satisfy the perception of loudness, it generates channel interaction so that we are no longer able to deliver frequency specific information to a discrete part of the cochlea. The perception will be of a blurred or fuzzy sound, particularly a problem when trying to listen to speech in the presence of competing noise.

An alternative situation is shown in **Figure 10c** where most electrodes have a sufficient local neural population but one electrode is located in a so-called dead region [32]. When electrode 2 is stimulated it can only recruit neurones from the population belonging to electrodes either side of it, delivering information about channel 2's frequency region to other parts of the cochlea, spreading stimulation widely and interfering with the otherwise discrete frequency information being delivered by the neighboring channels.

Unfortunately, it is not currently possible to determine what the number and distribution of neural elements is for any individual. The literature is not always helpful in this area. As shown in **Figure 1**, there is great variability in outcome for the identification of monosyllabic words. Since this task involves little top-down processing, much of the variability in outcome must come from the electro-neural interface. Beyond speech understanding, examining the ability to discriminate adjacent electrodes, or intra-electrode stimulation sites [33], also showed both

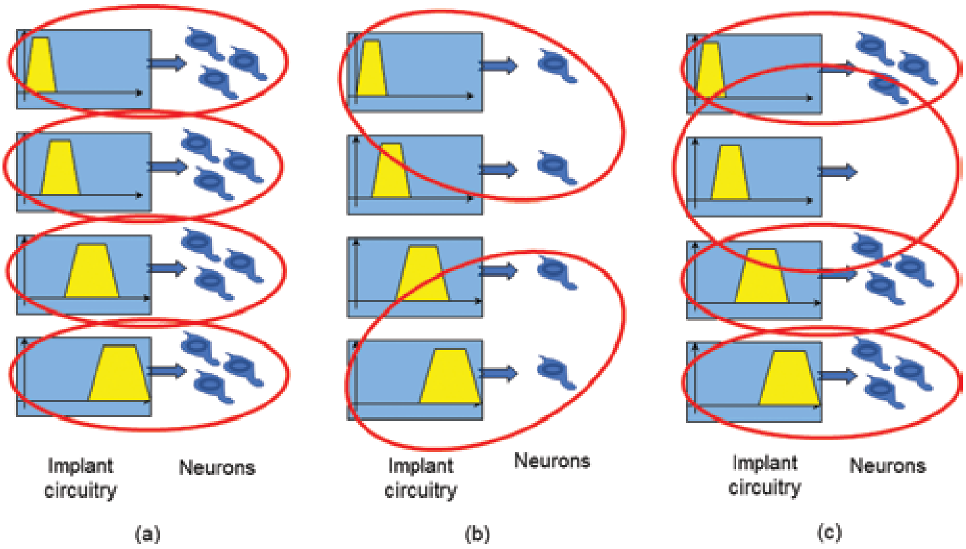

Figure 10.
A schematic representation of three different neural populations: (a) a full population exists for each channel, (b) a depleted population results in channel interaction and (c) a dead region requires recruitment from the populations rightly belonging to adjacent channels.

great variability between subjects and across the electrode array of individual subjects. This task, having no confound with cognitive processes related to speech understanding, further confirms the presence of peripheral variability and its likely contribution to variations in outcome.

It is unclear to what degree a loss of spiral ganglion cells (SGC) in humans will follow, even after years of severe to profound deafness. Histological studies of humans who had used a cochlear implant sometimes show a reasonable correlation between CNC word score and SGC count: for example R = 0.62 [34], R = 0.9 [35] but for a small group of only 6. However, the variability is such that the same SGC count can show variations of between 30% and 75% for CNC words, or the same CNC word score can be associated with 3000 or 18,000 SGCs. Examining the threshold current for detecting electrical stimulation in a group of 130 lateral wall electrode array users [36] showed significant differences between four groups: the increase in group mean threshold being associated with a reduction in monosyllabic word score. This works suggests that a higher SGC population (lower electrical threshold) is associated with better speech understanding.

The literature listed above indicates that there is a relationship between the number of spiral ganglion cells and the ability to identify monosyllabic words when using a cochlear implant. Contributions to speech understanding may also come from a large number of additional factors, some of which include: the distribution of SGCs, angular insertion of the electrode array, distance of the electrode contacts from the modiolar wall, presence or absence of peripheral processes, fibrous sheath formation and intrusion of new bone into the cochlea. How well a given implant user has had the parameters of their sound processor set, commonly referred to as their program, is another variable that we will examine next.

4.3 Programming the cochlear implant

As has been explained above, the small electrical dynamic range available to a CI user makes it necessary to carefully adjust the stimulation parameters to suit the requirements of each individual recipient. The most important adjustment is the amount of stimulation that will be delivered in response to acoustic activity. This must be done for each of the CI's separate channels. Each channel has two primary parameters that control its output. One will be typically called a most comfortable level, shortened to either M-Level or C-Level. The other is a threshold control, referred to as T-Level. The main CI manufacturers use these parameters slightly differently but to a good approximation T-Level sets the minimum stimulation level that the implant will deliver and M-Level will set the maximum stimulation level that can be delivered for an individual recipient. The sound processor will then arrange for the amount of acoustic range that it handles, somewhere between 40 and 80 dB depending on the user's setting and implant model, to be mapped to stimulation levels between T- and M-Level. In combination with the AGC of the system this will give the CI user access to their acoustic environment such that hearing levels of between 20 and 30 dB HL are achieved across the frequency range 250–8000 Hz. The combination of AGC and M-Level ensures that even high intensity sounds of 100 dB SPL do not produce uncomfortably loud sensations. Unlike acoustic hearing, it is generally possible to provide CI users with access to the full range of frequencies that are most important for speech understanding.

Which channels are activated is another important adjustment to make. Most audiologists are reluctant to deactivate channels, although sometimes a reduced set of channels can give a better outcome. In some cases an electrode array is not fully inserted into the cochlea, perhaps due to the cochlea being too small, or there being fibrosis tissue, or bone, that prevents a full insertion being obtained. Alternatively, electrode arrays can sometimes extrude from the cochlea [37, 38], either shortly after implantation or months to years later. In all these cases the more basal electrode contacts will need to be deactivated. Deleting electrodes from a program will lead to the frequency range being remapped across the remaining electrode contacts. There will be a coarser representation of frequency since fewer channels are now available. However, removing electrodes that are not inside the cochlea will produce a better outcome than simply leaving these electrodes active.

Beyond setting T- and M-Levels and defining an appropriate set of electrode contacts, there is sometimes adjustment made to the acoustic dynamic range mapped by the sound processor. This effectively controls the compression of acoustic sounds into the electrical dynamic range. It might seem logical to use as large an acoustic or input dynamic range (IDR) as possible, since this will maximize the range of sounds available to a CI user. However, it is the discrimination of different levels of sound in each channel that carries information. An excessively large IDR may squeeze these amplitude cues, reducing the ability of an implant recipient to understand speech. There are many parameters that can be adjusted in a CI system. However, it is common for the majority to remain at their default values. This may be through an inability to obtain user feedback, for example in young children, lack of time or knowledge on the part of the clinician, or a recommendation from the CI manufacturer.

How appropriate values are found for the T- and M-levels depends very much on the individual CI user. For a post-lingually deafened adult it is reasonably straightforward to find these. By presenting 200 ms bursts of stimulation and using a standard bracketing approach, the smallest detectable amount of stimulation for each channel can be found and this value set as the T-level. Similarly, progressively increasing the stimulation will allow an M-level to be found, the CI user often pointing to different categories on a loudness chart as the various levels of stimulation are presented. These measures can be made for each individual channel, channels can be programmed in groups of four, or only five or six channels across the electrode array measured with intermediate channels set to interpolated values.

For babies or young children and even for some adults, objective measures are often used to help set program levels. The most common measure used is the eCAP, the electrically elicited compound action potential [39]. The ability to record eCAPs is built into the fitting systems for all of today's major CI systems. Here masker-probe or alternate-subtraction techniques [40, 41] are used to reduce the large stimulus artifact. The amplitude of the remaining physiological signal, arising from synchronized activity on the auditory nerve, is then graphed against the stimulation level. A regression line extrapolates to intersect the stimulation axis which would correspond to a zero amplitude of eCAP. The stimulation value for which this occurs is then used as a guide for setting programming levels. Avoiding stimulus artifact and allowing sufficient neural synchronization, means that much lower stimulation rates are used when measuring eCAPs than for actual everyday stimulation. The means that the absolute eCAP values can fall at various parts of an individual's electrical range. Fortunately, it is the profile of values across the electrode array that it is important to determine. Once this is estimated a global change in level can be made to obtain appropriate loudness. In many cases the T-levels are set to 10% of the M-level since this is almost certainly not going to leave them set too high. Typically T-levels are measured at something like 25% of M-level [42]. When they can be measured and hence individually set, T-levels will tend to improve access to low intensity sounds. Often in clinical practice T-levels are set at a percentage of M-level even where they could be individually set: the additional benefit not being considered worth the additional effort needed for measurement.

Other objective measures are used to assist with programming, although less often due to these requiring additional equipment to be used in collaboration with the CI fitting system. There is a reasonable correlation between an electrically elicited stapedius reflex threshold (eSRT) and M-level [43]. Unlike eCAPs here the same stimulation rate can be used to measure eSRT as will be used in the everyday program. This simplifies the setting of levels and is partly behind why there is such a good correlation with M-level. Less commonly the electrically elicited auditory brainstem response (eABR) is used [44]. Again, eABR will require a lower stimulation rate to be used, so that the characteristic waveforms can be seen in up to 5 or 6 ms following stimulation. This tends to produce an extrapolated threshold for eABR quite high in an individual's electrical dynamic range. As with the eCAP and eSRT measures, it is the relative levels across channels that are important, the profile then being globally adjusted to determine the M-levels that will be used in the program.

Less frequently, some statistically based approaches are used for programming. Simple so-called "flat maps" are used where the T- and M-level is the same on each channel. These are justified by the spread of monopolar stimulation recruiting neurones form a larger section of the cochlea than associated with an individual electrode contact thus tending to produce a spatial averaging. Other approaches might use a template based on the statistical average of levels previously measured for earlier CI recipients. Approaches such as FOX [45] extend this technique, recommending a sequence of programs with progressively increasing levels that are used from the very beginning. For many CI users these techniques can work quite well, although numbers of outliers will require individually tailored programs to realize their potential outcome with comfortable stimulation and reasonable access to their acoustic environment.

Plasticity in the auditory system means that over time the M-levels will usually increase. The longer term M-levels might be typically double those that can be tolerated during the initial fitting. After the first 2 months of device use, neither T- or M-levels tend to change significantly over time [46]. Change in levels is highly individual requiring the initial program levels to be revised numbers of times during the first few months of implant use. Where a second (or third) fitting session is planned within around 2 weeks of the first fitting, most of the change can already be accommodated. Looking across large numbers of adult CI users, program levels will be stable by between 3 and 9 months following first fitting. Individual practice can result in pediatric levels being more slowly increased, leading to 6–12 months being needed to see stable levels.

5. Different patient groups

Cochlear implants were originally designed to help those suffering bilateral profound deafness who could not benefit from acoustic hearing aids. Traditionally candidacy would have required a loss of at least 90 dB HL across all of the audiometric frequencies from 125 to 8000 Hz. Over the past 30 years we have seen, improvements in outcome (speech understanding) through better sound coding strategies and electrode arrays, improvements in esthetics as the external equipment has moved from body worn to behind the ear or single piece processors, improvements in surgery with a skin to skin operating times of well under 1 hour, as well as much smaller incisions not requiring hair shaving and at least in some cases, the preservation of residual hearing. These developments have meant that a CI can now be considered for much more than the 0.2% of the population who suffer profound bilateral sensorineural deafness [47].

It is becoming more common for ears with useful low-frequency residual hearing to receive a CI. Candidacy can now include those with severe to profound levels of hearing loss above 1000–2000 Hz, but normal to moderate hearing loss for lower frequencies [48–50]: a group sometimes referred to as suffering partial deafness. Where the residual hearing can be preserved to within 10–20 dB of the pre-operative levels, many of these recipients use a combination of electrical and acoustic stimulation (EAS) in the same ear. Most CI manufacturers now make EAS processors so that a single instrument supports both modalities, offering comfort, convenience and allowing an EAS fitting to be made using a single piece of software.

Where there is some asymmetry in hearing, recent practice has seen only the poorer ear being implanted, while an acoustic hearing instrument is fitted to the contralateral ear. This is often referred to as bimodal hearing. Dedicated hearing instruments (HI) have been developed that match the compression characteristics and sound cleaning operations between the CI and HI, as well as offering wireless sharing of microphone and control signals. Such systems offer convenience for the user and can combine the natural acoustic low-frequency sound in the HI ear, with the high frequency information supporting speech understanding in the implanted ear.

Bilateral CI provision is now the standard of care in many healthcare systems, at least for children. Receiving two implants, either simultaneously or within a few months of each other, provides the best chance for the brain to have both sides work together. Redundancy, the countering of head shadow and a fuller sense of hearing are all advantages of bilateral implantation. It tends to be only considerations of cost that prevent bilateral CIs being offered universally to all those who could benefit from them. Again with wireless technology developing rapidly, the use of algorithms that combine microphones between the two CI sound processors can offer large improvements for listening in noise when beam formers are used to attenuate noise coming from directions other than directly ahead, particularly useful then the CI user is in a one-to-one conversation in a noisy location.

Where a second CI is not available and there is no aidable hearing on the contralateral ear a CROS, or strictly bi-CROS device can be used. Wireless CROS devices are available that essentially have their microphone pick up sound from the non-implanted side and wirelessly route it to the CI processor on the other side where it is mixed with the CI processor's microphone signal. This approach can reduce head shadow, although with stimulation only being delivered to one ear there is little ability to use the CROS device for localization. With the combination of HI and CI companies, for example Phonak and Advanced Bionics within the Sonova company, the migration of HI technology such as the ear-to-ear wireless technology has begun and will likely be more common in future.

In the German healthcare system, a CI is now available to those who suffer from single-sided deafness (SSD). Typically there may be also be some hearing loss on the better hearing side, making this a highly asymmetrical loss rather than a pure SSD. Those suffering with SSD would usually explore a CROS device and a bone conduction hearing aid before considering a CI. In the end around one third of SSD cases seen will elect to get used to hearing with only one ear, one third will use a bone conduction device and one third will receive a CI [51].

Tinnitus is another consideration that can influence treatment options, for SSD and beyond. Where the SSD is accompanied by intractable levels of tinnitus, a CI may provide relief [52]. The restoration of some input to the deafened ear can allow the tinnitus to either effectively disappear or at least be substantially reduced. In some cases, SSD in particular, the relief from tinnitus is found to be of much greater benefit than any hearing sensation arising from the implanted ear. The large majority of CI recipients report reduced amounts of tinnitus although in very rare cases tinnitus can be worsened through implantation.

5.1 Alternative stimulation sites

The cochlea is an attractive site for electrical stimulation, given that it presents tonotopic access to auditory nerve fibers with reasonably straightforward surgical access. However, where the cochlea has not formed properly or at all, due to some extreme malformation, a properly formed cochlea has been filled with bone or tissue, for example following bacterial meningitis, preventing all but minimal surgical access, or the auditory nerve is not available, either through malformation or following trauma, stimulation of the auditory system via the cochlea is not possible. In such cases alternative sites of stimulation may be used.

Auditory brainstem implants (ABIs) bypass the auditory nerve, targeting the next station of the auditory pathway: the cochlear nucleus located in the brainstem [53]. There is a tonotopic structure within the cochlear nucleus, although it is organized in the dimension of depth, so is not easy to access. Attempts to use a penetrating electrode array with a number of discrete needles has not been able to make better use of this tonotopic organization than a pad of flat electrodes placed on the surface of the cochlear nucleus [54]. Programming of an ABI device tends to be more difficult than that of a CI. Bone surrounding the cochlea usually keeps the CI's stimulation contained to auditory fibers. Only occasionally non-auditory stimulation of, for example, the facial nerve can be seen in muscle twitching around the mouth or eye. This is generally programmed around by deactivating electrode contacts or reducing stimulation levels. However, in the brain stem, functions such as respiration can be adversely effected by an ABI device. This calls for much more care when programming, leading to ABI devices being offered only by specialist centers. The surgery required to place an ABI is more invasive than that required for a CI, for example, requiring lifting of the cerebellum to gain sufficient surgical access. With ABIs being placed following removal of tumors there can be some distortion of brain structures. Some surgeons prefer to remove what are often sizable tumors, allow the brainstem and brain structures to settle into place again and then perform a second surgery during which the ABI is put into place. This two-stage approach is believed to provide less chance of the ABI's electrode pad moving out of position, risking substantial non-auditory stimulation. Outcomes with ABI devices are generally substantially poorer than with CIs. In some series there is essentially no open-set speech understanding possible [55], while in others the speech understanding is limited, with only occasional high levels of speech understanding [56]. The reasons for poor performance with ABIs are not fully understood. Beyond potential movements of the electrode pad, there are specialized auditory functions being carried out in the cochlear nucleus, meaning that simply assuming raw tonotopic stimulation patterns may not be sufficient. Additionally, those receiving ABI devices may have many other issues beyond deafness and these could also explain some of the difference in outcome.

Stimulation of even higher structures in the auditory system has been attempted through an auditory mid-brain implant (AMBI), where the electrode array is inserted into the inferior colliculus. Currently this is restricted to a pure research device [57]. Within the inferior colliculus it is possible to access a tonotopic organization, using a shortened version of a traditional CI electrode array, 10 mm long as opposed to 20–30 mm for most CI arrays. However, when accessing the auditory system at an even higher level than with an ABI device, the amount of pre-processing that should have already been done leaves a crude CI type coding strategy only able to support very limited outcomes. Already at this level higher stimulation rates are inappropriate, leaving limited sound coding strategy options [58].

While placement of an electrode array in the scala tympani, or where necessary in the scala vestibuli, leaves the electrode contacts quite close to their target neurones they are still some 1–3 mm away. This separation prevents discrete stimulation of local neural populations as discussed above. It has been proposed that an electrode array could be inserted directly into the auditory nerve, or failing this inside the modiolus. Promising results have been shown from acute experiments in cats [59]. Recording form electrodes placed in the inferior colliculus indicates that intra-neural simulation is more localized than stimulation using an electrode array placed in the scala tympani. There are considerable challenges to overcome before intra-neural stimulation could be considered for humans. The surgical access is not straightforward, risking losing all residual hearing. The human auditory nerve being in the order of 1 mm diameter would require very small stimulating contacts. While the stimulation currents required would be smaller than those needed for a traditional CI, charge density considerations require careful consideration. Also, how well an electrode array can be placed and be tolerated in the auditory nerve, without the destruction of auditory fibers or the formation of granulation tissue will need to be carefully studied. Finally, the structure of the auditory nerve is complex, with axons from different parts of the cochlea rolling into the tubular nerve, making tonotopic targeting an additional challenge.

6. Current research

With the CI field involving a wide range of professionals including, surgeons, nurses, audiologists, engineers, physicists, speech and language therapists, teachers of the deaf, hearing therapists, rehabilitation specialists, psychologists and health economists, research related to the field can cover a very wide range. Here some of the key topics that are most closely connected to extending current practice will be reviewed.

It is clear that when less trauma is inflicted during surgery that outcomes are better [60], this being the case whether or not there is residual hearing at risk [61, 62]. Hearing preservation is thus a key topic for surgeons. The design of less traumatic electrode arrays [48–50, 63] and the development of less traumatic surgical techniques [64], including the use of robot assisted insertion [65], are factors that can lead to reduced trauma. Providing real time feedback to the surgeon during electrode array insertion is a hotly researched area. Electrocochleography (ECochG), where acoustic stimulation of the ear produces a cochlear microphonic signal [66–68] that the surgeon can use to gauge proximity to structures, such as the basilar membrane, appears promising with clinical systems due to launch in 2019.

The use of drugs to reduce the body's reaction to implantation is also an area with some connection to minimizing trauma. Steroids such as Dexamethasone or antimitotic drugs [69] have been applied to suppress a fibrotic reaction during and immediately after surgery. Some longer term benefits have been shown but mainly in lower electrode contact impedances rather than significant outcome advantage [70]. Longer term deployment of steroids and other drugs has been proposed for some time [71] but has not yet seen clinical practice. Drugs such as neurotrophins [72] have been proposed to enhance the spiral ganglion but carry considerable risk of uncontrolled sprouting of new fibers that may not lead to any improvement in electrical stimulation [73]. Likewise the regrowth of hair cells or other cochlear structures [74] is an extremely challenging problem, although simply reconnecting peripheral processes that have been damaged while leaving intact hair cells [75] may be more manageable in the foreseeable future. While not a drug, near infrared light has been shown to promote tissue healing [76], helping reduce the extent of hearing loss following cochlear stress [77] and has been proposed as an approach that might also enhance the cochlea's ability to survive the traumas of electrode array insertion.

Deployment, development and assessment of sound coding strategies continues with a variety of goals. Optimizing compression parameters to maximize speech understanding [78], reviewing the effect of various parameters as well as individualized fitting approaches [79, 80] all promise improvements for strategies that are already available but could be fitted better. Likewise, tools to guide clinicians in fitting for performance, rather than simply for comfort [45] should also lead to substantial improvement in outcomes. Seeking improvement via limiting current spread, through tripolar stimulation [81, 82], phased array [83] or manipulation of field interactions [84] have so far not shown a general improvement, although benefits for some recipients have been demonstrated. New sound coding strategies that attempt to improve temporal information such as FSP [85], or reduce masking effects such as MP3000 [31] have been developed and introduced into clinical practice. There may be more benefit in reducing battery power from the likes of MP3000 or Optima than in any improvement in outcome. Connected to improving outcome, research into the listening effort required to understand speech has been increasing [86, 87].

The very nature of the speech tests used to evaluate CI systems is an active area of research. Standardization across languages has involved the use of matrix tests [88]. These tests involve a fixed syntax and a closed set of keywords so can be self-administered and used essentially indefinitely. However, the sentences are not fully representative of natural sentences. Avoiding fixed presentation levels, something necessary to evaluate the function of AGC has led to development of roving presentation level tests [89–91]. These tests provide insights into where a particular subject may have problems, so could be useful in supporting programming. As with the STARR tests, multiple speakers have been included in tests such as the AzBio [92] and taken much further with the coordinate response matrix test that can run on a multiple loudspeaker array and hence mimic a more realistic test environment [93].

Improved outcomes have been shown thorough use of the recently developed wireless technology supporting integrated bimodal [94, 95] and CROS [96]. EAS has also been shown to produce substantial improvements in outcome [97] although the test methodology used involves a questionable comparison of conditions.

With increasing numbers of CI recipients needing management by financially constrained health care systems, methods of improving patient management are being developed. These include the use of consumer devices, smart phones and tablets, to run Apps that can evaluate a CI user's speech understanding and qualitative condition as well as remotely analyzing the status of their implant and sound processor hardware [98] and delivering rehabilitation material usable by adult CI recipients and the families of young children. This whole area will necessarily see much development in the coming years.

The pressures on CI manufacturers to reduce size and improve comfort and ease of use continue, with cosmetic considerations playing a large role in the choice of which CI systems are selected. Reliability of both the implanted and external parts of the CI system needs to be continuously improved, underpinning consistent device use and reducing costs inherent in managing failures. Further pressures on cost are also critical to address so that the enormous unmet need for cochlear implants in developing countries can also be met.

Glossary of technical terms

Alternate-subtraction	a method of removing stimulus artifact that relies on delivering alternate phase stimulation pulses that are added to ideally cancel artifact and reinforce response
Anodic	refers to the positive going phase of a stimulation pulse that is generally assumed to provide change balance, hence avoiding the delivery of potentially harmful direct current
Automatic Gain Control (AGC)	a circuit that compresses the large acoustic dynamic range into a range that is more manageable for the restricted electrical dynamic range
Capacitor	an electrical component that does not allow direct current to pass so offers an additional level of protection to the body in the case of a fault occurring in an implant
Cathodic	refers to the negative going phase of a stimulation pulse that is generally assumed to depolarize neurones hence leading to electrical stimulation
Charge	an electrical measure formed by the product (strictly the integral) of current and time with units of Coulombs. The charge delivered determines the loudness perceived
Cochlear Implant (CI)	a surgically implantable prosthesis that bypasses a damaged cochlea providing hearing through direct electrical stimulation of the auditory nerve
Current source	an electrical circuit that delivers a programmed current, varying the amount of voltage necessary to achieve this depending on the electrical resistance offered by the body
Direct current	an electrical current that flows in one direction only that can be harmful to the body
Electrocochleography (ECochG)	the measurement of electrical signals produced by the cochlea in response to acoustic stimulation that indicates cochlear health

Electrolysis	an electro-chemical process whereby charged particles may be drawn towards an electrode with opposite charge, leading to electrode contact dissolution or tissue damage
Fast Fourier Transform (FFT)	an efficient software algorithm that evaluates the amount of sound energy at regularly spaced frequencies, so allows the channels to be calculated
Impedance	a measure of resistance to the flow of electrical current that takes account of frequency so is more general than resistance which strictly only applies to direct current
Masker-Probe	a method of stimulus artifact removal relying on introducing refraction to identify artifact and then using multiple stages of subtraction to isolate the neural response
Perilymph	a fluid contained within much of the cochlea that is electrically conductive
Resistance	a measure of how much the body will resist the flow of electrical current, strictly only considering direct current, impedance being the a more general factor
Spiral Ganglion Cell	the name given to hearing nerve neurones having a cell body in the cochlea’s modiolus, axon in the auditory nerve and dendrites innervating the inner hair cells
Voltage source	an electrical circuit that delivers a programmed voltage but that will allow current to vary depending on the resistance (impedance) offered by the body

Author details

Patrick J. Boyle
Advanced Bionics GmbH, European Research Center, Hannover, Germany

*Address all correspondence to: patrick.boyle@advancedbionics.com

References

[1] Volta A. On the electricity excited by the mere contact of conducting substances of different kinds. Proceedings of the Royal Society. 1800;**1**:27-29

[2] Djourno A. Eyries, Prosthese auditive par excitation electrique a distance du nerf sensoriel a' l'aide d'un bobinage inclus a demeure. La Presse Médicale. 1957;**65**:1417

[3] Simmons FB. Electrical stimulation of the auditory nerve in man. Archives of Otolaryngology. 1966;**84**(1):2-54

[4] Michelson RP. Electrical stimulation of the human cochlea. A preliminary report. Archives of Otolaryngology. 1971;**93**(3):317-323

[5] Chouard CH, MacLeod P. Implantation of multiple intracochlear electrodes for rehabilitation of total deafness: Preliminary report. Laryngoscope. 1976;**86**(11):1743-1751

[6] House WF. Goals of the cochlear implant. Laryngoscope. 1974;**84**(11):1883-1887

[7] Graham JM, Hazell JW. Electrical stimulation of the human cochlea using a transtympanic electrode. British Journal of Audiology. 1977;**11**(2):59-62

[8] Hochmair ES, Hochmair-Desoyer IJ, Burian K. Experience with implanted auditory nerve stimulator. Transactions—American Society for Artificial Internal Organs. 1979;**25**:357-361

[9] Clark GM, Tong YC, Martin LF. A multiple-channel cochlear implant: An evaluation using open-set CID sentences. Laryngoscope. 1981;**91**(4):628-634

[10] Shannon RV. Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. Hearing Research. 1983;**11**:157-189

[11] Frijns J, Kalkman R, Vanpoucke F, Bongers J, Briaire J. Simultaneous and non-simultaneous dual electrode stimulation in cochlear implants: Evidence for two neural response modalities. Acta Oto-Laryngologica. 2009;**129**(4):433-439

[12] Firszt JB, Koch DB, Downing M, Litvak L. Current steering creates additional pitch percepts in adult cochlear implant recipients. Otology & Neurotology. 2007;**28**(5):629-636

[13] McDermott HJ, McKay CM. Pitch ranking with nonsimultaneous dual-electrode electrical stimulation of the cochlea. The Journal of the Acoustical Society of America. 1994;**96**:155-162

[14] Shepherd RK, Linahan N, Xu J, Clark GM, Araki S. Chronic electrical stimulation of the auditory nerve using non-charge-balanced stimuli. Acta Oto-Laryngologica. 1999;**119**(6):674-684

[15] Wang JT, Wang AY, Psarros C, Da Cruz M. Rates of revision and device failure in cochlear implant surgery: A 30-year experience. Laryngoscope. 2014;**124**(10):2393-2399

[16] van der Marel KS, Briaire JJ, Verbist BM, Joemai RM, Boermans PP, Peek FA, et al. Cochlear reimplantation with same device: Surgical and audiologic results. Laryngoscope. 2011;**121**(7):1517-1524

[17] Leung RT, Shivdasani MN, Nayagam DA, Shepherd RK. In vivo and in vitro comparison of the charge injection capacity of platinum macroelectrodes. IEEE Transactions on Biomedical Engineering. 2015;**62**(3):849-857

[18] Stover T, Lenarz T. Biomaterials in cochlear implants. GMS Current Topics in Otorhinolaryngology—Head and Neck Surgery. 2009;**8**(10):10

[19] Shepherd RK, Carter P, Enke YL, Wise AK, Fallon JB. Chronic intracochlear electrical stimulation at high charge densities results in platinum dissolution but not neural loss or functional changes in vivo. Journal of Neural Engineering. 2018;**5**(10):1741-2552

[20] Zeng F-G. Compression and cochlear implants. In: Bacon SP, Popper AN, Fay RR, editors. Compression: From Cochlea to Cochlear Implants. New York: Springer; 2004. pp. 184-220

[21] Chatterjee M, Fu Q J, Shannon RV. Effects of phase duration and electrode separation on loudness growth in cochlear implant listeners. The Journal of the Acoustical Society of America. 2000;**107**(3):1637-1644

[22] Stöbich B, Zierhofer CM, Hochmair ES. Influence of automatic gain control parameter settings on speech understanding of cochlear implant users employing the continuous interleaved sampling strategy. Ear and Hearing. 1999;**20**:104-116

[23] Boyle P, Buchner A, Stone M, Lenarz T, Moore B. Comparison of dual-time-constant and fast-acting automatic gain control (AGC) systems in cochlear implants. International Journal of Audiology. 2009;**48**(4):211-221

[24] Khing PP, Swanson BA, Ambikairajah E. The effect of automatic gain control structure and release time on cochlear implant speech intelligibility. PLoS One. 2013;**8**(11):e82263. pp. 1-11

[25] Stone MA, Moore BCJ. Side effects of fast-acting dynamic range compression that affect intelligibility in a competing speech task. The Journal of the Acoustical Society of America. 2004;**116**:2311-2323

[26] Wilson BS, Finley CC, Lawson DT, Wolford RD, Eddington DK, Rabinowitz WM. Better speech recognition with cochlear implants. Nature. 1991;**352**:236-238

[27] Wilson BS, Finley CC, Lawson DT, Wolford RD, Zerbi M. Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. Journal of Rehabilitation Research and Development. 1993;**30**(1):110-116

[28] Buechner A, Frohne-Buechner C, Stoever T, Gaertner L, Battmer R, Lenarz T. Comparison of a paired or sequential stimulation paradigm with advanced bionics' high-resolution mode. Otology & Neurotology. 2005;**26**(5):941-947

[29] Patrick JF, Busby PA, Gibson PJ. The development of the nucleus(R) Freedom^{TM} cochlear implant system. Trends in Amplification. 2006;**10**(4):175-200

[30] Tye-Murray N, Lowder M, Tyler RS. Comparison of the F0F2 and F0F1F2 processing strategies for the Cochlear Corporation cochlear implant. Ear and Hearing. 1990;**11**(3):195-200

[31] Buchner A, Nogueira W, Edler B, Battmer RD, Lenarz T. Results from a psychoacoustic model-based strategy for the nucleus-24 and freedom cochlear implants. Otology & Neurotology. 2008;**29**(2):189-192

[32] Moore BCJ. Psychoacoustics of normal and impaired hearing. British Medical Bulletin. 2002;**63**:121-134

[33] Koch DB, Downing M, Osberger MJ, Litvak L. Using current steering to increase spectral resolution in CII and HiRes 90K users. Ear and Hearing. 2007;**28**(2 Suppl):38S-41S

[34] Kamakura T, Nadol JB Jr. Correlation between word recognition score and intracochlear new bone and fibrous tissue after cochlear implantation in the human. Hearing Research. 2016;**339**:132-141

[35] Seyyedi M, Viana LM, Nadol JB Jr. Within-subject comparison of word recognition and spiral ganglion cell count in bilateral cochlear implant recipients. Otology & Neurotology. 2014;**35**(8):1446-1450

[36] van der Beek FB, Briaire JJ, van der Marel KS, Verbist BM, Frijns JH. Intracochlear position of Cochlear implants determined using CT scanning versus fitting levels: Higher threshold levels at basal turn. Audiology & Neuro-Otology. 2016;**21**(1):54-67

[37] Dietz A, Wennstrom M, Lehtimaki A, Lopponen H, Valtonen H. Electrode migration after cochlear implant surgery: More common than expected? European Archives of Oto-Rhino-Laryngology. 2015;**12**:12

[38] van der Marel K, Verbist B, Briaire J, Joemai R, Frijns J. Electrode migration in cochlear implant patients: Not an exception. Audiology & Neuro-Otology. 2012;**17**(5):275-281

[39] Abbas PJ, Brown CJ. Electrically evoked auditory brainstem response: Growth of response with current level. Hearing Research. 1991;**51**(1):123-137

[40] Cohen LT, Richardson LM, Saunders E, Cowan RS. Spatial spread of neural excitation in cochlear implant recipients: comparison of improved ECAP method and psychophysical forward masking. Hearing Research. 2003;**179**(1-2):72-87

[41] Eisen MD, Franck KH. Electrically evoked compound action potential amplitude growth functions and HiResolution programming levels in pediatric CII implant subjects. Ear and Hearing. 2004;**25**(6):528-538

[42] Van der Beek FB, Frijns J. Population-based prediction of fitting levels for individual cochlear implant recipients. Audiology and Neurotology. 2014;**19**(20):1-16

[43] Hodges AV, Balkany TJ, Ruth RA, Lambert PR, Dolan-Ash S, Schloffman JJ. Electrical middle ear muscle reflex: Use in Cochlear implant programming. Otolaryngology—Head and Neck Surgery. 1997;**117**(3):255-261

[44] Allum JH, Shallop JK, Hotz M, Pfaltz CR. Characteristics of electrically evoked 'auditory' brainstem responses elicited with the nucleus 22-electrode intracochlear implant. Scandinavian Audiology. 1990;**19**(4):263-267

[45] Govaerts P, Vaerenberg B, De Ceulaer G, Daemers K, De Beukelaer C, Schauwers K. Development of a software tool using deterministic logic for the optimization of cochlear implant processor programming. Otology & Neurotology. 2010;**31**(6):908-918

[46] Gajadeera EA, Galvin KL, Dowell RC, Busby PA. The change in electrical stimulation levels during 24 months Postimplantation for a large cohort of adults using the nucleus(R) Cochlear implant. Ear and Hearing. 2017;**38**(3):357-367

[47] Davis A. Hearing in Adults. London: Whurr; 1995

[48] Skarzynski H, Lorens A, Matusiak M, Porowski M, Skarzynski P, James C. Partial deafness treatment with the nucleus straight research array cochlear implant. Audiology & Neuro-Otology. 2012;**17**(2):82-91

[49] Lenarz T, Stover T, Buechner A, Lesinski-Schiedat A, Patrick J, Pesch J. Hearing conservation surgery using the Hybrid-L electrode. Results from the first clinical trial at the Medical University of Hannover. Audiology & Neuro-Otology. 2009;**14**(Suppl 1):22-31

[50] Gantz BJ, Hansen MR, Turner CW, Oleson JJ, Reiss LA, Parkinson AJ. Hybrid 10 clinical trial: Preliminary results. Audiology & Neuro-Otology. 2009;**1**:32-38

[51] Arndt S, Laszig R, Aschendorff A, Hassepass F, Beck R, Wesarg T. Cochlear implant treatment of patients with single-sided deafness or asymmetric hearing loss. German version. Hals-Nasen-Ohrenheilkunde. 2017;**65**(7):586-598

[52] Mertens G, De Bodt M, Van de Heyning P. Cochlear implantation as a long-term treatment for ipsilateral incapacitating tinnitus in subjects with unilateral hearing loss up to 10 years. Hearing Research. 2016;**331**:1-6

[53] Shannon RV, Fayad J, Moore J, Lo WW, Otto S, Nelson RA, et al. Auditory brainstem implant: II. Postsurgical issues and performance. Otolaryngology and Head and Neck Surgery. 1993;**108**(6):634-642

[54] Otto SR, Shannon RV, Wilkinson EP, Hitselberger WE, McCreery DB, Moore JK, et al. Audiologic outcomes with the penetrating electrode auditory brainstem implant. Otology & Neurotology. 2008;**29**(8):1147-1154

[55] Shannon RV. Auditory implant research at the house ear institute 1989-2013. Hearing Research. 2015;**322**:57-66

[56] Colletti V, Shannon R, Carner M, Veronese S, Colletti L. Outcomes in nontumor adults fitted with the auditory brainstem implant: 10 years' experience. Otology & Neurotology. 2009;**30**(5):614-618

[57] Lenarz T, Lim HH, Reuter G, Patrick JF, Lenarz M. The auditory midbrain implant: A new auditory prosthesis for neural deafness-concept and device description. Otology & Neurotology. 2006;**27**(6):838-843

[58] Rode T, Hartmann T, Hubka P, Scheper V, Lenarz M, Lenarz T, et al. Neural representation in the auditory midbrain of the envelope of vocalizations based on a peripheral ear model. Frontiers in Neural Circuits. 2013;**7**(166):18

[59] Middlebrooks JC, Snyder RL. Intraneural stimulation for auditory prosthesis: Modiolar trunk and intracranial stimulation sites. Hearing Research. 2008;**242**(1-2):52-63

[60] Finley C, Holden T, Holden L, Whiting B, Chole R, Neely G, et al. Role of electrode placement as a contributor to variability in cochlear implant outcomes. Otology & Neurotology. 2008;**29**(7):920-928

[61] Aschendorff A, Kromeier J, Klenzner T, Laszig R. Quality control after insertion of the nucleus contour and contour advance electrode in adults. Ear and Hearing. 2007;**28**(2 Suppl):75S-79S

[62] Fitzpatrick DC, Campbell AP, Choudhury B, Dillon MT, Forgues M, Buchman CA, et al. Round window electrocochleography just before cochlear implantation: Relationship to word recognition outcomes in adults. Otology & Neurotology. 2014;**35**(1):64-71

[63] Dhanasingh A, Jolly C. An overview of cochlear implant electrode array designs. Hearing Research. 2017;**356**:93-103

[64] Todt I, Mittmann P, Ernst A. Hearing preservation with a Midscalar electrode comparison of a regular and steroid/pressure optimized surgical approach in patients with residual hearing. Otology & Neurotology. 2016;**37**:e349-e352

[65] Torres R, Jia H, Drouillard M, Bensimon JL, Sterkers O, Ferrary E, et al. An optimized robot-based technique for cochlear implantation to reduce array insertion trauma. Otolaryngology and Head and Neck Surgery. 2018;**159**(5):900-907

[66] Giardina CK, Brown KD, Adunka OF, Buchman CA, Hutson KA, Pillsbury HC, et al. Intracochlear electrocochleography: Response

patterns during cochlear implantation and hearing preservation. Ear and Hearing. 2019. (In Press)

[67] Bester CW, Campbell L, Dragovic A, Collins A, O'Leary SJ. Characterizing electrocochleography in cochlear implant recipients with residual low-frequency hearing. Frontiers in Neuroscience. 2017;**11**(141):1-8

[68] Harris MS, Riggs WJ, Koka K, Litvak LM, Malhotra P, Moberly AC, et al. Real-time intracochlear electrocochleography obtained directly through a cochlear implant. Otology & Neurotology. 2017;**38**(6):e107-e113

[69] Jia H, Francois F, Bourien J, Eybalin M, Lloyd RV, Van De Water TR, et al. Prevention of trauma-induced cochlear fibrosis using intracochlear application of anti-inflammatory and antiproliferative drugs. Neuroscience. 2016;**316**:261-278

[70] Wilk M, Hessler R, Mugridge K, Jolly C, Fehr M, Lenarz T, et al. Impedance changes and fibrous tissue growth after cochlear implantation are correlated and can be reduced using a dexamethasone eluting electrode. PLoS One. 2016;**11**(2):e0147552. p. 19

[71] Farahmand Ghavi F, Mirzadeh H, Imani M, Jolly C, Farhadi M. Corticosteroid-releasing cochlear implant: A novel hybrid of biomaterial and drug delivery system. Journal of Biomedical Materials Research. Part B, Applied Biomaterials. 2010;**94**(2):388-398

[72] Schindler JS, Schindler RA. Neurotrophin action in the cochlea: Implications for cochlear implants. Advances in Oto-Rhino-Laryngology. 1997;**52**:8-14

[73] Leake PA, Stakhovskaya O, Hetherington A, Rebscher SJ, Bonham B. Effects of brain-derived neurotrophic factor (BDNF) and electrical stimulation on survival and function of cochlear spiral ganglion neurons in deafened, developing cats. Journal of the Association for Research in Otolaryngology. 2013;**14**(2):187-211

[74] Taylor RR, Filia A, Paredes U, Asai Y, Holt JR, Lovett M, et al. Regenerating hair cells in vestibular sensory epithelia from humans. eLife. 2018;**18**(7):34817

[75] Wan G, Gomez-Casati ME, Gigliello AR, Liberman MC, Corfas G. Neurotrophin-3 regulates ribbon synapse density in the cochlea and induces synapse regeneration after acoustic trauma. eLife. 2014;**20**(3):03564

[76] Whelan HT, Smits RL Jr, Buchman EV, Whelan NT, Turner SG, Margolis DA, et al. Effect of NASA light-emitting diode irradiation on wound healing. Journal of Clinical Laser Medicine & Surgery. 2001;**19**(6):305-314

[77] Bartos A, Grondin Y, Bortoni ME, Ghelfi E, Sepulveda R, Carroll J, et al. Pre-conditioning with near infrared photobiomodulation reduces inflammatory cytokines and markers of oxidative stress in cochlear hair cells. Journal of Biophotonics. 2016;**9**(11-12):1125-1135

[78] Boyle PJ, Moore BCJ. Balancing cochlear implant AGC and near-instantaneous compression to improve perception of soft speech. Cochlear Implants International. 2015;**16**(S1):S9-S11

[79] Holden LK, Finley CC, Firszt JB, Holden TA, Brenner C, Potts LG, et al. Factors affecting open-set word recognition in adults with cochlear implants. Ear and Hearing. 2013;**34**(3):342-360

[80] Baudhuin J, Cadieux J, Firszt J, Reeder R, Maxson J. Optimization of programming parameters in children with the advanced bionics

cochlear implant. Journal of the American Academy of Audiology. 2012;**23**(5):302-312

[81] Fielden CA, Kluk K, Boyle PJ, McKay CM. The perception of complex pitch in cochlear implants: A comparison of monopolar and tripolar stimulation. The Journal of the Acoustical Society of America. 2015;**138**(4):2524-2536

[82] Srinivasan AG, Padilla M, Shannon RV, Landsberger DM. Improving speech perception in noise with current focusing in cochlear implant users. Hearing Research. 2013;**299**:29-36

[83] Smith ZM, Parkinson WS, Long CJ. Multipolar current focusing increases spectral resolution in cochlear implants. Conference Proceedings: Annual International Conference of the IEEE Engineering in Medicine and Biology Society. 2013;**9**(10):6610121

[84] Kalkman RK, Briaire JJ, Frijns JH. Stimulation strategies and electrode design in computational models of the electrically stimulated cochlea: An overview of existing literature. Network. 2016;**27**(2-3):107-134

[85] Muller J, Brill S, Hagen R, Moeltner A, Brockmeier SJ, Stark T, et al. Clinical trial results with the MED-EL fine structure processing coding strategy in experienced cochlear implant users. ORL: Journal for Otorhinolaryngology and Its Related Specialties. 2012;**74**(4):185-198

[86] Gagne JP, Besser J, Lemke U. Behavioral assessment of listening effort using a dual-task paradigm. Trends in Hearing. 2017;**21**:2331216516687287

[87] Firszt JB, Reeder RM, Holden LK, Dwyer NY. Results in adult cochlear implant recipients with varied asymmetric hearing: A prospective longitudinal study of speech recognition, localization, and participant report. Ear and Hearing. 2018;**39**(5):845-862

[88] Kollmeier B, Warzybok A, Hochmuth S, Zokoll MA, Uslar V, Brand T, et al. The multilingual matrix test: Principles, applications, and comparison across languages: A review. International Journal of Audiology. 2015;**2**:3-16

[89] Boyle P, Nunn T, O'Connor A, Moore B. STARR: A speech test for evaluation of the effectiveness of auditory prostheses under realistic conditions. Ear and Hearing. 2013;**34**(2):203-212

[90] Haumann S, Lenarz T, Buchner A. Speech perception with cochlear implants as measured using a roving-level adaptive test method. ORL: Journal for Otorhinolaryngology and Its Related Specialties. 2010;**72**(6):312-318

[91] Dincer D'Alessandro H, Ballantyne D, Boyle PJ, De Seta E, DeVincentiis M, Mancini P. Temporal fine structure processing, pitch, and speech perception in adult cochlear implant recipients. Ear and Hearing. 2018;**39**(4):679-686

[92] Spahr A, Dorman M, Litvak L, Van Wie S, Gifford R, Loizou P, et al. Development and validation of the AzBio sentence lists. Ear and Hearing. 2012;**33**(1):112-117

[93] Kitterick PT, Bailey PJ, Summerfield AQ. Benefits of knowing who, where, and when in multi-talker listening. The Journal of the Acoustical Society of America. 2010;**127**(4):2498-2508

[94] Veugen LC, Chalupper J, Snik AF, van Opstal AJ, Mens LH. Frequency-dependent loudness balancing in bimodal cochlear implant users. Acta Oto-Laryngologica. 2016;**136**(8):775-781

[95] Dorman MF, Cook Natale S, Agrawal S. The value of unilateral CIs, CI-CROS and bilateral CIs, with and without beamformer microphones, for

speech understanding in a simulation of a restaurant environment. Audiology & Neuro-Otology. 2018;**23**(5):270-276

[96] Snapp HA, Hoffer ME, Spahr A, Rajguru S. Application of wireless contralateral routing of signal technology in unilateral cochlear implant users with bilateral profound hearing loss. Journal of the American Academy of Audiology. 2018;**29**(10):17121

[97] Friedmann DR, Peng R, Fang Y, McMenomey SO, Roland JT, Waltzman SB. Effects of loss of residual hearing on speech performance with the CI422 and the hybrid-L electrode. Cochlear Implants International. 2015;**26**:26

[98] Cullington H, Kitterick P, Weal M, Margol-Gromada M. Feasibility of personalised remote long-term follow-up of people with cochlear implants: A randomised controlled trial. BMJ Open. 2018;**8**(4):2017-019640

Chapter 6

Endoscopic Ear Surgery in Children

Luca Oscar Redaelli de Zinis and Nader Nassif

Abstract

Endoscopic assistance is gradually gaining recognition in otology not only for office examinations but also during surgery. The first endoscopic surgical procedure that was started in our institution was endoscopic ventilation tube placement to manage children with stenotic and curved canals. Following this, endoscopy was used in all type I tympanoplasty and stage I cholesteatoma removals with the advantage of avoiding a postauricular or endaural approach. The last application of endoscopic assistance was to better visualize round window and scala tympani via posterior tympanotomy during cochlear implantation. There are several advantages in using endoscopes: the wide view obtained and the possibility to observe areas behind the angle with less invasiveness and its excellent resolution, in addition to its intense light and higher magnification that facilitates teaching and tutoring. The limits of endoscopic surgery are that one hand is always needed to hold the endoscope and the lack of a third dimension. Until miniaturization of 3D systems allow the possibility to work in the narrow external ear canal, in order to overcome the limitation that one hand is dedicated to the endoscope, we will describe the use of an endoscope holder in otologic procedures.

Keywords: endoscopy, ear surgery, pediatric otology, chronic otitis media, cholesteatoma, otitis media with effusion, cochlear implantation, myringoplasty, tympanic membrane perforation

1. Introduction

The impressive technological evolution of the last decades and the miniaturization of instruments have allowed to intervene with less invasiveness and better functional results in the field of surgical otology. The introduction of endoscopy with rigid endoscopes, coupled with video processors and high definition monitors, has opened new alternatives to conventional and validated otomicroscopic surgical approaches to the middle ear and avoid external access to various pathologies. Otoendoscopy is gradually taking hold on otology not only in diagnostic approaches but also in surgical aspects. It is well known that the endoscope has the advantage of offering a wider view compared to microscopes.

One of the first endoscopic surgical approaches to the middle ear was described by Thomassin [1]. Thereafter, many articles were published on otoendoscopic surgical procedures. Its limits are related to the lack of a third dimension, and in case of bloody fields, surgery becomes demanding and troublesome. In this chapter, we will discuss our experience developed over the years in endoscopic otology in children starting in 2004 as a diagnostic tool and then gradually as a surgical instrument starting from ventilation tube placement, myringoplasty, cholesteatoma removal, ossiculoplasty, cochlear implantation, and, finally, the application of a holder for two-handed endoscopic surgery.

2. Endoscopy as a diagnostic tool

Endoscopy provides an easy and comfortable means of examining the external auditory canal (EAC) and tympanic membrane in children for diagnostic purposes and follow-up as data is recorded and archived on a hard disk. Most children are uncooperative during office sessions for pre- and post-operative evaluations when they are laid on a bed under the microscope. With endoscopy, the child can sit on the lap of his/her parent, feeling safe and being more compliant, but otherwise by him/herself (**Figure 1**).

Figure 1.
On the left, a child sitting on the lap of his father during endoscopy; on the right, a boy lying on the examination bed by himself.

3. Ventilation tube placement

Myringotomy with placement of ventilation tubes (VT) is considered to be basic treatment for the ENT surgeon in certain pathologies of the middle ear such as otitis media with effusion (OME). The alternative to the use of a microscope in carrying out such a procedure is a rigid endoscope.

Endoscopic VT placement was started at our ENT Pediatric Department in 2008 by one colleague and was gradually followed by colleagues who shifted from a microscope to an endoscope until 2012, when all grommets were placed endoscopically. Nowadays, it is considered as standard technique, owing to the easy and straightforward nature of the procedure.

No special prior preparation is needed for endoscopic VT placement. In pediatric patients, the procedure is always performed under general anesthesia. The patient is placed in an otosurgical position, and cleansing of the surgical field is performed using an antiseptic product prior to removal of wax or debris when present. Wax is not a limitation in our hands, in contrast to what noted by other authors [2]; it is easily handled and cleaned endoscopically. Hair trimming was not needed in any patient. Antifogging liquid is needed to avoid blurred vision noted in another study [3].

The endoscopic procedure is stepwise starting with myringotomy in the anterior quadrants (**Figure 2A**) using a sickle knife; suctioning of secretions from the middle ear; positioning the rim of the grommet by forceps on the site of the incision (**Figure 2B**); and gentle slipping of the grommet inside the myringotomy with a needle or a pick (**Figure 2C**). We usually use two types of VTs: a Shepard grommet made of fluoroplastic and silicone Goode T-Tubes.

No age-related limitations were encountered in endoscopic VT placement and, in fact, the age of our patients at surgery ranged from 1 month to 15 years.

The **endoscope** to use in such procedures is 0° angled. At the beginning of our experience, surgical treatment was also performed with a 30° angled scope, but no significant benefit over the 0° endoscope was observed. Moreover, as described in other studies, different lengths of endoscopes, such as 6, 11, 14, 16, and 18 cm, can be used to carry out endoscopic management [2, 4]. In our hands, the use of 11 cm or longer scopes offers greater maneuvering space for other tools without limiting the excursion of the operative hand by touching the camera.

Endoscopes with different **diameters** (1.9, 2.7, 3, and 4 mm) are also available. Some authors have suggested the use of a small (2.7 mm) scope in pediatric patients [2]. In our experience, 3 or a 4 mm scopes can be used interchangeably except for stenotic, bending EAC or syndromic patients (i.e., Down S., Goldenhar S., Pallister-Killian S., etc.), and in that case, a 3.0 mm endoscope is recommended. A 2.7 mm, 18 cm long endoscope is avoided because in comparison with the 3.0 endoscope, it has less luminosity should slight pressure be exerted on the shaft and a black crescent appears laterally, reducing the operating field. As we all know, in narrow or oblique EACs, positioning of VTs with an operating microscope can be a troublesome task. The ear speculum itself occupies space in the membranous ear canal, further reducing its diameter. Secondly, in a curved EAC, it is often not possible to view the inferior quadrants of the eardrum where the tube should be placed. On the contrary, the use of an adequate endoscope in a narrow and curved canal offers a larger visual field. According to our experience, it is preferable to use an endoscope with a larger diameter that is compatible with the size of the EAC. There is no need for special **tools** for endoscopic placement of grommets. The same instrument set used in a standard microscopic approach is required. In our experience, sickle knives with the smallest tips, both straight and slightly curved, are the best choice. Additionally, delicate ear forceps having a working length of 8 cm are usually well handled. Suction tubes that fit well in the pediatric age are those with diameters from 3 to 5 French.

The standard **endoscopic technique** for grommet insertion has been described above, whereas T-tube placement is more articulated, especially in the presence of atelectasis. However, in these cases, the degree of difficulty is not different from conventional otomicroscopy procedures. After myringotomy, the T-tube is grasped

Figure 2.
(A) Myringotomy, (B) suction of secretions, and (C) positioning of the grommet.

either by both wings together or by one only and inserted inside the incision. The advantage of otoendoscopy compared to the operating microscope is that during placement of the T-tube, one can clearly evaluate the depth of insertion of the tube wings (**Figure 3**). The endoscope itself may be used to keep the T-tube in position while releasing and extracting the forceps.

Furthermore, an endoscopic surgical technique can be influenced by the size of EAC. Whenever the EAC diameter is stenotic, the surgical maneuver differs from standard cases. As performed in our patients, the tube is laid at the meatus and is then pushed medially and inserted in the myringotomy with a pick. The scope is not introduced deeply in the ear canal, but is stopped after its entrance. The impression of the endoscope on the membranous wall of the EAC can be noticed in **Figure 4A** and stenosis of the EAC with a diameter equal to that of the grommet is observed in **Figure 4B**.

In one of our patients, stenosis was not the only difficulty: the reduced space between the ear drum and the anterior wall of the EAC, that is, the acute anterior tympanomeatal angle, limited the surgical maneuvers was also problematic. Rotating the tube inside the incision was impossible due to the hindrance of the anterior wall. In this specific case, it was decided to switch to a T-tube to complete the procedure successfully (**Figure 5**). This was possible because in a very narrow space, holding the T-tube with a Hartmann forceps under endoscopic control allowed the flanges of the T to enter the myringotomy directly.

The type of **grommet** selected is another important issue in endoscopic myringotomies with VTs. Rigid materials such as fluoroplastic have some advantages over elastic ones (such as silicone or microgel). The reason lies in the need to push the grommet with a needle or a pick in smaller size EACs. The instruments would penetrate soft materials, thus not enabling correct insertion or getting stuck in it. In a couple of cases, the VT bounced back and we had to repeat the maneuver. Elastic materials might be preferred in wide EACs, where the grommet is applied directly by the ear forceps.

The **limitations** in handling instruments differ between left-handed and right-handed surgeons. For those right handed, tube placement in the right ear canal is more demanding, owing to the higher risk of excoriation of the inferior wall of the EAC and subsequent bleeding (**Figure 6**).

In left ears, the field of vision improves but there is a higher risk of skin lesions or hematoma of the anterior wall of the EAC. The opposite applies for those who are left handed; that is, excoriation of anterior canal wall in the right ear canal and inferior wall of the left ear canal (**Figure 6**).

Figure 3.
Positioning of T-tube grabbed by its wings.

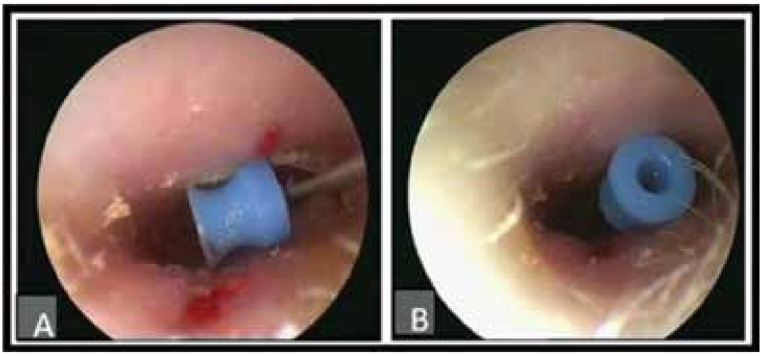

Figure 4.
Positioning of the grommet in stenotic EAC. (A) Notice the impression of the 2.7 mm endoscope greater in diameter w.r.t the EAC and (B) diameter of the grommet compared to that of the EAC.

According to the literature, a major limit of endoscopic ear surgery is bleeding. When it occurs, the surgeon is sometimes forced to interrupt or convert the procedure to an otomicroscopic, bi-manual technique [12, 13, 15]. Senior surgeons are accustomed to use the left hand for suction (if right-handed) and the right hand to maneuver the instruments within the EAC, due to traditional otosurgical training. This allows maintaining the tiny operative field within the ear canal clear from blood. In our series, bleeding was never a relevant issue: in no instance did the procedure become so demanding and troublesome that there was the need to convert it to the microscope. In case of bloody field due to inadvertent tearing of the canal skin, the simple application of a sponge soaked with a vasoconstricting agent such as epinephrine 1/1000 for a few minutes allowed adequate hemostasis. Following this, rinsing the ear canal with warm water or saline also helps to clear the field. It is clear that the surgical skills of the operator and his/her experience with the use of endoscopes play an important role in minimizing surgical trauma.

Single **handedness** in endoscopy is considered by different otologists as a limitation in ear surgery, even for simple procedures such as VT placement [2, 3]. According to our experience, this is not the case as long as the skin of the ear canal is respected by taking into consideration the above advice.

Unlike other studies in the literature [3], based on our experience, no special **training** for endoscopic grommet positioning is required. Special training is definitely required for more complex procedures such as tympanoplasties,

Figure 5.
T-tube placed in stenotic EAC.

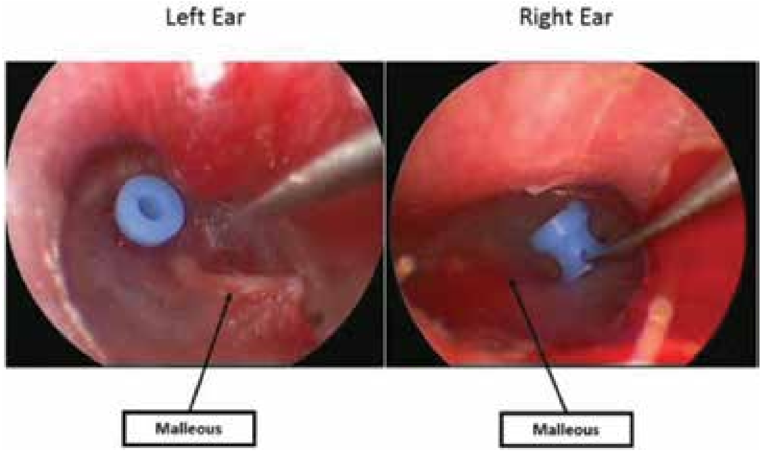

Figure 6.
Iatrogenic hematoma of the left ear anterior wall and excoriation of the right ear inferior wall for a right-handed surgeon.

ossiculoplasties, and stapedotomies. In fact, the learning curve for grommet insertion is very short, especially for colleagues who already perform endoscopic sinus surgery.

In conclusion, otoendoscopic VT placement is a valid and secure procedure. It is applicable in all patients independent of age, type of tube, Grommet tubes and T tubes, and anatomical conformation. The surgical approach is especially advantageous for grommet placement in narrow and curved EACs. In bloody fields, surgery may become more time consuming, but never compels the surgeon to abandon the technique. The learning curve is steep, especially for surgeons who are acquainted with endoscopic sinus surgery. Both cost-wise and in terms of logistic handling, the endoscope offers clear advantages. Nevertheless, the operating microscope must always be available in the operating room for rare cases of high jugular bulb or aberrant course of the internal carotid artery in the middle ear, which raises the hypothetical risk of bleeding that would not be controllable by an endoscopic approach alone. The endoscopic approach yields results that are comparable to traditional otomicroscopic techniques, but it is clearly superior in anatomically complex cases.

4. Tympanoplasty (myringoplasty)

Myringoplasty (MP) in children is one of the most common otologic procedures and can offer a success rate as high as 95% [5, 6]. It is considered a challenging procedure in children compared to adults due to narrowness of the EAC and generally smaller size of the ear [7, 8]. In a pediatric age, access to the tympanic membrane and elevation of tympanomeatal flap (TMF) to perform MP generally necessitates permeatal incision by employing an endaural or postauricular approach, especially in anterior and subtotal perforations, whereas a transmeatal approach is suitable only for small and posterior perforations [9–14]. In such cases, surgeons would not operate on children until the age of 10–14 years due to technical difficulties encountered in small anatomy, inability of the child to co-operate post-operatively, and increased risk of psychological trauma [10, 15–19]. In anterior perforation, surgery is more challenging as graft placement may be inaccurate [19], and the anterior aspect of the eardrum is more difficult to visualize, especially in children where the external ear canal dimensions are constraining [20, 21]. In our department, we have adopted an endoscopic technique since 2011 [22]. The use of endoscopy to perform myringoplasty may obviate such limiting factors. In agreement with our experience and that of others reported in the literature [23], an endoscopic approach can offer many advantages over a microscope approach in children. It provides the possibility to decrease morbidity by avoiding postauricular or endaural incision and applying a transcanal approach.

The **technique** consists of refreshing the margins of the perforation using a sickle knife or Rosen needle and grasping forceps. Two vertical incisions are performed at 12 and 6 o'clock on the skin, and at a horizontal one at about 0.2 cm from the annulus and the medial TMF is elevated. The graft is inserted under the anterior margin of the perforation, underlay fashion, and under or above the handle of the malleus depending of the extension of the perforation (**Figure 7A–L**); in the case that the perforation involves the anterior quadrants, it is applied over the malleolus (**Figure 8A, B**). Gelfoam is applied adequately in the middle ear, and then both the free part of the flap and the graft are repositioned.

According to literature reports, **bleeding** is considered as a limit of endoscopic ear surgery. When it does occur, the surgeon is sometimes forced to interrupt or convert to an otomicroscopic bi-manual procedure [4, 24, 25]. In our experience, bleeding was never a relevant issue: in no instance did the procedure become so demanding and troublesome that it had to be converted to a microscopic intervention.

Bleeding can be handled in several ways. Injection of the EAC with epinephrine 1:1000 solution and then waiting for a few moments before incising the skin. Another hint is that after the incision prominent bleeding will be noticed at the beginning, one may shift and start harvesting the tragal perichondrium in order to provide enough time for spontaneous hemostasis to take place. During elevation of the TMF, it could be controlled by irrigation with warm saline and local application of pledgets soaked with a vasoconstricting agent.

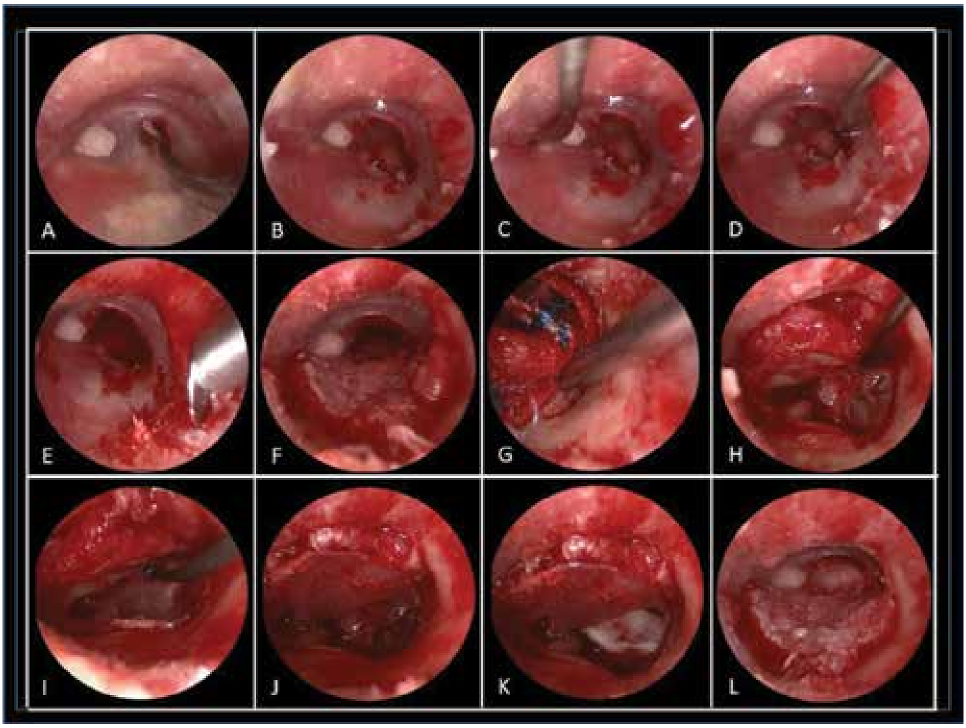

Figure 7.
Single-handed transcanal endoscopic myringoplasty, steps A through L.

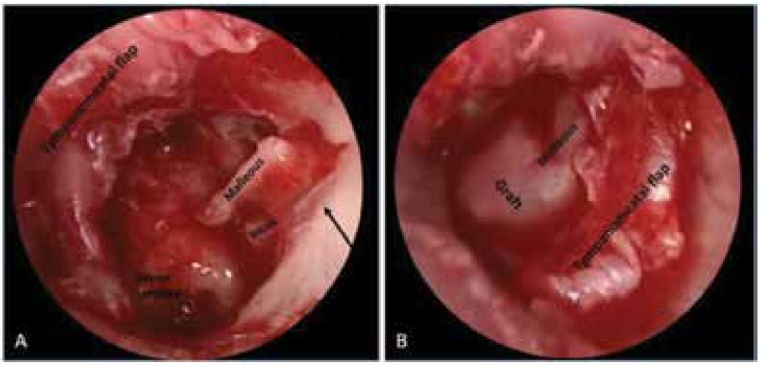

Figure 8.
(A) Tympanomeatal flap elevated and malleolus is denuded. (B) Graft positioned under the osseous annulus and over the malleolus.

A very helpful tool recently introduced as an otological instrument to overcome the presence of blood in the surgical field is the round knife with suction duct. During the dissection and elevation of the tympanomeatal, it offers the advantage of going through the dissection until reaching the annulus without the need to interrupt and use suction to aspirate blood from the surgical field (**Figure 9**).

The choice of the **graft** depends upon the preference of the surgeon. In our department, we either use the tragus (perichondrium ± cartilage) harvested and trimmed otherwise using biologic tissue. The advantage of biologic tissue is that one can reduce the surgical time by up to 15 min (**Figure 10**).

As described in other studies, endoscopes with different lengths, such as 6, 11, 16, and 18 cm, can be used to carry out endoscopic management [2, 4, 23]. According to our experience, the optimal length is between 11 and 16 cm in order to avoid impingement of instruments such as that noted on ventilation tube placement [22]. The rigid **endoscope** to use according to our experience is 3.0 mm and 16 cm long. This diameter and length fits all; both in a stenotic canal and a normal one, it provides good luminosity and adequate distance between the scope and otologic instruments. A 2.7 mm diameter, 18 cm long endoscope is not advisable due to its small field of vision, less luminosity, and inadequate rigidity of the shaft that will bend upon exertion of pressure, determining a crescent black spot field laterally (**Figure 11**).

In conclusion, endoscopic-assisted transcanal myringoplasty is feasible in all cases of tympanic perforation in children of any age and can be considered as a valid, alternative approach to the microscope. It is less invasive, especially in

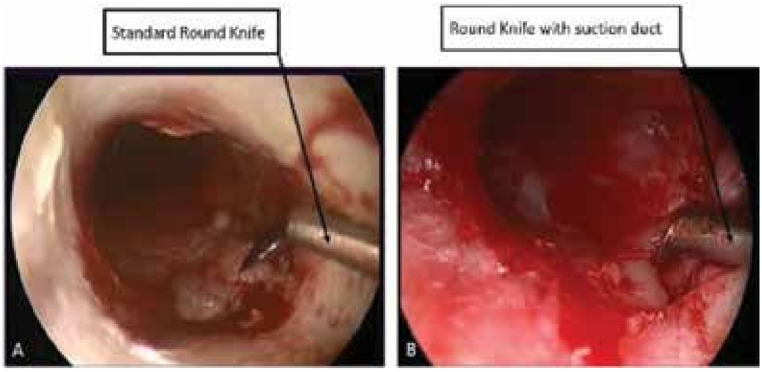

Figure 9.
(A) Standard round knife and (B) round knife with suction duct. Notice the difference of the surgical field in B, which is free of blood.

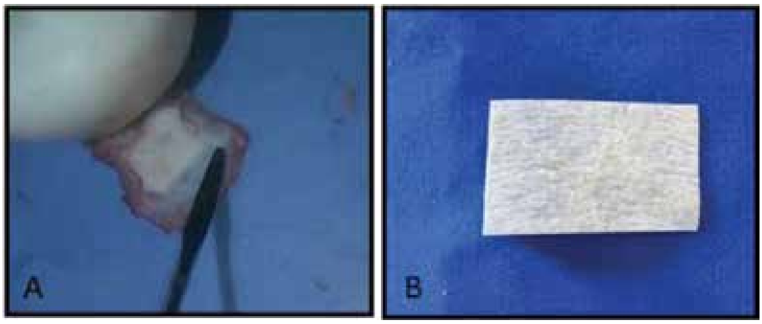

Figure 10.
(A) Autograft tragus perichondrium with cartilage isle and (B) xenograft acellular porcine small intestinal submucosa (Biodesign; Cook Medical Inc., Bloomington, IN).

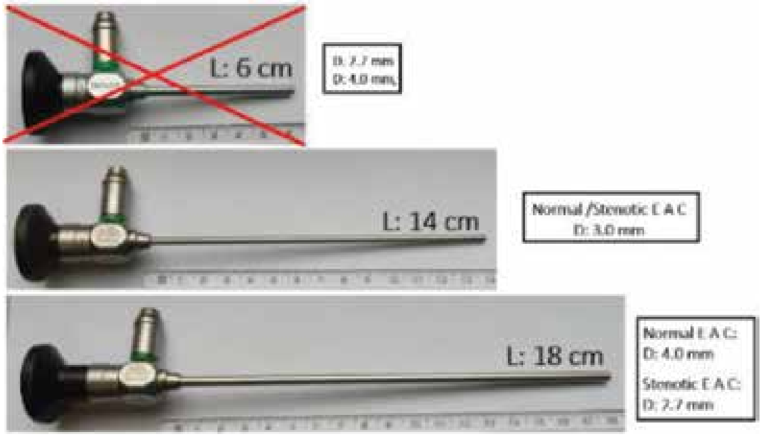

Figure 11.
Different lengths of endoscopes: 14 and 18 cm are indicated for ear surgery.

anterior perforations and in narrow and curved external canals where post-auricular or endaural approaches are otherwise required. An endoscopic approach grants better cosmetic outcomes and less psychological trauma with comparable anatomical and functional results vs. a traditional otomicroscopic technique.

5. Endoscopic approach to cholesteatoma

Endoscopy as a surgical tool in otosurgery is gradually evolving in daily practice when approaching middle and inner ear pathologies. Moreover, it is well known that cholesteatoma in a pediatric age, with an incidence that varies between 3 and 6 per 100,000 [26, 27], is an aggressive disease with respect to that in adults. The surgical approach is tailored according to the nature of cholesteatoma in being acquired or congenital, cystic or invasive. It often requires an extensive surgical approach, without undermining the preservation of hearing, as an attempt to eradicate the pathology due to its high rates of recidivism. The strategy to manage recidivism depends if it is residual, recurrent, or iatrogenic in nature. The standardized techniques at our disposition for such a destructive pathology, contrived upon the application of an operative microscope, are the transcanal (TC), canal wall-up (CWU), canal wall-down (CWD), and subtotal petrosectomy approaches. The surgeon should apply the indicated but least invasive and most effective approach, bearing in mind that the child will still have their entire life span ahead. Therefore, not only should eradication of the disease be considered, but functional outcomes such as hearing, balance, and stability of the cavity must be taken into consideration.

The application of otoendoscopy in recent years has allowed otosurgery to evolve and be less invasive in reconsidering certain standardized microscope operative techniques [28–30]. The major characteristic of endoscopy is to detect and dominate blind angles during surgery. Thus, in cholesteatoma surgery, endoscopy allows improved cholesteatoma removal while decreasing the rate of recidivism.

Cholesteatoma surgery is a step by step procedure that depends upon intraoperative findings. The approach cannot be decided a priori if it should be a TC, CWU, or CWD procedure. Radiologic work-up is another important element that is needed to plan the procedure. Nonetheless, during surgery, the extension of the disease may not be as expected due to the limits of radiology in defining the propagation of the pathology. Furthermore, the time elapsed between evaluating the result of radiology and the day of surgery may allow the disease to grow further and therefore invade and extend widely. For these reasons, a programmed, solely endoscopic surgical approach is not recommended, and an operative microscope is an indispensable tool to have next to the endoscope in the operating theater.

The advantages of endoscopy in ear surgery are to eradicate surgery and decrease morbidity, especially in children. According to our experience and that reported in the literature, the exclusive use of endoscopy for cholesteatoma removal should be limited when the disease involves only the tympanic cavity [31] and more precisely the mesotympanum and hypotympanum.

When cholesteatoma extends to the epitympanum, angled endoscopes offer a valid view to that area and for the aditus ad antrum. Therefore, the otosurgeon is at ease to decide the technical approach of the surgical procedure. Surgery may proceed as a TC approach, or extend to a combined endoscopic and microscopic one; therefore, mastoidectomy may become mandatory with a CWU or CWD approach in order to ensure that the pathology has been removed.

According to our experience, two major factors determine the choice of type or approach; the condition of the ossicular chain and nature and extension of the disease. In case of intact and healthy ossicles, and cystic cholesteatoma, such as congenital colesteatoma (**Figures 12–14**) being either lateral or medial to the ossicular chain, with an angled endoscope, there is a good chance of removing the cholesteatoma endoscopically.

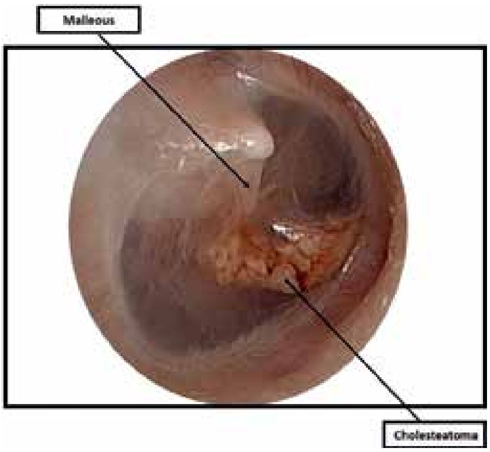

Figure 12.
Congenital cholesteatoma with an intact tympanic membrane.

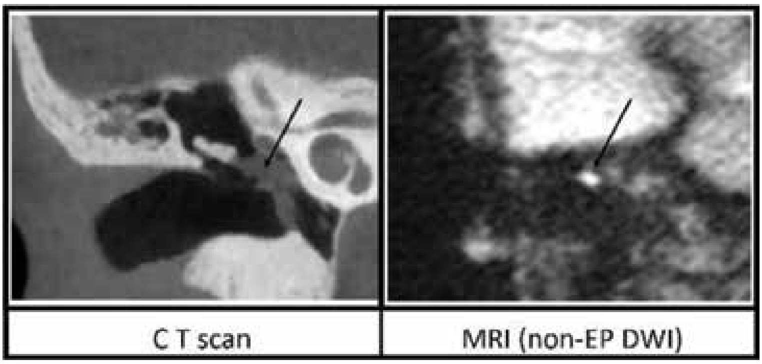

Figure 13.
Radiologic pre-operative work up: coronal CT and MRI showing the mass occupying the mesotympanum.

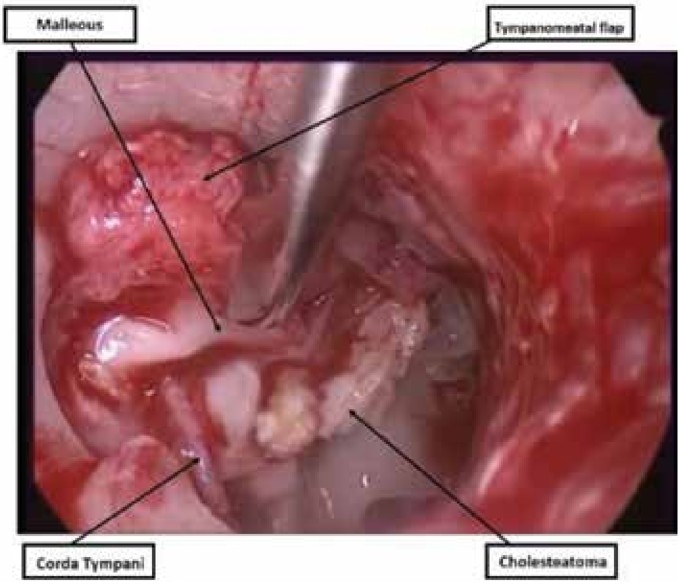

Figure 14.
Intraoperative endoscopic removal of cholesteatoma from the mesotympanum keeping the ossicles intact.

On the other hand, when the cholesteatoma is invasive and occupies the medial part of the head of the malleus and the body of the incus or extends to the antrum, it becomes a challenging task endoscopically. CWU mastoidectomy including antrostomy would be a plausible choice to remove the pathology and leave the ossicles intact. Conversely, in trying to approach such a pathology by an exclusive TC endoscopy, an extended atticotomy (reaching the antrum of the mastoid) and removal of the head of the malleus and the incus are required. This would determine the need to reconstruct both ossicles and the iatrogenic bony defect of the EAC with a piece of cartilage. In case of a small defect of the EAC, a small piece of cartilage may be adequate with good results in the future. If the defect is wide, a wedge of cartilage is needed to reconstruct the postero-superior part of the EAC. This closure would not be guaranteed in later years since the child is in growing phase. Therefore, an increase in the dimensions of the area of the EAC is inevitable with the consequence of a high probability of cholesteatoma recidivism, and CWD revision surgery would be necessary. As a consequence, otoendoscopy no longer has the advantages of being less invasive with reduced morbidity.

Children have an extremely active metabolism, and the risk of recidivism is higher than in adults. For this reason, children need closer follow-up and for a longer period, almost their entire lifespan. Non-EPI-DWI MRI is a very helpful tool in monitoring the disease and trying to avoid unnecessary surgical revisions. Most authors would consider MRI at 1 year postoperatively, otherwise explorative tympanoplasty in recommended [32].

In case of recurrence of disease, endoscopy plays an interesting role. If cholesteatoma is limited to the tympanic cavity such as the epitympanum, endoscopy grants fully transcanal removal of the disease without resorting to an post-aural approach (**Figures 15** and **16**).

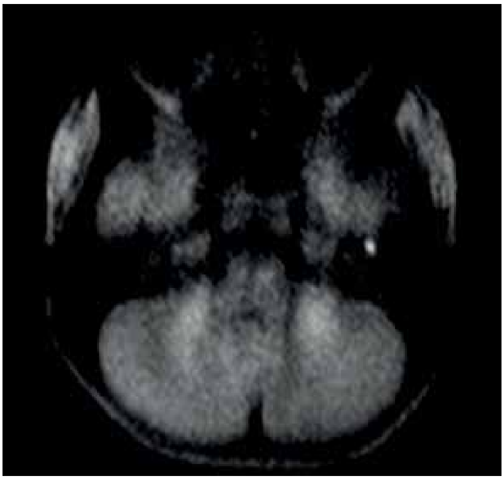

Figure 15.
Non-EPI-DWI MRI follow-up after 1 year of a 9-year-old child who underwent a CWU for cholesteatoma. Note the presence of the mass in the tympanic cavity.

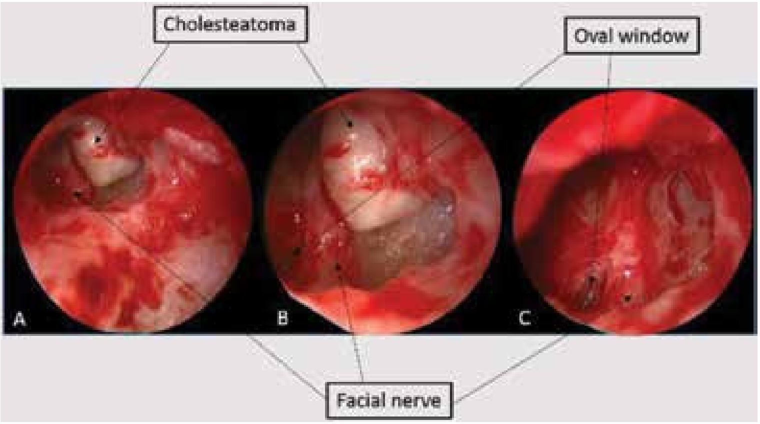

Figure 16.
Endoscopic transcanal of cholesteatoma removal for epitympanum.

In the case that cholesteatoma recurrence is in the mastoid cavity (**Figure 17**), endoscopy can also be an interesting tool for removal of cholesteatoma in a less invasive fashion, as shown in **Figure 18**.

Instrumentation for endoscopic cholesteatoma removal should include dedicated tools to better control its dissection and removal. The ones that really are helpful, besides being elongated and curved shafts, include a suction duct. They can be dissectors, knives, and suctions with sharp edges (**Figure 19**).

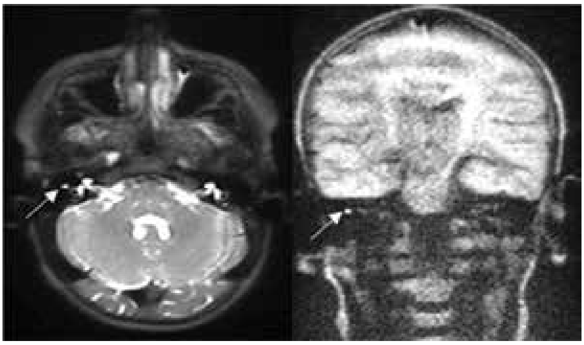

Figure 17.
MRI at 1 year follow-up in CWU in a 7-year-old child showing recurrence of disease in the mastoid cavity in proximity of the lateral semicircular canal.

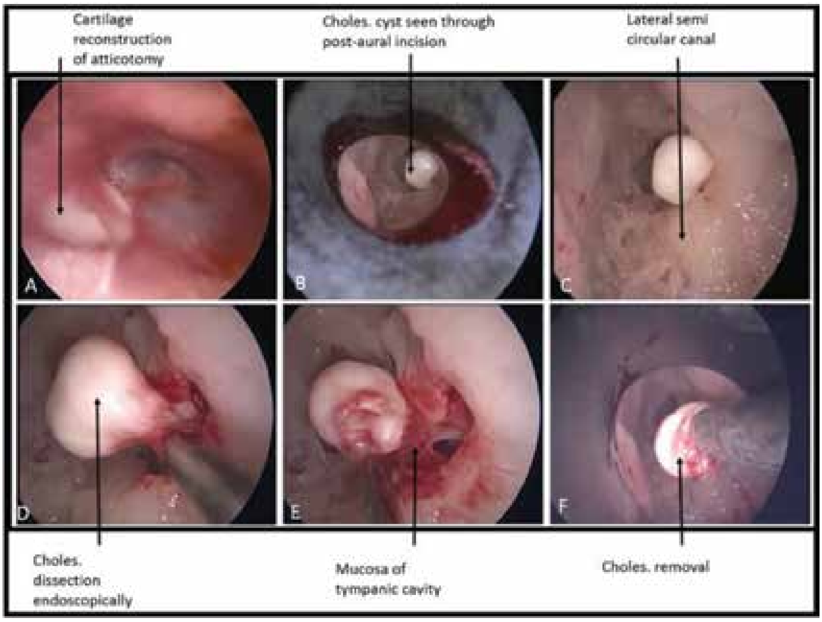

Figure 18.
Intraoperative removal of cholesteatoma through a small 1 cm post-aural incision to gain access to the mastoid after positioning the ear speculum.

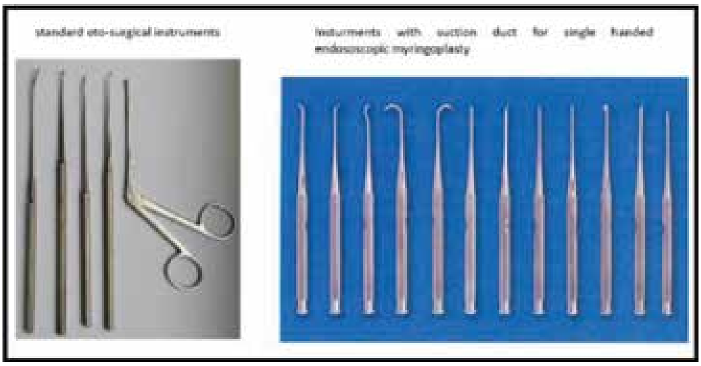

Figure 19.
On the left, standard surgical tools for otosurgery; on the right, Panetti Endoscopic Instrument Set (Spiggle and Theis, Overath, Germany) of dedicated tools containing curved sharp tips containing suction ducts.

6. Cochlear implantation

Nowadays, normal or near-normal low-frequency hearing threshold is present in many patients who are candidates for cochlear implantation, and combined electric acoustic stimulation can be proposed. Consequently, preservation of residual hearing is becoming a crucial issue. Avoiding trauma to inner ear structures when placing a multi-electrode array during the cochlear implant procedure is essential to preserve residual hearing, and the first important topic is the route of insertion of the multi-electrode array.

Scala tympani is the location for multi-electrode array cochlear implants when normal anatomy is present. The reasons for this are the greater sectional dimension compared to scala vestibuli, protection of cochlear duct by osseous spiral lamina and basilar membrane that are stronger than Reissner's membrane, close proximity to the spiral ganglion cell bodies and dendrites, and direct access through the round window (RW) [33, 34].

Access to the scala tympani to introduce the multi-electrode array during the cochlear implant procedure is obtained by opening the RWM, when the RW is sufficiently visible through a posterior tympanotomy or by performing a promontorial cochleostomy in the inferior-lateral wall of the scala tympani, when the RW exposure is not adequate. The advantages of RW access are: reduced risk of intracochlear trauma, wider stimulation surface, and facilitated perimodiolar position [35, 36].

Anatomical variations in the facial nerve, chorda tympani, and the RW niche may create obstacles to approach the RWM [37], and a classification of visualization to the RWM related to surgical approachability has been proposed. Following this classification, complete RWM exposure (considering the only one that guarantees a pure RW approach) was possible in 46% of children, while in 7% the RWM was not visible. [36].

Even if the RWM is completely visible, opening of the RWM does not provide straightforward access to the scala tympani because of the presence of a sharp bony crest in the anterior-inferior border of the niche called the "crista fenestrae," the morphology of which is highly variable [38].

When the RW is not visible and a promontorial approach is needed to consent opening of the scala tympani using a cochleostomy approach without damage to the osseous spiral lamina and basilar membrane, Adunka et al. [34] demonstrated, with an anatomical study, that drilling should proceed from the inferior to the RW annulus, with gradual progression toward the undersurface of the lumen. However, the most basal part of the scala tympani forms a fish hook-like curvature in three dimensions called the "hook" region of the cochlea [39, 40]. This portion contains the cul-de-sac of the endolymphatic space where the osseous spiral lamina, spiral ligament, and basilar membrane merge [40]. The lateral wall of the "hook" region shows large size variations with the consequence of the absence of a reliable landmark for safe access to the scala tympani when promontorial access is performed.

Finally, the angle formed by a line along the plane of the basal turn and the midline (a line between the nasal septum and the internal occipital protuberance) is variable, being inversely proportional with age, thus increasing the risk of damage of the cochlear duct in children. [41].

Applying rigid endoscopy through posterior tympanotomy consents the surgeon to gain both a view of the RW, in case it is invisible, and direction of the scala tympani in children, thereby allowing soft and less traumatic insertion of the multi-electrode array during cochlear implantation (**Figure 20**).

The use of an endoscope to assist cochlear implant surgery has been reported as a transcanal endoscopy by creating a tympanomeatal flap; this approach has been recently questioned because the basal turn of the cochlea align with a more posterior angle than that of the ear canal [42]. The idea to use a rigid endoscope

Figure 20.
(A) RW visible through posterior tympanostomy, (B) case where RW in invisible, and in (C) the advantage of the endoscope to visualize the RW.

to completely visualize the RW region from posterior tympanotomy has been recently reported: 0° (3 mm outer diameter, 14 cm length) and 0 and 30° (2.7 and 4 mm outer diameter, 11 and 6 cm in length) [43, 44]. Endoscopes of 2.7–4 mm diameter cannot likely enter a standard posterior tympanotomy and, that is, in case of posterior location of the RW, there is still suboptimal visualization. With a posterior tympanotomy of 2 mm and the introduction of a 1.9 mm outer diameter endoscope, the RW niche is always completely visible so that the thinning of the lips of the RW niche under endoscopic view until the projection of the lateral wall of the ST and the region of insertion of the basilar membrane are evident; this thus avoids damage to functionally relevant structures when detaching the annulus of the RWM.

The technique consists of a standard mini-invasive surgical approach under general anesthesia performed with a postauricolar access and transmastoid posterior tympanotomy of 2 mm (**Figure 21**).

The tip of a 0°, 1.9 mm diameter and 11 cm long endoscope is positioned in proximity of the upper part of the posterior tympanotomy to obtain a panoramic view of the inferior part of the medial wall of the tympanic cavity. The endoscope can be kept in place using a standard endoscope holder. The next steps are performed under direct endoscopic view. The bone overhanging the RW is lowered using a 1 mm microdrill under constant irrigation or irrigation/aspiration if an endoscope holder is used (**Figure 22**). At that point, exposure of the RW is now possible with a microscope through the posterior tympanotomy and the surgeon may proceed via the microscope approach.

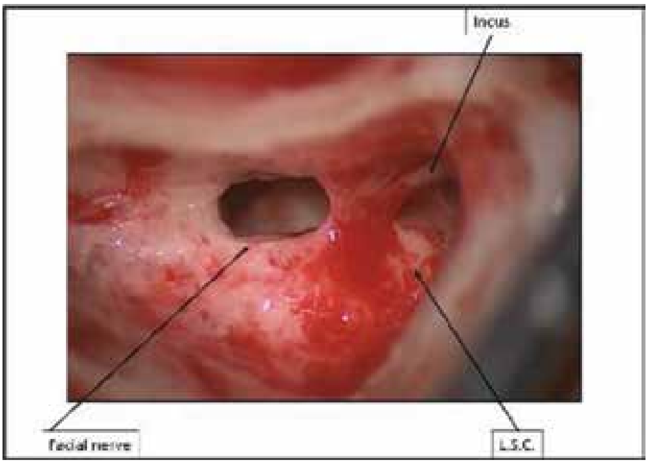

Figure 21.
RW not visible through posterior tympanotomy done by 2 mm burr.

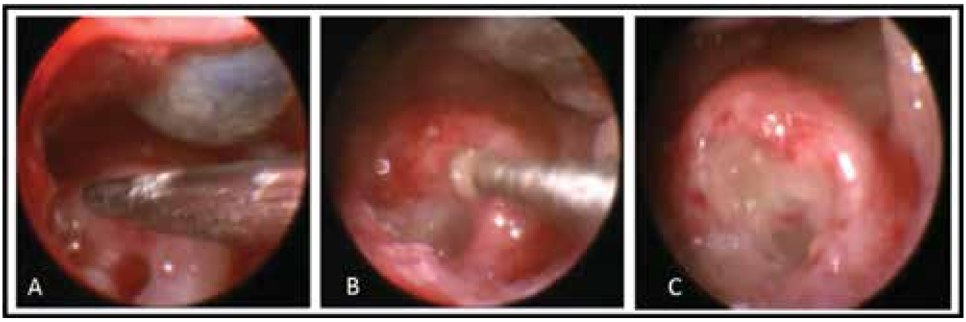

Figure 22.
(A) Visualization of the RW by 1.9 mm endoscope through the posterior tympanostomy, (B) drilling of overhanging bone, and (C) RWM exposed.

Otherwise, the surgeon may continue endoscopically and exploit the best use of endoscopy. The bone overhanging the RW in the anterior inferior aspect of promontory is thinned (**Figure 23A**) until a "blue line" inferiorly (corresponding to the ST) and a white line superiorly (corresponding to the insertion of the basilar membrane) are evident (**Figure 23B**). The anterior inferior part of the annulus of the RWM is detached from its bone insertion with a needle to open the ST (**Figure 23C**). The corresponding thinned bone, the crista fenestrae, is removed by a microcurette or a very low speed 0.5 mm microdrill until the direction of the canal is clearly evident. This allows gaining centered access to the ST and permits smooth linear introduction of the multi-electrode array into the basal turn (**Figure 23D**); attention must be given to avoid bone dust or fragments from entering into the ST. The diameter of the opening of the ST is not previously planned, provided that it is larger than the multi-electrode array diameter to leave perilymph coming out during insertion (**Figure 23**). Dexamethasone 4 mg/ml is gently flushed into the middle ear cavity and cochleostomy site. The electrode array is slowly inserted into the ST in a standard manner.

Endoscopic assistance through posterior tympanotomy has other advantages: in cases of particularly curved external auditory canal or small facial recess over-thinning of the posterior wall with potential breakdown, extensive drilling of the fallopian canal with potential facial damage is no longer necessary; finally, in case of cochlear malformation, a panoramic view of the middle ear medial wall helps to identify the site for cochleostomy with no need of transcanal opening of the middle ear.

In conclusion, endoscope-assisted round window cochleostomy is a practicable alternative to the classical microscopic approach. It allows exposure of hidden RW niche, avoids over-thinning of the posterior external auditory canal by extensive drilling, and is useful to prevent luxation or removal of the external auditory canal when cochlear drillout is indicated. Thus, it optimizes array orientation and introduction into the scala tympani and leads to better results for residual hearing.

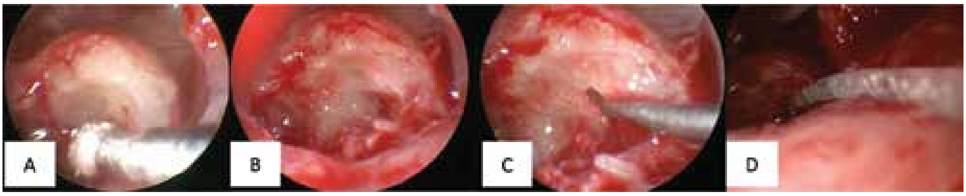

Figure 23.
Left ear. (A) Drilling of the RW lips, (B) exposure of the ST, (C) RW partially elevated, and (D) array insertion along the direction of the ST.

7. The holder, a two-handed technique

Otosurgeons are often skeptic and hesitant for endoscopic ear surgery for several reasons. First, single handedness in endoscopy is a limitation, especially in bleeding fields. When it occurs, bleeding is often a disturbing event, and frequent suction is needed so that the surgeon may be prone to interrupt the procedure and convert it to a traditional bimanual microscope technique [24, 45]. Second, otosurgeons are experienced with double-handed stereoscopic vision. Their teachings and therefore their maneuvers are based on two hands, whereas with endoscopy, otosurgeons have to manage maneuvers with one hand and lose the characteristic of the depth of vision. Differently, surgeons who practice sinus surgery are acquainted with a one hand procedure and for them approaching middle ear surgery is much preferred to operative microscope. This is why it seems to be more acceptable to nondedicated otosurgeons than to dedicated ones.

One of the drawbacks of the technique is being single handed. In case of a bloody field, especially in hyperplastic mucosa of the middle ear, surgery can become demanding and time consuming. Trying to overcome this limit, since January 2016, we started using the STORZ endoscope mechanical holding system followed a few months later by the Unitrack pneumatic holding system.

Different from other endoscopic procedures where a dynamic field is required, that is, cholesteatoma removal [46], during myringoplasty, the endoscope seldom needs to be moved to adjust the field of vision, so that the application of an endoscope holder is particularly favorable. The immediate advantage noticed is the rapidity of the procedure in elevating the tympanomeatal flap and fibrous annulus without frequently stopping to aspirate blood. Washing and suctioning simultaneously always guarantees optimal vision and cleaning of the endoscope. Another advantage is evident during introduction of the flap in case of liquid in the middle ear: suction by the second hand is promptly made. Positioning the graft underneath the anterior annulus with two hands is much easier by avoiding its wrinkling, and application of gelatin sponges under the graft itself is much easier. Finally, in one-handed surgery, the scope often has blurred vision due to blood clot or liquid left by hair in the EAC during the frequent introduction and extraction of the scope. The most important advantages of the use of a holding system are control of bleeding and shorter duration of surgery.

The endoscopic procedure consists of: (1) application of the endoscope holder on the operating table in front of the surgeon (**Figure 24**), (2) positioning of the endoscope at the mid level of the posterior part of the external auditory canal, (3) refreshing the margins of the perforation using a sickle knife and grasping forceps, (4) elevating a medial tympanomeatal flap with a semilunar incision at 12 and 6 o'clock, (5) inserting the graft under the malleus and the anterior margin of the perforation, and (6) applying gelatin sponges in the middle ear and, after repositioning the flap, in the ear canal. **Figure 25** shows the different steps of the endoscopic surgery and how it is handled bimanually, offering a clear advantage over a single-handed procedure.

The endoscopes used are 3 and 4 mm in diameter rigid 0° (Hopkins KARL STORZ GmbH & Co. Tuttlingen Germany), lengths 14 and 18 cm, respectively. The optic holder used is a mechanical articulating holding system (28,272 HC; 28,272 UGK; 28,172 HR: KARL STORZ GmbH & Co. Tuttlingen Germany) (**Figure 26**) or the Unitrack pneumatic holding system (Unitrac arm, RT040R, Aesculap AG, Tuttlingen Germany (**Figure 27**). All procedures are performed under general anesthesia.

The surgical maneuvers are managed better using a 3 mm vs. a 4 mm endoscope; according to our experience, we would recommend the 3 mm thanks to the greater space offered. The reason for using an 18 and 14 cm endoscope and not 6 or 11 cm

Figure 24.
Position of holder in front of the surgeon and screen applying the same concept of standard microscope surgery using both hands.

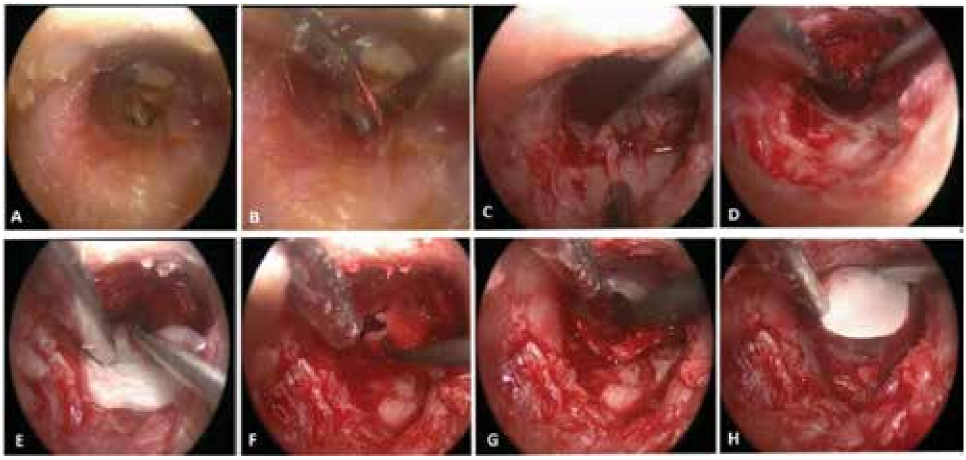

Figure 25.
(A–H) Steps of double-handed myringoplasty in an inferior perforation of the tympanic membrane.

Figure 26.
Mechanical holding system.

Figure 27.
Pneumatic holding system.

Figure 28.
Observe the working distance and instrument length by using 14 and 18 cm long endoscopes.

long is the possibility to maneuver both hands around the scope without encountering any obstacle by the camera and handle of the holder (**Figure 28**). The surgical instruments used for the microscope technique fit well with this technique.

There are some minor limitations of the technique: the endoscope is fixed in the canal, allowing a limited range of zooming and focusing. At the beginning of the procedure, the best view is to completely observe the ear canal, and then gradually magnify the middle ear throughout the surgery. Another limitation is that during the introduction of the graft in the EAC in some cases, the scope should be slightly pushed outward by the surgeon, which allows to see the graft lying completely on the posterior wall and therefore gliding it all the way through to the middle ear. It would be of a great help to have a camera with a foot pedal remote control in order to dynamically change the magnification and focus during the procedure and a motorized holder with fine movements to better manage the visual field.

8. Learning curve aspects

Learning curve varies upon both the complexity of otological procedure and prior experience of the otosurgeon with endoscopic sinus surgery. Furthermore, the profile of the learning curve is expressed through two factors.

The first factor is the variation of duration of the same surgery through time. According to our experience, surgical duration decreased with time in all types of endoscopic surgeries especially in ventilation tube positioning and tympanoplasty. For cholesteatoma endoscopic removal, since it was the last to introduce in our daily practice and therefore we were already acquainted with otoendoscopic surgery, duration of surgery did not undergo a tangible reduction.

The second factor is expressed by the anatomical results obtained through time. Concerning tympanoplasty, with time, we have experienced less reperforations in the reconstruction of the successive tympanic membranes. For cholesteatoma surgery, we become more confident in cholesteatoma removal from the hidden parts of the tympanic cavity, that is, sinus tympani and epitympanum. Such an experience permitted us to approach the middle ear less invasively such as creating less intracanal atticotomy or removing intact ossicles.

Concerning ossiculoplasties, stapedotomies, and cochlear implantation, it is preferable that they are advised after having obtained a certain confidence with otoendoscopic surgery.

9. Surgical training and young surgeons

Otosurgery belongs to the micro-surgery procedures. In order to progress through various techniques of otosurgery, the microscope is fundamental to adequately understand the anatomy in its three dimensions. For a beginner, the best choice is to proceed training with both techniques contemporary in order to understand the different perspective of the middle ear. Not all surgeries could be completely done by endoscope, and also in case of a complication or bleeding, a microscope surgery is necessary. For experienced otosurgeons, it is preferable to start their endoscopic surgery with ventilation tube placement then move to myringoplasty and therefore to cholesteatoma removal.

10. Conclusions

Endoscopy as a diagnostic and surgical tool in otology is now a valid instrument for different pathologies of the ear. Indications for solely endoscopic surgery should be well planned. Before starting endoscopic surgery, one should have good preparation in using the operating microscope and in certain surgeries should always be present in the operating theater. There are still some limits to endoscopy that in the future will be overcome thanks to technological advances such as miniaturization of 3D systems.

Conflict of interest

No conflict of interest is present.

Author details

Luca Oscar Redaelli de Zinis[1]* and Nader Nassif[2]

1 Department of Medical and Surgical Specialties, Radiological Sciences and Public Health, Section of Head and Neck Surgery, University of Brescia and Pediatric Otolaryngology Head Neck Surgery, Children's Hospital "ASST Spedali Civili", Brescia, Italy

2 Pediatric Otolaryngology Head Neck Surgery, Children's Hospital "ASST Spedali Civili", Brescia, Italy

*Address all correspondence to: luca.redaellidezinis@unibs.it

References

[1] Thomassin JM. Endoscopically Assisted Diagnostics and Surgery in Otology, Technique, Equipment and Indications. Vol. 1. Tuttlingen: Endo–Press; 1988

[2] Abou-Elhamd KE. Telescopic myringotomy and tube application. The Journal of Laryngology and Otology. 2000;**114**:581-583

[3] Lee FP. An alternative use of video-telescopic guidance for insertion of myringotomy tube. The Journal of Laryngology and Otology. 2006;**120**:e10

[4] Ayache S. Cartilaginous myringoplasty: The endoscopic transcanal procedure. European Archives of Oto-Rhino-Laryngology. 2013;**270**:853-860

[5] Vrabec JT, Deskin RW, Grady JJ. Meta-analysis of pediatric tympanoplasty. Archives of Otolaryngology Head and Neck Surgery. 1999;**125**:530-534

[6] Yoon TH, Park SK, Kim JY, Pae KH, Ahn JH. Tympanoplasty, with or without mastoidectomy, is highly effective for treatment of chronic otitis media in children. Acta Oto-Laryngologica, Supplement. 2007;**558**:44-48

[7] Caylan R, Titiz A, Falcioni M, De Donato G, Russo A, Taibah A, et al. Myringoplasty in children: Factors influencing surgical outcome. Journal of Otolaryngology Head & Neck Surgery. 1998;**118**:709-713

[8] Singh GB, Sidhu TS, Sharma A, Singh N. Tympanoplasty type I in children-an evaluative study. International Journal of Pediatric Otorhinolaryngology. 2005;**69**:1071-1076

[9] Pignataro L, Grillo Della Berta L, Capaccio P, Zaghis A. Myringoplasty in children: Anatomical and functional results. The Journal of Laryngology and Otology. 2001;**115**:369-373

[10] Umapathy N, Dekker PJ. Myringoplasty: Is it worth performing in children? Archives of Otolaryngology Head and Neck Surgery. 2003;**129**:1053-1055

[11] Uyar Y, Keleş B, Koç S, Oztürk K, Arbağ H. Tympanoplasty in pediatric patients. International Journal of Pediatric Otorhinolaryngology. 2006;**70**:1805-1809

[12] Al-Khtoum N, Hiari MA. Myringoplasty in children: Retrospective analysis of 35 cases. Brazilian Journal of Otorhinolaryngology. 2009;**75**:371-374

[13] Ribeiro JC, Rui C, Natercia S, Jose R, Antonio P. Tympanoplasty in children: A review of 91 cases. Auris, Nasus, Larynx. 2011;**38**:21-25

[14] Knapik M, Saliba I. Pediatric myringoplasty: A study of factors affecting outcome. International Journal of Pediatric Otorhinolaryngology. 2011;**75**:818-823

[15] Saliba I. Hyaluronic acid fat graft myringoplasty: How we do it. Clincal Otolaryngology. 2008;**33**:610-614

[16] Friedberg J, Gillis T. Tympanoplasty in childhood. The Journal of Otolaryngology. 1980;**9**:165-168

[17] Carr MM, Poje CP, Nagy ML, Pizzuto MP, Brodsky LS. Success rates in paediatric tympanoplasty. The Journal of Otolaryngology. 2001;**30**:199-202

[18] Lau T, Tos M. Tympanoplasty in children. An analysis of late results. The American Journal of Otology. 1986;**7**:55-59

[19] Koch WM, Friedman EM, McGill TJ, Healy GB. Tympanoplasty in children.

The Boston Children's Hospital experience. Archives of Otolaryngology Head and Neck Surgery. 1990;**116**:35-40

[20] Gersdorff M, Garin P, Decat M, Juantegui M. Myringoplasty: Long-term results in adults and children. The American Journal of Otology. 1995;**16**:532-535

[21] D'Eredità R, Lens MB. Anterior tab flap versus standard underlay myringoplasty in children. Otology and Neurotology. 2009;**30**:777-781

[22] Nassif N, Redaelli De Zinis LO, Berlucchi M, Zanetti D. Endoscopic ventilation tube placement in the pediatric age. Clincal Otolaryngology. 2014;**39**:50-53

[23] Dündar R, Kulduk E, Soy FK, Aslan M, Hanci D, Muluk NB, et al. Endoscopic versus microscopic approach to type 1 tympanoplasty in children. International Journal of Pediatric Otorhinolaryngology. 2014;**78**:1084-1089

[24] Tarabichi M. Endoscopic middle ear surgery. The Annals of Otology, Rhinology, and Laryngology. 1999;**108**:39-46

[25] Tarabichi M. Endoscopic management of limited attic cholesteatoma. The Laryngoscope. 2004;**114**:1157-1162

[26] Shohet JA, de Jong AL. The management of pediatric cholesteatoma. Otolaryngologic Clinics of North America. 2002;**35**(4):841-851

[27] Marchioni D, Soloperto D, Rubini A, Villari D, Genovese E, Artioli F, et al. Endoscopic exclusive transcanal approach to the tympanic cavity cholesteatoma in pediatric patients: Our experience. International Journal of Pediatric Otorhinolaryngology. 2015;**79**(3):316-322

[28] Badr-el-Dine M. Value of ear endoscopy in cholesteatoma surgery. Otology and Neurotology. 2002;**23**(5):631-635

[29] Ayache S, Tramier B, Strunski V. Otoendoscopy in cholesteatoma surgery of the middle ear: What benefits can be expected? Otology and Neurotology. 2008;**29**(8):1085-1090

[30] Barakate M, Bottrill I. Combined approach tympanoplasty for cholesteatoma: Impact of middle-ear endoscopy. The Journal of Laryngology and Otology. 2008;**122**(2):120-124

[31] Tarabichi M, Nogueira JF, Marchioni D, Presutti L, Pothier DD, Ayache S. Transcanal endoscopic management of cholesteatoma. Otolaryngologic Clinics of North America. 2013;**46**(2):107-130 (Review)

[32] Lecler A, Lenoir M, Peron J, Denoyelle F, Garabedian EN, Pointe HD, et al. Magnetic resonance imaging at one year for detection of postoperative residual cholesteatoma in children: Is it too early? International Journal of Pediatric Otorhinolaryngology. 2015;**79**(8):1268-1274

[33] Adunka OF, Buchman CA. Scala Tympani Cochleostomy I: Results of a survey. The Laryngoscope. 2007;**117**(12):2187-2194

[34] Adunka OF, Radeloff A, Gstoettner WK, Pillsbury HC, Buchman CA. Scala Tympani Cochleostomy II: Topography and histology. The Laryngoscope. 2007;**117**(12):2195-2200

[35] Paprocki A, Biskup B, Kozłowska K, Kuniszyk A, Bien D, Niemczyk K. The topographical anatomy of the round window and related structures for the purpose of cochlear implant surgery. Folia Morphologica. 2004;**63**(3):309-312

[36] Leong AC, Jiang D, Agger A, Fitzgerald-O'Connor A. Evaluation

of round window accessibility to cochlear implant insertion. European Archives of Oto-Rhino-Laryngology. 2013;**270**(4):1237-1242

[37] Hamamoto M, Murakami G, Kataura A. Topographical relationships among the facial nerve, chorda tympani nerve and round window with special reference to the approach route for cochlear implant surgery. Clinical Anatomy. 2000;**13**(4):251-256

[38] Angeli RD, Lavinsky J, Setogutti ET, Lavinsky L. The crista fenestra and its impact on the surgical approach to the scala tympani during cochlear implantation. Audiology and Neurotology. 2017;**22**(1):50-55

[39] Stidham KR, Roberson JB. Cochlear hook anatomy: Evaluation of the spatial relationship of the basal cochlear duct to middle ear landmarks. Acta Otolaryngologica. 1999;**119**(7):773-777

[40] Atturo F, Barbara M, Rask-Andersen H. On the anatomy of the "hook" region of the human cochlea and how it relates to cochlear implantation. Audiology and Neurotology. 2014;**19**(6):378-385

[41] Lloyd SK, Kasbekar AV, Kenway B, Prevost T, Hockman M, Beale T, et al. Cochlear rotation and its relevance to cochlear implantation. Cochlear Implants International. 2010;**11**(3):57-63

[42] Tarabichi M, Nazhat O, Kassouma J, Najmi M. Endoscopic cochlear implantation: Call for caution. The Laryngoscope. 2016;**126**:689-692

[43] Chen YH, Liu TC, Yang TH, Lin KN, Wu CC, Hsu CJ. Using endoscopy to locate the round window membrane during cochlear implantation: Our experience with 25 patients. Clincal Otolaryngology. 2018;**43**:357-362

[44] Güneri EA, Olgun Y. Endoscope-assisted cochlear implantation. Clinical and Experimental Otorhinolaryngology. 2018;**11**:89-95

[45] James AL. Endoscopic middle ear surgery in children. Otolaryngologic Clinics of North America. 2013;**46**:233-244

[46] Presutti L, Marchioni D, Mattioli F Villari D, Alicandri-Ciufelli M. Endoscopic management of acquired cholesteatoma: Our experience. Journal of Otolaryngology Head & Neck Surgery. 2008;**37**:481-487

Chapter 7

The Frequency Following Response: Evaluations in Different Age Groups

Milaine Dominici Sanfins, Michele Vargas Garcia,
Eliara Pinto Vieira Biaggio and Piotr Henryk Skarzynski

Abstract

In this chapter, recent data on the clinical application of the frequency following response (FFR) in different age groups will be presented. The chapter begins with the importance of using speech sounds in electrophysiological assessments. Then the FFR methodology is presented, giving normative data and the expected responses in different age groups: infants and young children, children and adolescents, and adults and the elderly. Finally, the unique responses of each age group are presented in order to show how this new technology can be an extremely useful tool for diagnosing hearing dysfunction.

Keywords: electrophysiology, frequency following response, speech perception, hearing, auditory evoked potential

1. Introduction

Until recently, electrophysiological evaluations were performed exclusively with nonverbal stimuli such as clicks and tone bursts which allow rapid and synchronous stimulation of neurons. However, the use of verbal stimuli, such as speech sounds, allows a more accurate analysis of the auditory system, especially if the aim is to investigate how the system decodes speech sounds involved in daily communication. Verbal and nonverbal stimuli are decoded in different ways and follow different trajectories through the central auditory nervous system.

Human communication consists predominantly of verbal stimuli, and it is important to understand how verbal sounds are coded at various levels of the auditory system. The need to develop research methods that are objective and accurately represent daily listening led to the development early this century of electrophysiological tests for measuring how speech sounds are perceived [1, 2]. Subsequently, a number of research groups have focused their efforts on using complex stimuli such as speech for diagnostic purposes [3–10].

The initial studies were performed in animal models [11] aiming to evaluate how the temporal and spectral properties of verbal stimuli were coded, and later human responses were also analyzed [12]. Among the electrophysiological procedures for investigating the processing and coding of verbal sounds, we highlight the frequency following response (FFR).

2. Frequency following response

Acquisition of an FFR is very similar to collecting an ABR with a click stimulus. However, interpretation of an FFR requires that the audiologist has a more sophisticated knowledge base. Speech stimuli allow a more complex analysis of the responses, such as their:

- timing;
- magnitude;
- frequency content and magnitude;
- frequency tracking;
- phase consistency;
- intrinsic factors; and
- difference between individual responses.

An FFR evaluation can be performed on different clinical populations and age groups, and below we give details of how the procedure varies depending on the patient's age. Because FFR is a relatively new procedure, initial work was done on adult subjects. Afterward, researchers turned their interest to the study of responses in infants and young children, children and adolescents, and the elderly.

In order for an FFR assessment to be useful in identifying auditory disorders at an early stage, normative values using different equipment and recording parameters need to be established and compared with language acquisition markers.

The distinctive features of FFRs in different age groups will be presented in three parts:

- evaluation in infants;
- evaluation in children and adolescents;
- evaluation in adults and the elderly.

3. Frequency following response: evaluation in infants

In clinical practice, a comprehensive hearing evaluation for infants and young children is essential, since the integrity of their auditory system is the basis for acquiring oral language. In this context, if one measures only the functioning of the peripheral auditory pathway, perhaps by recording and analyzing otoacoustic emissions and/or auditory brainstem evoked potentials, it significantly constrains one's knowledge of the patient's hearing status. Moreover, behavioral assessments of hearing in very young children are often inconclusive, considering the diversity of neuropsychomotor development in this age group.

The perception of speech is important for the development of receptive and expressive language [13]. Through auditory experiences, infants and toddlers acquire and master the linguistic elements necessary for effective communication. The experiences are associated with information from the other senses, and together they allow the acquisition and development of oral language. Through listening, the subject understands oral language and creates concepts, finally inter-relating them and expressing them through speech [14]. Thus, the importance of hearing for the acquisition and development of language is vital, and any disturbance to the auditory pathway has implications for oral communication as a whole [14].

FFR testing can be used with infants and young children as a predictor of the extent of future language appropriation—in other words as a way of identifying children who are at risk of deficits in oral language acquisition [2, 15]. Assessment by FFR of infants and young children is relatively recent, and published studies of its potential have only been done over the last decade. Before discussing what is known about FFR in this population, it is first necessary to clarify an important factor: maturation of the auditory pathway.

It is known that peripheral hearing is functional even before birth, whereas myelination and the organization of neural connections keep developing after birth [16, 17]. Indeed, the central structures, such as the subcortex and cortex, develop throughout the early years of human life. There is an ascending myelination of the auditory pathway, evidenced by magnetic resonance imaging. Up to the 13th week of life, there is an increase in myelination density of the cochlear nucleus, the superior olivary complex, and the lateral lemniscus, with the inferior colliculus demonstrating an increase in density around the 39th week of life [18]. This continuous process of myelination of the higher structures of the auditory pathway during the first year of life must be considered when evaluating the FFR, for it means that the lower the age of the evaluated subject, the greater the latency of the FFR waves [19, 20]. This increase in latency can also be seen in other auditory evoked potentials [21]. An FFR can be recorded from a neonate, but the responses only become readily apparent from the third month of age [15]. The existence of a series of FFR waves—V, A, C, D, E, F, and O—in neonates has been pointed out by several researchers [15, 19, 22–26]. FFR evaluations have been performed with the vowel /i/ [15, 24], the syllables /ba/ and /ga/ [26], and the syllable /da/ [23].

The FFR has been studied in neonates of different nationalities (Chinese, American) during the first days after birth, and the FFRs were nearly the same. This finding makes it possible to infer that, independent of the mother tongue, there is an innate capacity for speech coding in neonates at the subcortical level [22].

The evaluation of subcortical representation of speech coding was studied by evaluating FFRs in 28 healthy North American infants, 3–10 months of age. The study focused on the fundamental frequency (F0), the response time of the FFR, and the representation of harmonics. To analyze the data in the frequency domain, spectral amplitudes were calculated by fast Fourier transform (FFT) and divided into three frequency ranges: F0, 103–125 Hz; first formant (F1), 220–720 Hz; and high harmonics (HH), 720–1120 Hz. The F0 responses were more robust in infants 3 months of age and the amplitude of F0 did not show significant changes over the entire 6 months. For the F1 and HH frequencies, there was a rapid and systematic increase of amplitude from 3 to 6 months of age.

To analyze the data in the time domain, the peaks were identified manually and confirmed by a second observer. Waves I, III, and V were first identified in response to a click, and then, in the FFR, the same peak and following valley (V and A), the peaks (D, E, and F), and the displacement peak (O). Non-detectable peaks were marked as missing data points and were excluded from analysis. The latencies and amplitudes (baseline to peak) were extracted from the identified waves. The time domain analysis demonstrated a decrease in neural conduction time and an improvement in amplitude with increasing age. The latencies of A and O, the time interval between A and O, and the slope between V and A were shown to have a negative correlation between latency and age. In addition, there was an improvement in the morphology of all waves as age increased. It was also observed that infants 3–5 months of age had longer latencies, smaller intervals between A and O, and a lower V/A slope compared to those 6–10 months of age. This negative correlation between the latencies and the age of the infants, as well as the decrease of slope in the smaller children, is due to a maturational process occurring in the subcortical auditory system and shows that there is less neural synchrony in younger infants [23]. The authors also note that these findings indicate that at approximately 6 months of age, the coding of speech characteristics, both spectrally and temporally, becomes more like those of an adult, although the changes continue through to school age. These findings indicate that FFR evaluation can detect early disorders in the perception of speech sounds.

The researchers also investigated the development of subcortical speech processing in Chinese infants born in households in which the mother tongue was Mandarin. They recorded FFRs at two ages: 1–3 days of life and at 3 months. This prospective-longitudinal design study included only infants who had undergone auditory screening at birth, who had no obvious neurological disorders, and did not have any risk indicator for hearing loss. Initially, 44 newborns were tested by FFR during natural sleep. After that, the sample was divided into groups. For each group, the researchers selected different speech stimuli for the evaluation of FFR (monosyllables contrasting with Mandarin). Only 13 infants completed the follow-up protocol at the third month. The processing and tracking of the fundamental frequencies of human speech at the subcortical level, evidenced by the FFR, showed more robust responses when the babies were 3 months old. Researchers acknowledged the limitations of the study, including statistical analysis and data interpretation. A research weakness was the relatively low completion rate (i.e., 17/44 infants or 38.64%). This factor undermined the power of the conclusions and prevented the possibility of performing statistical analyses for each Mandarin tone used. Despite the limitations of the study, the findings fill a gap in understanding the developmental trajectory of subcortical processing during the first 3 months of life [25].

From the theoretical assumptions highlighted in the previous reference, it should be noted that the linguistic environment of a newborn has a substantial effect on the development of its speech perception. Even at birth, children are able to detect subtle differences in verbal sounds. Newborns can effectively differentiate all the features of human speech and most infants who participated in an FFR follow-up showed improvement in pitch tracking and response amplitudes at 3 months of age [25]. Such neural refinements observed by FFR are often highlighted in the literature for both infants [22, 24] and young infants [15, 23]. For example, in a longitudinal case report of one infant, the researchers obtained FFR records when the infant was 1, 3, 5, 7, and 10 months old. The results showed an evolving trajectory of development with a transition point of about 3 months [15].

Using FFR evaluation in preterm infants may also be an alternative for the early diagnosis of auditory disorders in this population related to the perception of speech sounds. Premature babies are at high risk of developing language disorders, so using FFR may be a way of measuring immature neural activity and predicting possible changes in the processing of verbal sounds. In order to do so, one study evaluated 12 premature Indian infants through FFR with the aim of exploring how an immature auditory system responds to complex acoustic stimuli such as speech [27]. Peaks V, A, C, D, E, and F were detected in almost all babies and with latencies and amplitudes similar to those reported in the literature. The waves could be replicated. The authors conclude that FFR may be a way of understanding how the human brain-stem receives speech signals and that such an assessment might be important for all high-risk babies. Although the findings of this study cannot be generalized, mainly due to the limited data (small sample and absence of a controls, among others), they point out the potential of FFR in evaluating infants from neonatal intensive care units.

More recently, studies that record FFRs in the presence of background noise have been published. It is known that competing noise can make speech comprehension more difficult in people of all ages. Speech-in-noise tests are clinically available but cannot be given to infants. Thus, the use of FFRs in noise may be an alternative for evaluating impaired speech perception in young children who are unable to respond to behavioral tests.

In this context, with the objective of examining the electrophysiological responses in the presence of noise, researchers have evaluated the FFR in 30 children with typical development under conditions with and without noise (a signal-to-noise ratio of +10 dB in the former) [28]. Babies were divided into two age groups: 7–12 and 18–24 months. For all infants, frequency analysis of the FFR with a Fourier transform was performed, analyzing the latency and amplitude of waves V, A, D, E, and F, and correlation tests were carried out. In both groups, the mean latency of all recorded waves was higher in the presence of noise. According to the authors, this suggests that, at least for infants up to 24 months, the presence of noise causes a delay in the appearance of FFR waves independent of age. In addition, they observed a greater amplitude of F0 in the noise condition in the group of older babies; this difference was not seen in the silent condition. Thus, the authors point out that, at 2 years of age, infants are less vulnerable to the degrading effects of noise compared to children younger than 12 months.

The development of phase lock and frequency representation has also been evaluated in infants. This was the focus of a study that included an initial sample of 56 typical babies, aged between 2 and 12 months, and evaluated the FFR with /ba/ and /ga/ stimuli presented in the right ear using the SmartEP equipment from Intelligent Hearing Systems [26]. These responses were also obtained in young adults to provide a reference for the course of development of neural synchrony (represented by phase lock) and response amplitude (represented by spectral magnitude). The results obtained in this study demonstrate that the strength of phase-lock in the fine structure at CV transition is higher in young adults compared to infants. However, phase lock for F0 was equivalent between adults and infants. The frequency of F0 was found to be higher in older infants compared to younger infants and adults. Thus, these data demonstrate that speech coding can be evaluated in infants from 2 months of age and that such data are of value in a clinical setting, since it is known that performing electrophysiological evaluation of hearing in young children is difficult because they are less able to remain still during a test. The data indicate that the FFR may be a way of testing babies who are at risk of developing a language disorder, examining the auditory coding mainly of the midbrain, but also reflecting contributions from the auditory nerve, brain stem, and cortex.

The most commonly used parameters in FFR evaluations are: monoaural stimulus, right ear stimulation, intensity of 80 dB SPL, syllable /da/ speech stimulus, alternating polarity, presentation rate of 10.9 stimuli per second, vertical placement of electrodes, insert headphones, and the subject sitting distracted or awake during recording [29].

Regarding the latency parameters, when FFR is done with the Navigator Pro AEP System (Natus Medical, Inc.) and a syllable stimulus, one group of researchers [19] pointed out that in 23 normal-hearing babies (0–12 months) the wave latencies were on average: V = 7.22 ms, A = 8.22 ms, D = 23.14 ms, E = 31.5 ms, F = 39.91 ms, and O = 49.64 ms. FFR wave latencies were also investigated in 53 children aged 3–5 years (**Tables 1** and **2**).

Parameters of FFR evaluation in infants and young children used in the Hearing Electrophysiology Service of the Federal University of Santa Maria, Brazil, are presented in **Table 3**.

	Waves					
	V	**A**	**D**	**E**	**F**	**O**
	Lat	**Lat**	**Lat**	**Lat**	**Lat**	**Lat**
∑	7.22	8.22	23.14	31.51	39.91	49.64
SD	0.42	0.43	0.66	0.49	0.45	1.32
Detect (%)	86.9	86.96	91.30	91.30	82.61	65.22

∑: average (ms), SD: standard deviation, Detect: the percent detectability for each peak.
Sample: 23 babies (0–1 years old).

Table 1.
FFR latency values using syllable /da/of 40-ms duration performed on babies with normal hearing (silent background) [19].

	Waves					
	V	**A**	**D**	**E**	**F**	**O**
	Lat	**Lat**	**Lat**	**Lat**	**Lat**	**Lat**
∑	6.59	7.56	22.36	30.90	39.34	48.14
SD	0.26	0.35	0.38	0.37	0.32	0.42
Detect	100	100	88.67	98.11	100	90.57

∑: average (ms), SD: standard deviation, Detect: the percent detectability for each peak.
Sample: 53 children (3–5 years old).

Table 2.
FFR latency values using syllable /da/ of 40-ms duration performed in children with normal hearing (in silence) [19].

Presentation parameters	**Setting**
Equipment	SmartEP, Intelligent Hearing Systems (IHS)
Transducer	Insert phones
Electrodes	Fz; Fpz; M1; M2 or Cz, M1, M2
Stimulation	Right ear
Stimulus	Syllable /da/
Duration of stimulus	40 ms
Presentation rate	10.9/s
Window	80–100 ms
Filter	Low pass of 100 Hz and high pass of 2000 Hz Low pass of 100 Hz and high pass of 3000 Hz
Polarity	Alternating
Intensity	80 dBnHL
Number of stimuli	6000
Reproducibility	2 × 3000 stimuli
Condition of evaluation	Awake and quiet
Impedance	3k Ohms
Artifact rejection	Acceptance if <10%

ms, millisecond; s, second; Hz, hertz; dB, decibel; HL, hearing level.

Table 3.
Parameters of FFR in infants and young children.

Source	Latency (ms)		Amplitude (μV)		VA measures	
	Σ	SD	Σ	SD	Σ	SD
V	6.61	0.25	0.31	0.15		
A	7.51	0.34	0.65	0.19		
C	17.69	0.48	0.36	0.09		
F	39.73	0.61	0.43	0.19		
Slope VA (μV/ms)					0.13	0.05
Area VA (μV × ms)					1.70	1.23

∑: average, SD: standard deviation.
Sample: 36 and 38 children and adolescents (8–12 years old) with normal hearing.

Table 4.
FFR latency and amplitude values using the syllable/da/of 40-ms duration, performed in children with normal hearing on the right ear (silent conditions) [12].

		Waves													
		V		A		C		D		E		F		O	
	Sex	Lat	Amp	Lat	Amp	Lat	Amp	Lat	Amp	Lat	Amp	Lat	Amp	Lat	Amp
Σ	M	6.53	0.10	7.53	0.19	18.43	0.08	22.29	0.17	30.86	0.21	39.31	0.17	48.02	0.13
	F	6.49	0.13	7.43	0.23	18.33	0.12	22.28	0.15	30.81	0.29	39.27	0.24	47.95	0.21
Med	M	6.49	0.10	7.53	0.18	18.28	0.07	22.24	0.09	30.86	0.21	39.28	0.7	48.11	0.13
	F	6.49	0.12	7.37	0.22	18.37	0.09	22.11	0.13	30.78	0.22	39.11	0.24	47.86	0.21
SD	M	0.19	0.05	0.32	0.04	0.44	0.05	0.32	0.07	0.53	0.07	0.44	0.08	0.45	0.07
	F	0.22	0.07	0.35	0.90	0.44	0.11	0.67	0.09	0.58	0.35	0.56	0.26	0.75	0.28

∑: average, Med: median, SD: standard deviation, M: male, F: female.
Sample: 40 children and adolescents (8–16 years old).

Table 5.
FFR latency and amplitude values for males and females using syllable /da/ of 40-ms duration performed in children with normal hearing (silent conditions) [30].

		Complex VA	
	Sex	Slope VA (ms/μV)	Area VA (ms × μV)
∑	M	0.31	0.29
	F	0.39	0.34
Med	M	0.29	0.31
	F	0.36	0.31
SD	M	0.11	0.09
	F	0.14	0.14

∑: average, Med: median, SD: standard deviation, M: male, F: female.
Sample: 40 children and adolescents (8–16 years old).

Table 6.
Complex VA (slope and area) values for males and females using syllable/da/of 40-ms duration performed in children with normal hearing (silent conditions) [30].

The early identification of hearing disorders through FFR evaluation allows a speech-language pathologist to intervene, lessening the damage that this disorder can have on the development of speech skills in early childhood [2, 20, 22, 31]. This

		Waves													
		V		**A**		**C**		**D**		**E**		**F**		**O**	
	Ear	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**
Σ	R	6.50	0.12	7.46	0.22	18.33	0.10	22.21	0.14	30.89	0.30	39.37	0.24	48.00	0.21
	L	6.51	0.11	7.48	0.21	18.41	0.11	22.36	0.13	30.78	0.23	39.20	0.19	47.95	0.16
Med	R	6.45	0.12	7.45	0.21	18.33	0.08	22.12	0.14	30.86	0.23	39.24	0.19	47.99	0.15
	L	6.53	0.11	7.41	0.21	18.33	0.09	22.28	0.11	30.78	0.21	39.07	0.18	48.03	0.15
SD	R	0.21	0.06	0.33	0.09	0.42	0.08	0.66	0.09	0.50	0.39	0.55	0.29	0.75	0.30
	L	0.21	0.06	0.36	0.07	0.46	0.10	0.44	0.08	0.61	0.09	0.47	0.09	0.54	0.12

∑: average, Med: median, SD: standard deviation, R: right, L: left.
Sample: 40 children and adolescents (8–16 years old).

Table 7.
FFR latency and amplitude values for right and left ears using syllable/da/of 40-ms duration performed on children with normal hearing (silent conditions) [30].

	Complex VA		
	Ear	**Slope VA (ms/µV)**	**Area VA (ms × µV)**
Σ	R	0.37	0.33
	L	0.34	0.31
Med	R	0.32	0.31
	L	0.32	0.31
SD	R	0.14	0.13
	L	0.13	0.13

∑: average, Med: median, SD: standard deviation, R: right, L: left.
Sample: 40 children and adolescents (8–16 years old).

Table 8.
Complex VA (slope and area) values for right and left ears using syllable/da/of 40 ms duration performed on children with normal hearing (silent conditions) [30].

		Waves													
		V		**A**		**C**		**D**		**E**		**F**		**O**	
	Age range	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**	**Lat**	**Amp**
Σ	8–11	6.53	0.12	7.44	0.22	18.37	0.11	22.26	0.15	30.80	0.25	39.34	0.21	47.95	0.17
	12–16	6.46	0.11	7.51	0.21	18.36	0.10	22.32	0.10	30.89	0.28	39.19	0.21	48.02	0.21
Med	8–11	6.53	0.11	7.45	0.21	18.37	0.09	22.20	0.14	30.78	0.23	39.28	0.20	47.95	0.15
	12–16	6.45	0.12	7.45	0.17	18.28	0.08	22.20	0.09	30.86	0.20	39.11	0.15	48.03	0.13
SD	8–11	0.23	0.06	0.32	0.10	0.46	0.09	0.53	0.08	0.62	0.19	0.56	0.11	0.75	0.14
	12–16	0.17	0.06	0.37	0.07	0.41	0.08	0.63	0.45	0.43	0.22	0.42	0.32	0.46	0.33

∑: average, Med: median, SD: standard deviation, R: right, L: left.
Sample: 40 children and adolescents (8–16 years old).

Table 9.
FFR latency and amplitude values for various age ranges using syllable/da/of 40-ms duration performed on children with normal hearing (silent conditions) [30].

assertion can be understood by appreciating the relationship between language development and the presence of stimulating auditory experiences in the first few months of life.

Future studies evaluating FFRs in infants will no doubt benefit from interdisciplinary collaboration which seeks to deepen understanding of the underlying mechanisms involved in the typical and atypical development of the auditory system during early childhood.

4. Frequency following response: evaluation in children and adolescents

Auditory impairment is almost invariably associated with language and communication deficits. Learning a spoken language depends on assimilating the acoustic and phonetic elements of a language [32]. The development of the central auditory nervous system begins in intrauterine life and continues until adolescence, over which time hearing abilities become more complex and elaborate.

Because of the close relationship between hearing, language, and learning, it is extremely important to monitor hearing over the course of life. Especially in children, be it pre-school or school age, the aim should be to monitor auditory function, either through behavioral or electrophysiological assessments. The ideal would be a combination of both behavioral and electrophysiological methods, so that with numerous evaluations there are crosschecks which allow a more accurate diagnosis to be made.

The electrophysiological procedure traditionally used in clinical practice is the click ABR. However, in evaluating children with language deficits, this type of sound stimulus is not ideal for making diagnoses. Assessments using verbal sound stimuli, such as used in FFR, appear to be more effective and reliable in cases of learning problems or school difficulties [6]. Evaluation via an FFR allows a detailed analysis of how verbal stimuli are encoded in the central auditory nervous system to be done.

The FFR allows fine-grained auditory processing deficits associated with real-world communication skills to be identified. As well as being used for the early identification of auditory processing, it can also be used to assess hearing across different clinical populations [33, 34]. This electrophysiological procedure can provide reliable and objective information about acoustic patterns such as timing, pitch, and timbre [35]. These three elements can be evaluated using different parts of the FFR, as follows:

- timing—via analysis of the onset and offset portions;
- pitch—by analysis of the fundamental frequency (F_0);
- timbre—from analysis of the harmonics of F_0.

Simplistically, it can be said that the FFR helps in understanding which speech sounds were spoken (their timing and harmonic cues) and who said it (pitch cues) [36]. In addition, an FFR test can be performed under two conditions: (i) in silence (presentation of verbal stimuli only), and (ii) in noise (presentation of verbal stimuli plus background noise).

In children and adolescents, studies have shown that FFRs change in latency as age increases. FFRs of children aged around 5 years appear to be very similar to the responses of children aged 8–12. However, the FFR pattern of children under 5 years has a somewhat different morphology and latency. According to Johnson et al. [33], the differences in children younger than 3 years are more evident in the initial portion of the responses (the onset), while in older children the change is more evident in the final portion (the offset) [3, 37].

Initial studies have focused on understanding the FFRs in children and adolescents under silent conditions and in subjects who have normal hearing and typical development. For the benefit of clinical audiologists, some of these studies are summarized below (**Tables 4–9**).

Table 10 shows the parameters used in children and adolescents at the Electrophysiology Department of the State University of Campinas using Biologic equipment and BioMARK software.

Because FFR is a new procedure, unstudied pathologies are gradually being added and, little by little, we are gaining new information about what effects the pathologies have on the responses of affected children and adolescents.

The FFRs of children diagnosed as poor readers frequently present as alterations in the timing and magnitude of timbre components [38]. The perception of the duration of a sound stimulus is essential for proficient reading, and the FFR can evaluate or monitor a decline in temporal and spectral precision. Children and adolescents with dyslexia commonly have difficulty perceiving speech sounds either in silence or in competing noise backgrounds. If a child has difficulty in perceiving speech sounds, their reading can be severely impaired [39]. Recently, Sanfins et al. [6] highlighted the importance of FFR as a biological marker in scholastic difficulties.

FFR evaluation in children who have suffered from secretory otitis media in the first 6 years of life, and who have undergone myringotomy for bilateral ventilation tube placement, exhibit changes in their FFR compared to normal children [5]. This study found that evaluating the FFR seems to be a promising method of identifying

Parameter	Settings
Equipment	Biologic Navigator Pro
Software	BioMARK
Electrode montage	Cz, M1, and M2
Stimulated ear	Right ear
Stimulus	Speech
Stimulus type	Syllable /da/
Stimulus duration	40 ms
Stimulus polarity	Alternating
Stimulus intensity	80 dB SPL
Stimulus rate	10.9/s
Number of sweeps	6000
Replicability	Twice for 3000 sweeps
Transducer	Insert
Assessment condition	Watching a movie
Impedance	1k Ohms
Window 85.33 ms	85.33 ms
Filter	100–2000 Hz
Artifact rejection	>10%

Cz: vertex, M1: left mastoid, M2: right mastoid, ms: millisecond, dB: decibel, SPL: sound pressure level, s: second, Hz: hertz.

Table 10.
Parameters of FFR in children and adolescents.

changes in the coding of speech stimuli in these children which might be undetected using traditional electrophysiological evaluation. The changes in their electrophysiological responses might serve as an alert to parents and educators, who can then adopt strategies to minimize the negative consequences on language development and academic achievement.

Another possibility for using FFR assessment may be in monitoring an auditory training program or even tracking the effect of therapeutic interventions. Studies have shown that children with learning disabilities can benefit from an auditory remediation program, and it might therefore be usefully accompanied by FFR examinations (because FFRs have good repeatability in test and retest) [40, 41]. In addition, bilingual children can also be monitored through FFR assessment. Researchers have confirmed that neural perception of speech seems to be more consistent in bilinguals than in monolinguals [42, 43]. Bilingual experience during childhood may favor plasticity in the neuronal coding of sound and improve fundamental frequency perception (F0).

Recently, the neurophysiological aspects of speech perception have been investigated in cases of autism spectrum disorder (ASD). The results showed that children with ASD tend to have changes in the sensation of pitch (frequency), which might explain a withdrawal from speech reception. The fundamental frequency (F0) and its harmonics contain speech information which is essential in conveying affect [44], so changes in FFRs are consistent with a defect in perceiving prosody. The inference is that prosody deficits in some ASD patients may derive from an inability to encode and transmit auditory information in the brainstem [45].

Traditionally, FFR testing is done by presenting verbal stimuli through an insert earphone with a silent background. However, the perception of speech in a noisy background is a much discussed topic. In the presence of noise, normally hearing individuals need to make constant adjustments in their central auditory nervous system to satisfactorily understand and process speech information. Of course, there are others who, in the presence of competing noise, experience great difficulty in understanding speech [46].

The evaluation of FFR in the presence of noise can be effectively used to diagnose children with learning disabilities [47]. Thus, identification of such children could lead to improvements in their reading and writing skills and in daily communication.

5. Frequency following response: evaluation in adults and the elderly

In the adult and elderly population, the need for detailed audiological investigation increases when the patient complains of hearing difficulties, even if auditory thresholds appear normal.

The evaluation of the FFR first involves time and prosody recordings, which provide important information about consonant and vowel discrimination and also aid in the perception of intonation [48]. For adults, but especially in the elderly, participation in these sorts of tests can assist in rehabilitation, either using a hearing aid or auditory training (or both).

The clinical usefulness of the FFR in gauging how well auditory information is being processed is unquestionable. In adults and the elderly, many studies have already been done to identify how the FFR can help in diagnosing complaints related to central auditory processing, thereby allowing better rehabilitation.

The latencies (mean and standard deviation) for adults and the elderly are presented in **Table 11**. The values come from Skoe et al. [19] who used Biologic and

Complex VA			
	Age range	Slope VA (ms/μV)	Area VA (ms × μV)
Σ	8–11	0.38	0.31
	12–16	0.33	0.34
Med	8–11	0.37	0.31
	12–16	0.28	0.31
SD	8–11	0.12	0.11
	12–16	0.16	0.16

Σ: average, Med: median, SD: standard deviation, R: right, L: left.
Sample: 40 children and adolescents (8–16 years old).

Table 11.
Complex VA (slope and area) values for age range using syllable/da/of 40-ms duration performed in children with normal hearing (silent conditions) [30].

Age (years)	Number	Latencies (maximum in milliseconds + 2 SD)					
		V	A	D	E	F	O
17–21	54	7.04	8.15	23.21	31.9	39.50	48.94
21–30	143	7.17	8.28	23.4	32.54	40.84	49.79
30–40	32	7.27	8.39	23.64	32.09	40.38	49.13
40–50	11	7.05	8.22	24.26	31.86	39.93	49.6
50–60	26	7.5	8.77	24.5	32.97	41.46	50.72
60–73	24	7.68	8.81	24.27	32.47	40.60	50.02

Data from [19].
SD: standard deviation.

Table 12.
FFR latency values based on mean values in ***Table 11*** *plus two standard deviations.*

Navigator Pro equipment. In this study, subjects aged between 18 and 72 years and distributed in 6 age brackets were used. In the case of adults, the authors list values for subjects aged 21–30 years (*n* = 143) and found that latency values tended to increase with age. Thus, the researchers emphasized the importance of conducting research on FFRs in different age groups, since normative values can be modified with the aging process.

In **Table 12** the maximum values of each wave are listed by adding two standard deviations to those in **Table 13**. Assuming the distribution is Gaussian means that this measure will cover 95% of the population.

Undoubtedly, the largest number of FFR studies have been performed using the Navigator Pro model from Biologic. Researchers tend to use this equipment together with the Intelligent Hearing Systems and SmartEP software [7, 49, 50].

One study aimed to assess the processing of auditory information in those with hearing loss through an evaluation of eight individuals, aged 46–58 years, with hearing loss [7]. FFRs (collected by SmartEP) were correlated with results from two auditory processing behavioral tests—the masking level difference test and the random gap detection test. No correlation was found between FFR and these tests. The researchers found that the generation of this potential is extremely complex and could encompass several functions and does not depend on just temporal resolution

Age	Number	Latency ∑ (mean in milliseconds)						Standard deviation											
		V	A	D	E	F	O	V (SD)	%	A (SD)	%	D (SD)	%	E (SD)	%	F (SD)	%	O (SD)	%
117–21	54	6.58	7.53	22.41	31.02	39.50	48.26	0.23	100	0.31	96.30	0.40	92.6	0.44	94.44	0.46	98.15	0.34	98.15
221–30	143	6.65	7.60	22.60	31.12	39.61	48.33	0.26	100	0.34	100	0.67	95.8	0.71	100	0.62	99.30	0.73	97.90
330–40	32	6.61	7.53	22.52	31.09	39.54	48.21	0.33	100	0.43	100	0.56	96.88	0.50	96.88	0.42	96.88	0.46	93.75
440–50	11	6.67	7.64	22.84	31.26	39.49	48.30	0.19	100	0.29	100	0.71	90.90	0.30	100	0.22	100	0.65	90.90
550–60	26	6.86	7.89	23.08	31.57	39.92	48.72	0.32	92.31	0.44	92.31	0.71	76.92	0.70	96.15	0.77	92.31	1.00	88.46
660–73	24	6.92	7.89	23.05	31.37	39.68	48.84	0.38	91.67	0.46	91.67	0.61	83.33	0.55	83.33	0.46	83.33	0.59	100

∑: Average (ms), SD: standard deviation, %: percent detectability for each peak.

Table 13.
FFR latency values for syllable /da/ of 40-ms duration, (silence) performed in adults and the elderly with normal hearing [19].

or selective attention [7]. Also seeking to correlate FFRs with hearing loss, Peixe et al. [49] evaluated 11 individuals, aged 23–59 years, with moderately severe hearing loss. They concluded that hearing loss may cause an increase in the FFR wave latency, but the waves are still present so long as the stimulus intensity is adjusted. In other words, the presence of FFR waves is related to the audibility of the signal.

Another interesting study was conducted with 30 young Indian adults aged 18–25 years [50]. The evaluation was carried out with the SmartEP equipment, and FFRs were present in all subjects evaluated. The latency and amplitude values of the analyzed elements were: wave V (lat = 6.81 ms and amp = 0.19 μV), wave C (lat = 16.82 ms and amp = 0.24 μV), wave D (lat = 24.75 ms and amp = 0.32 μV), wave E (lat = 31.36 ms and amp = 0.37 μV), and wave F (lat = 40.04 ms and amp = 0.29 μV).

Worldwide, there is a large increase in the number of elderly people. This entails providing better care for the elderly in all aspects of their health. With aging, there are structural changes in the peripheral and central auditory system which can lead to a decline in hearing. This, in turn, causes complaints of difficulty in understanding speech, especially in unfavorable environments [51, 52]. These impairments have a great impact on the life of the elderly, since in addition to causing social isolation, it can also lead to a depression and reduce cognitive function [53].

Only a few studies have focused on FFR in the elderly, with the most reported population being young adults [54]. Some researchers have pointed to the clinical applicability of FFR in different populations and with different pathologies [7, 19, 37, 55].

The effects of presbycusis on FFRs have been investigated in 18 individuals aged 61–78 years with hearing loss at frequencies of 2, 4, and 8 kHz (and compared with the responses of a control group of 19 young adults aged 20–26 years with normal hearing) [37]. The elderly group had lower amplitudes and increased latencies compared to the control group, demonstrating that the FFR can be affected by aging as well as hearing loss, but in different ways.

The effects of hearing loss on FFRs were described in a sample of 30 elderly individuals aged 60–71 years who were divided into two groups matched by gender and intelligence quotient: (i) normal hearing, and (ii) mild to moderate hearing loss [35]. With ABR clicks, all subjects had normal responses. FFR testing indicated that individuals with hearing loss could be assessed with this procedure, but there were changes in the frequency responses. In the elderly with hearing loss, there was a breakdown in the perception of the speech signal, which resulted in differences in signal parameters compared to the group with normal thresholds. This breakdown in neural synchrony may explain the greater difficulty subjects with hearing loss have in speech perception.

The evaluation of FFR in noisy environments is becoming more widespread, Thus, one study was carried out with 111 individuals between 45 and 78 years of age (mean 61.1 years) with normal to moderate hearing loss [56]. All subjects presented values within normal limits for the Montreal Cognitive Assessment (MoCA) and click ABR. In addition, they were tested on the SSQ (Speech, Spatial, and Qualities of Hearing Scale) which relates to auditory quality, as well as to the Quick Speech-in-Noise test (QuickSIN), in which phrases are presented binaurally with a verbal background babble. The FFR assessment demonstrated an increase in O-wave latency associated with speech comprehension difficulty in competing noise environments.

Supporting the observation that FFR traces are affected by increasing age, research on 34 individuals aged 22–77 years with normal hearing [57] found a decrease of the amplitude was associated with an increase in latency (**Figures 1** and **2**).

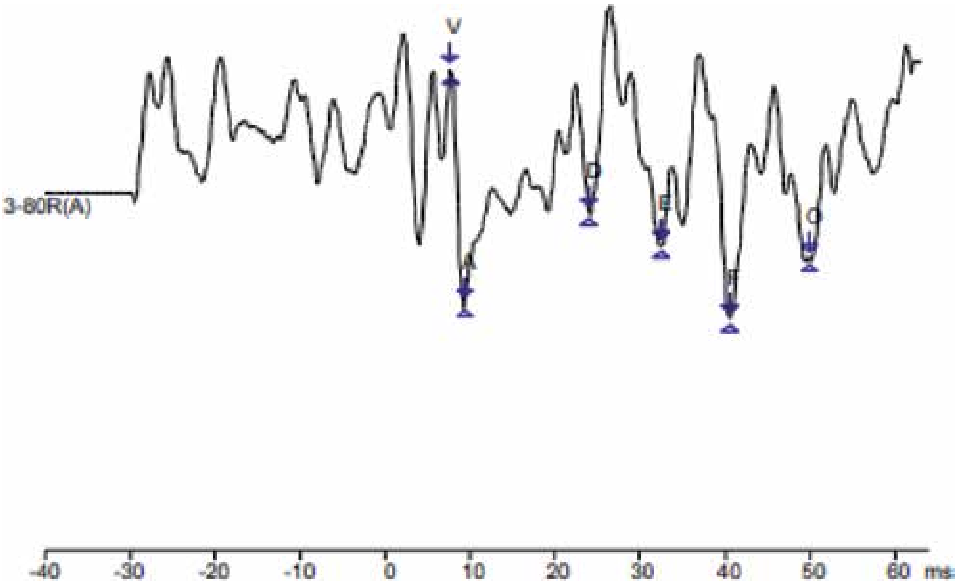

Figure 1.
FFRs of an infant 13 days old. Authors' data with FFR performed using SmartEP.

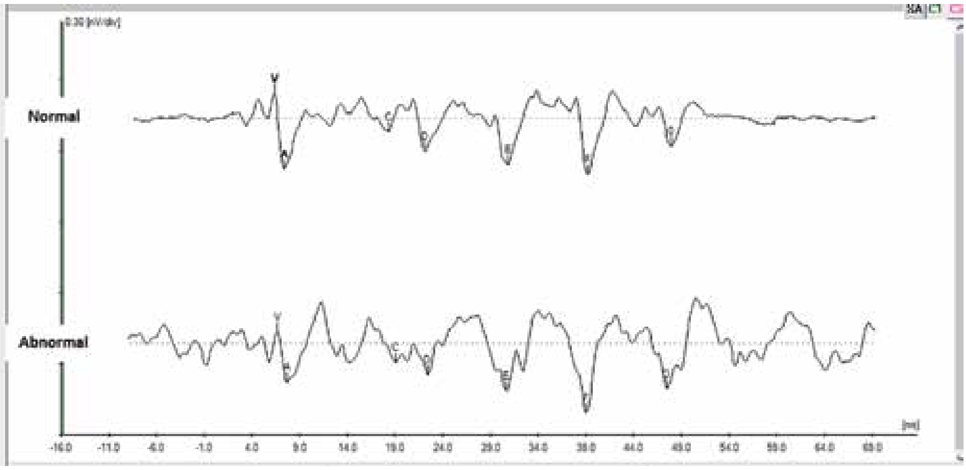

Figure 2.
FFRs of two 9-year-old-children. The top trace represents a normal response and the second represents an abnormal response. Authors' data using BioMARK software and Biologic equipment.

Figure 3 shows an FFR done on an adult aged 25 and on one aged 70. The shape of the FFR is similar in both, but there is an increase in latencies and some waves appear to be absent.

In these FFR tracings, it can be seen that the elderly subject had an increase in latency of all waves compared to the younger adult. Aging causes a progressive loss of structure or functioning of neurons, which can be seen as decreased auditory evoked potentials. Through the FFR, it is seen that there is also a reduction in the speed of neural activation from brainstem to cortical structures.

Our FFR evaluation in adults and the elderly used IHS equipment and the parameters are shown in **Table 14**.

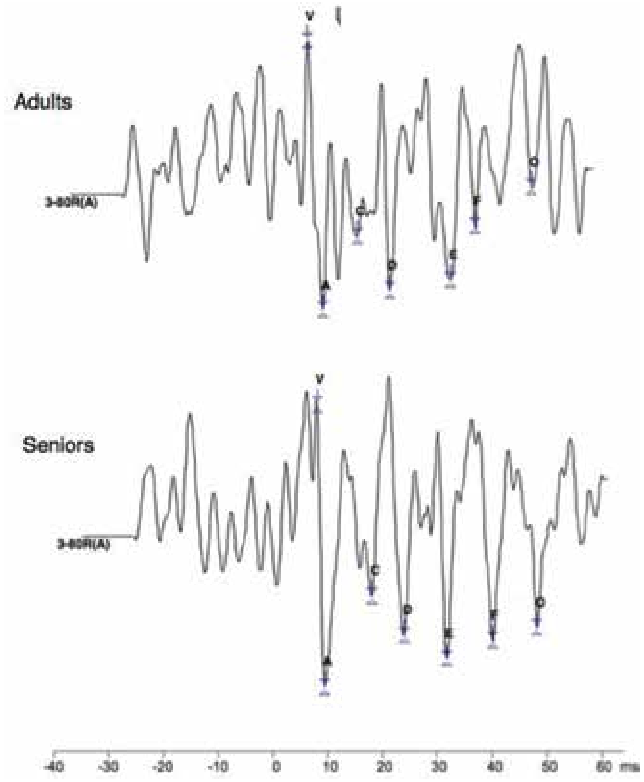

Figure 3.
FFRs of an adult aged 25 years (top) and another aged 70 (bottom). Note the increase in latency of the waves. Authors' data using SmartEP equipment.

Presentation parameters	Setting
Equipment	SmartEP Intelligent Hearing Systems (IHS)
Transducer	Insert phones
Electrodes	Fz, Fpz, M1, M2 or Cz, M1, M2
Stimulation	Right ear
Stimulus	Syllable /da/
Stimulus duration	40 ms
Presentation rate	10.9/s
Window	80–100 ms
Filter	Low pass of 100 Hz and high pass of 2000 Hz Low pass of 100 Hz and high pass of 3000 Hz
Polarity	Alternating
Intensity	80 dBnHL
Number of stimuli	6000
Reproducibility	2 × 3000 stimuli
Condition of evaluation	awake and quiet
Impedance	3k Ohms
Artifact rejection	>10%

Table 14.
Parameters of FFR in adults and the elderly.

6. Conclusion

FFR evaluations can be included as an extra examination in diagnostic testing and have an important role in crosschecking the results. It can also greatly assist making differential diagnoses in different clinical populations. However, each age group has FFRs with specific characteristics, so it is important that the audiologist has access to good normative values for the different age groups (infants and toddlers, young children, children and adolescents, adults and the elderly).

Terminology

10–20 International System	a standard system for electrode location
ABR	auditory brainstem response
AEP	auditory evoked potential. Evoked potential when using an auditory stimulus
BioMARK	Biological Marker of Auditory Processing is software that compares responses from a click to those from a synthetic syllable (usually /da/)
CANS	central auditory nervous system
CAP	central auditory processing
CAPD	central auditory processing disorder
CNS	central nervous system
CV syllable	a phoneme produced by a consonant and a vowel
FFR	frequency following response
Onset portion	the first part of an FFR that reflects the consonant
SAB	Scale of Auditory Behavior, a questionnaire for monitoring auditory processing skills
Sustained portion	the second part of an FFR that reflects the vowel
	Synthesized speech

Artificial human speech produced by a computer

Author details

Milaine Dominici Sanfins[1,2*], Michele Vargas Garcia[3], Eliara Pinto Vieira Biaggio[3] and Piotr Henryk Skarzynski[4,5,6]

1 Faculdade de Ciências Médicas, Universidade Estadual de Campinas, Campinas, Brazil

2 Centro de Eletrofisiologia e Neuroaudiologia Avançada, São Paulo, Brazil

3 Departamento de Fonoaudiologia, Universidade Federal de Santa Maria, Santa Maria, Brazil

4 Department of Teleaudiology and Hearing Screening, World Hearing Center, Institute of Physiology and Pathology of Hearing, Warsaw/Kajetany, Poland

5 Department of Heart Failure and Cardiac Rehabilitation, Medical University of Warsaw, Poland

6 Institute of Sensory Organs, Kajetany, Poland

*Address all correspondence to: msanfins@uol.com.br

References

[1] Johnson KL, Nicol TG, Kraus N. Brainstem response to speech: A biological marker of auditory processing. Ear and Hearing. 2005;**26**(5):424-434

[2] Skoe E, Kraus N. Auditory brainstem response to complex sounds: A tutorial. Ear and Hearing. 2010;**31**:320-324

[3] Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Aging affects neural precision of speech encoding. Journal of Neuroscience. 2012;**32**(41):14156-14164

[4] Bidelman GM. Multichannel recordings of the human brainstem frequency-following response: Scalp topography, source generators, and distinctions from the transient ABR. Hearing Research. 2015;**323**:68-80

[5] Sanfins M, Borges L, Donadon C, Hatzopoulos S, Skarzynski P, Colella-Santos M. Electrophysiological responses to speech stimuli in children with otitis media. Journal of Hearing Science. 2017;**7**(4):9-19

[6] Sanfins M, Borges L, Ubiali T, Colella-Santos M. Speech auditory brainstem response (speech ABR) in the differential diagnosis of scholastic difficulties. Brazilian Journal of Otorhinolaryngology. 2017;**83**(1):112-116

[7] Sanguebuche TR, Peixe BP, Bruno RS, Biaggio EPV, Garcia MV. Speech-evoked brainstem auditory responses and auditory processing skills: A correlation in adults with hearing loss. International Archives of Otorhinolaryngology. 2018;**22**(1):38-44

[8] Sanju HK, Mohanan A, Kumar P. Speech-evoked auditory brainstem response in individuals with diabetes mellitus type 2. The Journal of International Advanced Otology. 2017;**13**(1):77-82

[9] Tahaei AA, Ashayeri H, Pourbakht A, Kamali M. Speech evoked auditory brainstem response in stuttering. Scientifica (Cairo). 2014;**2014**:328646

[10] Elkabariti RH, Khalil LH, Husein R, Talaat HS. Speech evoked auditory brainstem response findings in children with epilepsy. International Journal of Pediatric Otorhinolaryngology. 2014;**78**(8):1277-1280

[11] Young E, Sachs M. Representation of steady-state vowels in the temporal aspects of the discharge patterns of populations of auditory-nerve fibers. The Journal of the Acoustical Society of America. 1979;**66**:1381-1403

[12] Russo N, Nicol T, Musacchia G, Kraus N. Brainstem responses to speech syllables. Clinical Neurophysiology. 2004;**115**:2021-2030

[13] Holt L, Lotto A. Speech perception as categorization. Attention, Perception, & Psychophysics. 2012;**72**:1218-1227

[14] McLaughlin S. Introduction to Language Development. Edmond: University of Central Oklahoma; 2007. 508 p

[15] Jeng F, Schnabel E, Dickman B, Ju J, Li X, Lin C, et al. Early maturation of frequency-following response to voice pitch in infants with normal hearing. Perceptual and Motor Skills. 2010;**111**:765-784

[16] Neville H, Bavalier D. Specificity and plasticity in neurocognitive development in humans. In: Gazzaniga M, editor. The New Cognitive Neurosciences. London: Cambridge Mass; 2000

[17] Bellis T. Assessment and Management of Central Auditory Processing Disorders in the Educational Setting. Thomson Delmar Learning: California; 2003

[18] Sano M, Kaga K, Kuan C, Ino K, Mima K. Early myelination patterns in the brainstem auditory nuclei and pathway: MRI evaluation study. International Journal of Pediatric Otorhinolaryngology. 2007;**71**:1105-1115

[19] Skoe E, Krizman J, Anderson S, Kraus N. Stability and plasticity of auditory brainstem function across the lifespan. Cerebral Cortex. 2015;**25**(6):1415-1426

[20] Jeng FC. Infant and childhood development: Intersections between development and language experience. In: Kraus N, Anderson S, White-Schwoch T, Fay R, Popper A, editors. The Frequency-Following Response: A Window into Human Communication. Switzerland: Springer; 2017. pp. 17-43

[21] Picton T, Woods D, Baribeau-Braun J, Healey T. Evoked potential audiometry. The Journal of Otolaryngology. 1977;**6**:90-119

[22] Jeng F, Hu J, Dickman B, Montgomery-Reagan K, Tong M, Wu G, et al. Cross-linguistic comparison of frequency-following responses to voice pitch in American and Chinese neonates and adults. Ear and Hearing. 2011;**32**:699-707

[23] Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Development of subcortical speech representation in human infants. The Journal of Acoustical Society of America. 2015;**137**:3346-3355

[24] Jeng F, Peris K, Hu J, Lin C. Evaluation of an automated procedure for detecting frequency-following responses in American and Chinese neonates. Perceptual and Motor Skills. 2013;**116**(2):456-465

[25] Jeng F, Lin C, Chou M, Hollister G, Sabol J, Mayhugh G, et al. Development of subcortical pitch representation in three-month-old chinese infants. Perceptual and Motor Skills. 2016;**122**:123-135

[26] Van Dyke K, Lieberman R, Presacco A, Anderson S. Development of phase locking and frequency representation in the infant frequency-following response. JSLHR. 2017;**60**:2740-2751

[27] Ayas M, Yaseen H, Rajashekhar B. Auditory brainstem processing of complex speech sounds in high risk infants-a preliminary study. International Journal of Current Research and Review. 2015;**7**:22-27

[28] Musacchia G, Ortiz-Mantilla S, Roesler C, Rajendra S, Morgan-Byrne J, Benasich A. Effects of noise and age on the infant brainstem response to speech. Clinical Neurophysiology. 2018;**129**:2623-2634

[29] Gabriel L, Vernier L, Ferreira M, Silveira A, Machado M. Parameters for applying the brainstem auditory evoked potential with speech stimulus: Systematic review. International Archives of Otorhinolaryngology. 2018;**22**(4):460-468

[30] Sanfins M. Electrophysiological Evaluation with Verbal and Non-verbal Sounds in Children with a History of Otitis Media. State University of Campinas; 2017. Available from: http://www.repositorio.unicamp.br/handle/REPOSIP/330747

[31] Hornickel J, Lin D, Kraus N. Speech-evoked auditory brainstem responses reflect familial and cognitive influences. Developmental Science. 2013;**16**(1):101-110

[32] Cacace AT, McFarland DJ. Central auditory processing disorder in school-aged children: A critical review. Journal of Speech, Language, and Hearing Research. 1998;**41**(2):355-373

[33] Johnson K, Nicol T, Zecker S, Kraus N. Developmental plasticity

in the human auditory brainstem. The Journal of Neuroscience. 2008;**28**(15):4000-4007

[34] Kraus N, Hornickel J. cABR: A biological probe of auditory processing. In: Geffner DS, Ross-Swain D, editors. Auditory Processing Disorders: Assessment, Management, and Treatment. 2a ed. San Diego: Plural Publishing; 2013. pp. 159-183

[35] Anderson S, Parbery-Clark A, White-Schwoch T, Drehobl S, Kraus N. Effects of hearing loss on the subcortical representation of speech cues. The Journal of the Acoustical Society of America. 2013;**133**(5):3030-3038

[36] Hornickel J, Kraus N. cABR can predict auditory-based communication skills. The Hearing Journal. 2012;**65**:28-30

[37] Vander Werff KR, Burns KS. Brain stem responses to speech in younger and older adults. Ear and Hearing. 2011;**32**(2):168-180

[38] Benasich A, Fitch R. Developmental Dyslexia: Early Precursors, Neurobehavioral Markers and Biological Substrates. Baltimore: Brookes Publishing; 2012

[39] Bogliotti C, Serniclaes W, Messaoud-Galusi SLSC. Discrimination of speech sounds by children with dyslexia: Comparisons with chronological age and reading level controls. Journal of Experimental Child Psychology. 2008;**101**:137-155

[40] Kumar P, Singh NK. BioMARK as electrophysiological tool for assessing children at risk for (central) auditory processing disorders without reading deficits. Hearing Research. 2015;**324**:54-58

[41] Song JH, Nicol T, Kraus N. Test-retest reliability of the speech-evoked auditory brainstem response. Clinical Neurophysiology. 2011;**122**(2):346-355

[42] Krizman J, Marian V, Shook A, Skoe E, Kraus N. Subcortical encoding of sound is enhanced in bilinguals and relates to executive function advantages. Proceedings of the National Academy of Sciences. 2012;**109**(20):7877-7881

[43] Krizman J, Skoe E, Kraus N. Bilingual enhancements have no socioeconomic boundaries. Developmental Science. 2016;**19**(6):881-891

[44] Patel A, Peretz I, Tramo M, Labrecque R. Processing prosodic and music pattern: A neurophsychological investigation. Brain and Language. 1998;**61**:123-144

[45] Russo N, Skoe E, Trommer B, Nicol T, Zecker S, Bradlow A, et al. Deficient brainstem encoding of pitch in children with autism spectrum disorders. Clinical Neurophysiology. 2008;**119**(8):1720-1731

[46] Wible B, Nicol T, Kraus N. Abnormal neural encoding of repeated speech stimuli in noise in children with learning problems. Clinical Neurophysiology. 2002;**113**:484-494

[47] Cunningham J, Nicol T, Zecker S. Neurobiologic responses to speech in noise in children with learning problems: Deficits and strategies for improvement. Clinical Neurophysiology. 2001;**112**:758-767

[48] Ladefoged P. A Course in Phonetics. Boston: Thomson Higher Learning; 2006

[49] Peixe BP, Silva DD, Biaggio EPV, Bruno RS, Sanguebuche TR, Garcia MV. Applicability of evoked auditory brainstem responses with complex stimuli in adults with hearing loss. International Archives of Otorhinolaryngology. 2018;**22**(3):239-244

[50] Sinha S, Basavaraj V. Speech evoked auditory brainstem responses: A new

tool to study brainstem encoding of speech sounds. Indian Journal of Otolaryngology and Head & Neck Surgery. 2010;**62**(4):395-399

[51] Schoof T, Rosen S. The role of auditory and cognitive factors in understanding speech in noise by normal-hearing older listeners. Frontiers in Aging Neuroscience. 2014;**6**:307

[52] Palmer SB, Musiek FE. Electrophysiological gap detection thresholds: Effects of age and comparison with a behavioral measure. Journal of the American Academy of Audiology. 2014;**25**(10):999-1007

[53] Lin F, Ferrucci L, Metter E, Um Y, Zonderman A, Resnick S. Hearing loss and cognition in the Baltimore longitudinal study of aging. Neuropsychology. 2011;**25**:763-770

[54] Sanfins M, Colella-Santos M. A review of the clinical applicability of speech-evoked auditory brainstem responses. The Journal of Health Science. 2016;**6**(1):9-16

[55] Ahadi M, Pourbakht A, Jafari AH, Jalaie S. Effects of stimulus presentation mode and subcortical laterality in speech-evoked auditory brainstem responses. International Journal of Audiology. 2014;**53**(4):243-249

[56] Anderson S, Parbery-Clark A, White-Schwoch T, Kraus N. Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. Journal of Speech Language and Hearing Research. 2013;**56**(1):31-43

[57] Clinard C, Tremblay K. Aging degrades the neural encoding of simple and complex sounds in the human brainstem. Journal of the American Academy of Audiology. 2013;**24**(7):590-599

Chapter 8

Role of Glucocorticoids in Hearing Preservation in Partial Deafness Treatment

Magdalena B. Skarżyńska

Abstract

During the last 15 years, cochlear implantation became available as a method of treatment for different types of hearing impairment. Leading, specialized centers have now introduced the analysis of the nonsurgical factors that could contribute to improve rates of hearing preservation in preoperative period, during surgery, or in postoperative period in patients who suffer from partial deafness. One of the approaches is using pharmacotherapy (glucocorticoids) as a factor that may improve hearing functions. Preservation of hearing in patients who suffered from partial deafness and underwent cochlear implantation by using two different regimes of corticosteroid therapy was the aim of the study carried out by the *World Hearing Center* (WHC). Forty-six patients were enrolled in the trial and divided into three subgroups. Hearing preservation (HP) was evaluated using pure tone audiometry (PTA) (11 frequencies ranging from 125 to 8000 Hz). The impact of administrated substances was evaluated by pure tone audiometry during six different periods: before cochlear implant surgery, during activation of audio processor, and 1, 6, 9, and 12 months after activation of audio processor in comparison with control group. According to *hearing preservation (HP)* classification, patients from the second group, to whom combined glucocorticoid therapy was administrated, achieved the best HP results. The complete hearing preservation index was observed in the highest percentage of patients from the second subgroup. The dispersion of measured values was lesser than in other subgroups. According to the results, administration of glucocorticoids (dexamethasone and prednisone or dexamethasone only) to the patients, who suffered from partial deafness and underwent cochlear implantation surgery, may be important in stabilization of hearing thresholds and in protection of hearing.

Keywords: glucocorticoids, dexamethasone, prednisone, partial deafness treatment, cochlear implant

1. Introduction

In the past cochlear implantation was a "gold standard" treatment method for patients who suffer from hearing impairment and dedicated only for patients totally deaf. During the last 15 years, cochlear implantation became available also as a treatment for different types of hearing impairment. Regardless of surgical technique for cochlear implantation (cochleostomy or an approach through the round window), specialized centers have now introduced the comprehensive analysis of the nonsurgical factors that could contribute to improve rates of hearing preservation in preoperative, during surgery, or in postoperative period in patients who suffer from partial deafness. One of the approaches of many leading research centers is using pharmacotherapy (glucocorticoids) as a factor which may improve hearing functions following cochlear implantation [1, 2]. Glucocorticoids play an important role in pharmacotherapy of many different otorhinolaryngological diseases, such as Meniere's disease, sudden sensorineural hearing loss (SSNHL), and tinnitus, and as a part of otorhinolaryngological procedures in surgeries (e.g., cochlear implantation) [3, 4]. The effects of treatment listed diseases are different and mainly depend on treatment results, adverse effects of used medications, and additional pharmacological treatment that was used during treatment. Unfortunately, the side effects of glucocorticoids are serious, and as the result, sometimes pharmacotherapy has to be stopped, and discontinuation of therapy is the only solution.

On the one hand, insertion of specific electrode of cochlear implant requires perfection in surgical techniques, but on the other it is difficult to do perfectly. Clinically approved algorithm of corticosteroid therapy (local or systematic) is discussed as one factor in reducing oxidative stress, an inflammatory reaction, and as a result apoptosis of hearing cells. A major challenge in effective administration and delivery medicines is the *blood-labyrinth barrier* (BLB) and physical inaccessibility of the inner ear, especially the apical part of the cochlea. It seems to be crucial especially for patients who suffer from partial deafness (PD) (apical hair cells are responsible for receiving low frequencies).

1.1 Characteristic: glucocorticoids

The adrenal cortex synthesizes two classes of steroids: corticosteroids (mineralocorticoids and glucocorticoids) and androgens. One of the differences is the number of carbon atoms. Corticosteroids have 21 carbon atoms, and androgens have 19 carbon atoms. In the human body, the main glucocorticoid is cortisol, and the main mineralocorticoid is aldosterone [5].

Glucocorticoid receptor (GR) is located in the cytoplasm in inactive form until it binds with the molecule of glucocorticoid. This action results in activation of receptor and translocation complex: *glucocorticoid receptor for glucocorticoid* to the nucleus. Activation of receptor is based on dissociation from the associated proteins. After translocation to the nucleus, a complex of glucocorticoid receptor interacts with specific, short DNA sequences with the regulatory regions. The regions are termed *glucocorticoid responsive elements* (GREs) and allow induction of the gene transcription by glucocorticoids. This process is very complicated because of interaction with specific cofactors and proteins and still not well and completely understood by scientists and researchers [5]. Not only positive response to glucocorticoid is possible. According to Webster and Cidlowski, genes negatively regulated by glucocorticoids were also identified [6]. An example of downregulation (negative regulation) is to repress the expression of gene responsible for encoding cytokines or enzymes (e.g., collagenase). Both play an important role in inflammatory and immune reactions. According to the information provided, this negative expression appears to play a key function in anti-inflammatory and immunosuppressive effects of the glucocorticoids. Anti-inflammatory activity of representative glucocorticoids is presented below (**Table 1**). Dexamethasone and betamethasone are two glucocorticoids with the highest anti-inflammatory activity. If cortisol has anti-inflammatory activity defined as 1, then prednisone, prednisolone, triamcinolone, and 6α-methylprednisolone have 4–5 times stronger anti-inflammatory properties, with longer half-life than cortisol. Examples of representative glucocorticoids and their properties are shown in **Table 1**.

	Anti-inflammatory activity	**Biological half-life $t_{1/2}$ (h)**
Cortisol	1	Short: $t_{1/2}$ = 8–12
Cortisone	0.8	Short: $t_{1/2}$ = 8–12
Fludrocortisone	10	Intermediate: $t_{1/2}$ = 12–36
Prednisone	4	Intermediate: $t_{1/2}$ = 12–36
Prednisolone	4	Intermediate: $t_{1/2}$ = 12–36
6α-methylprednisone	5	Intermediate: $t_{1/2}$ = 12–36
Triamcinolone	5	Intermediate: $t_{1/2}$ = 12–36
Betamethasone	25	Long: $t_{1/2}$ = 36–72
Dexamethasone	25	Long: $t_{1/2}$ = 36–72

Table 1.
Characteristics of representative corticosteroids [5].

1.2 Functions and activity of glucocorticoids

The two key roles glucocorticoids play as biological and pharmaceutical compounds are anti-inflammatory and immunosuppressive roles. Glucocorticoids also affect:

1. Carbohydrate and protein metabolism. Glucocorticoids stimulate the liver to form glucose in biochemical reaction (gluconeogenesis) from amino acids and glycerol or/and stimulate the liver to release glucose from glycogen. At the same time, the diminishing of glucose is reduced, reaction of lipolysis and protein breakdown increases, and as a result the blood glucose level rises. Patients suffering from diabetes or other forms of hyperglycemia during glucocorticoid therapy should be under special control. Glucocorticoids induce increased protein metabolism and deliver compounds such as amino acids for further reactions.

2. Lipid metabolism. One effect of therapy with corticosteroids is redistribution of fat tissue known as Cushing's syndrome.

3. Water and electrolyte balance. Glucocorticoids exert negatively on metabolism of Ca^{2+} due to reduction of absorption from the digestive system and increased excretion via kidneys. Prophylaxis of osteoporosis requires supplementation of Ca^{2+} ions and physical activity adequate to possibilities of patient. Additionally, glucocorticoids reduce activity of osteoclasts and stimulate the activity of osteoblasts. Glucocorticoid therapy sometimes may cause increased retention of Na^{+} ions and decrease in concentration of K^{+} ions because of interaction with receptor for mineralocorticoids. The deterioration in ion level may affect the cardiovascular system.

4. Impaired wound healing. Due to reduction of synthesis of collagen, glycosaminoglycans and disturbance in fibroblast function problem with healing wound may occur.

5. Anti-inflammatory and immunosuppressive activity. Glucocorticoids can suppress or prevent inflammatory reactions in different ways: reduction in diapedesis of granulocytes and proliferation of lymphocytes Th, inhibition/

reduction of activation of macrophages, neutrophils, mast cells and cytokines (interleukins 1, 2, 3, 4, 5, 6, 8), and tumor necrosis factor alpha (TNF-α); reduction in the expression of cyclooxygenase 2 (COX-2) resulting in dropping of production of a few prostanoids; intensification of activity of catechol amines; and reduction in production of histamine by basophils [7, 8].

Sometimes glucocorticoid treatment in the otorhinolaryngological diseases requires high doses or long time of therapy. It may cause adverse effects. Some of them are listed below:

1. Repression in responding on infections and injuries

2. Propensity to opportunistic infections

3. Propensity to hyperglycemia

4. Muscular dystrophy

5. Cushing's syndrome

6. Glaucoma (mostly in patients genetically predisposed)

7. Osteoporosis

Acute discontinuation of treatment may cause adrenocortical insufficiency, especially when therapy was long-term. It is important to reduce the dose of glucocorticoids slowly, not suddenly [5].

1.3 Pharmacokinetics of glucocorticoids and studies on animal model

Many factors have an impact on pharmacokinetics of drugs. Pharmacokinetics is described by acronym *LADME (liberation, absorption, distribution, metabolism, and elimination*). Firstly, the therapeutic (or its carrier) must be water-soluble, because of distribution in the blood. The protein binding of drugs is one of the key factors in initial parts in pharmacokinetic process. The greater the protein binding of drug is, the longer the activity of therapeutic due to its function as a stock of drug in the organism. Absorption depends on lipophilicity and solubility of drugs. According to the data and publications, only a few medical substances can effectively be used in otorhinolaryngological practice due to achieving sufficient concentration in the inner ear [9]. Two main groups of drugs are used in clinical practice: aminoglycoside (mainly gentamicin) in pharmacotherapy of Meniere's disease [10] and corticosteroids (dexamethasone, triamcinolone and dexamethasone) in pharmacotherapy of *idiopathic sudden sensorineural hearing loss* (ISSHL) and other cases of acute hearing loss [11]. The inner ear from a pharmacokinetic point of view is a multicompartment model [12, 13] with stable fluids and balance between them (due to the presence of *blood-labyrinth barrier* (BLB)). Distribution process depends on many different factors such as route of administration, model of administration (single or repeated administration), dose of medicine, ionic composition, and pH or osmolarity of solution. The same factors of drug chemical and physical properties influence the elimination of drug from the organism (clearance, rate of removal).

In a study published in 2017, authors in animal model (guinea pig) compared dexamethasone with saline. Both substances were administrated intravenously 60 min before implantation. As a final conclusion, authors stressed that dexamethasone could reduce scarring process as the electrode negotiated the hook region or near the electrode tip, but they did not observe the relation between dexamethasone and reduction of fibrosis relating to cochleostomy [13]. In vitro studies showed the correlation between reduction (loss) of auditory cells after exposure to *tumor necrosis factor* alpha (TNF-alpha) and dexamethasone-releasing polymer used to coat electrode of cochlear implant carries.

Research carried out on animal model proved that prolonged steroid therapy could significantly improve hearing preservation rate (including pharmacokinetic and morphological analysis) when the electrode of cochlear implants was covered with dexamethasone (special formulation with controlled drug release) [14]. However, Honeder et al. did not confirm that steroids could have a positive impact on residual hearing in a guinea pig model. One reason why both authors gain different results may lie in different types of steroid therapy. In the first study, dexamethasone was used on the contrary to the second study where triamcinolone was administrated [15]. Douchement et al. investigated the effects of steroids using a gerbil animal model. Animals were implanted with an electrode with controlled dexamethasone delivery (1 and 10% concentration of dexamethasone) on one side and a conventional electrode on the contralateral side. Hearing levels were established based on the tone bursts on auditory brain stem responses at 4–6-week postimplantation and at 1-year postimplantation period for older gerbils. A 1-year observation period showed significantly improved results obtained for the high auditory frequencies, but the results for the low frequencies were ambiguous [16].

Cho et al. analyzed the efficacy of preoperative and intraoperative schemes of administration of steroids for hearing preservation. Dexamethasone was administrated systematically at the dose of 5 mg/ml in the preoperative period and then topically (off-label) during cochlear implantation surgery. Pure tone audiometry (PTA) was measured in four frequencies: 250, 500, 1000, and 2000 Hz. Statistically significant differences were observed between the steroid group and the control group, supporting the observation and beneficial impact of administration steroid treatment [17].

1.4 Glucocorticoid administration: the possibilities

During treatment of otologic diseases, two routes of administration of glucocorticoids are possible: local and systemic. Local administration (e.g., transtympanic injection) allows to achieve high concentration of glucocorticoid in the middle ear, but due to presence of Eustachian tube, the medication may be partly evacuated. Local drug administration to the middle and inner ear avoids "the first pass effect." The main advantages of local drug delivery are as follows:

1. Reduction dose of medicine
2. Achieving high concentration
3. Better effects of treatment
4. Reducing possibility of adverse effects
5. Bypassing of the blood-labyrinth barrier (BLB)

Local drug administration may involve intracochlear administration (e.g., stem cell, gene therapy) or extracochlear administration (e.g., intratympanic injection). A combination of both routes of drug delivery to the ear is also possible [13]. According to publications, systemic administration in treatment of otorhinolaryngological diseases is known as noninvasive route of drug delivery, due to lack of damaging of tympanic membrane. Adverse effects of systemic delivery may be one of the purposes of discontinuation of the therapy. The presence of *blood-labyrinth barrier* (BLB) in the inner ear is one of the causes of problems with reaching high concentration of drug.

1.5 The experience in using glucocorticoids in PDT patients

This study was the first study which was carried out in the Institute of Physiology and Pathology of Hearing World Hearing Center and will be continued with different groups of patients and different implants and algorithms of glucocorticoid administration. The aim of the study was to evaluate different regimes of administration of glucocorticoids: dexamethasone and dexamethasone/prednisone to *partial deafness* treatment (PDT) patients who underwent cochlear implantation on the hearing preservation. Implant used in the study was the MED-EL implant with an electrode length of 28 mm (Flex 28). The impact of administrated glucocorticoids on hearing was measured in six different periods:

1. Preoperatively

2. During activation of audio processor

3. One month after activation of audio processor

4. Six months after activation of audio processor

5. Nine months after activation of audio processor

6. One year (12 months) after activation of audio processor

Forty-six patients were enrolled to the trial and then divided randomly into three subgroups. Patients from the first subgroup underwent intravenous steroid therapy (**Figure 1**). According to the scheme, 30 min before implantation, dexamethasone at the dose of 0.1 mg/kg body mass was administrated intravenously to patients from the first subgroup. For the next 3 consecutive days in every 12 h, dexamethasone was administrated intravenously at the same dose to each patient.

Patients from the second subgroup underwent prolonged (combined) steroid therapy: oral and intravenous. Three days prior to the surgery, prednisone was administrated orally at a dose of 1 mg/kg body mass/day. Then, 30 min before the cochlear implantation surgery, dexamethasone was administrated intravenously at a dose of 0.1 mg/kg body mass. For the next 3 consecutive days in every 12 h, dexamethasone was administrated intravenously (the same as in the first subgroup). For the next 3 days, prednisone was administrated orally at the dose of 1 mg/kg body mass/day. After this period, the dose of prednisone was reduced (10 mg per every

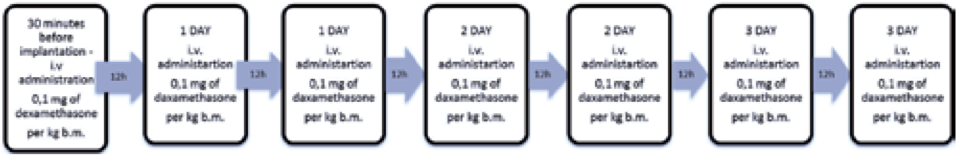

Figure 1.
Scheme of steroid administration in the first subgroup of patients.

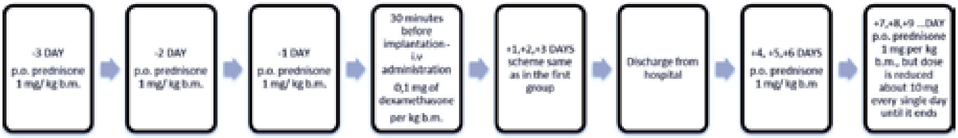

Figure 2.
Scheme of steroid administration in the second subgroup of patients.

day) till complete reduction of dose, due to reducing the risk of adverse effects. The algorithm of administration of glucocorticoids to patients form the second subgroup is presented below (**Figure 2**).

The third subgroup was a control group and underwent standard cochlear implantation procedure [18].

1.6 Administrated glucocorticoids

According to the protocol of this study, two different algorithms of administration with two different glucocorticoids were proposed in the study. Although both substances belong to the same pharmacological group, both of them have different pharmacokinetics and pharmacodynamic properties. Dexamethasone is a synthetic glucocorticoid (*molecular weight 392.46 g/mol*) with anti-inflammatory, anti-allergic, and immunomodulating activity. In common practice dexamethasone is administrated *intravenously* or off-label, e.g., *transtympanic injections*. After *intravenous* administration the mean time to peak concentration (C_{max}) is between 10 and 30 min, and the half-life ($t_{1/2}$) is from 2.2 to 3.8 h. Transport proteins are responsible for transport and distribution of dexamethasone in blood. Dexamethasone is mainly metabolized by the liver and eliminated with the bile. Only 2.6% of the chemically unchanged dose is eliminated via kidneys.

Prednisone is a synthetic glucocorticoid (derivative of cortisone) and classified according to the *Anatomical Therapeutic Chemical (ATC) Classification System* as H02 AB 07. Prednisone is *prodrug* which converts into *active metabolite—prednisolone* (higher anti-inflammatory activity). According to characteristic of medical product and literature data, bioavailability of prednisone administrated orally is between 70 and 90%. The mean time to peak concentration (C_{max}) is between 1 and 2 h. Half-life ($t_{1/2}$) is between 3.4 and 3.8 h in plasma and 18–36 h in tissue. Binding prednisone with plasma proteins is between 70 and 73% (binding prednisolone (active metabolite) to the plasma proteins is higher (90–95%)). Similar to dexamethasone, prednisone is metabolized mainly by the liver and eliminated with the bile. Pharmacodynamic and pharmacokinetic data were based on characteristics of medical products: dexamethasone and prednisone.

1.7 Methodology of the study: primary and secondary outcomes and inclusion and exclusion criteria

The primary outcome variables were mean values of hearing thresholds averaged across all 11 frequencies (125–8000 Hz). The secondary outcome variable was hearing preservation (HP). Hearing preservation (HP) was calculated by comparing hearing thresholds in the 1-year postoperative period with the preoperative hearing thresholds, according to the hearing preservation (HP) formula (below) and converted to three levels: minimal, partial, and complete hearing preservation.

$$\mathrm{HP} = \left(1 - \frac{PTApost - PTApre}{PTAmax - PTApre}\right) * 100\,[\%] \tag{1}$$

In this equation, *PTApre* is the pure tone average measured preoperatively, *PTApost* is the pure tone average measured postoperatively, and *PTAmax* is the maximal sound intensity generated by a standard audiometer, usually 120 dB hearing level (HL), and *HP* is the rate of hearing preservation in percentage [19].

The protocol of this prospective clinical trial was approved by the Bioethics Commission. Patients enrolled to the study suffered from severe-to-profound hearing loss and were classified according to Prof. H. Skarżyński *partial deafness treatment (PDT) classification* into two groups: *partial deafness treatment-electrical* stimulation (PDT-EC) and *partial deafness treatment-electroacoustic* stimulation (PDT-EAS) (**Figure 3**) [20, 21].

Inclusion and exclusion criteria were in accordance with the consensus of the international *HEARRING* group on hearing preservation in cochlear implant. Study eligibility criteria included participants ≥18 years of age with a cochlear duct length of ≥27.1 (measured by computer tomography), with:

1. Hearing sound levels in the range of 10–120 decibels (dB) and sound frequencies of 125–250 hertz (Hz)

2. Hearing sound levels of 35–120 dB and frequencies of 500–1000 Hz

3. Hearing sound levels of 75–120 dB and frequencies of 2000–8000 Hz [18]

Exclusion criteria included suffering from severe diseases when the steroid treatment could worsen the patient's condition or when there would be a possibility of interaction between medication intake by patients and steroids. Nonparametric tests were used in the study due to discrepancies in the number of participants between all subgroups, small number of participants in the study, and violation of normal distribution of pure tone audiometry results [18].

1.8 Main results and observations

Preoperative hearing threshold levels of patients from the first, the second, and the control subgroup were similar. The difference between patients from the three subgroups was not statistically significant, which means that hearing

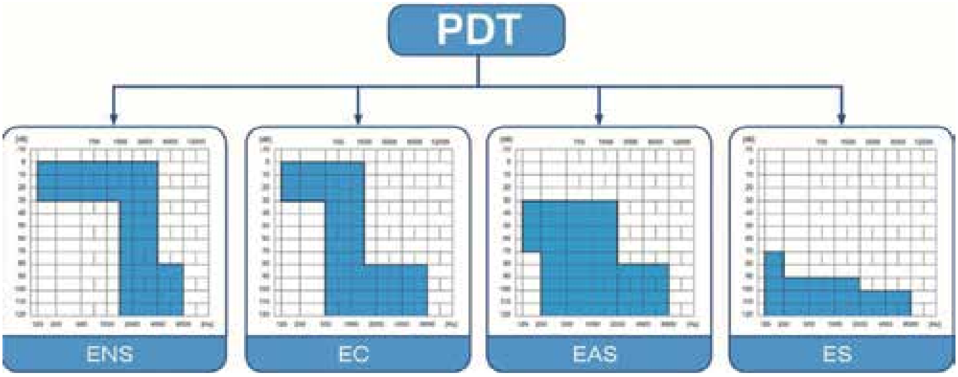

Figure 3.
Partial deafness treatment groups for cochlear implantation. ENS, electro-natural stimulation; EC, electrical complement; EAS, electrical-acoustic stimulation; ES, electrical stimulation.

thresholds in preoperative period of all participants, who were enrolled to the study, were similar.

The deterioration of average hearing thresholds (measured by pure tone audiometry) was observed from the first point of observation—the activation period. A significant difference was observed between two groups: patients from the second subgroup (combined steroid therapy, prednisone + dexamethasone) and the control subgroup. Patients from the second subgroup had better pure tone audiometry (PTA) results considering low frequencies in comparison with the results of patients from the control group. Similar observation was done in 1, 6, 9, and 12 months after activation follow-up periods. The results of the study may be even more a promise and beneficial for patients. The hearing of participants of the study, to whom combined (prolonged) glucocorticoid therapy was administrated, remained stable during all observed follow-up periods (*activation, 1-month, 6-month, 9-month, and 12-month post-activation follow-ups*), and they did not vary significantly (**Figure 4**).

The hearing preservation (HP) rate is calculated using hearing preservation formula by comparing hearing threshold in the 12-month postoperative period with the preoperative hearing thresholds. Then the results were divided, according to the hearing preservation (HP) formula, into minimal HP, partial HP, and complete HP according to **Table 2**. The smallest variability of results was observed in the second subgroup (patients to whom prednisone and dexamethasone were administrated) as well as the highest overall HP rate. Patients from the second subgroup (prolonged steroid therapy) and nearly 69% of the patients from the first subgroup had partial or complete hearing preservation. The majority of patients from the control group had minimal hearing preservation (**Table 2** and **Figure 5**).

1.9 Final conclusion

This study is the first study to report the results of two different regimes of steroid administration in human subjects who underwent cochlear implantation in comparison with the control group. As it was said in the previous paragraph, the findings of this study have shown that glucocorticoid therapy not only stabilizes hearing thresholds but also preserves hearing ability in adult patients. The combination of *intravenously* administrated dexamethasone and *orally* administrated prednisone in one scheme of administration seems to be the optimal treatment regimen.

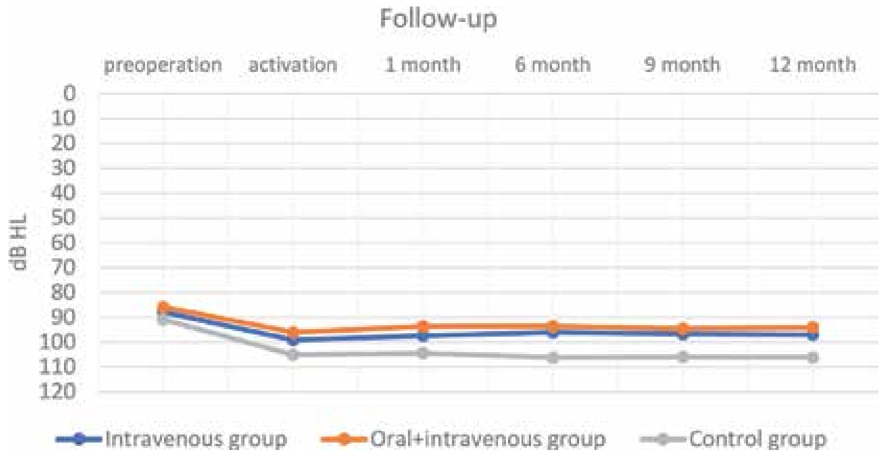

Figure 4.
Mean hearing thresholds of patients with standard steroid therapy (subgroup No 1), patients with prolonged steroid therapy (subgroup No 2), and control patients (subgroup No 3) in the preoperative period, upon activation, and at 1-month, 6-month, 9-month, and 12-month follow-up after CI implantation.

	Minimal (0–25%)	Partial (26–75%)	Complete (≤75%)
Subgroup No 1	5 (31.2)	7 (43.8)	4 (25.0)
Subgroup No 2	0 (0.0)	8 (61.5)	5 (38.5)
Control group	12 (70.6)	3 (17.6)	2 (11.8)

Table 2.
Hearing preservation 12 months after CI implantation, according to the type of treatment (data are given as the number of patients (percentage in brackets).

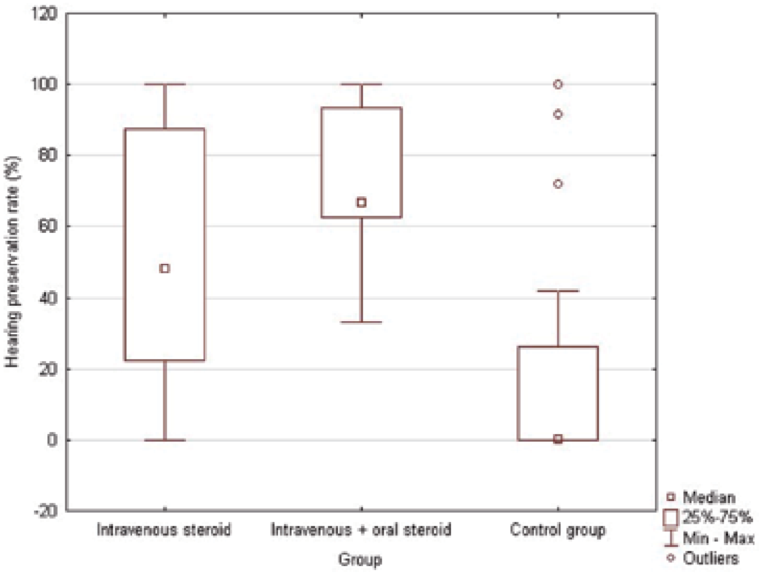

Figure 5.
Hearing preservation (HP) rate in three subgroups.

Previously published studies have shown that there have been new directions in the development and use of electrodes and cochlear implant surgery in recent years. Currently, researchers, clinicians, and commercial companies are working on developing modern steroid-eluting electrodes or electrodes with controlled drug delivery. The results of the preliminary study described in this chapter suggested that combined glucocorticoid administration (according to scheme of administration in the second subgroup) is beneficial in preserving and stabilizing hearing thresholds in patients undergoing cochlear implantation surgery. The findings of this study are supported by the results of similar studies [1, 17]. However, the present study adds to the findings of previous studies by having a relatively long follow-up period, of 12 months after activation, with study analysis conducted during six different follow-up periods. According to the results, administration of glucocorticoids (dexamethasone and prednisone or dexamethasone only) to the patients, who suffered from partial deafness and underwent cochlear implantation surgery, may be important in stabilization of hearing thresholds and in protection of hearing. The dispersion of measured values in the second group (the second subgroup) was lesser than in the first and the control group.

Author details

Magdalena B. Skarżyńska[1,2,3]

1 Institute of Sensory Organs, Kajetany, Polnad

2 Center of Hearing and Speech Medincus, Kajetany, Poland

3 World Hearing Center, Institute of Physiology and Pathology of Hearing, Warsaw/Kajetany, Poland

*Address all correspondence to: m.skarzynska@csim.home.pl

IntechOpen

References

[1] Sweeney AD, Carlson ML, Zuniga MG, Bennett ML, Wanna GB, Haynes DS, et al. Impact of perioperative oral steroid use on low-frequency hearing preservation after cochlear implantation. Otology & Neurotology. 2015;**36**(9):1480-1485

[2] Rah YC, Lee MY, Kim SH, Kim DH, Eastwood H, O'Leary SJ, et al. Extended use of systemic steroid Is beneficial in preserving hearing in guinea pigs after cochlear implant. Acta Oto-Laryngologica. 2016;**136**(12):1213-1219

[3] Chrousos GP. Adrenocorticosteroids and adrenocortical antagonists. In: Basic and Clinical Pharmacology. The McGraw-Hill Companies Inc.; Chapter 39, edition 13. 2015

[4] Alles MJRC, der Gaag MA, Stokroos RJ. Intratympanic steroid therapy for inner ear diseases, a review of the literature. European Archives of Oto-Rhino-Laryngology. 2006;**263**(9):791-797

[5] Brunton L, Knollman B, Hilal-Dandan R. Goodman and Gilman's The Pharmacological Basis of Therapeutics. T. 1. 13. wyd. McGraw-Hill Education; 2017

[6] Webster JC, Cidlowski JA. Mechanisms of glucocorticoid-receptor-mediated repression of gene expression. Trends in Endocrinology and Metabolism. 1999;**10**(10):396-402

[7] Munck A, Guyre PM, Holbrook NJ. Physiological functions of glucocorticoids in stress and their relation to pharmacological actions. Endocrine Reviews. 1984;**5**(1):25-44

[8] Lo J, Campbell L, Sale P, Chambers S, Hampson A, Eastwood H, et al. The role of preoperative steroids in atraumatic cochlear implantation surgery. Otology & Neurotology. 2017;**38**(8):1118-1124

[9] Salt AN, Plontke SK. Pharmacokinetic principles in the inner ear: Influence of drug properties on intratympanic applications. Hearing Research. 2018;**368**:28-40

[10] Lange G. Gentamicin and other ototoxic antibiotics for the transtympanic treatment of Menière's disease. Archives of Oto-Rhino-Laryngology. 1989;**246**(5):269-270

[11] Hamid M, Trune D. Issues, indications, and controversies regarding intratympanic steroid perfusion. Current Opinion in Otolaryngology & Head and Neck Surgery. 2008;**16**(5):434-440

[12] Liebau A, Plontke SK. Lokale Medikamententherapie bei Innenohrschwerhörigkeit. HNO. 2015;**63**(6):396-401

[13] Plontke SK, Götze G, Rahne T, Liebau A. Intracochlear drug delivery in combination with Cochlear implants. HNO. 2017;**65**(1):19-28

[14] Liu Y, Jolly C, Braun S, Stark T, Scherer E, Plontke SK, et al. In vitro and in vivo pharmacokinetic study of a dexamethasone-releasing silicone for Cochlear implants. European Archives of Oto-Rhino-Laryngology. 2016;**273**(7):1745-1753

[15] Honeder C, Zhu C, Schöpper H, Gausterer JC, Walter M, Landegger LD, et al. Effects of sustained release dexamethasone hydrogels in hearing preservation cochlear implantation. Hearing Research. 2016;**341**:43-49

[16] Douchement D, Terranti A, Lamblin J, Salleron J, Siepmann F, Siepmann J, et al. Dexamethasone eluting electrodes for Cochlear implantation: Effect on residual hearing. Cochlear Implants International. 2015;**16**(4):195-200

[17] Cho HS, Lee K-Y, Choi H, Jang JH, Lee SH. Dexamethasone is one of the factors minimizing the inner ear damage from electrode insertion in cochlear implantation. Audiology & Neuro-Otology. 2016;**21**(3):178-186

[18] Skarżyńska MB, Skarżyński PH, Król B, Kozieł M, Osińska K, Gos E, et al. Preservation of hearing following cochlear implantation using different steroid therapy regimens: A prospective clinical study. Medical Science Monitor: International Medical Journal of Experimental and Clinical Research. 2018;**24**:2437

[19] Skarżyński H, Lorens A, Skarżyński PH. Electro-natural stimulation (ENS) in partial deafness treatment: A case study. The Journal of Health Science. 2014;**4**(4):67-71

[20] Skarzynski H, van de Heyning P, Agrawal S, Arauz SL, Atlas M, Baumgartner W, et al. Towards a consensus on a hearing preservation classification system. Acta Oto-Laryngologica. Supplementum. 2013;(564):3-13

[21] Skarzynski H, Lorens A, Dziendziel B, Skarzynski PH. Expanding pediatric cochlear implant candidacy: A case study of electro-natural stimulation (ENS) in partial deafness treatment. International Journal of Pediatric Otorhinolaryngology. 2015;**79**(11):1896-1900

Edited by Stavros Hatzopoulos,
Andrea Ciorba and Piotr H. Skarzynski

This book presents the latest findings in clinical audiology with a strong emphasis on new emerging technologies that facilitate and optimize a better assessment of the patient. The book has been edited with a strong educational perspective (all chapters include an introduction to their corresponding topic and a glossary of terms). The book contains material suitable for graduate students in audiology, ENT, hearing science and neuroscience.

Published in London, UK

IntechOpen

ISBN 978-1-83968-453-1