# Time-Locked Auditory Cortical Responses in the High-Gamma Band: A Window into Primary Auditory Cortex

**Jonathan Z. Simon[1, 2, 3*], Vrishab Commuri[1], Joshua P. Kulasingham[4]**

[1]Department of Electrical and Computer Engineering, University of Maryland, College Park MD, United States

[2]Department of Biology, University of Maryland, College Park MD, United States

[3]Institute for Systems Research, University of Maryland, College Park MD, United States

[4]Department of Electrical Engineering, Linköping University, Linköping, Sweden

[*]**Correspondence:**
Jonathan Z. Simon
jzsimon@umd.edu

**Keywords: phase locked response, medial geniculate body, high frequency, envelope following response, cortical FFR**

## Abstract

Primary auditory cortex is a critical stage in the human auditory pathway, a gateway between subcortical and higher-level cortical areas. Receiving the output of all subcortical processing, it sends its output on to higher-level cortex. Non-invasive physiological recordings of primary auditory cortex using electroencephalography (EEG) and magnetoencephalography (MEG), however, may not have sufficient specificity to separate responses generated in primary auditory cortex from those generated in underlying subcortical areas or neighboring cortical areas. This limitation is important for investigations of effects of top-down processing (e.g., selective-attention-based) on primary auditory cortex: higher-level areas are known to be strongly influenced by top-down processes, but subcortical areas are often assumed to perform strictly bottom-up processing. Fortunately, recent advances have made it easier to isolate the neural activity of primary auditory cortex from other areas. In this perspective, we focus on time-locked responses to stimulus features in the high gamma band (70-150 Hz) and with early cortical latency (~40 ms), intermediate between subcortical and higher-level areas. We review recent findings from physiological studies employing either repeated simple sounds or continuous speech, obtaining either a frequency following response (FFR) or temporal response function (TRF). The potential roles of top-down processing are underscored, and comparisons with invasive electrocorticography (ECoG) and animal model recordings are made. We argue that MEG studies employing continuous speech stimuli may offer particular benefits, in that only a few minutes of speech generates robust high gamma responses from bilateral primary auditory cortex, and without measurable interference from subcortical or higher-level areas.

## Contribution to the Field

In this perspective, we investigate non-invasive physiological auditory responses from human primary auditory cortex, obtained using electroencephalography (EEG) and magnetoencephalography (MEG) in the high gamma band (70-150 Hz). These responses also have a short cortical latency (~40

ms), later than subcortical areas but earlier than higher-level cortical areas. Recent findings from EEG and MEG studies are reviewed, that employ either repeated simple sounds or continuous speech, and analyzed using frequency following responses (FFRs) or temporal response functions (TRFs), respectively. Different approaches give different insight as to the role of human primary auditory cortex, and especially how it is affected by top-down cortical processing. In particular, MEG studies employing continuous speech stimuli offer a “sweet spot”, whereby using only a few minutes of speech stimulus can generate robust high gamma responses from primary auditory cortex, without measurable interference from either subcortical, or other higher-level cortical areas.

## 1 Introduction

Primary auditory cortex plays a key role in the human brain’s processing of sounds, being a major gateway between auditory subcortical areas, including the inferior colliculus (midbrain) and thalamus, and higher order auditory cortical areas, including secondary auditory areas, associative auditory areas, and language areas. While the neurophysiology of primary auditory cortex has been studied for decades in animal models, there are still many unanswered questions. One of the hallmarks of primary auditory cortex in animal models is its sluggishness compared to subcortical areas, since its typical neurons time-lock[1] to acoustic modulations only up to a few tens of Hz (Lu et al., 2001; Joris et al., 2004), though at the same time it does respond very reliably (temporally) to brief acoustic features, with a precision of milliseconds (Elhilali et al., 2004).

Less is known about temporal processing in *human* primary auditory cortex, where neurophysiological recording techniques for healthy subjects are restricted to non-invasive methods, primarily electroencephalography (EEG) and magnetoencephalography (MEG). Neither EEG nor MEG has very fine spatial resolution (typically a few centimeters) and so may not be able to distinguish different neural sources based purely on their anatomical origin. Both, however, have sufficient temporal resolution to distinguish typical response latencies of primary auditory cortex (~40 ms) from subcortical (shorter latency) and non-primary (longer latency) auditory areas.

Beyond these commonalities, EEG and MEG have distinctive strengths and weaknesses. EEG is sensitive to neural sources throughout the brain at both low frequencies (tens of Hz) and high frequencies (hundreds of Hz) (Kraus et al., 2017; White-Schwoch et al., 2019). It is therefore relatively straightforward to record time-locked activity from any auditory area of the brain, but it may be difficult to distinguish contributions from multiple areas, at least without additional information (e.g., response latency, which can be used to distinguish between the sources giving rise to the auditory P1 and N1 components). In contrast, MEG is insensitive to subcortical neural sources (Hämäläinen et al., 1993), though not entirely unresponsive, as seen below. Perhaps counterintuitively, this insensitivity gives MEG an advantage over EEG, by allowing recordings from auditory cortical sources without substantial subcortical interference (Ross et al., 2020). Nevertheless, MEG responses from different auditory cortical areas can still interfere with each other.

Another consideration is that EEG’s sensitivity to most auditory sources holds for both low and high frequencies, but because of MEG’s cortical bias and because cortical responses are usually sluggish, MEG typically only captures cortical sources at low frequencies. An important counterexample, however, is the case of fast (~100 Hz) auditory time-locked cortical responses (Hertrich et al., 2012;

---

[1] We employ the term “time-locked” neural responses rather than “phase-locked” since phase is only defined when the coupled stimulus/response is analyzed in a narrow frequency band. The term “time-locking”, sometimes called “neural tracking” when applied to low frequency responses to speech, applies equally well to narrowband and broadband cases.

Coffey et al., 2016). At these frequencies there are few, if any, cortical sources aside from primary auditory cortex. In this sense, MEG recordings of fast time-locked auditory cortical responses act as an exquisite window into primary auditory cortex, without interference from subcortical or other cortical areas. Therefore, it may be especially suited for questions regarding how primary auditory cortical responses are affected by cognitive processes, whether modulated by top-down neural activity (e.g., selective attention or task-specific processing) or supplemented by super-auditory aspects of the stimulus (e.g., processing of speech sounds using language-based information).

One newly established method to analyze neural responses to continuous speech (Hamilton & Huth, 2018) is temporal response function (TRF) analysis (Lalor et al., 2009; Ding & Simon, 2012). TRFs are an effective tool to disambiguate neural sources based on their characteristic latencies, as will be discussed below.

## 2 Results

Fast (~100 Hz) cortical time-locked auditory responses are typically investigated using one of two different stimulus paradigms. The more time-honored paradigm is the frequency following response (FFR)(Kraus et al., 2017), for which a typical stimulus is either acoustically simple, such as click trains or amplitude modulated tones (e.g., Gorina-Careta et al., 2021), or consists of many repetitions of a short but more complex stimulus, such as a single syllable (e.g., Coffey et al., 2016).

The well-established FFR paradigm (or really, family of paradigms, including the envelope following response; EFR) has been used to great effect with EEG to investigate midbrain responses to acoustic stimuli. Near 100 Hz, midbrain sources dominate the EEG FFR over cortical sources, and well above 100 Hz there is little to no cortical EEG FFR contribution at all (Coffey et al., 2019). Until the MEG FFR investigations of Coffey et al. (2016), however, it was not widely appreciated how substantial the cortical FFR contributions might be near 100 Hz. In this seminal paper, the investigators presented the 120-ms syllable /da/, synthesized with a 98 Hz fundamental frequency in the vowel portion, for 14,000 repetitions (sufficient to also obtain responses from subcortical sources despite the cortical bias of MEG). The cortical responses, whose sources were consistent with primary auditory cortex, were prominent and showed a significant lateralization to the right-hemisphere, with a longer latency profile compared to subcortical components. This work firmly established the measurability of distinct cortical contributions to the FFR near 100 Hz. In comparison, Gorina-Careta et al. (2021) demonstrated that the MEG FFR at the much higher frequency of 333 Hz (15,200 tone-burst repetitions) originated solely from subcortical sources (Figure 1). Note that both these studies demonstrate that, while MEG is not incapable of measuring high frequency FFR from subcortical

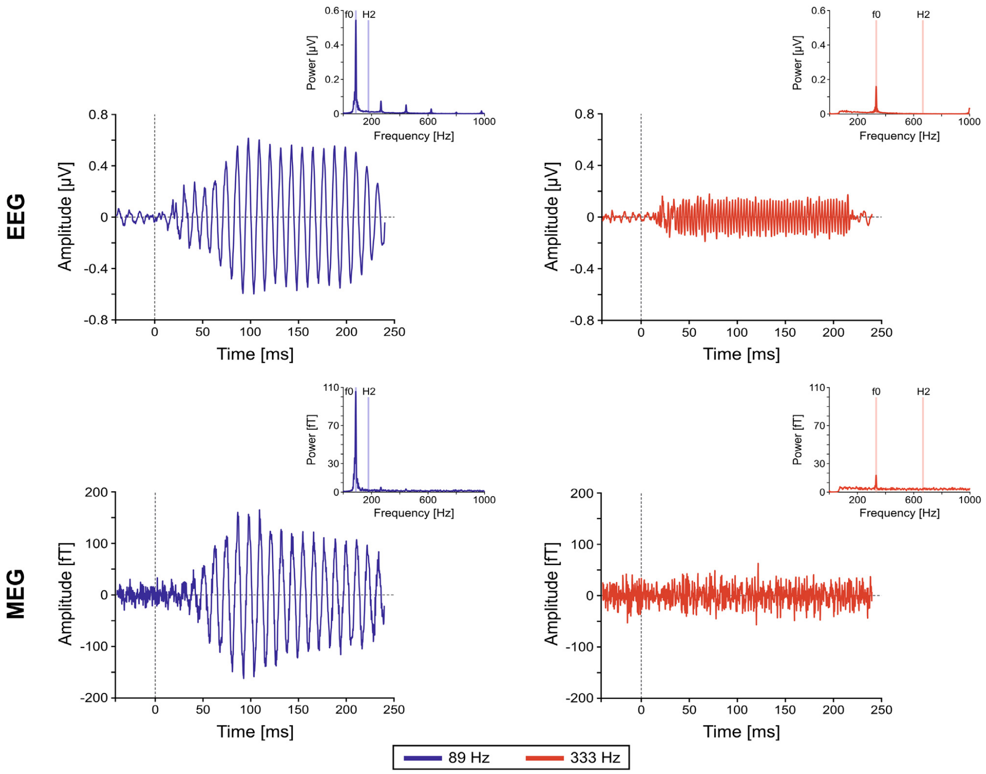

Figure 1. Grand-averaged FFR time course and spectral representations (insets) of single-channel EEG and MEG elicited in the high gamma frequency range (89 Hz; blue) and the very high gamma range (333 Hz; red). It can be shown that the very high gamma frequency (333 Hz; red) FFR is almost entirely subcortical for both EEG and MEG. In contrast, the high gamma frequency (89 Hz; blue) FFR is almost entirely cortical for MEG and a mix of cortical and subcortical for EEG. [From Gorina-Careta et al. (2021), Fig. 1.]

sources, the number of repetitions required is considerable, with an associated experimental design cost (e.g., limited to a small number of stimulus types).

One of the limitations of the FFR paradigm is that accessing the different latencies of distinct sources may not be straightforward, since the FFR is ultimately just the evoked response to a sustained stimulus: a linear sum of overlapping responses from multiple sources with different latencies (Teichert et al., 2022). A more recently developed paradigm uses neural responses to continuous speech, such as individual sentences (e.g., Hertrich et al., 2012) or longer narrated story passages (e.g., Kulasingham et al., 2020). The use of the continuous speech stimulus paradigm, combined with TRF analysis, sidesteps this temporal overlap issue by deconvolving the sustained response from the stimulus, which often allows direct comparison of neural source peak latencies. Though typical uses of TRF analysis employ the slow (< 10 Hz) acoustic envelope as the stimulus feature with which to deconvolve (Di Liberto et al., 2015; Cervantes Constantino & Simon, 2018), the TRF methodology generalizes well to other stimulus features (Brodbeck & Simon, 2020). This includes responses from high frequency stimulus features processed in subcortical areas (Maddox & Lee, 2018; Polonenko & Maddox, 2021).

High frequency (70 Hz – 200 Hz) MEG TRFs were first investigated by Kulasingham et al. (2020) using only six minutes of continuous speech as the stimulus. Responses source-localized to bilateral primary auditory cortex, with a small but significant lateralization to the right hemisphere (Figure 2A). The peak latency of the cortical response, 40 ms, is consistent with a primary auditory cortical origin. Analysis additionally revealed that frequencies contributing to time-locking fell off substantially above 100 Hz. This demonstration that such a short recording can reveal responses localized to primary auditory cortex serves several purposes. It allows future experiments to include multiple stimulus conditions (e.g., presenting stimuli under different task conditions or at different SNRs), and at the same time ensures that the responses do not contain measurable subcortical interference.

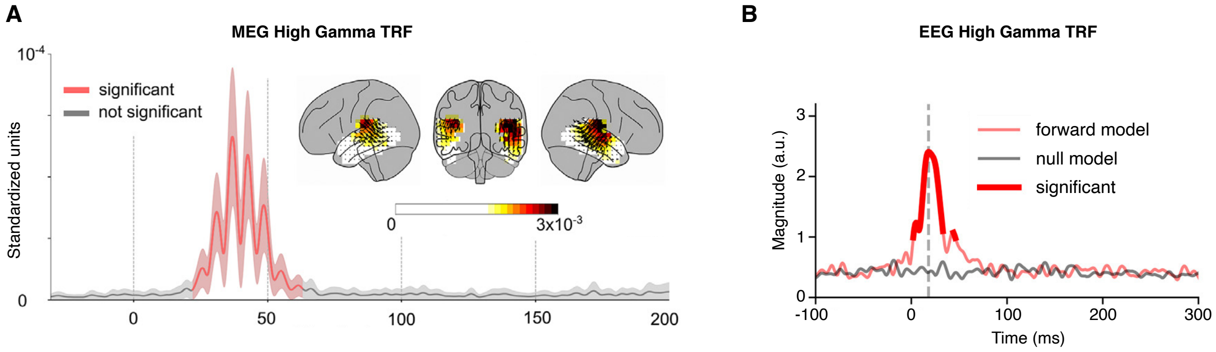

Figure 2. High Gamma TRFs. **A**. MEG high frequency (70 Hz – 200 Hz) TRF from six minutes of continuous speech. The figure shows the grand-averaged amplitude of TRF source-localized current-dipole vectors, averaged across voxels in the cortical ROI (± standard error across subjects; red indicates amplitude significantly greater than noise). The TRF has a peak latency of ~40 ms and oscillates with a frequency of ~80 Hz (note that since only the TRF amplitude is shown, and not signed current values, signal troughs and peaks both appear as peaks). Inset: the distribution of TRF current-dipole vectors in the brain at each voxel at the moment of the maximum response; color represents response strength and arrows represent TRF current-dipole orientations. [Modified from Kulasingham et al. (2020), Fig. 3.] **B.** EEG high frequency (70 Hz – 200 Hz) TRF from 40 minutes of continuous speech. The figure shows the grand-averaged Hilbert envelope of the TRF averaged across channels; bright red indicates magnitude significantly greater than the null model. The TRF magnitude significantly exceeds that of the null model in two latency ranges: between 2 and 33 ms with a peak at 18 ms (dominantly subcortical), and between 44 ms to 46 ms with a peak at 45 ms (dominantly cortical). [Modified from Kegler et al. (2022), Fig. 3.]

High frequency (70 Hz – 200 Hz) EEG TRFs with cortical contributions have also been recently investigated by Kegler et al. (2022). These TRFs show a pair of peaks with distinguishable latencies allowing inference of separate sources, each with a separate anatomical origin and auditory processing role (analogous to traditional P1 and N1 peaks arising from separate cortical sources). In this case, the earlier peak at 18 ms is consistent with a subcortical origin, and the later peak at 45 ms is consistent with a dominantly cortical origin (Figure 2B).

It should not be surprising that earlier invasive ECoG recordings had already demonstrated similar high gamma time-locked cortical responses almost a decade earlier (Brugge et al., 2009; Steinschneider et al., 2013), using click trains and isolated speech sounds. What is surprising is that such responses could be seen even non-invasively. The most robust time-locked high gamma ECoG responses are seen in primary auditory cortex, specifically posteromedial Heschl's gyrus (Nourski, 2017), but smaller time-locked high gamma responses are also seen in other auditory cortical areas. As such, ECoG remains a premiere electrophysiological method for obtaining responses known to originate in primary auditory cortex, but only for a fraction of subjects relative to those eligible for MEG or EEG recordings.

## 3 Discussion

As indicated above, a physiological window into human primary auditory cortex allows the investigation of the extent to which primary auditory cortex is influenced by higher order cortical areas. How, and under which circumstances, are primary auditory cortical responses modulated by top-down neural activity, or affected by language-specific non-auditory features of the stimulus? A related question is to what extent subcortical auditory areas might be influenced by cortical processing. Neither can be answered without first identifying the specific sources of neural activity (e.g., midbrain vs. thalamus vs. primary auditory cortex) being modulated by distant cortical activity.

Using MEG, Hartmann & Weisz (2019) demonstrated that the FFR near 100 Hz from right hemisphere primary auditory cortex is modulated by intermodal (auditory vs. visual) attention. Most FFR investigations use EEG, which is well-suited to separate responses from primary auditory cortex from those originating in other cortical areas, but, as indicated above, has difficulty in separating auditory subcortical and primary auditory cortical contributions. Intriguing results include: modulation of the EEG FFR by selective attention for frequencies near 100 Hz but not above 200 Hz (Holmes et al., 2018); modulation by overall level of attention near 150 Hz (Price & Bidelman, 2021); and, at 100 Hz, modulation by whether a continuous-speech masker is in a known vs. unknown (but acoustically similar) language (Presacco et al., 2016; Zan et al., 2019). There has also been a report of selective attentional modulation of subcortical auditory responses to continuous speech (Forte et al., 2017); the result has not yet been replicated, however, and due to the specialty of the analysis method it is as yet difficult to rule out entirely whether the result might be due to cortical response leakage.

More recently, using EEG with a continuous speech stimulus, Kegler et al. (2022) demonstrated that the high gamma EEG TRF arising from a combination of subcortical and primary auditory cortical sources (illustrated in Figure 2B) is modulated by word-boundary effects. This is strong evidence that a linguistic (super-acoustic) feature can modulate either primary auditory cortical or auditory subcortical processing (or both). Commuri et al. (in preparation) have also recently demonstrated that the high gamma MEG TRF, originating solely from bilateral primary auditory cortex, is indeed modulated by selective attention, using re-analysis of previously published data (Kulasingham et al., 2021).

There is additional evidence that human primary auditory cortical responses exhibit modulation arising from other cortical areas, but the effects are subtle. Using ECoG and employing selective attention to one of two competing talkers, O'Sullivan et al. (2019) did not observe modulation of cortical responses in Heschl’s gyrus (the anatomical location of primary auditory cortex), while, in contrast, they did find modulation in non-primary areas, as expected. Using a similar paradigm to investigate the role of selective attention on MEG low frequency cortical TRFs, Brodbeck et al. (2018), did see evidence of significant TRF modulation at short latencies consistent with a primary auditory cortex origin (in addition to the expected strong modulation at longer latencies), but only under limited conditions.

In animal studies, top-down (task-dependent) modulation of neural activity in primary auditory cortex has been seen as far back as two decades ago (Fritz et al., 2003). Despite the robustness and reproducibility of these results, however, the effect size is nevertheless small, and it has not been clear until recently whether such modulations would ever be observable non-invasively.

What is the physiological origin of the high gamma time-locked responses from primary auditory cortex? Two theories have been put forward. The first concerns the physics underlying the generators of EEG and MEG signals, which are dominantly driven by dendritic currents produced by synaptic inputs (Hämäläinen et al., 1993; Buzsaki et al., 2012), i.e., the same mechanisms that also give rise to the local field potential (LFP). For primary auditory cortex, the most significant neural input is the spiking output of the medial geniculate body (MGB) of the thalamus, whose spiking rates can reach up to 100 Hz (Miller et al., 2002). A second theory, strongly tied to the first, is that the spikes of primary auditory cortex, which can only fire at rates well below 100 Hz, can nevertheless fire with temporal precision of the order of milliseconds (Elhilali et al., 2004). It has been recently shown by Downer et al. (2021) that these precise but infrequent spikes are actually highly synchronous across the local population, even to the point of acting as a time-locked *population* model for fast acoustic features (almost up to 200 Hz). Indeed, Gnanateja et al. (2021) recently demonstrated a connection between both these explanations, using intracortical FFR (90-140 Hz) recordings from multiple species, to show both an LFP FFR and a multi-unit (spiking) FFR, in the thalamorecipient layers of primary auditory cortex.

In conclusion, recent advances in auditory neuroscience have opened up new non-invasive windows into the neurophysiology of primary auditory cortex. Using EEG FFR techniques, responses are dominantly subcortical but also contain strong contributions from primary auditory cortex at frequencies near 100 Hz. Using MEG FFR techniques, responses are dominantly from primary auditory cortex for frequencies near 100 Hz (though at higher frequencies subcortical responses can also be detected given sufficient recording time). EEG TRF studies have the potential to show both auditory subcortical and primary auditory cortical contributions to the time-locked high gamma responses to continuous speech, but, unlike FFR, segregated in time/latency. Finally, MEG time-locked high gamma TRF studies may hold great promise in isolating primary auditory cortical responses from other areas, due to its insensitivity to subcortical sources and its ability to differentiate competing cortical sources in both time and anatomical location.

## 4 Author Contributions

JZS wrote the initial draft of the manuscript. All authors contributed to the interpretations of results and discussions, were involved in manuscript revision, and approved the final version.

## 5 Funding

This work was supported by grants from the National Institutes of Deafness and Other Communication Disorders (R01-DC019394), the National Institute on Aging (P01-AG055365), the National Science Foundation (SMA-1734892) and the William Demant Foundation (20-0480).

## 6 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.

## References

Brodbeck C, Hong LE, Simon JZ (2018) *Rapid Transformation from Auditory to Linguistic Representations of Continuous Speech*, Curr Biol 28 3976-3983 e3975. DOI: 10.1016/j.cub.2018.10.042

Brodbeck C, Simon JZ (2020) *Continuous speech processing*, Curr Opin Physiol 18 25-31. DOI: 10.1016/j.cophys.2020.07.014

Brugge JF, Nourski KV, Oya H, Reale RA, Kawasaki H, Steinschneider M, Howard MA, 3rd (2009) *Coding of repetitive transients by auditory cortex on Heschl's gyrus*, J Neurophysiol 102 2358-2374. DOI: 10.1152/jn.91346.2008

Buzsaki G, Anastassiou CA, Koch C (2012) *The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes*, Nat Rev Neurosci 13 407-420. DOI: 10.1038/nrn3241

Cervantes Constantino F, Simon JZ (2018) *Restoration and Efficiency of the Neural Processing of Continuous Speech Are Promoted by Prior Knowledge*, Front Syst Neurosci 12 56. DOI: 10.3389/fnsys.2018.00056

Coffey EB, Herholz SC, Chepesiuk AM, Baillet S, Zatorre RJ (2016) *Cortical contributions to the auditory frequency-following response revealed by MEG*, Nat Commun 7 11070. DOI: 10.1038/ncomms11070

Coffey EBJ, Nicol T, White-Schwoch T, Chandrasekaran B, Krizman J, Skoe E, Zatorre RJ, Kraus N (2019) *Evolving perspectives on the sources of the frequency-following response*, Nat Commun 10 5036. DOI: 10.1038/s41467-019-13003-w

Commuri V, Kulasingham JP, Simon JZ (in preparation) *Cortical Responses Time-Locked to Continuous Speech in the High-Gamma Band Depend on Selective Attention*,

Di Liberto GM, O'Sullivan JA, Lalor EC (2015) *Low-Frequency Cortical Entrainment to Speech Reflects Phoneme-Level Processing*, Curr Biol 25 2457-2465. DOI: 10.1016/j.cub.2015.08.030

Ding N, Simon JZ (2012) *Emergence of neural encoding of auditory objects while listening to competing speakers*, Proc Natl Acad Sci U S A 109 11854-11859. DOI: 10.1073/pnas.1205381109

Downer JD, Bigelow J, Runfeldt MJ, Malone BJ (2021) *Temporally precise population coding of dynamic sounds by auditory cortex*, J Neurophysiol 126 148-169. DOI: 10.1152/jn.00709.2020

Elhilali M, Fritz JB, Klein DJ, Simon JZ, Shamma SA (2004) *Dynamics of precise spike timing in primary auditory cortex*, J Neurosci 24 1159-1172. DOI: 10.1523/JNEUROSCI.3825-03.2004

Forte AE, Etard O, Reichenbach T (2017) *The human auditory brainstem response to running speech reveals a subcortical mechanism for selective attention*, Elife 6 DOI: 10.7554/eLife.27203

Fritz J, Shamma S, Elhilali M, Klein D (2003) *Rapid task-related plasticity of spectrotemporal receptive fields in primary auditory cortex*, Nat Neurosci 6 1216-1223. DOI: 10.1038/nn1141

Gnanateja GN, Rupp K, Llanos F, Remick M, Pernia M, Sadagopan S, Teichert T, Abel TJ, Chandrasekaran B (2021) *Frequency-Following Responses to Speech Sounds Are Highly Conserved across Species and Contain Cortical Contributions*, eNeuro 8 DOI: 10.1523/ENEURO.0451-21.2021

Gorina-Careta N, Kurkela JLO, Hamalainen J, Astikainen P, Escera C (2021) *Neural generators of the frequency-following response elicited to stimuli of low and high frequency: A* *magnetoencephalographic (MEG) study*, Neuroimage 231 117866. DOI: 10.1016/j.neuroimage.2021.117866

Hämäläinen M, Hari R, Ilmoniemi RJ, Knuutila J, Lounasmaa OV (1993) *Magnetoencephalography—theory, instrumentation, and applications to noninvasive studies of the working human brain*, Reviews of Modern Physics 65 413-497. DOI: 10.1103/RevModPhys.65.413

Hamilton LS, Huth AG (2018) *The revolution will not be controlled: natural stimuli in speech neuroscience*, Language, Cognition and Neuroscience 0 1-10. DOI: 10.1080/23273798.2018.1499946

Hartmann T, Weisz N (2019) *Auditory cortical generators of the Frequency Following Response are modulated by intermodal attention*, Neuroimage 203 116185. DOI: 10.1016/j.neuroimage.2019.116185

Hertrich I, Dietrich S, Trouvain J, Moos A, Ackermann H (2012) *Magnetic brain activity phase-locked to the envelope, the syllable onsets, and the fundamental frequency of a perceived speech signal*, Psychophysiology 49 322-334. DOI: 10.1111/j.1469-8986.2011.01314.x

Holmes E, Purcell DW, Carlyon RP, Gockel HE, Johnsrude IS (2018) *Attentional modulation of envelope-following responses at lower (93-109 Hz) but not higher (217-233 Hz) modulation rates*, J Assoc Res Otolaryngol 19 83-97. DOI: 10.1007/s10162-017-0641-9

Joris PX, Schreiner CE, Rees A (2004) *Neural processing of amplitude-modulated sounds*, Physiol Rev 84 541-577. DOI: 10.1152/physrev.00029.2003

Kegler M, Weissbart H, Reichenbach T (2022) *The neural response at the fundamental frequency of speech is modulated by word-level acoustic and linguistic information*, Front Neurosci 16 915744. DOI: 10.3389/fnins.2022.915744

Kraus N, Anderson S, White-Schwoch T (2017) The frequency-following response: a window into human communication. In: *The frequency-following response*, 1-15: Springer.

Kulasingham JP, Brodbeck C, Presacco A, Kuchinsky SE, Anderson S, Simon JZ (2020) *High gamma cortical processing of continuous speech in younger and older listeners*, NeuroImage 222 117291. DOI: 10.1016/j.neuroimage.2020.117291

Kulasingham JP, Joshi NH, Rezaeizadeh M, Simon JZ (2021) *Cortical Processing of Arithmetic and Simple Sentences in an Auditory Attention Task*, J Neurosci 41 8023-8039. DOI: 10.1523/JNEUROSCI.0269-21.2021

Lalor EC, Power AJ, Reilly RB, Foxe JJ (2009) *Resolving precise temporal processing properties of the auditory system using continuous stimuli*, J Neurophysiol 102 349-359. DOI: 10.1152/jn.90896.2008

Lu T, Liang L, Wang X (2001) *Temporal and rate representations of time-varying signals in the auditory cortex of awake primates*, Nat Neurosci 4 1131-1138. DOI: 10.1038/nn737

Maddox RK, Lee AKC (2018) *Auditory Brainstem Responses to Continuous Natural Speech in Human Listeners*, eNeuro 5 DOI: 10.1523/ENEURO.0441-17.2018

Miller LM, Escabi MA, Read HL, Schreiner CE (2002) *Spectrotemporal receptive fields in the lemniscal auditory thalamus and cortex*, J Neurophysiol 87 516-527. DOI: 10.1152/jn.00395.2001

Nourski KV (2017) *Auditory processing in the human cortex: An intracranial electrophysiology perspective*, Laryngoscope Investig Otolaryngol 2 147-156. DOI: 10.1002/lio2.73

O'Sullivan J, Herrero J, Smith E, Schevon C, McKhann GM, Sheth SA, Mehta AD, Mesgarani N (2019) *Hierarchical encoding of attended auditory objects in multi-talker speech perception*, Neuron 104 1195-1209 e1193. DOI: 10.1016/j.neuron.2019.09.007

Polonenko MJ, Maddox RK (2021) *Exposing distinct subcortical components of the auditory brainstem response evoked by continuous naturalistic speech*, Elife 10 DOI: 10.7554/eLife.62329

Presacco A, Simon JZ, Anderson S (2016) *Effect of informational content of noise on speech representation in the aging midbrain and cortex*, J Neurophysiol 116 2356-2367. DOI: 10.1152/jn.00373.2016

Price CN, Bidelman GM (2021) *Attention reinforces human corticofugal system to aid speech perception in noise*, Neuroimage 235 118014. DOI: 10.1016/j.neuroimage.2021.118014

Ross B, Tremblay KL, Alain C (2020) *Simultaneous EEG and MEG recordings reveal vocal pitch elicited cortical gamma oscillations in young and older adults*, Neuroimage 204 116253. DOI: 10.1016/j.neuroimage.2019.116253

Steinschneider M, Nourski KV, Fishman YI (2013) *Representation of speech in human auditory cortex: is it special?*, Hear Res 305 57-73. DOI: 10.1016/j.heares.2013.05.013

Teichert T, Gnanateja GN, Sadagopan S, Chandrasekaran B (2022) *A linear superposition model of envelope and frequency following responses may help identify generators based on latency*, Neurobiology of Language 3 441-468.

White-Schwoch T, Anderson S, Krizman J, Nicol T, Kraus N (2019) *Case studies in neuroscience: subcortical origins of the frequency-following response*, J Neurophysiol 122 844-848. DOI: 10.1152/jn.00112.2019

Zan P, Presacco A, Anderson S, Simon JZ (2019) *Mutual information analysis of neural representations of speech in noise in the aging midbrain*, J Neurophysiol 122 2372-2387. DOI: 10.1152/jn.00270.2019

## Figure Captions

Figure 1. Example FFRs. Grand-averaged FFR time course and spectral representations (insets) of single-channel EEG and MEG elicited in the high gamma frequency range (89 Hz; blue) and the very high gamma range (333 Hz; red). It can be shown that the very high gamma frequency (333 Hz; red) FFR is almost entirely subcortical for both EEG and MEG. In contrast, the high gamma frequency (89 Hz; blue) FFR is almost entirely cortical for MEG and a mix of cortical and subcortical for EEG. [From Gorina-Careta et al. (2021), Fig. 1.]

Figure 2. Example High Gamma TRFs. **A.** High frequency (70 Hz – 200 Hz) MEG TRF from six minutes of continuous speech. The grand-averaged amplitude of TRF source localized current-dipole vectors, averaged across voxels in the cortical ROI, is shown (± standard error across subjects; red indicates amplitude significantly greater than noise). The TRF has a peak latency of ~40 ms and oscillates with a frequency of ~80 Hz (note that since only the TRF amplitude is shown, and not signed current values, signal troughs and peaks both appear as peaks). Inset: the distribution of TRF current-dipole vectors in the brain at each voxel at the moment of the maximum response; color represents response strength and arrows represent TRF current-dipole orientations. [Modified from Kulasingham et al. (2020), Fig. 3.] **B.** High frequency (70 Hz – 200 Hz) EEG TRF from 40 minutes of continuous speech. The grand-averaged magnitude of the Hilbert transform of the TRF averaged across channels, is shown; bright red indicates magnitude significantly greater than the null model). The TRF magnitude significantly exceeds that of the null model in two latency ranges: between 2 and 33 ms with a peak at 18 ms (dominantly subcortical), and between 44 ms to 46 ms with a peak at 45 ms (dominantly cortical). [Modified from Kegler et al. (2022), Fig. 3.]