Hearing Research 335 (2016) 25–32

Contents lists available at ScienceDirect

Hearing Research

journal homepage: www.elsevier.com/locate/heares

Research paper

# Functional magnetic resonance imaging confirms forward suppression for rapidly alternating sounds in human auditory cortex but not in the inferior colliculus

Christian Harm Uhlig*, Andrew R. Dykstra, Alexander Gutschalk**

*Department of Neurology, University of Heidelberg, INF 400, 69120 Heidelberg, Germany*

ARTICLE INFO

*Article history:*
Received 19 November 2015
Received in revised form
8 February 2016
Accepted 15 February 2016
Available online 17 February 2016

*Keywords:*
Forward suppression
Selective adaptation
Stimulus-specific adaptation
Inferior colliculus
Stream segregation
Interaural time differences
Functional magnetic resonance imaging

ABSTRACT

Forward suppression at the level of the auditory cortex has been suggested to subserve auditory stream segregation. Recent results in non-streaming stimulation contexts have indicated that forward suppression can also be observed in the inferior colliculus; whether this holds for streaming-related contexts remains unclear. Here, we used cardiac-gated fMRI to examine forward suppression in the inferior colliculus (and the rest of the human auditory pathway) in response to canonical streaming stimuli (rapid tone sequences comprised of either one repetitive tone or two alternating tones). The first stimulus is typically perceived as a single stream, the second as two interleaved streams. In different experiments using either pure tones differing in frequency or bandpass-filtered noise differing in inter-aural time differences, we observed stronger auditory cortex activation in response to alternating vs. repetitive stimulation, consistent with the presence of forward suppression. In contrast, activity in the inferior colliculus and other subcortical nuclei did not significantly differ between alternating and monotonic stimuli. This finding could be explained by active amplification of forward suppression in auditory cortex, by a low rate (or absence) of cells showing forward suppression in inferior colliculus, or both.



## 1. Introduction

Separating sound sources in complex environments is a critical function of the auditory system that allows humans and other animals to hear out and selectively attend sources of interest. A widely used paradigm in the study of such source segregation is auditory 'stream-segregation', or 'streaming' (Bregman, 1990; Miller and Heise, 1950; van Noorden, 1975), in which a repetitive sequence of alternating tones (A and B) can be perceived either as one integrated stream or two distinct, segregated streams. Most previous studies have utilized frequency differences (ΔF) between the A and B tones to examine streaming, but differences along other dimensions – e.g. pitch (Vliegen et al., 1999), amplitude modulation (Grimault et al., 2002) or spatial lateralization (Boehnke and Phillips, 2005) – can also produce streaming.

Microelectrode studies in animal models suggest that streaming based on ΔF could be subserved by frequency-specific forward suppression in auditory cortex (AC) (Bee and Klump, 2004; Brosch and Schreiner, 1997; Fishman et al., 2004, 2001; Scholes et al., 2015, 2011). Similarly, the spatial tuning of AC neurons is known to be sharpened in streaming contexts that can be well explained by a stimulus-specific forward suppression model (Middlebrooks and Bremen, 2013). On a macroscopic scale, analogous results have been obtained in human listeners with magnetoencephalography (Gutschalk et al., 2005), electroencephalography (Snyder et al., 2006), and functional magnetic resonance imaging (fMRI) (Gutschalk et al., 2007; Schadwinkel and Gutschalk, 2010; Wilson et al., 2007). However, one question that remains is whether the streaming-related forward suppression that has been observed in AC emerges there or instead reflects the output of subcortical processes.[1]

Forward suppression has been observed in the inferior colliculus

*Abbreviations:* ΔF, Frequency differences; ΔITD, Differences of Interaual Time Differences; AC, Auditory Cortex; CN, Cochlear Nucleus; fMRI, functional magnetic resonance imaging; HG, Heschl's gyrus; ITD, Interaural Time Difference; MGB, Medial geniculate nucleus; NLL, Nuclei of the lateral lemniscus; PT, Planum Temporale; ROI, Region of Interest; S.D., Standard Deviation; S.E.M, Standard Error of the Mean; SOC, Superior Olivary Complex

* Corresponding author.
** Corresponding author.
*E-mail addresses:* christian@uhlig.email (C.H. Uhlig), alexander.gutschalk@med.uni-heidelberg.de (A. Gutschalk).

[1] Long-term adaptation, which has been suggested to subserve stream segregation independent of forward suppression (Micheyl et al., 2005), has in fact been observed in the cochlear nucleus (CN) of anaesthetized guinea pigs (Pressnitzer et al., 2008).

http://dx.doi.org/10.1016/j.heares.2016.02.010
0378-5955/

(IC) of awake marmosets using a two-tone paradigm (Nelson et al., 2009). Furthermore, stimulus-specific adaptation (Pérez-González and Malmierca, 2014), a phenomenon akin to forward suppression, has been observed in the rodent IC and medial geniculate nucleus (MGB) within the context of the auditory oddball paradigm, where rare deviant sounds are presented amidst frequent standards (Anderson et al., 2009; Malmierca et al., 2009; Patel et al., 2012; Pérez-González et al., 2005; Zhao et al., 2011). Whether forward suppression and stimulus-specific adaptation reflect a common or different underlying mechanisms remains unclear. Both are simply defined by reduction of neural responses when a stimulus is repeated, which is stronger for identical stimuli and weaker for more dissimilar stimuli. Two of us have previously used another synonym – selective adaptation – in the context of streaming (Gutschalk et al., 2005; Gutschalk and Dykstra, 2014) and consider the three terms interchangeable. In the present paper we use forward suppression in the context of streaming and stimulus-specific adaptation in the context of the oddball paradigm, as that is how previous, disparate lines of research have used them. This also allows for the possibility that different mechanisms or anatomical centres may be recruited by the two paradigms.

Using sparse-sampled (Edmister et al., 1999; Hall et al., 1999), cardiac-gated (Guimaraes et al., 1998) fMRI, we examined whether streaming-related forward suppression observed in AC is inherited from subcortical structures, particularly the IC. We hypothesized that stimulus sequences with alternating frequencies (ΔF, experiment 1) or interaural time differences (ΔITD, experiment 2) would produce larger blood-oxygenation level-dependent (BOLD) activity than monotone control sequences, reflecting the presence of forward suppression in both the IC and AC.

## 2. Materials and methods

### 2.1. Participants

30 healthy listeners participated in the study, 15 per experiment. The mean age was 25 ± 3 years (standard deviation: S.D.) in experiment 1 and 23 ± 2 years (S.D.) in experiment 2. Ten participants of experiment 1 and twelve participants of experiment 2 were female. An additional five participants (three from experiment 1 and two from experiment 2) were excluded from analysis due to excessive movement in the scanner. All participants had clinically normal pure-tone audiograms with threshold less than 15 dB HL between 0.125 and 12.5 kHz. Each participant provided written informed consent prior to their participation in the experiments. All experiments were approved by the ethics committee of Heidelberg University Medical School and conform with the in accordance with The Code of Ethics of the World Medical Association (Declaration of Helsinki).

### 2.2. Stimuli and procedures

All stimuli were generated with MATLAB (The Mathworks, Natick, MA, USA) and stored as wave files with a sample-rate of 48 kHz. The wave files were presented via a D/A converter and headphone amplifier (MR Confon; MR confon GmbH, Magdeburg, Germany) with Sensimetrics S14 in-ear headphones (Sensimetrics Corporation, Malden, MA, USA). The non-linear transfer function of the in-ear headphones was corrected using software provided by the manufacturer.

#### 2.2.1. Experiment 1: frequency separation – ΔF

The stimuli consisted of sequences of pure tones labelled A and B, with frequencies of 600 and 849 Hz, respectively (i.e., the B-tone was 6 semitones higher than the A tone), at 74 dB sound-pressure level (SPL). The three conditions were: CONST1 (AAAA …), CONST2 (BBBB …), and ALT (ABAB …). The two CONST conditions were used in order to equalize the long-term power spectra between the ALT and CONST conditions so as to avoid such a confound in the ALT-versus-CONST contrast. The length of each tone was 100 ms with 15 ms on- and off-raised-cosine ramps. There were no silent interstimulus intervals between subsequent tones. Each tone sequence comprised 320 tone repetitions, amounting to a duration of 32 s. The ALT condition was repeated 32 times and the CONST1 and CONST2 conditions were presented 16 times each. The sequences were presented in pseudo-random order with a variable interstimulus interval of 24–32 s. This was done in order to stagger the BOLD volume acquisitions with respect to stimulus onset to allow for reconstruction of the BOLD signal timecourse (see Section 2.3.). The presentation of each 32-s tone sequence was started with a delay of 0, 2, 4, or 6 s relative to the trigger sent at the onset of the preceding scanner acquisition. Because of the cardiac gating procedure (see below), there is additional variability of the interstimulus interval in the sub-second range.

#### 2.2.2. Experiment 2: spatial separation – ΔITD

In experiment 2, stimuli were composed of band-limited (0–2 kHz) noise bursts, lateralized to the left (L) or right (R) with an ITD of ±500 μs. Sound level was 75 dB SPL. In analogy to experiment 1, three conditions were used: CONST1 (RRRR …), CONST2 (LLLL …), and ALT (RLRL …). Parameters and procedures were otherwise similar to experiment 1. Conditions CONST1 and 2 were presented 12 times each, and ALT was presented 24 times.

#### 2.2.3. Procedures for both experiments

The task was explained one day before the fMRI session, including one or two training runs with circumaural headphones connected to a desktop computer. The training ended after the experimenter was confident that the participants understood the task, although some participants nonetheless made a small number of erroneous button presses. To determine if the participants perceived the stimuli as one or two separate streams in the scanner, they were instructed to evaluate the stimuli by pressing one of two buttons after the end of each 32 s sequence. One button indicated that they had heard one sequence most of the time, while the other button indicated that they had heard two separate streams most of the time.

### 2.3. Imaging

All MRI data were acquired in a Siemens 3T Magnetom Tim Trio scanner (Siemens, Erlangen, Germany) with a 32-channel, phased-array head coil. Two T1-weighted magnetization-prepaired rapid gradient echo sequences (MPRAGE) with a dimension of 256 × 256 × 192 voxel, an isovoxel resolution of 1 $mm^3$, a TR of 1570 ms, a TE of 2.63 ms, a TI of 900 ms, and a Flip Angle of 9° in one frame for each participant was collected. These scans were used to place the functional volume. To cover the complete ascending auditory pathway, the functional volume was placed in a near-coronal orientation orthogonal to the Sylvian fissure, such that AC and the brainstem were inside the volume. The functional volume comprised 21 slices (2.1 mm thickness, distance 10%) with a field of view of 204 × 204 mm (120 × 120 voxel, resolution 1.7 × 1.7 mm) and included the brain-stem as well as most of AC. The parameters for BOLD imaging were echo time (TE) = 42 ms, inversion time (TI) = −1 ms, Flip Angle 90° the phase coding was chosen from feet to head. In-house software was used for stimulus presentation and collection of participants' responses with a LUMItouch optical response keypad (Photon Control, Burnaby, BC, Canada). A modified sparse-imaging (Edmister et al., 1999; Hall et al., 1999) and cardiac-gated (Guimaraes et al., 1998) paradigm with an average TR of 8.0 s