

# Temporal Context Invariance Reveals Neural Processing Timescales in Human Auditory Cortex

**Sam V. Norman-Haignere (sn2776@columbia.edu)**

Zuckerman Institute, Columbia University, 3227 Broadway  
New York, NY 10027 United States

**Laura K. Long (lkl2124@columbia.edu)**

Zuckerman Institute, Columbia University, 3227 Broadway  
New York, NY 10027 United States

**Orrin Devinsky (Orrin.Devinsky@nyulangone.org)**

Comprehensive Epilepsy Center, NYU Langone Medical Center, 223 East 34th Street  
New York, NY 10016 United States

**Werner Doyle (Werner.Doyle@nyulangone.org)**

Comprehensive Epilepsy Center, NYU Langone Medical Center, 530 1st Avenue  
New York, NY 10016 United States

**Guy M. McKhann (gm317@cumc.columbia.edu)**

Department of Neurology, Columbia University Medical Center, 710 West 168th Street  
New York, NY 10032 United States

**Catherine A. Schevon (cas2044@cumc.columbia.edu)**

Department of Neurology, Columbia University Medical Center, 710 West 168th Street  
New York, NY 10032 United States

**Adeen Flinker (Adeen.Flinker@nyulangone.org)**

Department of Neurology, NYU Langone Medical Center, 145 East 32nd Street  
New York, NY 10016 United States

**Nima Mesgarani (nima@ee.columbia.edu)**

Zuckerman Institute, Columbia University, 3227 Broadway  
New York, NY 10027 United States

## Abstract:

Natural stimuli like speech and music are structured at many timescales. But it remains unclear how these diverse timescales are neurally coded. Do neural processing timescales increase along the cortical hierarchy? Are there distinct timescales for particular stimulus categories? What information is coded at each timescale? Answering these questions has been challenging because there is no general method for estimating sensory integration periods: the temporal window within which stimulus features alter the neural response. Here, we introduce a simple experimental paradigm for inferring the integration period of any time-varying response. We present segments of natural stimuli in a sequence, such that same segment occurs in two different contexts (different surrounding segments). We then measure how long the segments need to be for the response to become invariant to the context. We apply this paradigm to map temporal integration periods in human auditory cortex using electrocorticography

data from epilepsy patients. Our map reveals a clear gradient in which integration periods grow as one moves away from primary auditory cortex, providing support for hierarchical models. We also show that selectivity for sound categories first emerges at timescales of ~200 ms, approximately the duration of speech syllables and musical notes.

**Keywords:** temporal integration; auditory cortex; electrocorticography; ECoG; natural stimuli

## Extended abstract

Natural stimuli are structured at timescales from milliseconds (e.g. phonemes) to seconds (e.g. words) and minutes (e.g. narrative structure). Understanding and modeling how these diverse timescales are



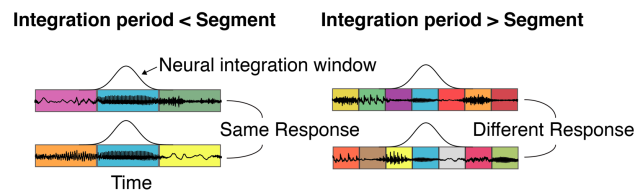
neurally coded is a central goal of sensory and computational neuroscience (Honey et al., 2012).

Sensory timescales are often defined in terms of their temporal integration period (Theunissen and Miller, 1995): the time window when stimuli alter the response. Integration periods are central to many theories and models of sensory coding. Hierarchical models often posit that integration periods grow as one ascends the sensory hierarchy (Honey et al., 2012; Overath et al., 2015). Other theories posit that different hemispheres (Zatorre et al., 2002) or cortical regions (Overath et al., 2015) have distinct integration periods, or that certain stimulus classes might require dedicated processing timescales.

There are two well-known approaches for estimating sensory integration periods. One approach is to derive an explicit model relating the stimulus to the response. In the auditory system, it is common to estimate a “spectrotemporal receptive field” (STRF): a linear mapping between a spectrogram and the neural response. This approach is effective if the response is linear with respect to a spectrogram. But cortical responses are known to be highly nonlinear (Sahani and Linden, 2003), particularly in non-primary regions (Norman-Haignere and McDermott, 2018), and STRFs could thus yield misleading results. Estimating nonlinear models of neural responses remains a challenging task particularly in higher-order sensory regions.

A second approach is to temporally scramble natural sounds, and measure the power or reliability of the neural signal as a function of the scrambling window (Honey et al., 2012). A common finding is that putatively higher-order brain regions respond more strongly to stimuli that have intact temporal structure at longer timescales. However, because scrambling paradigms simply measure the power in response to more or less scrambled stimuli they cannot detect selectivity for stimulus features that are similar on average between intact and scrambled stimuli. This fact helps explain why primary auditory regions often show no effect of scrambling (Overath et al., 2015), since they plausibly respond to features such as frequency and modulation that are not greatly altered by scrambling.

To address these limitations, we introduce a novel paradigm, which we term “temporal context invariance” or “TCI”, for estimating the integration period of any time-varying response. The TCI paradigm is effective because it directly tests the core idea of an integration period: that the response should be invariant to any stimuli falling outside the integration period. We measure responses to stimulus segments of different



**Figure 1.** Schematic of Temporal Context Invariance (TCI) Paradigm. Segments of natural stimuli (here sounds) are presented in a random order such that the same segment occurs in two different contexts (different surrounding segments). If the neural integration period is less than the segment duration, there should be a moment when the response is the same across the two contexts.

size (similar to scrambling paradigms) and test whether the response to a given segment is affected by what came before and after (our innovation) (**Figure 1**). If the integration period is less than the segment size, then there should be a moment at which the response to the current segment is unaffected by the surrounding segments. We can thus estimate the integration period by varying the segment size, and measuring at what point the response becomes context invariant.

Our study makes three contributions: (1) We show that our TCI method is effective at measuring integration periods throughout auditory cortex, unlike standard paradigms. (2) We show that integration periods increase substantially from primary to non-primary regions. (3) We show short-timescale electrodes are best predicted by spectrotemporal features of sound while long-timescale electrodes are best predicted by a sound’s category (e.g. whether it is speech or music). Using decoding analyses, we show that selective responses to speech and music first emerge at timescales of ~200 ms, suggesting selectivity for syllables or musical notes.

## References

- Honey CJ, Thesen T, Donner TH, Silbert LJ, Carlson CE, Devinsky O, Doyle WK, Rubin N, Heeger DJ, Hasson U (2012) Slow Cortical Dynamics and the Accumulation of Information over Long Timescales. *Neuron* 76:423–434.
- Norman-Haignere SV, McDermott JH (2018) Neural responses to natural and model-matched stimuli reveal distinct computations in primary and nonprimary auditory cortex. *PLoS Biol* 16:e2005127.
- Overath T, McDermott JH, Zarate JM, Poeppel D (2015) The cortical analysis of speech-specific temporal structure revealed by responses to sound quilts. *Nat Neurosci* 18:903–911.

Sahani M, Linden JF (2003) How linear are auditory cortical responses? In: *Advances in neural information processing systems*, pp 125–132.

Theunissen F, Miller JP (1995) Temporal encoding in nervous systems: A rigorous definition. *J Comput Neurosci* 2:149–162.

Zatorre RJ, Belin P, Penhune VB (2002) Structure and function of auditory cortex: music and speech. *Trends Cogn Sci* 6:37–46.