



# Neural Representation of Complex Sounds in the Human Auditory Brainstem

Ananthanarayan A. Krishnan

**JAIISH**(2020)  
Vol 39 pp. 1-22

## Affiliations

Department of Speech Language  
Hearing Sciences, Auditory  
Electrophysiology Laboratory, Purdue  
University, 715 Clinic Drive,  
Lyles-Porter Hall, Room 3060, West  
Lafayette, 47907-2122, Indiana, USA

## Corresponding Author

Ananthanarayan A. Krishnan  
Department of Speech Language  
Hearing Sciences, Purdue University  
West Lafayette  
47907-2122  
Indiana  
USA.

## Key Words

*Phase-locking  
temporal scheme  
envelope  
temporal fine structure  
pitch shifts*

## Abstract

*Phase-locked neural activity in the brainstem generating the human frequency following response (FFR) has been shown to preserve neural information about certain spectrotemporal attributes of complex sounds that likely contributes to the perception of speech, language, music, and segregation of concurrent sound streams. Here we review some evidence in support of the view that the temporal pattern of neural activity in the FFR does preserve information relevant to the representation of spectra and pitch of complex sounds. Specifically, FFR spectra for both steady state and time variant complex sounds show clear peaks at formant related harmonics that follow the changes in formant frequencies for time-variant sounds. Similarly, neural information relevant to steady- and time-variant pitch, pitch shifts, and changes in pitch salience are well preserved in the FFR and exhibits a strong correlation with behavioral measures. For inharmonic, frequency shifted, and unresolved harmonics stimuli that produce pitch shifts or multiple pitches, neural activity relevant to these perceived changes in pitch is primarily contained in the neural representation of the temporal fine structure (TFS). Both reverberation and noise degrade the neural representation of envelope and TFS with phase locking to ENV showing greater resilience. Overall, these results clearly suggest that the FFR provides a robust physiologic window to evaluate the nature of neural representation of spectra and pitch of complex sounds in normal and impaired ears, age related changes in neural encoding, and to understand the role of experience in shaping subcortical processing and its application to re-training and perceptual learning.*

©JAIISH, All Rights Reserved

## INTRODUCTION

Periodic complex sounds including speech represent the sum of multiple, harmonically related frequency components and are characterized by a slowly varying periodic envelope (ENV) and rapidly varying frequency components contained in the envelope referred to as the temporal fine structure (TFS) (Figure 1, top trace). Using these complex sounds, two distinct types of sustained brainstem neural activity can be elicited- one reflecting phase-locked neural activity that follows the envelope periodicity (Envelope Following Response- EFR, 4<sup>th</sup> trace from the top) and the other, phase-locked neural activity that follows the TFS (Frequency Following Response - FFR, 5<sup>th</sup> trace from the top). The EFR with a waveform largely representing the low frequency envelope periodicity and the FFR with a waveform representing phase-locking to the relatively higher frequencies contained in the TFS of complex sounds are shown in spectral data in Figure 1. It should be noted here that the phase-locked activity elicited by a single onset polarity stimulus is a combination of neural phase-locking to the envelope periodicity and the temporal fine structure, although the response waveform is dominated by the more robust phase-locking

to the envelope periodicity. Both the EFR (to modulation frequencies higher than 90 Hz) and the FFR above 100 Hz is thought to reflect sustained phase-locked neural activity primarily in the inferior colliculus (Krishnan, 2007; Bidelman, 2018).

The phenomenon of neural phase-locking in auditory neurons has been well demonstrated at all levels along the auditory pathway and serves as the fundamental basis for the robust temporal encoding scheme in the auditory system. However, the decrease in the upper frequency limit for neural phase-locking from about 5000 Hz in the auditory nerve, 1500-1700 Hz in the inferior colliculus, and less than about 150 Hz in the auditory cortex together with the decrease in neural phase-locking ability in the entire population of auditory nerve fibers with increasing frequency (the membrane time constant of the IHC being a limiting factor (Palmer & Russell, 1986)) poses a serious challenge for the effective operation of temporal encoding schemes at higher levels along the auditory pathway, and at higher frequencies. Since both EFR and FFR overlap to produce a complex response waveform in response to a fixed onset polarity stimulus, averaging the responses to alternating polarity will effectively isolate the EFR

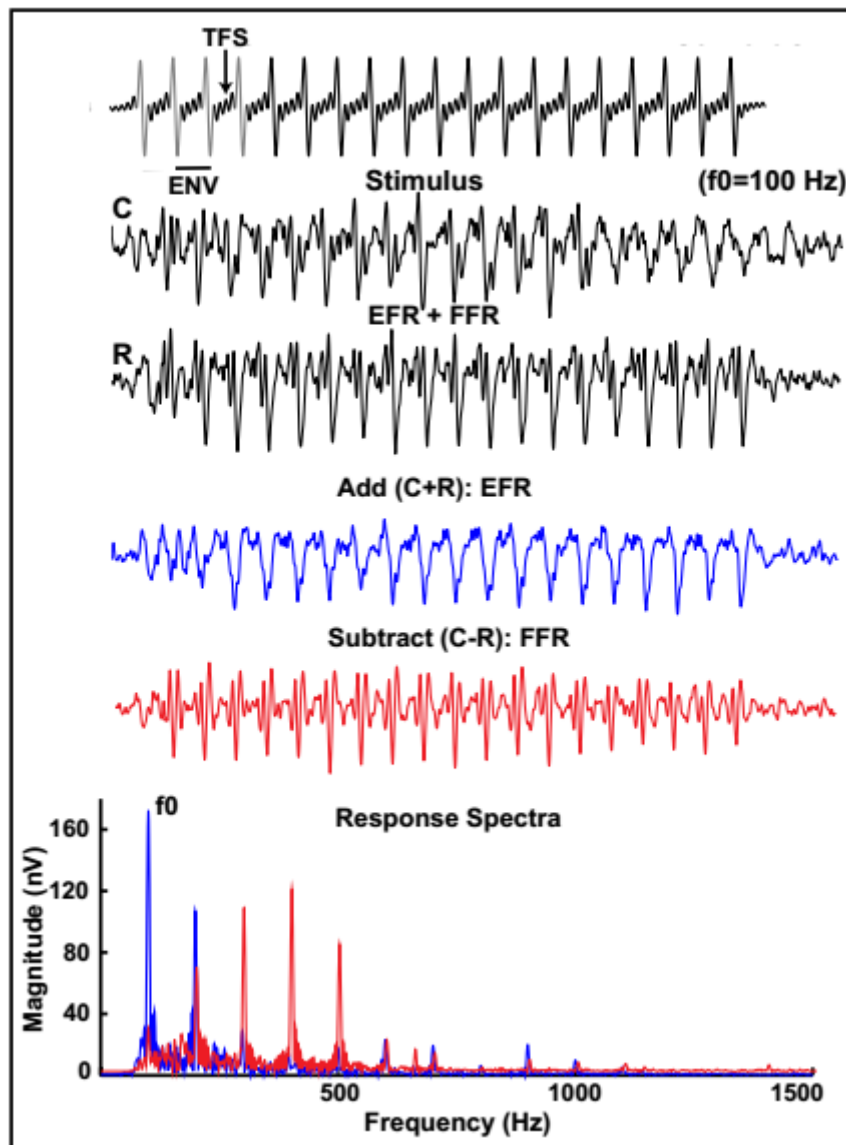


Figure 1: Extraction of the EFR and the FFR using addition and subtractions of responses to condensation and rarefaction onset polarities. The top waveform is the stimulus waveform of a complex sound ( $f_0 = 100$  Hz) illustrating the envelope periodicity and the TFS. Second and the third response represents FFR/EFR to condensation (C) and rarefaction (R) onset polarities, respectively. Addition of these two responses (4<sup>th</sup> waveform from the top) results in the EFR, and subtraction (C-R) results in the FFR (5<sup>th</sup> waveform from the top). Note the lower frequency periodicity of the EFR compared to the FFR. The spectra of the EFR (blue) and FFR (red) to a complex tone is shown at the bottom of the waveforms. The FFR components extends to several higher frequency harmonics while the EFR shows a prominent peak at  $f_0$ .

from the FFR as well as the cochlear microphonic and stimulus artifact since the EFR does not change with polarity change (Chandan & Krishnan, 2018; Ananthkrishnan & Krishnan, 2018; Aiken & Picton, 2008; Krishnan, 2002; Greenberg et al., 1987). Thus, EFR alone is summed and preserved while the other components are subtracted out (Figure 1). To extract the FFR (Figure 1), the responses to opposite polarities are subtracted (Chandan & Krishnan, 2018; Ananthkrishnan & Krishnan, 2018; Aiken & Picton, 2008; Krishnan, 2002). This is similar to the compound histogram technique developed by Arthur et al. (1971) and Gobllick and Pfeifer (1969) for the

temporal analysis of single-unit discharge patterns.

While different labels have been used to identify these responses (for example, EFR has been referred to as Auditory steady-state responses when amplitude modulated stimuli have been used, speech ABR when speech stimuli have been used, or as  $FFR_{ENV}$ ,  $FFR_{SPEC}$  or  $FFR_{TFS}$  when complex sounds like speech and/or complex tones have been used), we have chosen to use the more appropriate operational descriptor: EFR for the envelope following response since the response reflects neural phase-locking to the envelope periodicity regardless of the type of complex sound used to elicit them; and FFR for

responses representing neural phase-locking to the spectral components of the stimulus (i.e. TFS). Both EFRs and FFRs have been shown to preserve information about neural timing and/or neural representation of certain acoustic features relevant to speech- and pitch perception (Suresh & Krishnan 2019; Ananthakrishnan & Krishnan, 2018; Krishnan & Gandour, 2017; Krishnan & Plack, 2011; Aiken & Picton, 2008; Krishnan, 2007; Krishnan, 2002) and as such provides an effective physiologic window to assess certain aspects of temporal processing in normal and impaired auditory systems (Easwar et al., 2015b; Anderson et al., 2013; Ananthakrishnan, Krishnan, & Bartlett, 2016; Plyler & Krishnan, 2001). This review is deliberately limited to studies examining neural encoding of spectra and pitch of complex sounds. The specific aim here is to demonstrate that the phase-locked neural activity in a population of neural elements in the rostral brainstem generating the FFR does indeed preserve robust information about certain acoustic and temporal features that may contribute to the development of speech- and pitch perception. The objective clearly is to highlight the potential utility of the FFR as a neural metric to understand neural representation of behaviorally relevant features contained in complex sounds in mostly normal ears.

## Neural representation of spectra of complex sounds

### Steady state speech-like and speech sounds

It is well established that encoding of the first two formants of most vowels is sufficient for their identification (Brown, 1958; Carlson, Fant, & Granstrom, 1975; Peterson & Barney, 1952). Neural phase locking based temporal-place encoding scheme has been shown to play a dominant role in the neural representation of the spectrum of steady-state and time-variant sounds in the population response of single neurons in the auditory nerve and ventral cochlear nucleus (Blackburn & Sachs, 1990; Caspary, Rupert & Moushegian, 1977; Keilson, Richards, Wyman, & Young, 1997; Miller & Sachs, 1983; 1984; Palmer, Winter, & Darwin, 1986; Recio & Rhode, 2000; Rhode, 1998; Sachs & Voight, 1983; Young & Sachs, 1979). The scalp-recorded FFR has also been shown to preserve spectral and pitch relevant information contained in speech and non-speech complex stimuli (Krishnan, 1999; Krishnan, 2002; Ananthakrishnan & Krishnan, 2018; Krishnan, Gandour & Suresh, 2016; Krishnan & Gandour, 2014; Bidelman, Gandour & Krishnan, 2011; Bidelman & Krishnan, 2011; Krishnan & Agarwal, 2010; Bidelman & Krishnan, 2009; Krishnan & Plack, 2011; Smalt et al., 2012; Krishnan, Gandour & Bidelman, 2010; Ananthakrishnan, Luo & Krishnan, 2017; Ananthakrishnan, Krishnan, & Bartlett, 2016; Suresh, Krishnan, & Luo, 2018; Swaminathan, Krishnan, & Gandour, 2008; Galbraith, Jhaveri, & Kuo, 1997; Greenburg,

1980; Krishnan, Xu, Gandour, & Cariani, 2004; 2005; Aiken & Picton 2008a).

There has been a steady increase in the number of studies that have used speech elicited FFRs to examine neural encoding in normal and impaired ears to speech and speech-like stimuli (Ananthakrishnan & Krishnan, 2018; Krishnan, Gandour & Suresh, 2016; Krishnan & Gandour, 2014; Bidelman, Gandour & Krishnan, 2011; Bidelman & Krishnan, 2011; Krishnan & Agarwal, 2010; Bidelman & Krishnan, 2009; Krishnan & Plack, 2011; Smalt et al., 2012; Krishnan, Gandour & Bidelman, 2010; Ananthakrishnan, Luo & Krishnan, 2017; Ananthakrishnan, Krishnan, & Bartlett, 2016; Suresh, Krishnan, & Luo, 2018; Swaminathan, Krishnan, & Gandour, 2008; Galbraith, Jhaveri, & Kuo, 1997; Greenburg, 1980; Krishnan, Xu, Gandour, & Cariani, 2004; 2005; Anderson, 2013). However, to date only a few studies have specifically addressed neural encoding of spectra of speech sounds (Krishnan, 1999; 2002; Aiken & Picton 2008). Krishnan (1999) recorded FFRs elicited by three different two-tone approximations of English back vowels (/u/, /), and /a/) at 88, 75, 65, and 55 dB nHL to evaluate the neural representation of spectra of these sounds. Spectral analyses of the FFRs revealed distinct peaks at frequencies corresponding to the first and the second formants across all levels suggesting that phase-locked activity among two distinct populations of neurons are indeed preserved in the FFR. A follow up study (Krishnan, 2002) using steady state synthetic speech versions of these three vowels showed clear peaks at harmonics proximal to the first and second formant of the vowel /u/ (Figure 2, left panel) and vowel /a/ (Figure 2, middle panel) at all levels with F1, F2 amplitude increasing with increasing intensity. However, the amplitude growth functions for the higher smaller F2 components were shallower compared to F1 (Figure 2, right panel). These results were interpreted to suggest that the robust responses at the two formant related harmonics represents phase-locked activity in two distinct populations of brainstem auditory neurons driven by their characteristic places along the cochlear partition. Support for this view is derived from the following observations: FFR responses to moderate intensity tone-bursts are place specific (Anathanarayan & Durrant, 1992); distinct response peaks at F1, F2, and at 2f-f2 for the two-tone vowel stimuli (Krishnan, 1999); and FFRs ability to accurately track the frequency change presented in either an upward, or downward swept tonal glide (Krishnan & Parkinson, 2000). In this framework, each harmonic in the complex stimuli would engage a specific place on the cochlea (therefore a specific frequency) which would produce phase-locked activity in two distinct population of brainstem neurons with characteristic frequencies proximal to F1, and F2. That is, the spectral peaks in the FFR harmonics close to F1 and F2 likely reflect phase-locked activity from distinct populations of neurons. Krishnan (2002) also observed that for each vowel, the F1 response was

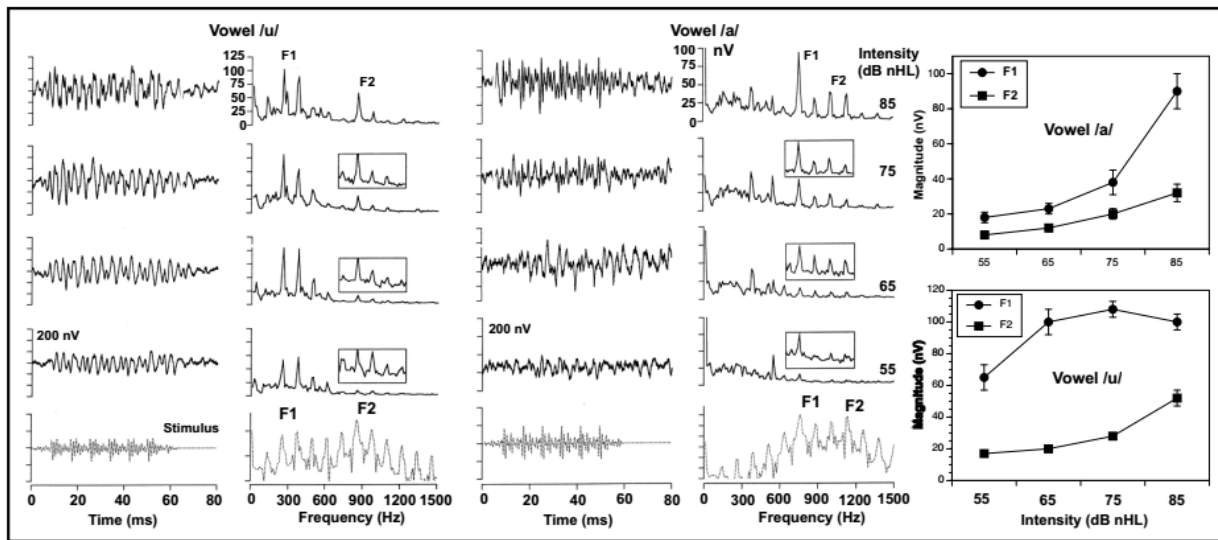


Figure 2: FFR response waveforms, spectra (plotted as a function of intensity) and amplitude growth functions for steady state vowels /u/ and /a/. The inset in the spectral data shows the magnified smaller F2 related spectral peaks. For both vowels spectral components can be detected down to 55 dB nHL. For both vowels and for both formants amplitude grows with intensity with different characteristics. The stimulus waveforms and spectra (dotted traces) are shown at the bottom of the response waveforms and spectral for each vowel. Data replotted from Krishnan (2002).

dominant at high stimulus levels consistent with the phenomenon of formant capture. Young & Sachs (1979) demonstrated this in the responses of the auditory nerve fibers to speech stimuli. Specifically, as intensity increases, not only does the response at the F1 place increase but also progressively reflects increasing contributions from units with higher characteristic frequencies due to upward spread of excitation (synchrony spread). This process likely accounts for the observation of the dominant spectral peak at F1 in the FFR data. Finally, the observation of smaller FFR response amplitudes for harmonics between the formants seen in the FFR data suggests selective synchrony suppression to enhance spectral peaks at the formant frequencies (Krishnan, 2002). Taken together, these results clearly suggest that the phase-locked activity underlying the FFR is able to preserve the spectral representation of formant and non-formant harmonics of speech sounds when their first two formant frequencies are below the upper frequency limit of human FFR -about 1700 Hz. Indeed, from a translational application perspective, comparison of FFRs elicited by a steady state vowel in normal and hearing impaired ears show a clear degradation of the representation of TFS with little change in the representation of the envelope periodicity (Ananthakrishnan, Krishnan, & Bartlett, 2016) suggesting that the reduced speech discrimination ability in individuals with sensorineural hearing loss may, at least in part, be a consequence of degraded representation in the brainstem of certain acoustic features important for speech perception.

### Time-variant speech-like and speech sounds

Most natural speech sounds are inherently time-variant. That is, their spectrotemporal acoustic features change over time. There is compelling evidence that these time-varying acoustic features (for example, formant transitions) of speech sounds play an important role in speech perception (Jacobson, Fant, & Halle, 1963). Thus, in order to be useful, phase-locked activity generating the FFR should be able to preserve some of these time-variant features. As an initial step to examine this, Krishnan & Parkinson (2000) evaluated the encoding of simple linearly rising (400–600 Hz) and falling (600–400 Hz) tonal glides, grossly approximating formant trajectories in real speech. The results of this study demonstrated that the human FFR does indeed follow the trajectory of the rising (Figure 3, left panel) and falling tonal glides (Figure 3, right panel) in a robust fashion. The authors proposed that the changing frequency in the stimulus was encoded by a progressive shift in the population of neurons phase-locked to the changing stimulus frequency. Also, the decreasing amplitude with increasing frequency observed for the rising glide supports the view that neural phase locking deteriorates with increasing frequency. Similar FFR findings have been reported by Billings et al., (2019); Clinard & Cotter (2016) in adults, and Madhavi, Krishnan, & Weber-Fox (2009) in children. Interestingly, results in the Madhavi, Krishnan, & Weber-Fox study, comparing neural representation of tonal glides in children with normal language and children with specific language impairment (SLI), revealed that a subset of children with SLI showed poor neural tracking of the frequency change as the

rate of change of frequency was increased suggesting a disruption in temporal processing-again, pointing to the sensitivity of the FFR to evaluate aspects of auditory temporal processing. The observation of smaller overall amplitude for the falling glide compared to the rising glide has been interpreted to suggest that the neural activity for rising frequency is more synchronous than for a falling frequency (Dau, 2003; Janssen et al., 1991). However, Billings et al., (2019) failed to observe this differential sensitivity to direction in their study. Plyler & Krishnan (2001) also demonstrated that the phase-locked activity reflected in the FFR elicited by CV syllables was able to track the formant transition in normal hearing adults but not in individuals with sensorineural hearing loss. Shorter duration formant transitions (50 ms or less) in shorter CV syllables does not appear to show a clear following in the FFR, probably obscured by the overlapping onset components.

More recently, Suresh & Krishnan (2020) demonstrated that the FFR is able to follow the trajectory of decreasing frequencies of the first and second formant in a diphthong as the vowel moves from /a/ to /au/ to /u/ (Figure 4). Both the spectrogram (Figure 4, middle panels) and the instant spectral slices representing the earlier /a/ segment, the transition /au/ segment and the later /u/ segment (Figure 4, bottom panel) clearly show that the phase-locked activity is able to follow the formant frequency changes over time. Collectively, these results suggest that the phase-locked activity is able to represent the time varying spectrotemporal attributes of the stimulus and therefore may be used to assess aspects of auditory temporal processing in different clinical populations.

### Speech in adverse listening conditions: Reverberation and Noise

Speech communication almost always occurs in the presence of competing background sounds and/or in a reverberant environment. These adverse listening conditions challenge the auditory system's ability to extract certain acoustic features of sounds that are important for speech perception. While background noise effects are usually explained in terms of masking of the target, in reverberant conditions, the original sound waves are reflected from walls, floors, and ceilings therefore producing a spectrotemporal distortion that degrades perception. Reflected sound waves arrive at the listener's ear delayed relative to the original. The temporal overlap between incident and delayed attenuated signals results in a noisier signal compromised by multiple delays, attenuation, and spectrotemporal distortions. With its ability to preserve  $f_0$  and formant information, the FFR serves as an effective analytic tool to study the neural representation of certain acoustic features in quiet and how they may be degraded in adverse listening conditions like noise and reverberation.

### Reverberation

Psychophysical studies have shown that reverberation can have deleterious effects on an individual's ability to identify and discriminate acoustic features of consonants (Gelfand Silman, 1979); vowels (Drgas & Blaszak, 2009; Nabelek & Letowski, 1988); and time-varying formant cues (Nabelek & Dagenias, 1986b). The consequent reduction in identification and discrimination reflects both forward masking, where preceding segments mask subsequent segments, and self-masking, where temporal smearing occurs within each phoneme (Nabelek et al., 1989; Wang & Brown, 2006). Physiological data from single neurons in the ventral cochlear nucleus (Sayles & Winter, 2008; Sayles, Stasiak, & Winter., 2015; 2016), and the scalp-recorded brainstem FFR data (Krishnan, Suresh, & Gandour, 2019; Bidelman, 2017; Bidelman & Krishnan, 2011) show that neural representation of pitch-relevant fine structure information based on neural phase-locking is degraded in the presence of reverberation. Bidelman & Krishnan (2011) using the vowel /i/ with time-varying  $f_0$ , observed that speech evoked FFRs show an overall reduction in response magnitude due to reverberation induced desynchronization (Figure 5 left and middle panels showing disruption of phase-locking in both the autocorrelogram (left) and the spectrogram (middle), particularly pronounced at the higher formant related harmonics, while maintaining the representation of  $f_0$  with only a gradual decrease in magnitude until at least the most severe reverberation condition (Figure 5, top right). However, they did show a progressively decreasing correlation between the response in the dry condition (no reverberation) and the responses in reverberant conditions (Figure 5, bottom right). In contrast, Bidelman, Davis, & Pridgen (2018) showed little or no change in the  $f_0$  component of the FFR elicited by a vCv speech token in presence of mild and medium levels of reverberation. Sayles, Stasiak & Winter (2016) showed a significant degradation in the periodicity tagged discharge for only stimuli with time-varying pitch with little or no change for stimuli with steady state pitch. It is possible that that the absence of an appreciable reverberation induced degradation in Bidelman, Davis & Pridgen (2018) may be due to the very gradual change in pitch over time in their stimulus. More recently, Krishnan, Suresh & Gandour (2019), using a lexical tone with time-varying pitch (Mandarin tone 2 with a curvilinear rising pitch contour) preceded by a long noise segment, showed that the neural representation of pitch relevant information (as reflected in the ability to follow the  $f_0$  changes) progressively deteriorated with increasing reverberation (slight to moderate levels) with greater disruption of the later rapidly accelerating pitch segment. This latter result is consistent with animal studies that show neural encoding of pitch based on timing information is severely degraded in the presence of reverberation in the caudal brainstem (Sayles, Stasiak, & Winter,

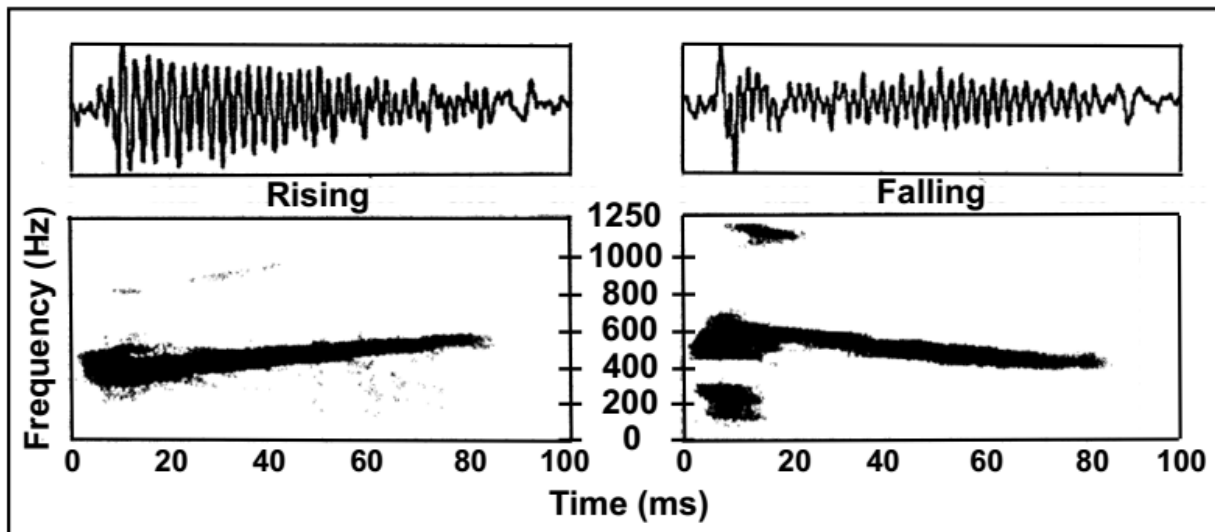


Figure 3: FFR waveforms and spectrograms to an upward swept (400-600 Hz) on the left, and downward swept (600-400 Hz) tone on the right. The rising and falling bands in spectrogram clearly show the ability of the phase-locked activity to follow frequency change. From Krishnan & Parkinson, 2000.

2016; Sayles, Stasiak, & Winter 2015; Sayles & Winter, 2008-ventral cochlear nucleus). It is likely that reverberation degrades the sustained fine-grained, phase-locked temporal discharge pattern of brainstem neural responses following the pitch contour. This degradation is likely due to the loss of robust periodicity and smearing of dynamic changes over time in the fine structure of the stimulus (Houtgast and Steeneken, 1973, 1985). Another possible explanation for the relative resilience of  $f_0$  is because the FFR is dominated by phase-locking to low frequency resolved components rather than phase-locking to the weaker envelope modulation cue resulting from interaction of unresolved higher harmonics that has been shown to degrade markedly with increasing reverberation (Sayles & Winter, 2008). Sayles & Winter (2008) suggest that both the more robust neural phase-locking in the low frequency channels in general and the more salient responses to resolved components increases their resistance to temporal smearing resulting from reverberation.

### Background noise

Reduced amplitude has been observed in the presence of noise for brainstem EFR/FFR (Bidelman, Davis, & Pridgen, 2018; Prevost et al., 2013; Smalt et al., 2012; Song et al., 2011; Li & Jeng, 2011; Parbery-Clark et al., 2009; Russo et al., 2004). Smalt et al. (2012), using complex tone with unresolved harmonics (12<sup>th</sup>-17<sup>th</sup>) of a 90 Hz  $f_0$  and low pass masking to eliminate distortion products in the lower frequencies, observed robust peaks at  $f_0$  (envelope) and lower harmonics (EFR due to modulation at 90 Hz resulting from the periodicity of the complex stimulus and the distortion produced by the interaction of the unresolved harmonics). While the cubic distortion

component showed a systematic decrease in periodicity strength with increasing noise level,  $f_0$  remained largely unaltered even at high noise levels (Figure 6). However, as noise level was increased, phase-locked components at higher harmonics (2 $f_0$ -to-8 $f_0$ -overlapping the F1 and F2 harmonics of English back vowels) reduced in magnitude which in turn may contribute to the difficulties in speech perception in noise. There is also some evidence showing enhancement of  $f_0$  in the presence of noise (Bidelman, Davis, & Pridgen, 2018; Prevost et al. 2013; Smalt et al. 2012). One possible explanation for this enhancement is the presence of stochastic resonance in the auditory system (Cunningham et al. 2002; Henry, 1999). That is, addition of noise presumably improves (the underlying neural mechanism is not known) the neural entrainment of a weak periodic signal. Given the high stimulus presentation levels used to elicit these responses it is more likely that low frequency tails of higher frequency neurons are recruited at higher intensities so that a broad array of fibers are phase locking to the  $f_0$  and consequently reducing the effects of noise. The resilience of at least the  $f_0$  component of the FFR to degradative effects of noise and reverberation suggests that neural synchronization to  $f_0$  is robust (at least for steady state pitch where enough cycles of pitch information may be preserved to extract pitch) and not readily susceptible to degradation by noise or reverberation. This resilience to degradation by noise or reverberation is fortuitous since  $f_0$  plays a critical role in the perception of speech, music and the ability to segregate concurrent sounds

### Reverberation vs Noise

Noise and reverberation are fundamentally different in terms of how they distort the spectrotempo-

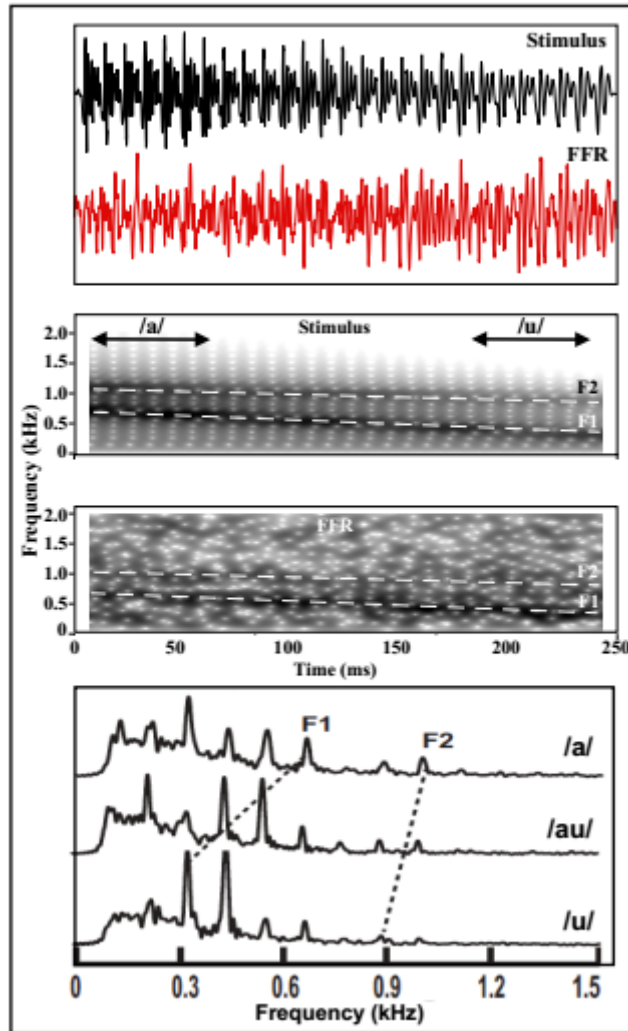


Figure 4: Stimulus (top) and FFR (bottom) waveforms for the diphthong /au/ (top panel); spectrograms of the stimulus and FFR (middle panel); and instant spectral slices of the FFR illustrated in a waterfall fashion (bottom panel) to show the changes in formant related peaks in the FFR that follow the transition from vowel /a/ (top), to diphthong /au/ (middle), to vowel /u/ (bottom). The FFR clearly shows the neural phase-locking reflected in the FFR is tracking the time-varying F1 and F2 decrease in frequency (indicated by the dotted white line).

ral attributes of complex stimuli, therefore it follows that their effects could also manifest differently in the phase-locked neural activity. Bidelman, Davis, & Pridgen (2018) demonstrated that speech-evoked brainstem FFR ( $f_0$ ) and early cortical responses were degraded more by noise than reverberation. Previous physiological studies (Al Osman et al., 2017; Bidelman and Krishnan, 2011; Sayles and Winter 2008; Sayles, Stasiak, & Winter, 2016) have also shown better preservation of neural representation in reverberation compared to noise (Bidelman, 2017, review). A qualitative comparison of changes in cortical pitch response (CPR) due to reverberation (Krishnan, Suresh, & Gandour, 2019), and background noise (Suresh, Krishnan, & Gandour, 2017) reveal that both types of adverse listening conditions result in degradation of CPR neural activity. In reverberation, language dependent enhancement is maintained across conditions; in background noise, it is maintained only for the most favorable SNR (+5

dB) condition. Behavioral measures have also shown different patterns of vowel confusions in noise and reverberation (Nabelek and Dagenais, 1986). Taken together, these results are consistent with Bidelman, Davis, & Pridgen (2018) suggesting differential effects of noise and reverberation. Sayles, Stasiak, & Winter (2016) also observed differential effects for noise and reverberation. Specifically, reverberation significantly impaired segregation of concurrent vowels which had a pitch contour, but not vowels with steady state pitch. In contrast, noise impaired segregation of vowels with both steady state, and time-variant pitch contour. It remains unclear if these differences simply reflect fundamental differences in the extent to which spectrotemporal pitch-relevant acoustic features are disrupted and/or the relative effectiveness of masking mechanisms under both noise and reverberation. Surprisingly, the differential effects of noise and reverberation on speech representations is not observed at the cortical level

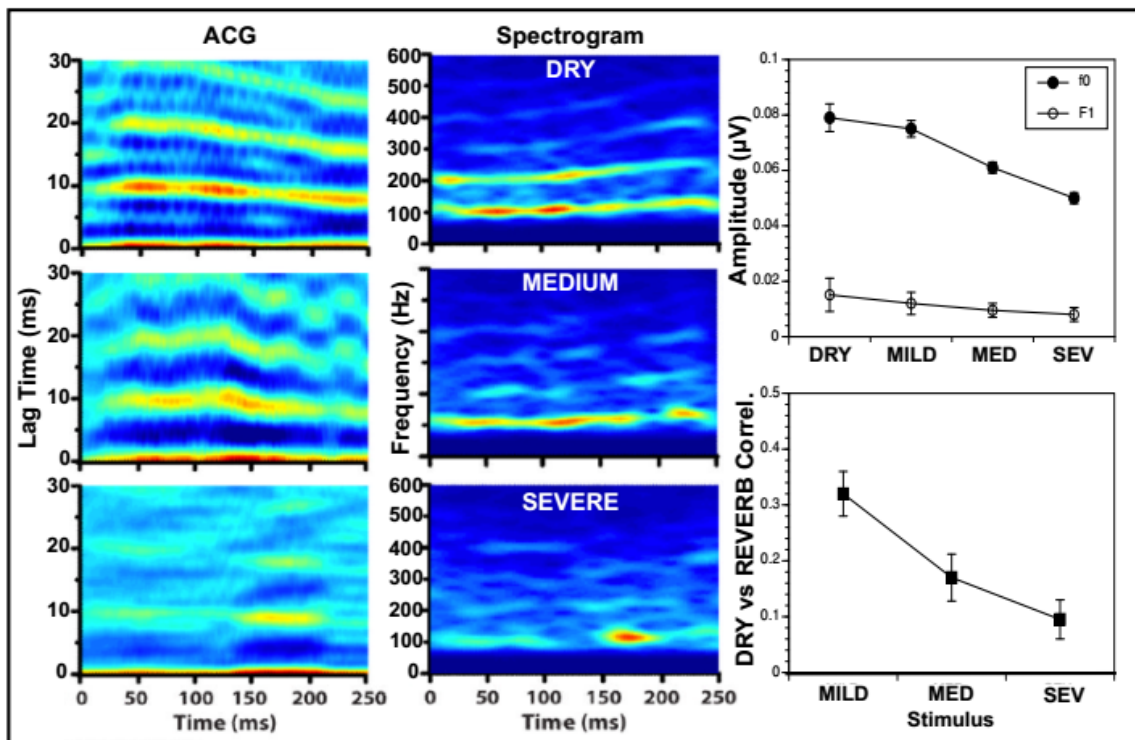


Figure 5: Autocorrelograms (left), spectrograms (middle), mean  $f_0$  degradation in and F1 amplitude (right top), and change in  $f_0$  amplitude (bottom right) plotted as a function of increasing degree of reverberation. Note the reverberation related degradation in  $f_0$  and F1 representation particularly at the medium and severe degrees of reverberation. Replotted using data from From Krishnan & Bidelman (2011).

(Mesgarani et al. 2014) suggesting operation of central compensatory mechanisms mitigating the effects of reverberation on the neural representation of time varying periodicity at the subcortical levels. We cannot rule out interplay between bottom up and top down mechanisms (Suga et al., 2000, 2003) to aid signal selection in adverse listening conditions. Such interplay, however, does not explain why signal selection would be selectively aided in reverberation but not noise. Clearly, more research is needed to provide a more complete, quantitative characterization of the how noise and reverberation differentially affect the neural encoding of complex sounds. Development of optimal signal processing strategies implemented in conventional amplification devices will have to consider appropriate strategies to mitigate the differential deleterious effects of noise, and reverberation on the neural representation of certain acoustic features of speech important for speech perception.

### Neural representation of pitch-relevant information of complex sounds

#### Neural correlates of pitch of harmonic, inharmonic, and frequency shifted sounds

For a range of harmonic complex sounds that produce a low pitch percept, the dominant interval interspike interval present in the population of auditory

nerve fibers always correspond to the pitch heard in perceptual experiments (Cariani & Delgutte, 1996a—the dominant interval hypothesis). Psychoacoustic studies have demonstrated that inharmonic and frequency-shifted complex tones (Shouten, 1940; de Boer, 1956) produce pitches that differ from that suggested by the envelope periodicity/harmonic spacing (pitch shift and/or pitch ambiguity). Consistent with this, auditory nerve physiological data using similar stimuli indicate that these pitch shifts also have direct correlates in the population interspike interval distribution (ISID) of the auditory nerve (Cariani, 1996b). That is, pooled interval distributions in response to inharmonic stimulus segments show multiple maxima corresponding to the multiple pitches heard by human listeners (pitch ambiguity). Similarly, for frequency shifted complex sounds where octave drops in pitch are expected, there is corresponding patterns of peaks in the pooled interspike interval distribution.

Given that FFR reflects sustained phase-locked activity in a population of neural elements in the brainstem, Krishnan & Plack (2011b) reasoned that the temporal pattern of phase-locked activity generating the FFR may preserve information consistent with pitch ambiguity of inharmonic complex sounds, and pitch shifts of frequency shifted complex sounds similar to what has been demonstrated in the pooled interval distribution at the auditory nerve



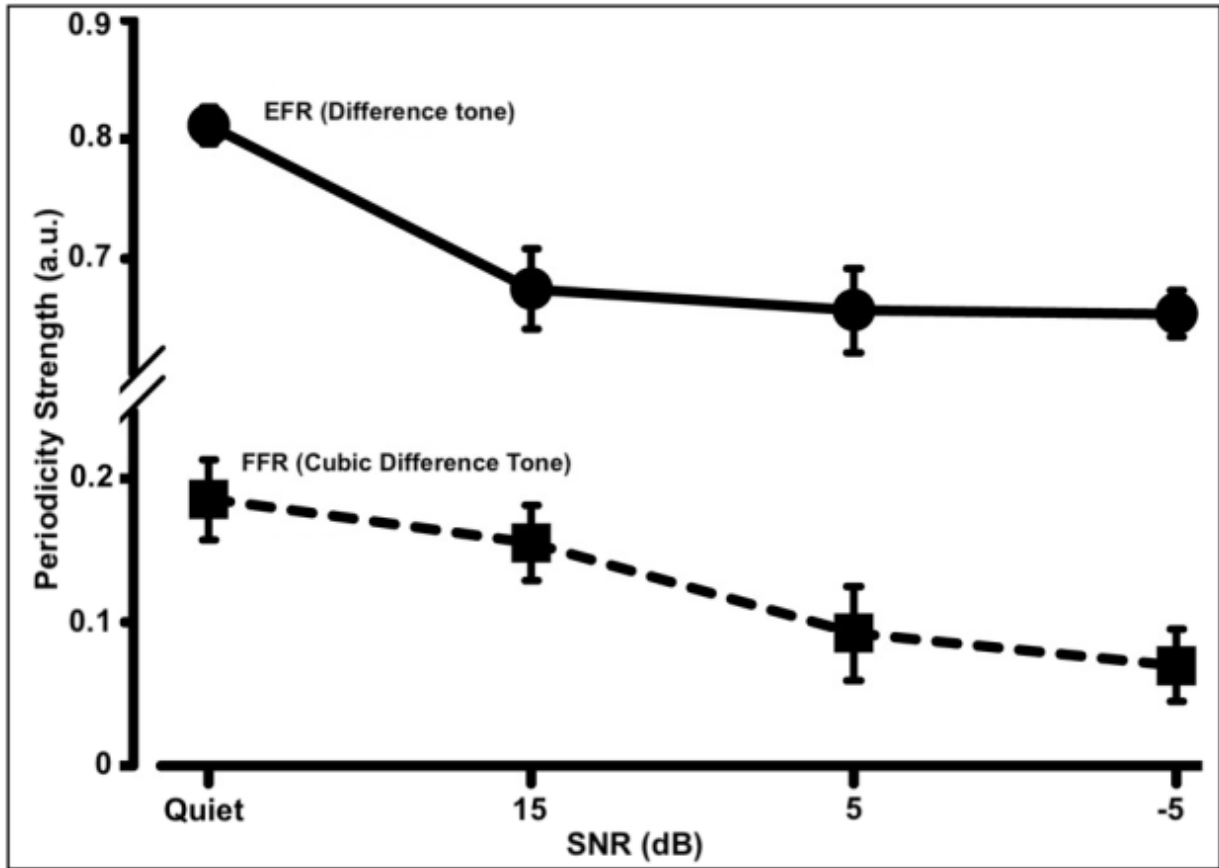


Figure 6: Mean periodicity strength plotted for the EFR (solid circle) and for the FFR (cubic difference tone) for the quiet and increasing levels of noise. Replotted from Smalt et al. (2012).

level (Cariani & Delgutte, 1996a; 1996b). To evaluate pitch ambiguity FFRs were recorded to amplitude modulated (AM) tones with fixed modulation frequency ( $F_m = 125$  Hz) and different carrier frequencies (625, and 750 Hz-harmonic since  $F_c/F_m$  is an integer; 687 and 733 Hz- inharmonic and aperiodic since  $F_c/F_m$  is not an integer). To evaluate pitch shift, FFRs were recorded using a three-component harmonic complex comprised of the 2<sup>nd</sup> (488 Hz), 3<sup>rd</sup> (732 Hz), and 4<sup>th</sup> (976 Hz) harmonic of a 244 f<sub>0</sub>, and another three-component complex where the three components was shifted downwards in frequency by 122 Hz (that is, 366 Hz, 610 Hz, and 854 Hz) (Figure 8, first two panels on the left). For all stimuli (harmonic, inharmonic, and frequency shifted) EFRs showed auto-correlations functions (ACF) peaks that corresponded to the invariant envelope periodicity (Figure 7, left panel, and Figure 8, third panel from the left). Consistent with this, the EFR spectral data showed peaks at the harmonic spacing, and at integer multiples of this value (Figure 7, second panel, and Figure 8, right panel). In contrast, FFR (remember, reflects phase-locking to the fine structure components) showed ACF peaks corresponding to the F<sub>0</sub> for the harmonic stimuli, and single or multiple ACF peaks (which did not correspond to f<sub>0</sub>) for inharmonic/frequency shifted stimuli (Figure 7 left panel and Figure 8, right panel). FFR ACF peaks corre-

sponding to pitch(es) for harmonic and inharmonic AM tones (Figure 7) are identical to the peaks in the pooled interspike interval distribution for auditory nerve (Cariani & Delgutte, 1996b) (Figure 7, third panel). The pitches estimated from the population-interval distribution for these stimuli closely correspond to the pitch shifts (first period effect of pitch shift) that have been observed for human listeners and are in close agreement with de Boer's rule (de Boer, 1976). For inharmonic conditions estimated pitch was 121 Hz for AM 733; and 114.5 Hz and 139 Hz for AM 687. For the frequency shifted conditions, estimated pitches were 203 and 305 Hz -all approximating with de Boer's rule ( $p = F_c/n$ ), where n is an integer near  $F_c/F_m$  values. Although the spectral data for the FFR showed peaks (spectral data in Figures 7, and 8) at the frequency components and at lower harmonics (presumably cubic difference distortion products) as expected, no clear peaks corresponding to the pitch were observed for the inharmonic/shifted stimuli. Based on these results, the authors concluded that FFR (fine structure) results are consistent with the auditory nerve response to AM tones (Cariani et al., 1996b) and with the FFRs to frequency shifted complex sounds (Greenberg et al., 1987). These authors suggest that the multiple maxima in the ACFs for inharmonic/frequency shift stimuli correspond to the multiple pitches heard by

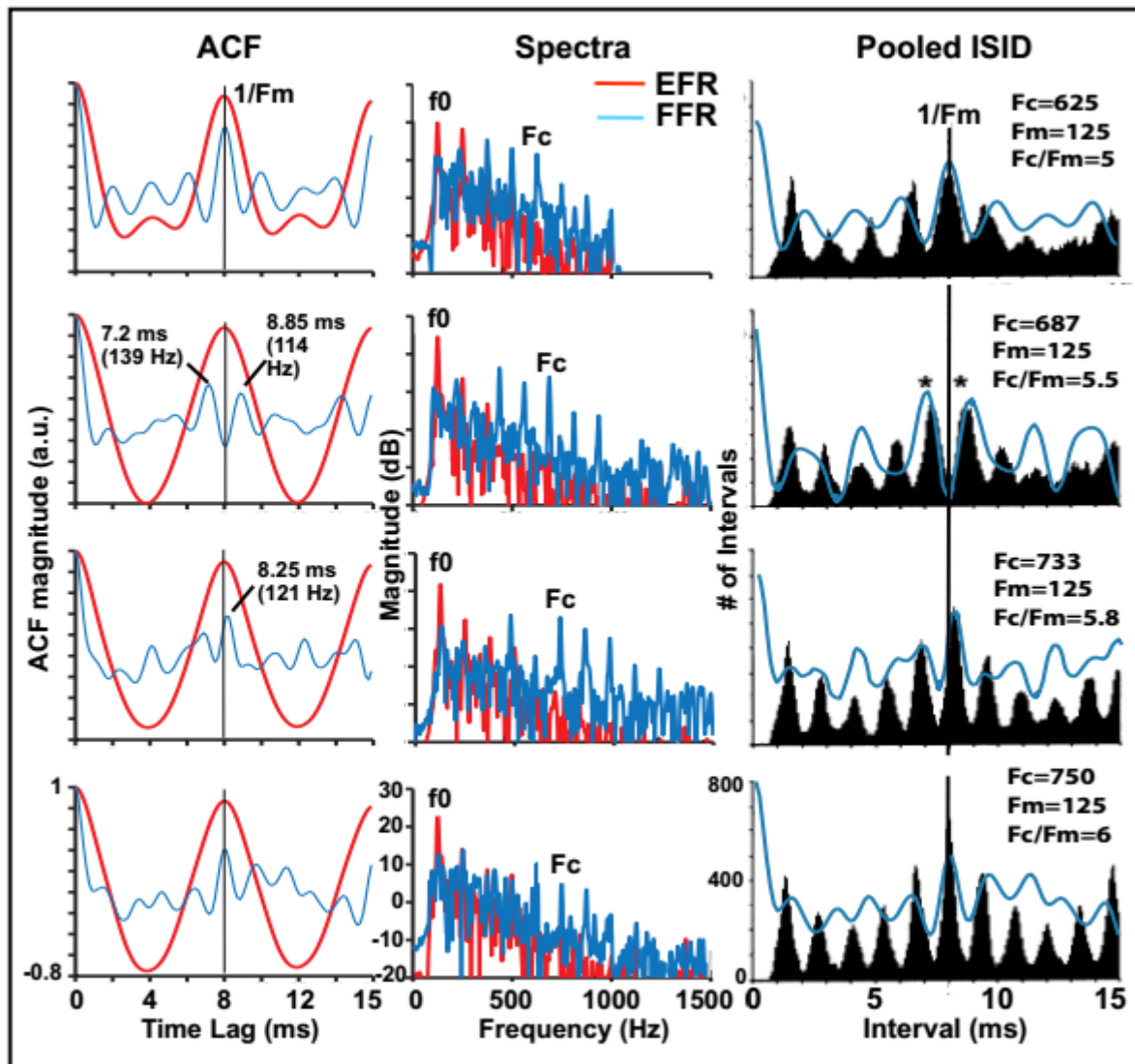


Figure 7: Autocorrelation functions (ACF) (left column) and spectra (middle column) of the EFRs (red/dark) and FFRs (blue/light) plotted for the harmonic (625 Hz and 750 Hz) and inharmonic (687 Hz and 733 Hz) amplitude modulated stimuli with an  $f_0 = 125$  Hz. **EFR** ACFs show a clear peak at  $1/F_m$  for all stimuli and correspondingly the spectral data shows peaks at  $f_0$  and at a few higher harmonics. **FFR** ACFs show a clear peak at  $1/F_m$  only for the harmonic stimuli and shifted or multiple peaks for the inharmonic stimuli (the pitch periods are identified). The FFR spectra show clear peaks at  $F_c$ , its sidebands and at cubic difference tones. EFR ACFs overlaid on the auditory nerve pooled interspike intervals are remarkably similar across stimuli. From Krishnan & Plack (2011), Unpublished data presented at the 2011 midwinter meeting of the Association for Research in Otolaryngology.

human listeners (pitch shift/ambiguity). Consistent with these views, the results presented here suggest that the FFR does not merely reflect neural phase locking to waveform envelope (Hall, 1979). Rather information relevant to pitch shift and pitch ambiguity is preserved in the temporal distribution of neural activity in the midbrain phase locked to the fine structure. Similar results for frequency shifted complex tones have been reported by Gockel, Carlyon, Mehta, & Plack (2011). The similarity of these FFR results with auditory nerve data raises the question whether phase-locked activity in the midbrain, as reflected in the FFR, represents local pitch encoding or merely a passive reflection of pitch relevant information preserved in the neural activity that has

been transmitted from the auditory nerve. Based on their failure to observe any pitch-relevant information in the FFRs to three tone harmonic stimuli presented dichotically, Gockel, Carlyon, Mehta, & Plack (2011) concluded that there was no additional pitch-relevant processing at the level of the brainstem. Several arguments may be presented to counter this inference. First, if the temporal code for pitch available at the brainstem level also utilizes autocorrelation-like analysis to determine the global distribution of interspike intervals from the temporal pattern of neural activity across a population of neurons, it would necessarily share certain fundamental attributes of the same temporal code operating at the level of the auditory nerve. Second, it is not clear

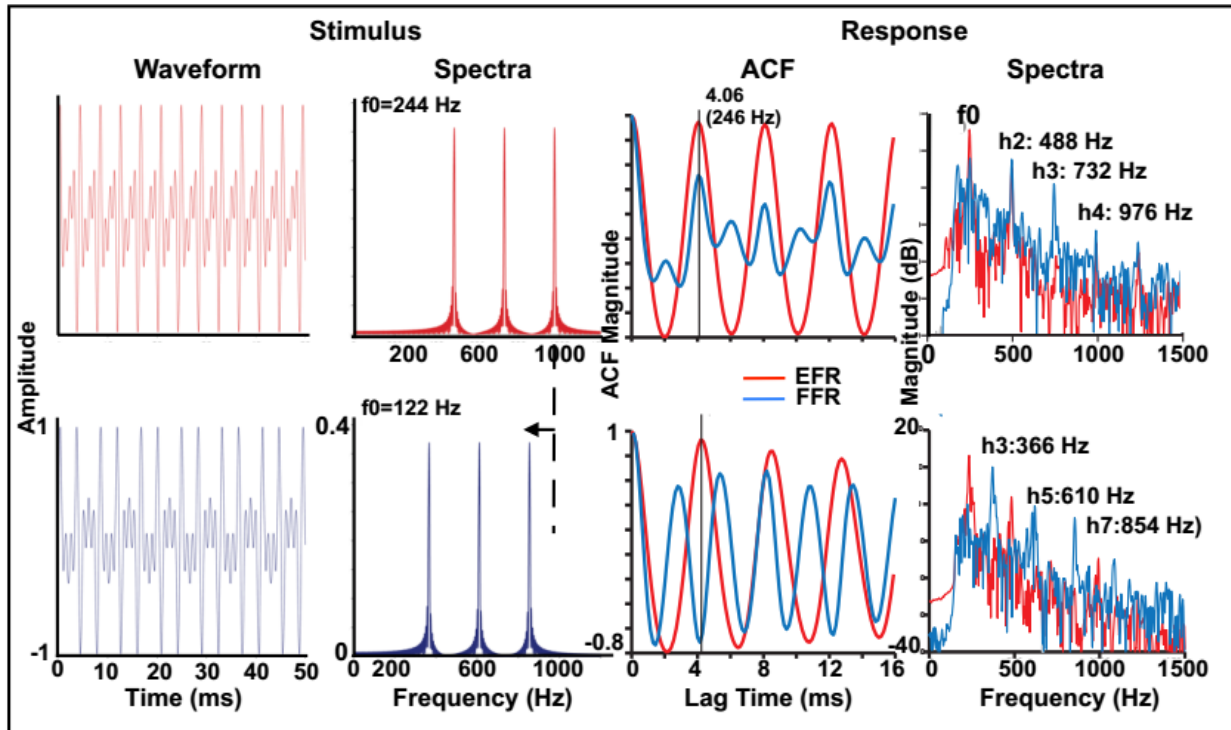


Figure 8: Waveforms and spectra of harmonic (red/dark), and inharmonic and frequency shifted (blue/light) complex tones. Column three and four illustrated the ACFs and spectra of the EFR (red/dark) and FFR (blue/light) responses. No clear peak proximal to perceived pitch is discernible in the FFR spectral data for the frequency shifted stimulus. Also note that the multiple peaks of FFR ACF for the frequency shifted tone surrounds the dominant EFR ACF peak at 4 ms. From Krishnan & Plack (2011b), Unpublished data presented at the 2011 midwinter meeting of the Association for Research in Otolaryngology.

that their dichotic stimuli produced the same pitch as when all harmonics are presented to the same ear. Notwithstanding, the pitch salience of their stimuli would be weak. It is possible that FFR-related neural activity is not sufficiently robust to preserve the less salient pitch for their stimuli. In our own experience, we have failed to measure FFR correlates of the less salient dichotic Huggins pitch.

### Neural correlates of resolved vs unresolved complex sounds

Psychoacoustic studies have also shown that complex tones containing resolved harmonics evoke stronger pitches than complex tones with only unresolved harmonics (for example, Houtsma and Smurzynski, 1990; Carlyon and Shackleton, 1994). Also, unresolved harmonics presented in alternating sine and cosine (ALT) phase produce a doubling of pitch, presumably the result of harmonic interaction in the basilar membrane (Shackleton and Carlyon, 1994a). Physiological studies show a consistent correlate of the pitch doubling for ALT stimuli as harmonic resolution decreases. That is, the interspike interval distributions (ISIDs) not only show peaks at  $F_0$  and  $2F_0$  but also, the interval based measure of pitch strength is almost as large at the envelope frequency  $2F_0$  as at the  $F_0$  for alternating phase stimuli with unresolved harmonics based

on both period histograms and ACGs (Horst et al., 1990; Palmer and Winter, 1992, 1993; Cedolin and Delgutte, 2005; Shackleton et al., 2009). Krishnan & Plack (2011) examined whether the temporal pattern of phase-locked neural activity reflected in the scalp recorded human frequency following response (FFR) preserves information relevant to pitch strength and pitch doubling for ALT stimuli. FFR and behavioral discrimination measures were obtained for complex tone burst stimuli with harmonics added in either sine phase (SIN) or alternating phase (ALT) phase. SIN-phase complexes had an  $F_0$  of 90 Hz (SIN 90) or 180 Hz (SIN 180). ALT phase complexes had an  $F_0$  of 90 Hz (ALT 90). For each of the complexes, harmonics were filtered into one of four spectral regions: 360-900 Hz, 720-1260 Hz, 1080-1620 Hz, and 1440-1980 Hz. These spectral regions were chosen to include stimuli with low-order harmonics that are completely resolved, and stimuli with higher order harmonics that are completely unresolved, in the cochlea. A low-pass Gaussian noise with cutoff frequencies set 180 Hz below the start of the complex passband (180 Hz, 540 Hz, 900 Hz, and 1260 Hz for the four spectral regions) was used to mask combination tones. The grand average FFRs for the three stimuli (SIN90, ALT90, and SIN180) for the resolved and unresolved spectral conditions are shown in Figure 9. The dotted box grouping the SIN 90 and ALT 90 for the resolved conditions, and the ALT90 and the SIN180 unre-

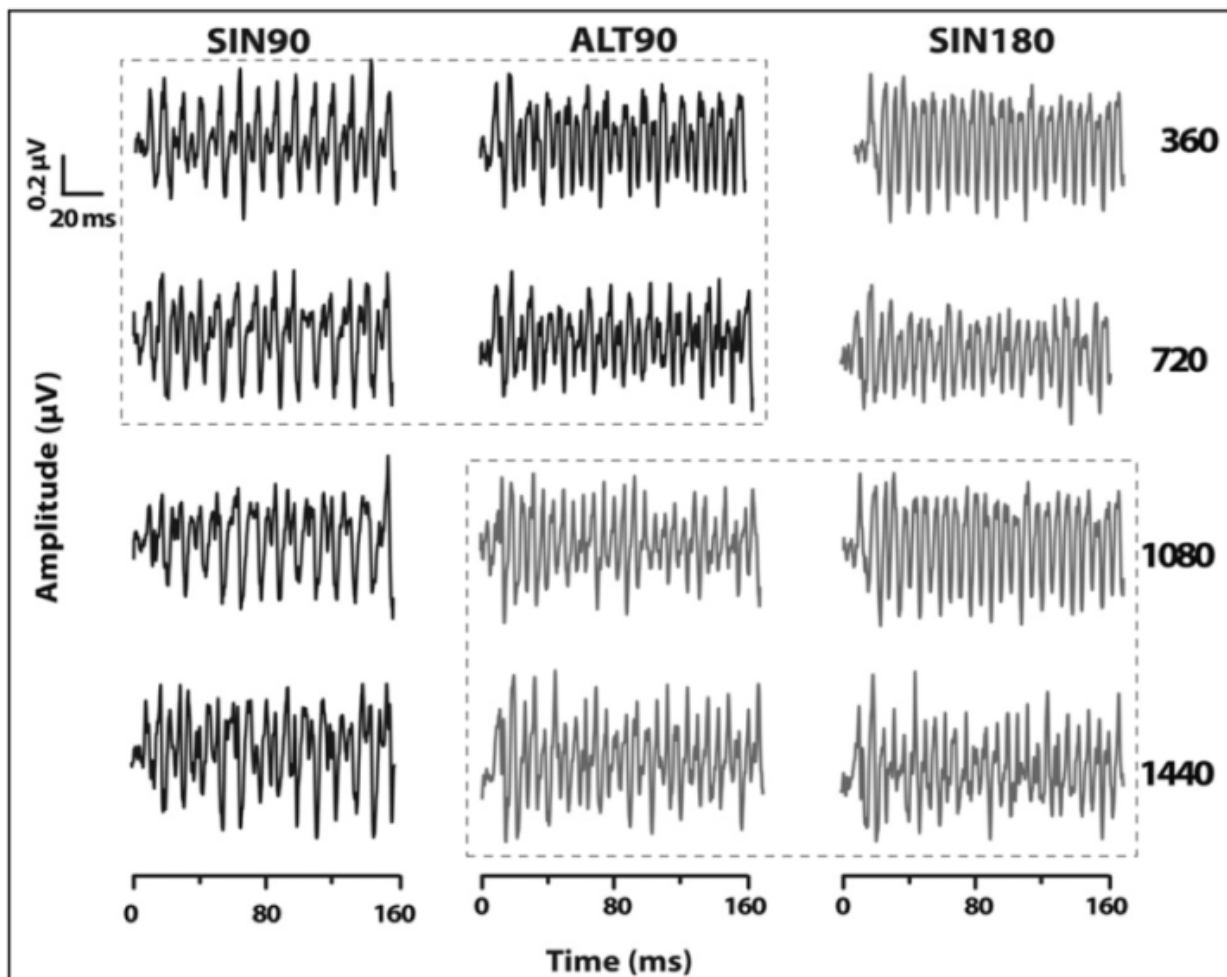


Figure 9: Grand averaged FFR waveforms across spectral regions (identified on the right) for SIN 90, ALT 90, and SIN 180 stimuli. The dashed boxes enclosing the SIN 90 and ALT 90 for the resolved stimuli, and the ALT 90 and SIN 180 for the unresolved stimuli illustrates the doubling in periodicity in the FFR to ALT 90 stimuli moving from resolved to unresolved stimuli. Replotted from Krishnan & Plack (2011).

solved conditions, respectively illustrates the change in periodicity of the FFR from 90 Hz for the resolved conditions to 180 Hz for the unresolved conditions—consistent with a doubling of perceived pitch. Furthermore, their results showed that FFR periodicity strength decreased as harmonic resolution decreased (ACF functions in Figure 10), qualitatively consistent with previous behavioral measures but appreciably smaller compared to the large changes in  $f_0$  difference limen (FODL) observed in behavioral measures. Also, FFR spectra (Figure 10) indicated a different pattern of phase-locked neural activity for ALT stimuli with resolved and unresolved harmonics consistent with the doubling of pitch observed in their behavioral estimates. Specifically, the shift in the relative prominence of the autocorrelation peaks in their FFR data for ALT 90 stimuli, and the clear shift in spectral pattern (from one consistent with a 90Hz  $f_0$  for the resolved ALT 90 stimuli to one more consistent with a 180 Hz  $f_0$  for unresolved ALT 90 stimuli) clearly suggest that the temporal pattern of neural activity relevant to perceptual doubling of pitch is indeed preserved in the neural activity

underlying the FFR. Lastly, the similarity between the autocorrelation analyses performed at the level of the auditory nerve (model response) and at the level of the IC (FFR) appears to suggest that a temporal representation of pitch based on pooled neural ISIDs is still potentially available at the level of the midbrain. Finally, the strong correlation between their FFR data and behavioral estimates of pitch suggests that the phase-locked neural activity reflected in the scalp recorded FFR preserves sensory level pitch information that may contribute to pitch perception. Thus, the scalp recorded FFR may provide for a non-invasive analytic tool to evaluate neural encoding of complex sounds in humans.

#### Relative roles of envelope and temporal fine structure in pitch

Periodic complex sounds (for example, amplitude modulated tones, complex tones, synthetic vowels) that are commonly used to evaluate temporal representation of pitch relevant information in the auditory system contain strong envelope (ENV) mod-

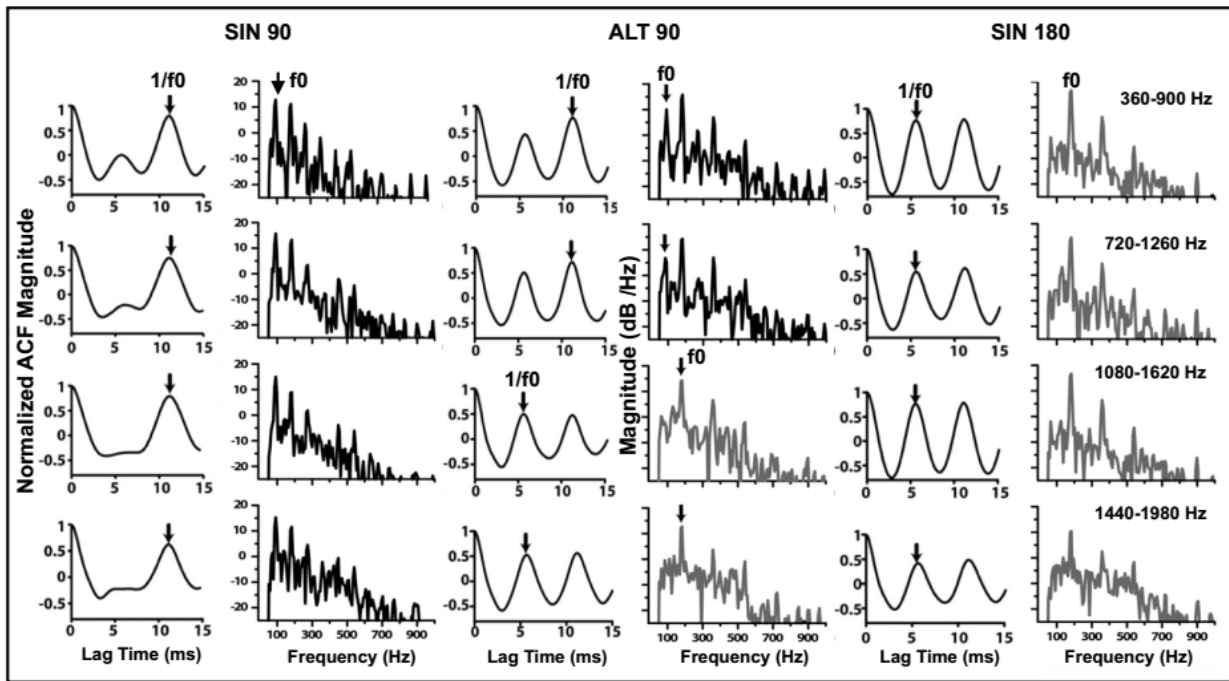


Figure 10: Mean FFR autocorrelation functions (ACF) and spectra for SIN 90 (Left two panels), ALT90 (middle two panels), and SIN180 (right two panels) plotted across the four spectral regions. The spectral regions are identified the right of each row of plots. For all three stimuli the ACF magnitude tends to decrease moving from resolved to unresolved stimuli; The different ACF peak delays for SIN90 and SIN180 remains unchanged across the spectral bands (i.e. going from resolved to unresolved stimuli); and both the ACF peak location and the spectral pattern for the ALT90 changes from resolved to unresolved stimuli consistent with the doubling of perceived pitch. From Krishnan & Plack (2011, Hearing Research).

ulation in addition to the waveform temporal fine structure (TFS). The relative roles of ENV and TFS in the temporal encoding of stimulus features relevant to pitch have been evaluated in the cochlear nucleus using iterated rippled noise (IRN) with positive (IRN<sub>p</sub>) and negative (IRN<sub>n</sub>) gain (Shofner, 1991, 1999; Neuert et al., 2005; Verhey and Winter 2006; Sayles and Winter 2007). IRN is generated

using wideband noise (WBN) that is delayed, attenuated, and then added/subtracted to the original WBN noise in an iterative manner. Unlike other pitch producing periodic complex sounds, IRN stimuli do not have highly modulated envelopes but impart a temporal regularity in the temporal fine structure. The autocorrelation functions (ACF) for IRN<sub>p</sub> and IRN<sub>n</sub> are identical for waveform ENV (peak at time lag corresponding to the delay ( $d$  ms) but different for the TFS (Shofner, 1991, 1999; Neuert et al., 2005). For TFS, the ACF of IRN<sub>p</sub> shows a peak at time lag corresponding to the delay, and the ACF of IRN<sub>n</sub> shows a peak at twice the delay ( $2*d$ ). The pitch of IRN<sub>p</sub> corresponds to the reciprocal of the delay (Bilsen and Ritsma, 1970; Yost, 1996; Sayles and Winter 2008) and for IRN<sub>n</sub>, pitch corresponds to  $1/(2*d)$  Hz (Yost, 1996; Shofner, 1999). That is, the perceived pitch of IRN<sub>n</sub> is an octave lower than that of IRN<sub>p</sub>. For IRN<sub>p</sub> and IRN<sub>n</sub> ENV response is the same, but TFS response is different and closely in line with the perceived pitch change.

Shofner (1999), showed that the neural autocorrelations in response to IRN<sub>p</sub> and IRN<sub>n</sub> indicated that the temporal discharge of primary-like units in the CN reflected the TFS of the stimulus, whereas the temporal discharge patterns of chopper units reflected the stimulus ENV.

Shofner (1999) concluded that the primary-like neurons are more likely to preserve temporal information relevant to pitch. These findings have been corroborated by both Verhey and Winter (2006), and Sayles and Winter (2007). Robust FFRs preserving pitch-relevant information have been recorded using IRN stimuli with both constant, and dynamic pitch (Krishnan, Gandour & Suresh, 2017a; b; Krishnan et al., 2016; Krishnan & Gandour, 2014). Ananthakrishnan & Krishnan (2018) examined whether the differential sensitivity to ENV and TFS in response to differences in pitch produced by IRN<sub>p</sub> and IRN<sub>n</sub> stimuli is preserved in the temporal pattern of ensemble phase-locked neural activity at more rostral levels in the brainstem as reflected in the human FFR. They reasoned that if FFRs simply reflect neural phase-locking to the waveform envelope and not pitch per se, then the FFRs to both IRN<sub>p</sub> and IRN<sub>n</sub> should be nearly identical. However, if the temporal pattern of activity reflected in the FFR aligns with the TFS for IRN<sub>p</sub>, and IRN<sub>n</sub>, it would support the view that the temporal pattern of FFR neural activity phase-locked to the TFS does indeed con-

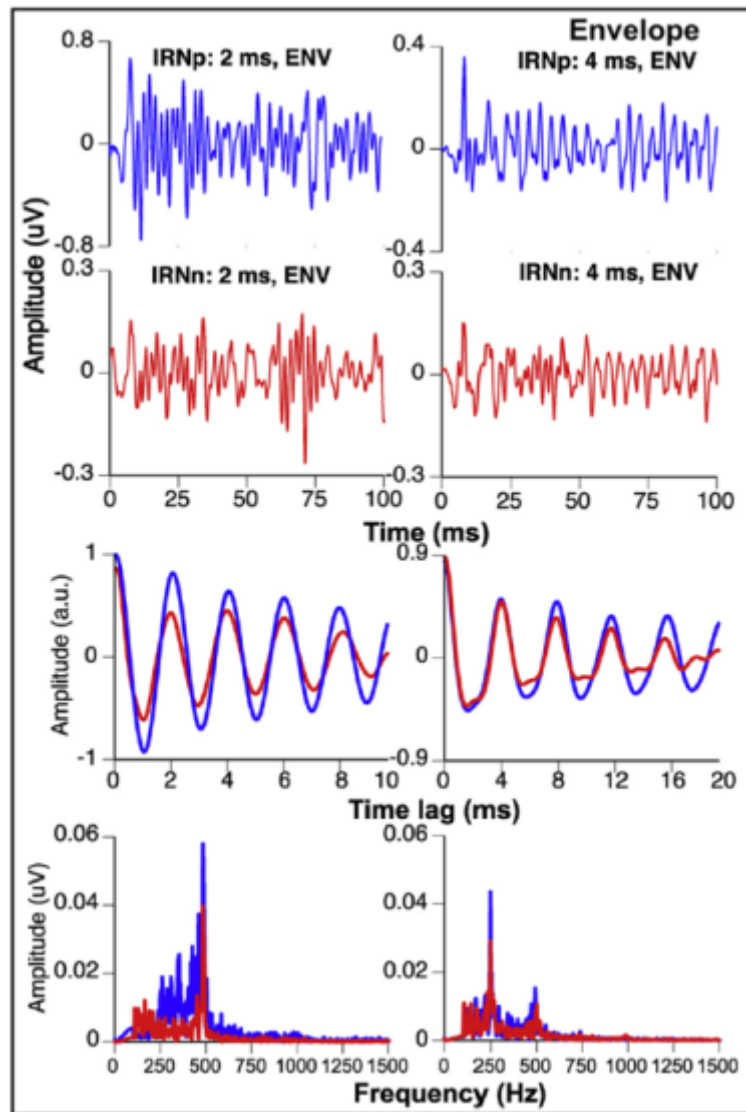


Figure 11: Grand average FFRENV waveform segments (top panels), ACFs (middle panel) and spectra (bottom panel) representing neural phase-locking to ENV for the 2ms delay (left panels) and 4ms delay (right panels) stimuli. While the ACF peak delay shifts with change in the periodicity (2 ms vs 4 ms) the ACF functions (and the spectra) remain the same for both IRNp and IRNn. IRNp is indicated by the blue/light tracing and IRNn by the red/dark tracing; FFR indicates frequency following response; ENV, envelope; IRN, iterated rippled noise; p, positive gain; n, negative gain. Replotted from Ananthakrishnan & Krishnan (2018).

tain temporal information relevant to pitch. FFRs were obtained in response to IRNp and IRNn stimuli with 2 ms, and 4 ms delay. Comparison of the independent analysis of the phase-locked activity to ENV, and TFS revealed that only the phase-locked activity to the TFS showed differences in both spectra and ACF that closely matched the pitch difference between the two stimuli (Figure 12) but not the phase-locked responses to the envelope periodicity (Figure 11). That is, both the ACF peak locations and the spectral patterns were consistent with the behaviorally observed pitch for both IRNp and IRNn stimuli (Figure 12). The temporal pattern of phase-locking to the envelope likely preserves information relevant to pitch change due to changes in delay (2 ms vs 4 ms) (Figure 11). This is consistent

with previous results from Shofner (1999) that indicate that the envelope responses to IRNp and IRNn for a given delay are identical and therefore cannot index pitch change. Only the temporal pattern of phase-locked neural activity to the TFS accounts for the behaviorally observed pitch and pitch lowering for all stimuli. Shofner (1991, 1999) showed that the ACFs for IRNp and IRNn were identical for waveform ENV (with peak corresponding to the delay) but for TFS, IRNp showed a peak at the delay (similar to the envelope ACF) and IRNn showed a peak at twice the delay (Shofner, 1991, 1999; Neuert et al., 2005). Also, for IRNp, behavioral measures of pitch correspond to the reciprocal of the delay (Bilsen and Ritsma, 1970; Yost, 1996; Sayles and Winter 2008), and for IRNn, pitch corresponds to  $1/(2*d)$  Hz (Yost,

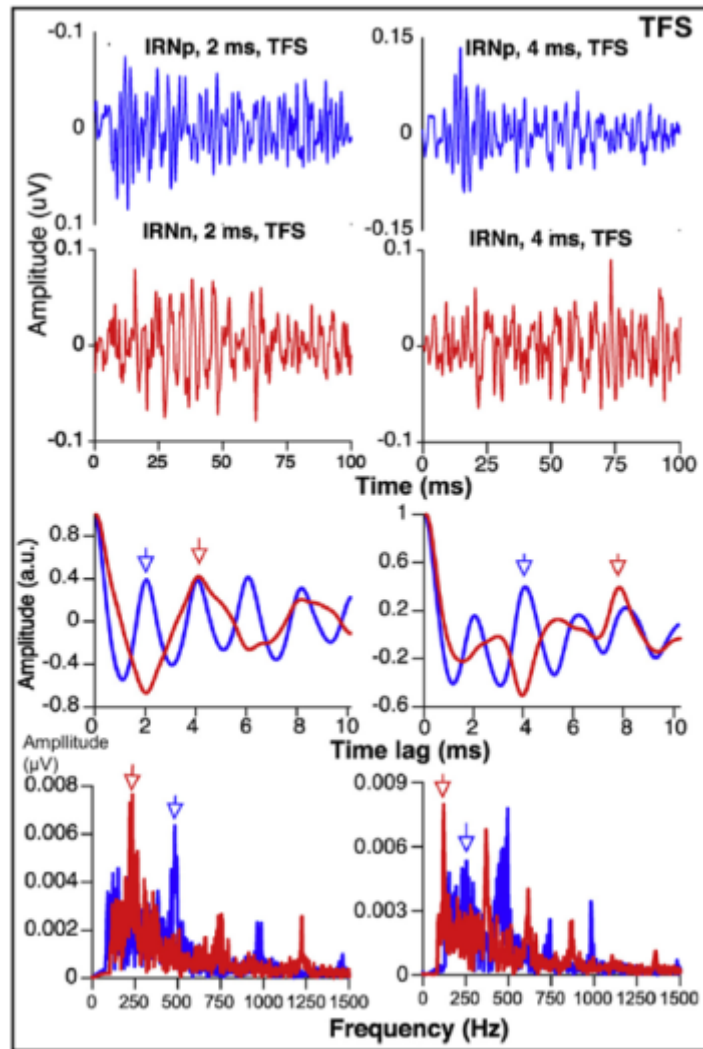


Figure 12: Grand average  $FFR_{TFS}$  waveform segments (top panels), ACFs (middle panel) and spectra (bottom panel) representing neural phase-locking to TFS for the 2ms delay (left panels) and 4ms delay (right panels) stimuli. Both the ACF and the spectra patterns are different for IRNp and IRNn-consistent with the perceived pitch. IRNp is indicated by the blue/light tracing and IRNn by the red/dark tracing; FFR indicates frequency following response; TFS, temporal fine structure; IRN, iterated rippled noise; p, positive gain; n, negative gain. Replotted from Ananthakrishnan & Krishnan (2018).

1996; Shofner, 1999). Shofner (1999), based on comparison of the neural autocorrelograms in response to IRNp and IRNn, concluded that the primary-like neurons in the cochlear nucleus reflecting responses to the waveform temporal fine-structure are more likely to preserve temporal information relevant to pitch whereas envelope related activity is preserved in chopper neurons more involved in encoding the envelope modulation. These findings have been corroborated by both Verhey and Winter (2006), and Sayles and Winter (2007). In addition, current models that explain pitch encoding of IRN stimuli rely on the temporal processing of waveform fine-structure information (Patterson et al., 1996; Yost, 1996; Yost et al., 1996). Taking the results from the above physiologic studies and the results of FFR study described in this section (Ananthakrishnan & Krishnan, 2018) it is clear that the temporal pattern of neural activity encoding “pitch change” are primarily driven by

the waveform fine structure and appears to be available in the phase-locked neural activity in the mid-brain. Finally, these results also suggest that the use of alternating polarity alone will not be optimal to evaluate neural representation of pitch since they represent phase-locking to only the envelope periodicity.

### Neural correlates of pitch salience

For a variety of complex sounds, including speech or music, perceived pitch and its salience is closely related to the periodicity strength of the stimulus waveform (Shofner & Selas, 2002; Fastl & Stoll, 1979; Yost, 1996b). IRN stimuli allows systematic manipulation of the waveform temporal regularity and temporal fine structure, and therefore pitch salience. Perceptually, IRN produces a pitch corresponding to the reciprocal of the delay, and its

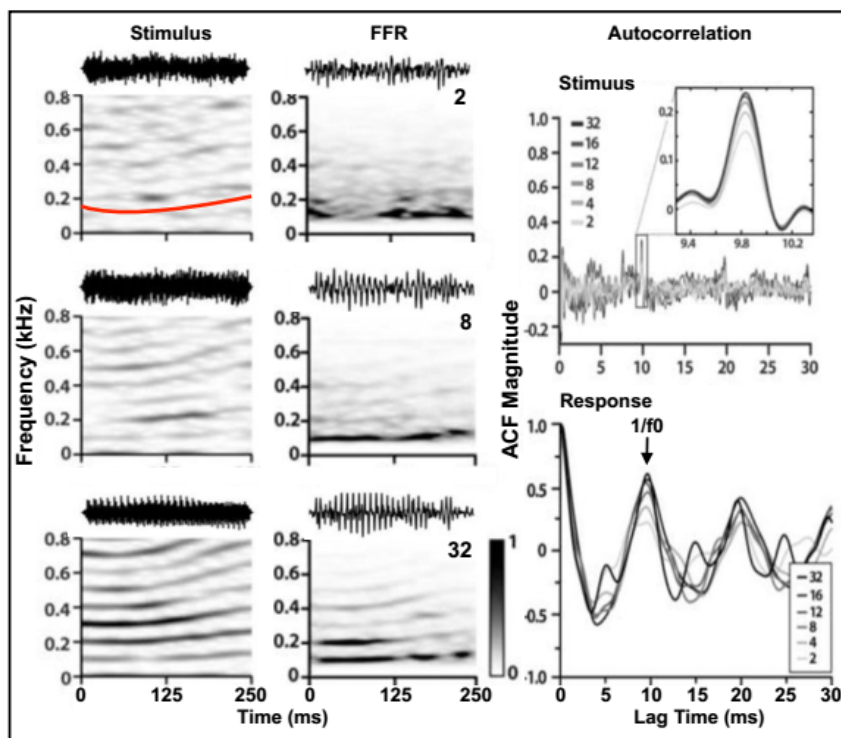


Figure 13: IRN stimulus (left) and FFR (middle) waveforms plotted at the top of spectrogram for each stimulus, and the autocorrelation functions for the stimulus (top right) and the FFR (bottom right) are all plotted as a function of iteration steps (2, 8, 32). The pitch contour is shown in red/black in the stimulus spectrogram for iteration 2. Consistent with the stimulus, the response spectrograms and the ACF functions shows clearly the steady growth of the response at the pitch period as the increase in temporal regularity of the stimulus with increase in iteration steps increases the perceived salience of the pitch. Replotted from Krishnan et al., (2011).

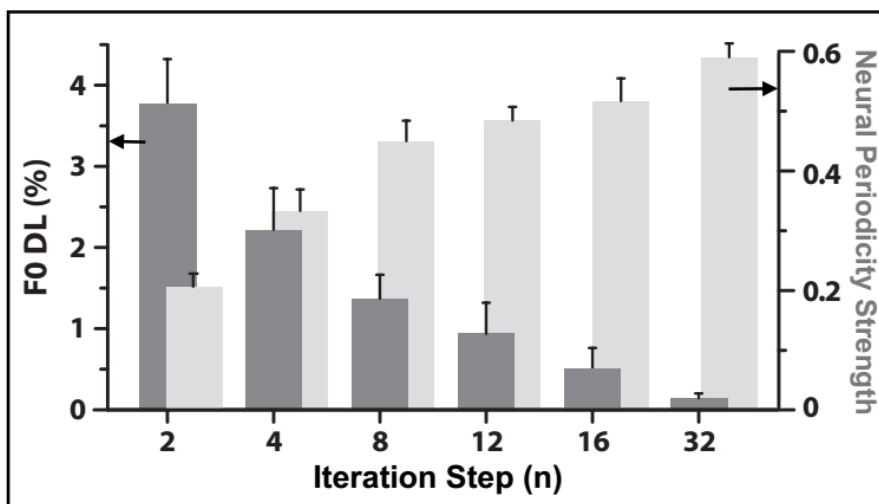


Figure 14: Comparison of the changes in the fundamental frequency difference limen (FODL) and the neural periodicity strength as a function of IRN iteration steps. Note the inverse relationship between FODL and Neural Periodicity Strength. Data from Krishnan et al., (2011) replotted.



pitch salience grows with increasing number of iterations (Patterson et al., 1996; Yost, 1978, 1996a; Yost and Hill, 1979). Physiologically, recordings of responses to static (i.e., single pitch) and time-varying (i.e., dynamic pitch) IRN stimuli from auditory nerve fibers (Fay et al., 1983; ten Kate & van Bakkum, 1988) and cochlear nucleus neurons (Bilsen et al., 1975; Sayles and Winter, 2007; Shofner, 1991; Shofner, 1999; Winter et al., 2001) show that the pitch of harmonic IRN is represented in the firing patterns of action potentials locked to either the temporal fine structure or envelope periodicity. Krishnan, Bidelman, & Gandour (2010) sought to demonstrate that the pitch relevant information preserved in the phase-locked neural activity generating the FFR is sensitive to changes in pitch salience and is correlated with corresponding changes in perceptual pitch salience. They recorded FFRs using an IRN version of lexical tone with time varying pitch that varied only in degree of pitch salience (the number of iterations of the IRN stimulus was systematically (Figure 13, stimulus panels) varied to produce a continuum from no pitch to very strong pitch). They also measured behavioral frequency difference limens (F0DLs) obtain a perceptual estimate related to pitch salience. Their results showed that neural periodicity strength increased systematically with increase in temporal regularity in stimulus periodicity, suggesting that the FFR pitch-relevant neural activity is indeed sensitive to changes in pitch salience (Figure 14, response panels). Consistent with this, F0DLs decreased with increasing stimulus temporal regularity (Figure 14). The negative correlation between neural periodicity strength and behavioral F0DL suggests that increasing strength in the representation of pitch salience may, at least in part, contribute to improvement in the perceptual ability to discriminate pitch with increasing temporal regularity in the stimulus periodicity. This finding is consistent with electrophysiological studies that show a predictable relationship between neural and behavioral measures of pitch salience. For example, the latency and amplitude of the, pitch onset response (a cortical pitch specific response) varies systematically with the pitch salience of an IRN stimulus (Krumbholz et al., 2003; Soeta et al., 2005), suggesting that the neural activity underlying the generation of the pitch onset response is involved in extracting an initial estimate of the pitch salience of the sound; the strong correspondence between neural pitch strength of complex sounds and their pitch salience in auditory nerve responses (Cariani and Delgutte, 1996a; Cariani and Delgutte, 1996b); and the more robust neural periodicity strength in the FFR for consonant intervals compared to dissonant pitch intervals in music (Bidelman and Krishnan, 2009). Overall, these results showing growth in FFR neural periodicity strength with increasing iteration steps likely reflects an improvement in neural encoding of pitch relevant periodicities in the brainstem. Thus, neural information relevant to pitch salience may already

be emerging in pre-attentive, early sensory-level processing in the midbrain.

## CONCLUSIONS

There has been a growing interest in the utilizing of the FFR/EFR to address clinical/research questions as reflected by an appreciable increase in published reports on EFR and FFR using a variety of complex sounds in the last decade. The temporal pattern of phase-locked activity underlying the scalp recorded FFR preserves information about certain acoustic features of complex sounds that presumably contribute to speech and pitch perception. While it is likely that this neural information reflects neural processing locally in the midbrain (that may be subject to top-down modulation), more conclusive studies are needed to determine if FFR elicited by complex sounds reflects output of local neural encoding processes or merely a propagated version of neural encoding at the auditory nerve level. The question about the locus of the process notwithstanding there is compelling evidence to suggest that the FFR can indeed be used as a robust analytic tool to examine neural encoding of stimulus features that may contribute to speech, and pitch perception and how degraded neural representation (as reflected in the FFR) consequent to peripheral-, central auditory disorders and/or adverse listening conditions could negatively alter these percepts. The focus clearly has shifted, rightfully so, in applying FFRs as a metric of auditory temporal processing in normal and impaired ears, age related changes in neural encoding, and to understand the role of experience-dependent plasticity in shaping subcortical processing and its application to training and perceptual learning. The determination of a strong relationship between neural representation of acoustic features and perception will motivate the development of objective outcome measures for hearing prosthetic devices and facilitate development of individual centered optimal signal processing strategy based on the FFR measure. However, caution should be exercised in the over interpretation of the information contained in the FFR.

## REFERENCES

- Aiken, S. J., & Picton, T. W. (2008). Envelope and spectral frequency-following responses to vowel sounds. *Hearing Research, 245*, 35–47.
- Ananthakrishnan, S., & Krishnan, A. (2018). Human frequency following responses to iterated rippled noise with positive and negative gain: Differential sensitivity to waveform envelope and temporal fine structure. *Hearing Research, 367*, 113-123.
- Ananthakrishnan, S., Krishnan, A., & Bartlett, E. (2016). Human frequency following response: Neural representation of envelope and temporal fine structure in listeners with normal hearing and sensorineural hearing loss. *Ear & Hearing, 37*, e91-e103.

- Ananthkrishnan, S., Lou, Xin., & Krishnan, A. (2017). Human frequency following responses to vocoded speech. *Ear & Hearing, 38*, e256-e267.
- Ananthanarayan, A. K., & Durrant, J. J. (1992). The frequency following response and the onset response: Evaluation of frequency specificity using a forward-masking paradigm. *Ear & Hearing, 13*, 228-232.
- Anderson, S., Parbery-Clark, A., White-Schwoch, T., Drebohl, S., & Kraus, N. (2013). Effects of hearing loss on the subcortical representation of speech cues. *The Journal of the Acoustical Society of America, 133*, 3030-38.
- Arthur, R. M., Pfeifer, R. R., & Suga, N. (1971). Properties of two-tone inhibition in primary auditory neurons. *Journal of Physiology, 212*, 593-609.
- Bidelman, G., & Krishnan, A. (2009). Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *Journal of Neuroscience, 29*, 13165-13171.
- Bidelman, G., Gandour, J., & Krishnan, A. (2011). Cross-domain effects of music and language experience on the representation of pitch in the human auditory brainstem. *Journal of Cognitive Neuroscience, 23*, 425-434.
- Bidelman, G. M., Davis, M. K., & Pridgen, M. H. (2018). Brainstem-cortical functional connectivity for speech is differentially challenged by noise and reverberation. *Hearing Research, 367*, 149-160.
- Bidelman, G. M. (2017). Communicating in challenging environments: Noise and reverberation. In: N. Kraus, S. Anderson, T. White-Schwoch, R.R. Fay, & A.N. Popper AN (Eds.), *Springer Handbook of Auditory Research: The frequency-following response: A window into human communication*. New York, N.Y: Springer Nature.
- Bidelman, G. M. (2018). Subcortical sources dominate the neuroelectric auditory frequency-following response to speech. *NeuroImage, 175*, 56-69.
- Bidelman, G. M., & Krishnan, A. (2011). Effects of reverberation on brainstem representation of speech in musicians and non-musicians. *Brain Research, 1355*, 112-125.
- Billings, C. J., Bologna, W. J., Muralimanohar, R. K., Madsen, B., & Molis, M. R. (2019). Frequency following responses to tone glides: effects of frequency extent, direction, and electrode montage. *Hearing Research, 375*, 25-33.
- Bilsen, F. A., & Ritsma, R. J. (1970). Some parameters influencing the perceptibility of pitch. *The Journal of the Acoustical Society of America, 47*, 469-475.
- Bilsen, F. A., ten Kate, J. H., Buunen, T. J. F., & Raatgever, J. (1975). Responses of single units in the cochlear nucleus of the cat to cosine noise. *The Journal of the Acoustical Society of America, 58*, 858-866.
- Blackburn, C. C., & Sachs, M. B. (1990). The representations of the steady-state vowel sound /e/ in the discharge patterns of cat anteroventral cochlear nucleus neurons. *Journal of Neurophysiology, 63*, 1191-1212.
- Brown, R. (1958). *Words and Things*. Glencoe, IL: Free Press, 36-42.
- Cariani, P. A., & Delgutte, B. (1996a). Neural correlates of the pitch of complex tones. I. Pitch and pitch salience. *Journal of Neurophysiology, 76*(3), 1698-1716.
- Cariani, P. A., & Delgutte, B. (1996b). Neural correlates of the pitch of complex tones. II. Pitch shift, pitch ambiguity, phase invariance, pitch circularity, rate pitch, and the dominance region for pitch. *Journal of Neurophysiology, 76*, 1717-1734.
- Carlson, R., Fant, G., & Granstrom, B. (1975). Two-formant models, pitch and vowel perception. In G. Fant, & M.A.A. Tatham (Eds.), *Auditory Analysis and Perception of Speech*. London: Academic, pp 55-82.
- Carlyon, R. P., & Shackleton, T. M. (1994). Comparing the fundamental frequencies of resolved and unresolved harmonics: Evidence for two pitch mechanisms. *The Journal of the Acoustical Society of America, 95*, 3541-3554.
- Casparly, D. M., Rupert, A. L., & Moushegian, G. (1977). Neuronal coding of vowel sounds in the cochlear nuclei. *Experimental Neurology, 54*, 414-431.
- Cedolin, L., & Delgutte, B. (2005). Pitch of complex tones: rate-place and interspike interval representations in the auditory nerve. *Journal of Neurophysiology, 94*, 347-362.
- Clinard, C. G., & Cotter, C. M. (2015). Neural representation of dynamic frequency is degraded in older adults. *Hearing Research, 323*, 91-98.
- Cunningham, J., Nicol, T., King, C., Zecker, S. G., & Kraus, N. (2002). Effects of noise and cue enhancement on neural responses to speech in auditory midbrain, thalamus and cortex. *Hearing Research, 169*, 97-111.
- Dau, T. (2003). The importance of cochlear processing for the formation of auditory brainstem and frequency following responses. *The Journal of the Acoustical Society of America, 113*, 936-950.
- De Boer, E. (1956). On the Residue in Hearing. *Academic Thesis, University of Amsterdam*.
- De Boer, E. (1976). On the "Residue" and auditory pitch perception. *Auditory System, 479-583*.
- Drgas, S., & Blaszkak, M. A. (2009). Perceptual consequences of changes in vocoded speech. *Ear & Hearing, 41*, 300-311.

- Easwar, V., Purcell, D. W., Aiken, S. J., Parsa, V., & Scollie, S. D. (2015b). Evaluation of speech-evoked envelope following responses as an objective aided outcome measure: Effect of stimulus level, bandwidth, and amplification in adults with hearing loss. *Ear and Hearing, 36*, 635–652.
- Fastl, H., & Stoll, G. (1979). Scaling of pitch strength. *Hearing Research, 1*, 293–301.
- Fay, R. R., Yost, W. A., & Coombs, S. (1983). Psychophysics and neurophysiology of repetition noise processing in a vertebrate auditory system. *Hearing Research, 12*, 31–55.
- Galbraith, G., Jhaveri, S. P., & Kuo, J. (1997). Speech-evoked brainstem frequency-following responses during verbal transformations due to word repetition. *Electroencephalography and Clinical Neurophysiology, 102*, 46–53.
- Gelfand, S. A., & Silman, S. (1979). Effects of small room reverberation upon the recognition of some consonant features. *The Journal of the Acoustical Society of America, 66*, 22–29.
- Goblick, T. J., & Pfeifer, R. R. (1969). Time-domain measurements of the cochlear nonlinearities using combination click stimuli. *The Journal of the Acoustical Society of America, 46*, 924–938.
- Gockel, H. E., Carlyon, R. P., Mehta, A., & Plack, C. J. (2011). The frequency following response (FFR) may reflect pitch-bearing information but is not a direct representation of pitch. *Journal of the Association for Research in Otolaryngology, 12*, 767–782.
- Greenberg, S. (1980). Neural temporal coding of pitch and vowel quality. *UCLA Working Papers in Phonetics, 52*, (doctoral dissertation). Los Angeles.
- Greenberg, S., Marsh, J. T., Brown, W. S., & Smith, J. C. (1987). Neural temporal coding of low pitch. I. Human frequency-following responses to complex tones. *Hearing Research, 25*, 91–114.
- Hall, J. W. III. (1979) Auditory brainstem frequency following responses to waveform envelope periodicity. *Science, 205*, 1297–1299.
- Henry, K. R. (1999). Noise improves transfer of near-threshold, phase-locked activity of the cochlear nerve: Evidence for stochastic resonance? *Journal of Comparative Physiology A: Neuroethology, Sensory, Neural, and Behavioral Physiology, 184*, 577–584.
- Horst, W., Javel, E., & Farley, G.R. (1990). Coding of spectral fine structure in the auditory nerve. II. Level-dependent nonlinear responses. *The Journal of the Acoustical Society of America, 88*, 2656–2681.
- Houtgast, T. & Steeneken, H. J. M. (1985). A review of the MTF-concept in room acoustics, *The Journal of the Acoustical Society of America, 77*, 1069–1077.
- Houtgast, T., & Steeneken, H. J. M. (1973). The modulation transfer function in room acoustics as a predictor of speech intelligibility. *Acustica, 28*, 66–73.
- Houtsma, A., & Smurzynski, J. (1990). Pitch identification n and discrimination for complex tones with many harmonics. *The Journal of the Acoustical Society of America, 87*, 304–310.
- Jacobson, R., Fant, G., & Halle, M. (1963). Preliminaries to Speech Analysis. Cambridge, MA: MIT.
- Janssen, T., Steinhoff, H.J., & Bohnke, F. (1991). Zum Entstehungs Mechanisms der Frequenzfolgepotentiale. *Otorhinolaryngology Nova, 1*, 16–25.
- Keilson, S. E., Richards, V. M., Wyman, B. T., & Young, E. D. (1997). The representation of concurrent vowels in the cat anesthetized ventral cochlear nucleus: evidence for a periodicity-tagged spectral representation. *The Journal of Acoustical Society of America, 102*, 1056–1071.
- Krishnan, A., Gandour, J., & Bidelman, G. (2010). Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hearing Research, 268*, 60–66.
- Krishnan, A., Gandour, J., & Bidelman, G. (2010b). Brainstem pitch representation in native speakers of Mandarin is less susceptible to degradation of stimulus temporal regularity. *Brain Research, 1313*, 124–133.
- Krishnan, A. (1999). Human frequency following responses to two-tone approximations of steady-state vowels. *Journal of Audiology and Neuro-Otology, 4*, 95–103.
- Krishnan, A. (2002). Human frequency-following responses: representation of steady-state synthetic vowels. *Hearing Research, 166*, 192–201.
- Krishnan, A. (2006). Human frequency following response. In R.F. Burkard, M. Don, & J.J. Eggermont (Eds.), *Auditory evoked potentials: Basic principles and clinical application*. (pp. 313–335). Baltimore, MD: Lippincott Williams & Wilkins.
- Krishnan, A., & Agrawal, S. (2010). Human frequency-following response to speech-like sounds: correlates of off-frequency masking. *Journal of Audiology & Neuro-Otology, 15*, 221–228.
- Krishnan, A., & Gandour, J. T. (2017). Brainstem representation of pitch relevant information is shaped by language experience. In N. Kraus, S. Anderson, T. White-Schwoch, A.N. Popper, & R.R.Fay, (Eds.), *The Frequency-following Response: A Window into Human Communication*. (pp. 45–73). Auditory Research Series, ASA Press, Springer.
- Krishnan, A., & Gandour, J. T. (2014). Language experience shapes processing of pitch relevant information in the human brainstem and auditory cortex: Electrophysiological evidence. *Acoustics Australia, 42*, 187–199.

- Krishnan, A., & Parkinson, J. (2000). Human frequency following response: Representation of tonal sweeps. *Journal of Audiology & Neuro-Otology*, 5, 312-321.
- Krishnan, A., & Plack, C. (2011b). Human frequency following responses: Correlates of the pitch of complex stimuli with inharmonic and frequency-shifted components. Data presented at the Mid-winter meeting of the Association for Research in Otolaryngology (ARO), Baltimore.
- Krishnan, A., & Plack, C. (2011). Neural encoding in the human brainstem relevant to the pitch of complex tones. *Hearing Research*, 275, 110-119.
- Krishnan, A., Bidelman, G. M., & Gandour, J. T. (2010). Neural representation of pitch salience in the human brainstem revealed by psychophysical and electrophysiological indices. *Hearing Research*, 268, 1-2. doi: 10.1016/j.heares.2010.04.016.
- Krishnan, A., Gandour, J. T., & Chandan, S. (2016). Language-experience plasticity in neural representation of changes in pitch salience. *Brain Research*, 1637: 102-117.
- Krishnan, A., Gandour, J. T., & Chandan, S. (2017b). Differential sensitivity to changes in pitch acceleration in the auditory brainstem and cortex. *Brain & Language*, 169, 22-27.
- Krishnan, A., Gandour, J. T., Xu, Y., & Chandan, S. (2016). Language-dependent changes in pitch-relevant neural activity in the auditory cortex reflect differential weighting of temporal attributes of pitch contours. *Journal of Neurolinguistics*, 41, 38-49.
- Krishnan, A., Suresh, C., & Gandour, J. T. (2019). Tone language experience-dependent advantage in pitch representation in brainstem and auditory cortex is maintained under reverberation. *Hearing Research*, 177, 63-71.
- Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. (2005). Encoding of pitch in the human brainstem is sensitive to language experience. *Cognitive Brain Research*, 25, 161-168.
- Krishnan, A., Xu, Y., Gandour, J. T., & Cariani, P. A. (2004). Human frequency-following response: representation of pitch contours in Chinese tones. *Hearing Research*, 189, 1-12.
- Krumbholz, K., Patterson, R. D., Seither-Preisler, A., Lammertmann, C., & Lütkenhoner, B. (2003). Neuromagnetic evidence for a pitch processing center in Heschl's gyrus. *Cerebral Cortex*, 13, 765-772.
- Li, X., & Jeng, F. C. (2011). Noise tolerance in human frequency-following responses to voice pitch. *The Journal of the Acoustical Society of America*, 129, 21-26.
- Madhavi, B., Krishnan, A., & Weber-Fox, C. (2009). Brainstem correlates of temporal auditory processing in children with specific language impairment. *Developmental Neuroscience*, 13, 77-91.
- Mesgarani, N., David, S. V., Fritz, J. B., & Shamma, S. A. (2014). Mechanisms of noise robust representation of speech in primary auditory cortex. *Proceedings of the National Academy of Science, USA* 111, 6792-6797.
- Miller, M. I., & Sachs, M. B. (1983). Representation of stop consonants in the discharge patterns of auditory-nerve fibers. *The Journal of the Acoustical Society of America*, 74, 502-517.
- Miller, M. I., & Sachs, M. B. (1984). Representation of voice pitch in discharge patterns of auditory-nerve fibers. *Hearing Research*, 14, 257-279.
- Nabelek, A. K., & Dagenais, P. A. (1986). Vowel errors in noise and in reverberation in hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 80, 741-748.
- Nabelek, A. K., & Letowski, T. R. (1988). Similarities of vowels in nonreverberant and reverberant fields. *The Journal of the Acoustical Society of America*, 83, 1891-1899.
- Nabelek, A. K., Letowski, T. R., & Tucker, F. M. (1989). Reverberant overlap- and self-masking in consonant identification. *The Journal of the Acoustical Society of America*, 86, 1259-1265.
- Neuert, V., Verhey, J. L., & Winter, I. M. (2005). Temporal representation of the delay of iterated rippled noise in the dorsal cochlear nucleus. *Journal of Neurophysiology*, 93, 2766-2776.
- Palmer, A. R., & Russell, I. J. (1986). Phase-locking in the cochlear nerve of the guinea-pig and its relation to the receptor potential of inner hair-cells. *Hearing Research*, 24, 1-15.
- Palmer, A. R. and Winter, I. M. (1993). Coding of the fundamental frequency of voiced speech sounds and harmonic complexes in the cochlear nerve and ventral cochlear nucleus. In M. A. Merchán, J. M. Juiz, D. A. Godfrey, & E. Mugnaini, (Eds.), *The mammalian cochlear nuclei: Organization and Function*. pp. 373-384, New York, NY: Plenum Press.
- Palmer, A. R., & Winter, I. M. (1992). Cochlear nerve and cochlear nucleus responses to the fundamental frequency of voiced speech sounds and harmonic complex tones. *Auditory Physiology Perception*, 231-239.
- Palmer, A. R., Winter, I. M., & Darwin, C. J. (1986). The representation of steady-state vowel sounds in the temporal discharge patterns of the guinea pig cochlear nerve and primary like cochlear nucleus neurons. *The Journal of the Acoustical Society of America*, 79, 100-113.
- Parbery-Clark, A., Skoe, E., & Kraus, N. (2009). Musical experience limits the degradative effects of background noise on the neural processing of sound. *The Journal of Neuroscience*, 29, 14100-14107.
- Patterson, R. D., Handel, S., Yost, W. A., & Datta, A. J. (1996). The relative strength of the tone

- and noise components in iterated ripple noise. *The Journal of the Acoustical Society of America*, *100*, 3286–3294.
- Peterson, G. E., & Barney, H. L. (1952). Control methods used in a study of vowels. *The Journal of the Acoustical Society of America*, *24*, 175–184.
- Plyler, P., & Krishnan, A. (2001). Human frequency following response: Representation of CV formant transitions in normal-hearing and hearing-impaired listeners. *Journal of the American Academy of Audiology*, *12*, 523–533.
- Prevost, F., Laroche, M., Marcoux, A. M., & Dajani, H. R. (2013). Objective measurement of physiological signal-to-noise gain in the brainstem response to a synthetic vowel. *Clinical Neurophysiology*, *124*, 52–60.
- Recio, A., & Rhode, W.S. (2000). Representation of vowel stimuli in the ventral cochlear nucleus of the chinchilla. *Hearing Research*, *146*, 167–184.
- Rhode, W. (1998). Neural encoding of single-formant stimuli in the ventral cochlear nucleus of the chinchilla. *Hearing Research*, *117*, 39–56.
- Russo, N., Nicol, T., Musacchia, G., & Kraus, N. (2004). Brainstem responses to speech syllables. *Clinical Neurophysiology*, *115*, 2021–2030.
- Sachs, M. B., Voigt, H. F., & Young, E. D. (1983). Auditory nerve representation of vowels in background noise. *Journal of Neurophysiology*, *50*, 27–45.
- Sayles, M., & Winter, I. M. (2007). The temporal representation of the delay of dynamic iterated rippled noise with positive and negative gain by single units in the ventral cochlear nucleus. *Brain Research*, *1171*, 52–66.
- Sayles, M., & Winter, I. M. (2008). Reverberation challenges the temporal representation of the pitch of complex sounds. *Neuron*, *58*, 789–801.
- Sayles, M., Stasiak, A., & Winter, I. M. (2015). Reverberation impairs brainstem temporal representations of voiced vowel sounds: challenging “periodicity-tagged” segregation of competing speech in rooms. *Frontiers of Systems Neuroscience*, *8*, doi: 10.3389/fnsys.2014.00248. 00248.
- Sayles, M., Stasiak, A., & Winter, I. M. (2016). Neural segregation of concurrent Speech: Effects of background noise and reverberation on auditory scene analysis in the ventral cochlear nucleus. In *Physiology, Psychoacoustics and Cognition in Normal and Impaired Hearing*, pp 389–397. Part of the *Advances in Experimental Medicine and Biology book series (AEMB, volume 894)*.
- Schouten, J. F. (1940). The residue, a new component in the subjective sound analysis. *Proceedings of the Koninklijke Nederlandse Akademie van Wetenschappen*, *43*, 356–365.
- Shackleton, T. M., & Carlyon, R. P. (1994a). The role of resolved and unresolved harmonics in pitch perception and frequency modulation discrimination. *The Journal of the Acoustical Society of America*, *95*, 3529–40.
- Shackleton, T. M., Liu, L. F., & Palmer, A. R. (2009). Responses to diotic, dichotic, and alternating phase harmonic stimuli in the inferior colliculus of Guinea pigs. *Journal of the Association for Research in Otolaryngology*, *10*, 76–799.
- Shofner, W. P. (1999). Responses of cochlear nucleus units in the chinchilla to iterated rippled noises: analysis of neural autocorrelograms. *Journal of Neurophysiology*, *81*, 2662–2674.
- Shofner, W. P. (1991). Temporal representation of rippled noise in the anteroventral cochlear nucleus of the chinchilla. *The Journal of the Acoustical Society of America*, *90*, 2450–66.
- Shofner, W. P., & Selas, G. (2002). Pitch strength and Stevens’s power law. *Perception Psychophysics*, *64*, 437–50.
- Smalt, C. J., Krishnan, A., Bidelman, G. M., Ananthakrishnan, S., & Gandour, J. T. (2012). Neural correlates of cochlear distortion products and their influence on representation of pitch relevant information in the human brainstem. *Hearing Research*, *292*, 26–34.
- Soeta, Y., Nakagawa, S., & Tonoike, M. (2005). Auditory evoked magnetic fields in relation to iterated rippled noise. *Hearing Research*, *205*, 256–61.
- Song, J. H., Skoe, E., Banai, K., & Kraus, N. (2011). Perception of speech in noise: Neural correlates. *Journal of Cognitive Neuroscience*, *23*, 2268–2279.
- Suresh, C., Krishnan, A., & Gandour, J. T. (2017). Language experience-dependent advantage in pitch representation in the auditory cortex is limited to favorable signal-to-noise ratios. *Hearing Research*, *355*, 42–53.
- Suresh, C. H., Krishnan, A., & Luo, X. (2020). Human frequency following responses to vocoded speech with amplitude modulation alone, and amplitude modulation plus frequency modulation.
- Swaminathan, J., Krishnan, A., & Gandour, J. T. (2008). Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. *Neuroreport*, *19*, 1163–1167.
- ten Kate J. H., van Bakkum M. F. (1988). Synchrony-dependent autocorrelation in eighth-nerve fiber response to rippled noise. *Journal of the Acoustical Society of America*, *84*, 2092–2102.
- Verhey, J. L., & Winter, I. M., (2006). The temporal representation of the delay of iterated rippled noise with positive or negative gain by chopper units in the cochlear nucleus. *Hearing Research*, *216*, 43–51.

Wang, D., & Brown, G. J. (2006). In D. Wang, & G.J. Brown (Eds.), *Computational Auditory Scene Analysis: Principles, Algorithms, and Applications* (pp. 395). New York, NY: Wiley/IEEE Press.

Winter, I. M, Wiegrebe, L., & Patterson, R. D. (2001). The temporal representation of the delay of iterated rippled noise in the ventral cochlear nucleus of the guinea-pig. *Journal of Physiology*, 537, 553–66.

Yost, W. A. (1996). Pitch of iterated rippled noise. *The Journal of the Acoustical Society of America*, 100, 511-518.

Yost, W. A. (1996b). Pitch of iterated rippled noise. *The Journal of the Acoustical Society of America*, 100, 511–518.

Yost, W. A., Patterson, R., & Sheft, S. (1996). A time domain description for the pitch strength of iterated rippled noise. *The Journal of the Acoustical Society of America*, 99, 1066-1078.

Young, E. D., & Sachs, M. B. (1979). 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*, 66, 1381–1403.