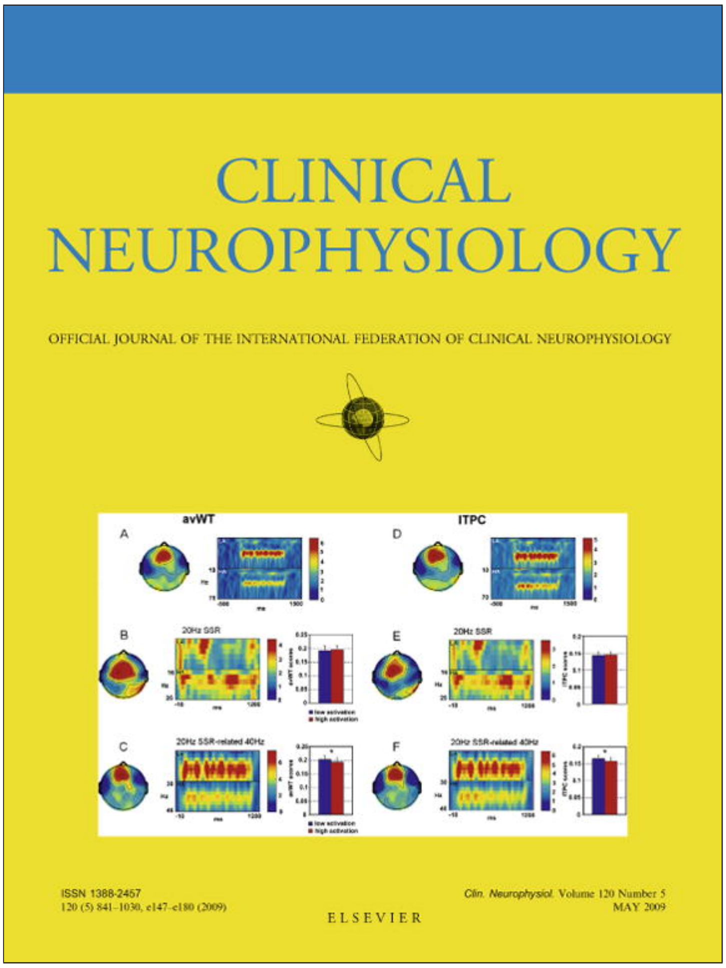
CLINICAL
NEUROPHYSIOLOGY
OFFICIAL JOURNAL OF THE INTERNATIONAL FEDERATION OF CLINICAL NEUROPHYSIOLOGY
avWT
ITPC
A
B
C
D
E
F
20Hz SSR
20Hz SSR-related 40Hz
ISSN 1388-2457
120 (5) 841–1030, e147–e180 (2009)
Clin. Neurophysiol. Volume 120 Number 5
MAY 2009
ELSEVIER

Clinical Neurophysiology 120 (2009) 959–966

Contents lists available at ScienceDirect

Clinical Neurophysiology

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

# Exploring the relationship between physiological measures of cochlear and brainstem function

S. Dhar [a,b,*], R. Abel [a], J. Hornickel [a], T. Nicol [a], E. Skoe [a], W. Zhao [a], N. Kraus [a,b,c]

[a] Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA
[b] The Hugh Knowles Center for Basic and Applied Hearing Research, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA
[c] Departments of Neurobiology and Physiology, Otolaryngology, Northwestern University, Evanston, IL 60208, USA

ARTICLE INFO

*Article history:*
Accepted 24 February 2009
Available online 5 April 2009

*Keywords:*
Otoacoustic emissions
Cochlea
Speech-evoked auditory brainstem response
Speech encoding

ABSTRACT

*Objective:* Otoacoustic emissions and the speech-evoked auditory brainstem response are objective indices of peripheral auditory physiology that are used clinically for assessing hearing function. While each measure has been extensively explored, their interdependence and the relationships between them remain relatively unexplored.
*Methods:* Distortion product otoacoustic emissions (DPOAEs) and speech-evoked auditory brainstem responses (sABRs) were recorded from 28 normal-hearing adults. Through correlational analyses, DPOAE characteristics were compared to measures of sABR timing and frequency encoding. Data were organized into two DPOAE (*Strength* and *Structure*) and five brainstem (*Onset*, *Spectrotemporal*, *Harmonics*, *Envelope Boundary*, and *Pitch*) composite measures.
*Results:* DPOAE *Strength* shows significant relationships with sABR *Spectrotemporal* and *Harmonics* measures. DPOAE *Structure* shows significant relationships with sABR *Envelope Boundary*. Neither DPOAE *Strength* nor *Structure* is related to sABR *Pitch*.
*Conclusions:* The results of the present study show that certain aspects of the speech-evoked auditory brainstem responses are related to, or covary with, cochlear function as measured by distortion product otoacoustic emissions.
*Significance:* These results form a foundation for future work in clinical populations. Analyzing cochlear and brainstem function in parallel in different clinical populations will provide a more sensitive clinical battery for identifying the locus of different disorders (e.g., language based learning impairments, hearing impairment).

## 1. Introduction

Distortion product otoacoustic emissions (DPOAEs) and speech-evoked auditory brainstem responses (sABRs) are objective measures of peripheral auditory physiology. These tools are used in both clinical and research applications, often in tandem for differential diagnosis. It is important, therefore, to understand issues related to their overlap and independence. The current study is our maiden attempt to explore these relationships in normal hearing young adults. The underlying objective is to examine and document the links between cochlear and brainstem function and ultimately improve the clinical power of these instruments by using them together for specific clinical purposes.

### 1.1. Otoacoustic emissions

Otoacoustic emissions (OAEs) are signals generated in the cochlea that are detectable in the ear canal (Kemp, 1978, 1979). These acoustic signals are considered a byproduct of physiological processes necessary for normal hearing, specifically outer hair cell function (Brownell, 1982). Otoacoustic emissions can be generated spontaneously (SOAEs) and can also be evoked by clicks (transient-evoked otoacoustic emissions, TEOAEs), single tones (stimulus-frequency otoacoustic emissions, SFOAEs), or tone pairs (distortion product otoacoustic emissions, DPOAEs).

Emissions evoked by tone pairs, or DPOAEs, are equally popular in the clinic and the laboratory. They are measured by stimulating the cochlea simultaneously with two pure tones ($f_1$ and $f_2$, $f_1 < f_2$). Distortion products at various frequencies arithmetically related to the stimulus frequencies are generated in the cochlea. The DPOAE at $2f_1 - f_2$ is the most robust in human ears under certain stimulus conditions and is used routinely in clinical practice.

---

* Corresponding author. Address: Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Northwestern University, 2240 Campus Drive, Evanston, IL 60208, USA. Tel.: +1 847 491 2470; fax: +1 847 491 2523.
*E-mail address:* s-dhar@northwestern.edu (S. Dhar).

1388-2457/$36.00 © 2009 International Federation of Clinical Neurophysiology. Published by Elsevier Ireland Ltd. All rights reserved.
doi:10.1016/j.clinph.2009.02.172

DPOAEs such as the one at $2f_1 - f_2$ are lower in frequency than the stimulus tones making their characteristic frequency (CF) place apical to $F_1$ and $F_2$ on the basilar membrane. There is now irrefutable evidence that for apical DPOAE that, the signal measured in the ear canal is a mixture of two components, one from the overlap region between the traveling wave patterns of the stimulus tones and the other from the CF region of the DPOAE itself (Mauermann et al., 1999a; Talmadge et al., 1999). In many current theories of OAE generation, these two DPOAE components are modeled to arise from fundamentally different mechanisms resulting in significantly different phase-frequency functions of each component (Shera and Guinan, 1999). The phase of the overlap component (also called wave-fixed or distortion component in the literature) is relatively invariant with frequency. On the other hand, the phase of the DP CF component (also called the place-fixed or the reflection component in the literature) varies rapidly with frequency.

As these two components with different phase gradients are mixed in the ear canal, the interaction between them causes a pattern of alternating maxima and minima in the level-frequency function known as fine structure (Dhar et al., 2002). The presence of fine structure in a given ear reflects the presence of the two components and their relative magnitudes determine the depth of fine structure. Two equal components would lead to the deepest fine structure while complete domination by either component would result in little or no fine structure. There is initial evidence that fine structure characteristics could be a more sensitive indicator of alterations in cochlear status than the currently-used metric of overall level of DPOAEs (Wagner et al., 2008). Thus, the origin of fine structure in basic mechanical properties of the cochlea makes it interesting to examine its relationship with other physiological phenomena in the auditory system. Here we report such an exploration of the relationship between distortion product fine structure and speech-evoked brainstem responses.

### 1.2. Speech-evoked auditory brainstem response

The auditory brainstem, a conglomerate of nuclei belonging to the efferent and afferent auditory systems, receives and processes the output of the cochlea en route to the higher centers of auditory processing. The function of the brainstem nuclei can be assessed using stimulus-evoked electrophysiology. Evoked brainstem responses, often using click stimuli, can be diagnostic of clinical populations because of their temporal precision. When evoked by a periodic stimulus, such as speech or music, a frequency-following response (FFR) results. The FFR is driven by neural phase locking and reflects the fundamental periodicity of the stimulus and its harmonics. It is likely generated in the inferior colliculus and lateral lemniscus (Hoormann et al., 1992; Marsh et al., 1970; Moushegian et al., 1973; Smith et al., 1975; Worden and Marsh, 1968). There is also a growing body of literature showing that the human brainstem response is malleable with lifelong linguistic (Krishnan et al., 2005; Swaminathan et al., 2008) and musical experience (Kraus et al., in press; Musacchia et al., 2007; Strait et al., 2009; Wong et al., 2007), as well as short-term auditory training (Russo et al., 2005; Song et al., 2008).

The speech-evoked brainstem response to a consonant–vowel syllable, such as the voiced syllable [da] used in this study, contains both an onset, similar to the click-evoked response, due to the initial noise burst marking the onset of the consonant, and an FFR corresponding to the periodic, voiced formant transition. In the response, the acoustic properties of the stimulus are represented by discrete response peaks representing both transient events in the stimulus, such as voicing onset, and a sustained FFR to the fundamental periodicity (i.e., glottal pulsing) of the vowel.

Latency delays in transient sABR peaks have been found in children with language impairments relative to normal learning children (Banai et al., 2005, 2009; Banai and Kraus, 2008; Cunningham et al., 2001; Johnson et al., 2005; King et al., 2002; Wible et al., 2004) and in this population, peak latencies are particularly affected by the stimulus presentation rate and background noise (Basu et al., in press; Wible et al., 2004). The FFR peaks track the fundamental frequency ($F_0$) of the stimulus; yet, the raw peak latencies are also likely modulated by the high frequency content of the stimulus (Johnson et al., 2008; Hornickel et al., 2009b) which is important for determining phonemic identity. The latencies of FFR peaks have been shown to differ depending on ear of stimulation (Hornickel et al., 2009a) and their timing is related to reading ability (Banai et al., 2005, 2009).

Frequency-domain analyses of the sABR reveal energy at the fundamental frequency and harmonics of the voiced syllable. Measures of harmonics have been found to differ for right and left ear presentation (Hornickel et al., 2009a), between reading impaired and normal learning children (Cunningham et al., 2001; Johnson et al., 2005; Wible et al., 2004; Banai et al., 2009), and to be significantly correlated with measures of reading ability (Banai et al., 2009). These effects are likely due to the importance of harmonics in determining and distinguishing speech sounds. In contrast, measures of $F_0$ representation have not been found to be significantly related to reading (Banai et al., 2009), nor does its encoding differ between reading impaired and normal learning children (Johnson et al., 2005; Kraus and Nicol, 2005; Wible et al., 2004; Banai et al., 2009) or ear of stimulation (Hornickel et al., 2009a).

While there is a vast literature on DPOAEs and brainstem responses, little is known about the relationship between these measures despite their widespread use in the assessment of hearing. Only a few studies have examined both measures in the same subjects (Cone-Wesson et al., 2000; Elsisy and Krishnan, 2008; Oswald et al., 2006; Purcell et al., 2006) and even fewer have related the function of both despite their common relationship to clinical populations and efferent control (de Boer and Thornton, 2008; Hall, 1992; Russo et al., 2005; Song et al., 2008). The current study proposes to identify relationships between DPOAEs and speech-evoked brainstem responses recorded in normal hearing young adults. Deeper understanding of the aspects that do and do not overlap between the two responses will allow for a more detailed knowledge of hearing function and better inform clinical practice.

## 2. Methods

### 2.1. Participants

Participants were 28 adults (ages 19–30, mean = 25; 17 women) who were right handed. All participants had normal (less than 20 dB HL) audiometric thresholds for octaves from 250 to 8000 Hz, no conductive loss as evidenced by a lack of an air-bone threshold gap, and normal click-evoked brainstem response, as measured by wave V latency. All OAE and sABR results reported are from the right ear.

### 2.2. Procedure

In order to encourage participants to remain as still as possible, they were allowed to watch a movie of their choice during data collection. In the case of the sABR, the movie soundtrack was played at ~40 dB SPL in soundfield, while only subtitles were presented for the DPOAE recording. The sABR and DPOAE test sessions occurred within 4 months of each other and often on the same day. All participants were monetarily compensated for their time. The Internal Review Board of Northwestern University approved all procedures.

### 2.3. DPOAE recording and processing

Level and phase estimates of DPOAEs were obtained at closely spaced frequencies (2-Hz apart) using stimulus tones ($f_1$ and $f_2$, $f_2 > f_1$, $f_2/f_1 = 1.22$) swept in frequency over a $(2f_1 - f_2)$ range of 500–10,200 Hz. The stimulus tones, presented at 65 and 55 dB SPL, respectively, were swept in frequency at the rate of 8 s per octave below 6 kHz and 24 s per octave above 6 kHz, while keeping the frequency ratio between them constant. Signal generation and recording were controlled by an Apple Macintosh computer and custom software via a MOTU 828MkII Firewire audio interface. Signals generated by the MOTU were passed through a Behringer headphone amplifier and delivered to the ear canal from MBQuartz speakers coupled through an Etymotic ER10B probe. The ER10B microphone was used to pick up the signal in the ear canal which was amplified by the ER10B pre-amplifier and stored on disk for analysis. Digitization and recording used a sampling rate of 44,100 Hz with a 24 bit converter. The level and phase of the resultant DPOAE at $2f_1 - f_2$ were estimated using a least-squares-fit algorithm. The raw data were screened to have a minimal signal-to-noise ratio of 6 dB before analysis. Data points where the noise floor was above 0 dB SPL were also rejected. The preliminary output of this analysis was fitted with a smoothing function before maxima and minima were identified using change in slope direction. A threshold of −10 dB SPL was set following careful inspection of the data and DPOAE levels were normalized to this threshold, yielding positive values for average DPOAE level. Two composite measures were then created, *Strength* and *Structure* (Fig. 1). These new measures were created specifically to give us a condensed metric to use in comparison with measures of brainstem function and to the best of our knowledge have not been used previously in the literature.

### 2.4. Strength

The measure *Strength* used here is the product of the overall level (in dB) normalized to a threshold of −10 dB SPL over the 1–3 kHz frequency range[1] of the DPOAE and the average fine structure depth (in dB) over the same range. Thus ears with greatest overall DPOAE level and distinct fine structure had largest values of *Strength*. Lack of fine structure could be offset by large overall DPOAE level in a given ear. Similarly, small overall DPOAE level could be offset by exceptionally deep fine structure. Because of the normalization procedure described above, all participants had positive *Strength* scores.

### 2.5. Structure

The second measure of DPOAE used here is termed *Structure*. It is calculated by dividing the number of fine structure periods identified in the 1–3 kHz range by the average spacing over that range. Spacing was computed by dividing the absolute spacing in Hz between adjacent fine structure periods by the center-frequency. Because the phase of the DPOAE component from the overlap region of the stimulus tones does not vary substantially with frequency (Dhar and Shaffer, 2004; Mauermann et al., 1999a,b), the spacing of fine structure periods in a given ear is dependent on the slope of the phase of the DP CF component. Specifically, ears with the steepest slope of the phase of the DP CF component have the most closely spaced fine structure periods, and therefore the highest *Structure* values.

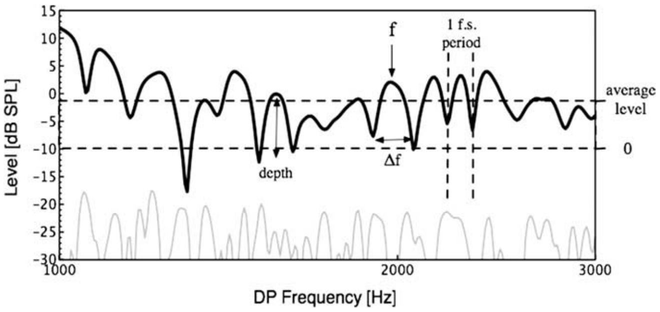

**Fig. 1.** A typical DPOAE response from a normal-hearing adult subject, with all fine structure variables defined. The black line indicates the ear canal response, and the gray line shows the level of the noise floor. Depth is defined by the difference in level between a maximum and the geometric mean of its surrounding minima. Fine structure spacing is characterized by the ratio between $f$, the center frequency, and $\Delta f$, the distance in Hz between adjacent minima. Average DPOAE level is normalized to a −10 dB SPL threshold, yielding positive values for all data points.

### 2.6. Brainstem response recording and processing

The voiced [da] stimulus is a 40 ms synthesized speech syllable, containing a release burst and voiced formant transition (Fig. 2, top left). It was produced in KLATT (Klatt, 1980) with a fundamental frequency ($F_0$) that linearly rises from 103 to 125 Hz with voicing beginning at 5 ms and an onset release burst during the first 10 ms. While the utterance is short, and there is no steady-state vowel, the [da] is perceived as a consonant–vowel syllable.

Stimuli were presented through an insert earphone (ER-3; Etymotic Research, Elk Grove Village, IL) at 80.3 dB SPL using alternating polarities at 10.9 Hz. Responses were recorded with a vertical montage of three Ag–AgCl electrodes (central vertex (Cz), forehead ground, and ipsilateral earlobe reference). Responses were recorded with the Bio-logic Navigator Pro system, BioMARK module (Bio-logic Systems Corp., a Natus Company, Mundelein, IL). Responses were online bandpass filtered from 100 to 2000 Hz (12 dB/octave) and digitally sampled at 6857 Hz. Six thousand artifact-free trials were collected in two blocks and averaged using a 74.67 ms time window (−15.8 ms pre-stimulus). Trials with artifact exceeding ±23 μV were excluded from the average.

Data analyses followed published reports using a similar stimulus and recording parameters. The characteristic seven peaks of the response (V, A, C, D, E, F, O, Fig. 2, left, bottom) were manually identified by the experimenter, and confirmed by a second rater. See Russo et al. (2004) and Johnson et al. (2005) for an in depth review of these peaks. Peaks V, A, D, E and O were 100% detectable in all subjects. Peak C was not detectable in one subject, and peak F was likewise absent in another subject. Spectral analyses using fast Fourier transforms over the time range 21.9–40.6 ms (Fig. 2, right), encompassing D, E and F, were computed using routines coded in Matlab 7 (The MathWorks, Inc., Natick, MA). Five composite measures of the sABR to [da] are used in this paper. In order to create them, $Z$ scores of each constituent measure were calculated ([sample-mean]/standard deviation) and the composite was an average of the constituent $Z$ scores, i.e., $(Z_{\text{measure}1} + Z_{\text{measure}2} + \cdots + Z_{\text{measure}N})/N$.

### 2.7. Onset

A composite was created from the latencies of the two onset peaks, V and A, which mark the onset of sound, and are comparable to the click-evoked peak V and Vn.

### 2.8. Spectrotemporal

This composite was created from the latencies of peaks D, E, and F, which arise in response to the fundamental periodicity of the

---

[1] In preliminary analyses DPOAE data between 1 and 3 kHz, which encompasses the sABR [da]'s second and third formants, demonstrated the strongest relationship with the brainstem measures. All results reported in this manuscript use DPOAE data between the frequencies of 1 and 3 kHz.

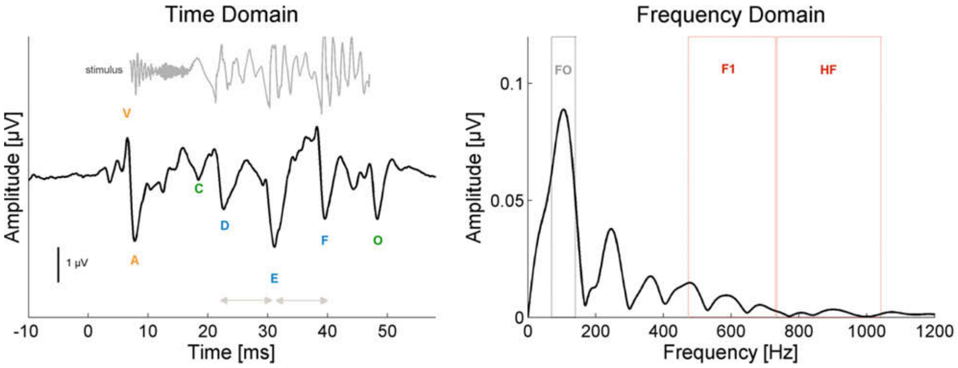

**Fig. 2.** Schematic of the brainstem response to the speech syllable [da] in the time (left) and frequency (right) domains. The time domain waveform of the stimulus is plotted in gray above sABR waveform. The stimulus is time-shifted 8 ms in order to facilitate its visual comparison with the response and to account for the neural conduction lag. *Onset* measures (V, A latencies) are labeled in orange. *Spectrotemporal* (D, E, F latencies) elements are labeled in blue. *Envelope Boundary* measures (C, O latencies) are labeled in green. *Pitch* measures ($F_0$ amplitude, E–D, F–E inter-peak latencies) are labeled in gray and *Harmonic* measures ($F_1$, HF amplitudes) are labeled in red. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this paper.)

stimulus (glottal pulsing in the case of speech), but are also affected by the higher harmonic information in the speech signal.

### 2.9. Envelope Boundary

The latencies for peaks C and O were combined to form the *Envelope Boundary* composite. Peak C marks the start of voicing while peak O is thought to signal the end of voicing, and the two bookend the periodic portion of the response, thus forming the boundary of the response to the stimulus envelope. While the term *envelope* may be used to describe the response over a range of temporal scales, in the present study it refers to that associated with the voiced portion of the stimulus.

### 2.10. Pitch

Average spectral amplitude was calculated for a range encompassing the fundamental frequency ($F_0$), 103–120 Hz. *Pitch* is a composite of the amplitude at $F_0$ and the interpeak intervals between peaks D and E, and E and F. These interpeak intervals mimic the fundamental periodicity of the stimulus and are suggestive of $F_0$ encoding. While other aspects of speech are certainly important for the perception of pitch, we focus here on the fundamental frequency which has major contributions to the percept (Cruttenden, 1997).

### 2.11. Harmonics

The *Harmonics* measure is a composite of the average spectral energy from two frequency bands: first formant ($F_1$) 455–720 Hz, and high frequency (HF) 721–1154 Hz. $F_1$ includes the harmonics of the stimulus that make up the most prominent frequencies of the first formant range. The HF range is composed of harmonics between the first and second formants ($F_1$ and $F_2$, respectively). Because $F_2$ and higher formants are above the phaselocking limits of the brainstem, no higher frequency ranges were included. See Table 1 for the means of all individual measures.

### 2.12. Statistical analyses

In order to determine relationships between the OAE measures of *Strength* and *Structure* and the five sABR composites, while also taking into account relatedness among the brainstem measures, two multivariate regressions were run in SPSS (SPSS Inc., Chicago, IL). While the ratio of cases (28) to independent variables (5 composite sABR measures) was small, the measures were normally distributed, did not show evidence of collinearity, and inspection of the residual plots indicated the data met the assumptions of normality, linearity, and homoscedasticity (Tabachnick and Fidell, 2007). sABR composites were added to the model using the Enter method and then removed from the model using a backward Stepwise method if they did not significantly contribute to prediction of the variance in the dependent measure ($p < 0.1$). Additionally, Pearson's correlations among the composites and the constituent measures were conducted in SPSS for additional support of multiple regression results. If we were to correct for multiple comparisons by adopting an $\alpha$ level of 0.01, the majority of conducted correlations would not be significant, however, the pattern of significant results with an $\alpha$ of 0.05 indicates that the brainstem and OAE measures are indeed moderately to strongly related in absolute terms.

**Table 1**
Composite variables and the means of their component measures.

| Composite | Measure | Mean (standard deviation) |
|---|---|---|
| Harmonics | F1 | 0.732 (0.34) µV |
| | HF | 0.343 (0.14) µV |
| Spectrotemporal | D | 22.915 (0.80) ms |
| | E | 31.068 (0.62) ms |
| | F | 39.520 (0.42) ms |
| Onset | V | 6.718 (0.26) ms |
| | A | 7.644 (0.37) ms |
| Envelope | C | 18.525 (0.64) ms |
| | O | 48.392 (0.52) ms |
| Pitch | D–E | 8.153 (0.88) ms |
| | E–F | 8.467 (0.57) ms |
| | F0 | 5.605 (1.67) µV |
| Strength | Normalized amplitude | 12.294 (5.03) dB |
| | Depth | 5.520 (0.98) dB |
| Structure | Fine structure periods | 12.607 (7.20) |
| | Frequency spacing | 10.570 (3.05) Hz |

## 3. Results

### 3.1. Relationships between DPOAE Strength and sABR measures

The strongest model predicting variance in DPOAE *Strength* comprised the *Spectrotemporal* and *Harmonics* composite measures only ($R = 0.611$, adjusted $R^2 = 0.323$, $F(2,27) = 7.432$, $p < 0.01$). No other brainstem measures significantly contributed unique variance to the predictive model (see Fig. 3, which schematically

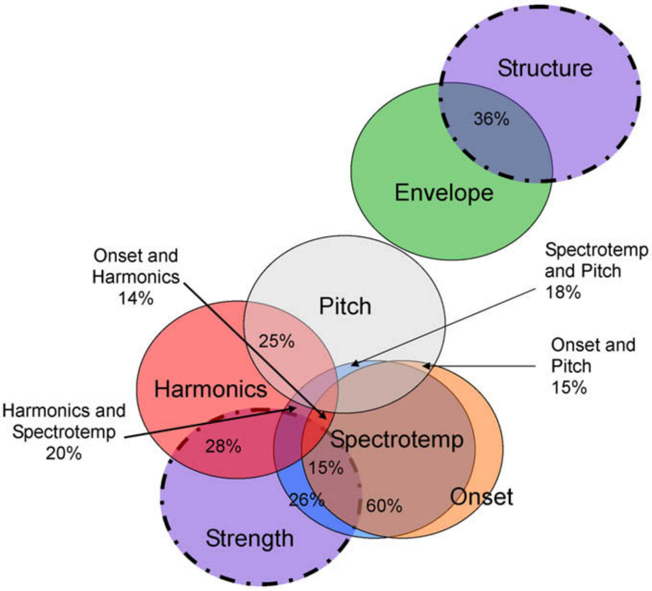

**Fig. 3.** Venn diagram of the relationships between composite sABR and DPOAE measures. sABR circles are solid; DPOAE circles are dashed. The percentage overlap between the variables ($R^2$; bivariate relationships) is marked. While all measures are in some minor way related, and all circles in the figure should be touching, any overlap not represented was less than 0.5%. *Spectrotemporal* was shortened to *Spectrotemp* for space constraint reasons. Note that only a weak relationship was observed between the two DPOAE composite measures, Strength and Structure ($R^2 = 0.0025$).

shows overlap among all the measures and indicates $R^2$ values for each pairing). The *Spectrotemporal* and *Harmonics* composite sABR measures were weighted similarly in the model with standardized $\beta$ coefficients of −0.341 and 0.381, respectively. The negative standardized coefficient for *Spectrotemporal* indicates that individuals with greater DPOAE *Strength* also demonstrated earlier sABR peak latencies. *Harmonics*' positive standardized coefficient suggests that greater DPOAE *Strength* was related to increased spectral amplitude at the $F_1$ and HF frequency ranges in the sABR. See Fig. 4 for the plot of *Strength* against *Harmonics*, its best predictor.

### 3.2. Relationships between DPOAE Structure and sABR measures

In the predictive model of DPOAE *Structure*, only *Envelope Boundary* was a significant predictor ($R = -0.587$, adjusted $R^2 = 0.319$, $F(1,27) = 13.662$, $p = 0.001$, Fig. 3). The *Envelope Boundary* composite standardized $\beta$ coefficient was −0.587, suggesting increases in *Structure* were related to decreases in the latencies of the *Envelope Boundary*. See Fig. 4 for the plot of *Structure* against *Envelope Boundary*.

### 3.3. sABR Pitch and Onset are not unique predictors of either DPOAE measure

The brainstem *Onset* and *Pitch* composites did not significantly contribute to either model. *Onset* was found to be related to *Strength* ($r = 0.391$, $p < 0.05$), but did not contribute significantly to the predictive model. The inclusion of *Onset* in the *Strength* model was predicted based on previously established relationships between brainstem response latency and factors affecting stimulus integrity, such as hearing impairment and stimulus level (Hall, 1992). The overlap in variance between *Onset* and *Spectrotemporal* composites was found to be quite large and this great overlap likely forced the exclusion of sABR *Onset* from the DPOAE *Strength* model because the measure did not predict any unique variance beyond that predicted by *Spectrotemporal* (see Fig. 3). Similarly, *Pitch* was somewhat related to *Harmonics*, *Onset*, and *Spectrotemporal* (see Fig. 3), but was not significantly correlated with *Strength* or *Structure* ($r = 0.199$, $p$ = n.s.; $r = -0.234$, $p$ = n.s., respectively). These results indicate that encoding of the fundamental frequency is not captured by the DPOAE measures utilized in this study.

## 4. Discussion

The results of the current study show clear and significant relationships between speech-evoked auditory brainstem responses and cochlear function assessed via distortion product otoacoustic emissions. In exploring these relationships we have organized the data to represent specific aspects of brainstem and cochlear function. Two DPOAE (*Strength* and *Structure*) and five sABR (*Onset*, *Spectrotemporal*, *Harmonics*, *Envelope Boundary*, *Pitch*) composite measures were created. Relationships with the DPOAE measures were found for sABR measures *Spectrotemporal*, *Harmonics*, and *Envelope Boundary*, but not uniquely for *Onset* nor *Pitch*.

### 4.1. Relationships between DPOAE Strength and sABR measures

DPOAE *Strength* reflects the mechanical activity of the basilar membrane, with a higher *Strength* value arguably signifying a stronger cochlear amplifier. *Strength* was found to be related to

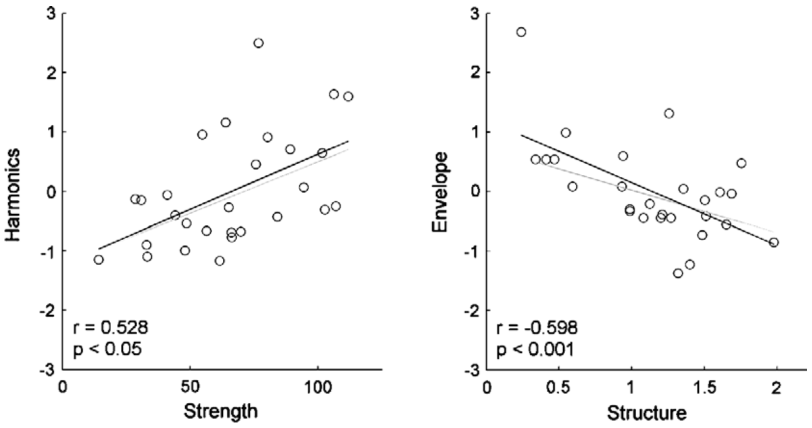

**Fig. 4.** Relationships between DPOAE measures and their predictors. DPOAE *Strength* (left panel) is positively related to its strongest predictor, *Harmonics*, while (right panel) *Structure* is negatively related to its strongest predictor, *Envelope Boundary*. All brainstem composite measures are plotted as $Z$-scores. Black lines (and reported statistics) are regressions using all data points, gray lines are the regressions omitting individuals with $Z > 2$.

the latency of transient and spectrotemporal elements, and the amplitude of the harmonics composite of $F_1$ and HF as well as the individual measures. Relationships between spectral amplitudes and DPOAE *Strength* are greatest for the HF range, followed by the $F_1$ range, and lastly the $F_0$ range, which showed a weak relationship.

The range of harmonics through HF represents spectral content roughly up to, but not including, the second formant of the [da], and is encoded through phase locking. A healthier cochlea, arguably represented by greater DPOAE *Strength*, leads to greater cochlear activity at the formant frequencies possibly resulting in reduced latency of peaks D, E, and F. At lower frequencies, physiological noise increases at the cochlear apex and neural synchrony decreases. This could lead to weaker relationships between strength and sABR spectral amplitude at lower frequency regions. This trend also suggests that as the stimulus frequency increases and approaches the phase-locking limits of the brainstem, the contribution of cochlear mechanics increases in importance, implying that the relationship between DPOAE *Strength* and spectral amplitude would only continue to strengthen for frequencies higher than those analyzed here.

The progression of increasing overlap between *Strength* and $F_0$, $F_1$, and HF may also be due to the increasing similarity in frequency range between the measures. Perhaps if OAE responses at the $F_1$ range and the $F_0$ ranges (i.e., less than 1 kHz) were more reliable (i.e., significantly above the noise floor), the same type of relationship seen for as *Strength* and HF would be exhibited for DPOAE and brainstem measurements from comparable frequency ranges.

Peak latencies of the sABR *Spectrotemporal* composite were also found to be correlated with DPOAE *Strength*. While the peaks that comprised the former occur roughly at the fundamental periodicity of the stimulus, the absolute latencies of the peaks are modulated by spectrotemporal movement of formant transitions between a consonant and vowel (Johnson et al., 2008). If the latency of the peaks is modulated by the robustness of harmonic encoding, as we believe it is, then more defined mechanical activity on the basilar membrance, as suggested by increased *Strength*, would lead to more robust harmonic encoding, and result in greater effects on peak latency.

Relationships between sABR *Onset* (as well as click-evoked peak V) and DPOAE *Strength* are predicted given the previously established clinical relationship of response latency increasing with hearing loss and decreasing with increasing stimulus level (Hall, 1992). Healthier cochleae with more active amplification processes will arguably have greater mechanical activity resulting in greater DPOAE *Strength*. Our results indicate that these ears also tend to demonstrate reduced latencies of the onset-related sABR waves. The inverse relationship between stimulus level and latency as well as the direct relationship between hearing threshold and latency are usually explained on the basis of latency being driven by the most basal activity on the basilar membrane. However, the relationship between DPOAE *Strength* and sABR *Onset* latency may not be as straightforward. A more active cochlear amplifier leads to greater mechanical activity in a more limited area along the basilar membrane as the mechanical response is more sharply tuned (Robles and Ruggero, 2001). Thus earlier onset latencies are either a result of more synchronized neural activity from this limited area of the basilar membrane or both DPOAE *Strength* and sABR latency are driven by a different, but common mechanism.

### 4.2. Relationships between DPOAE Structure and sABR measures

*Structure* reflects both the presence and spacing of DPOAE fine structure. The presence of a DPOAE component from the $2f_1 - f_2$ region of the cochlea results in fine structure. The steeper the phase slope of the DPCF component the closer the peaks of DPOAE fine structure and the greater the *Structure* value in a given ear.

Structure was found to be correlated with brainstem Envelope Boundary measures only. The *Envelope Boundary* composite was composed of peak C and O latencies. DPOAE *Structure* was correlated with the sABR *Envelope Boundary* and the two constituent peak latencies (C and O) individually. Peak C is thought to signal the onset of voicing in the speech stimulus, while peak O corresponds to the cessation of voicing (Kraus and Nicol, 2005). Together these response peaks demarcate the most prominent envelope cues in the stimulus. Envelope cues are important for speech perception (Shannon et al., 1995), and are especially crucial for speech recognition in cochlear implant users. Fine structure spacing in DPOAEs is inversely related to the steepness of the slope of the phase of the DP CF component. A steeper phase/frequency function of stimulus frequency OAEs has been shown to be related to psychophysical tuning measured using forward masking (Shera et al., 2002). However, this finding has proven to be highly controversial and is being actively debated in the literature (Siegel et al., 2005). The relationship between *Structure* and *Envelope Boundary* may be even more complicated to interpret as the neural generators of the low-frequency response characteristics classified as *Envelope Boundary* may be higher in the auditory midbrain or the cortex.

The mutual exclusivity of relationships with *Strength* and *Structure* suggests that they assess different constructs. *Strength* reflects the gain of the cochlear amplifier while *Structure* represents the relative equality of multiple DPOAE components. It appears that the gain of the cochlea is important for brainstem encoding of higher frequency information, crucial for distinguishing phonemes, while the phase of the DPOAE component from the DP CF region is related to the encoding of stimulus envelope by the brainstem.

### 4.3. sABR pitch is not related to either DPOAE measure

Unlike the other brainstem measures, sABR *Pitch* was not found to contribute to the predictive models for the DPOAE measures or correlate with them individually. This is in line with previous work which found pitch to be distinct from brainstem timing and harmonics (Russo et al., 2004; Kraus and Nicol, 2005; Hornickel et al., 2009a), and that only the latter are impaired in reading-impaired children (Banai et al., 2005, 2009; Cunningham et al., 2001; Johnson et al., 2005; King et al., 2002; Wible et al., 2004).

While this dissociation may explain the results of the current study, the insignificant relationships with *Pitch* could also result from difficulties in recording OAEs in the frequency range of the $F_0$. Physiological noise is high at typical speech fundamental frequency ranges and OAE responses tend to be below, or close to, the noise floor. Future work using OAE recording techniques that can overcome the physiological noise in this frequency range may be instructive.

### 4.4. Role of the efferent system

Both OAE and brainstem measures are influenced by the efferent system. In our measurements of DPOAEs, there was no induced activation of the efferent system. However, there is evidence that the stimuli used to evoke OAEs themselves activate the efferent olivocochlear system and cause a change in OAE magnitude and phase (Guinan et al., 2003). These effects are operational in the tens-of-milliseconds time range and are arguably mediated at the level of the brainstem via the superior olivary complex (Guinan, 2006). The general effect is a reduction in DPOAE magnitude in the first few milliseconds after the stimuli are activated (Guinan, 2006). In contrast, the effect of the efferent system on brainstem measures is arguably more long term and mediated by higher centers in the cortico-thalamic pathway, via corticofugal mechanisms (Suga and Ma, 2003). Given the two contrasting time lines of efferent activity, we argue that the relationships between brainstem and OAE measures are not merely manifestations of the same efferent phenomenon affecting the peripheral auditory system.

## 5. Conclusion

The results of the present study show that certain aspects of the speech-evoked auditory brainstem responses to speech (*Harmonics*, *Spectrotemporal*, and *Envelope Boundary*) are related to, or covary with, cochlear function as measured by *Strength* and *Structure* of DPOAEs. As such, these results form a foundation for future work in clinical populations, such as patients with hearing loss, individuals with language-based learning impairments, and patients with speech in noise perception deficits. Brainstem responses in children with language-based learning impairments have delays in timing and reductions in harmonic encoding but normal pitch encoding (Banai et al., 2005, 2009; Basu et al., in press; Cunningham et al., 2001; Johnson et al., 2005; King et al., 2002; Wible et al., 2004) and it is possible that the relationships found in the current study could vary as a function of language (e.g., reading) ability and listening in noise performance. Similarly, alterations in these relationships may be observed in the aging system, as well as other clinical conditions. Analyzing cochlear and brainstem function in parallel in different clinical populations will provide a more sensitive clinical battery for identifying the locus of different disorders. It is also possible that these relationships can be enhanced with proper auditory training (de Boer and Thornton, 2008).

## Acknowledgments

Work reported here was supported by Grants DC008420 and DC01510 from the NIH/NIDCD. We wish to thank the members of the Auditory Research Lab and the Auditory Neuroscience Lab, specifically Judy H. Song for her help with data collection and other aspects of this work. We also thank Professor Steven Zecker for his advice on appropriate statistical treatment of these data.

## References

Banai K, Kraus N. The dynamic brainstem: implications for APD. In: McFarland D, Cacace A, editors. Current controversies in central auditory processing disorder. San Diego: Plural Publishing Inc.; 2008. p. 269–89.

Banai K, Nicol T, Zecker SG, Kraus N. Brainstem timing: implications for cortical processing and literacy. J Neurosci 2005;25:9850–7.

Banai KHJ, Skoe E, Nicol T, Zecker S, Kraus N. Reading and subcortical auditory function. Cereb Cortex, 2009; in press, doi:10.1093/cercor/bhpo24.

Basu MKA, Weber-Fox C. Brainstem correlated of temporal auditory processing in children with specific language impairment. Dev Sci, in press.

Brownell WE. Cochlear transduction: an integrative model and review. Hear Res 1982;6:335–60.

Cone-Wesson B, Vohr BR, Sininger YS, Widen JE, Folsom RC, Gorga MP, et al. Identification of neonatal hearing impairment: infants with hearing loss. Ear Hear 2000;21:488–507.

Cruttenden A. Intonation. Cambridge: Cambridge University Press; 1997.

Cunningham J, Nicol T, Zecker SG, Bradlow A, Kraus N. Neurobiologic responses to speech in noise in children with learning problems: deficits and strategies for improvement. Clin Neurophysiol 2001;112:758–67.

de Boer J, Thornton AR. Neural correlates of perceptual learning in the auditory brainstem: efferent activity predicts and reflects improvement at a speech-in-noise discrimination task. J Neurosci 2008;28:4929–37.

Dhar S, Shaffer LA. Effects of a suppressor tone on distortion product otoacoustic emissions fine structure: why a universal suppressor level is not a practical solution to obtaining single-generator DP-grams? Ear Hear 2004;25:573–85.

Dhar S, Talmadge CL, Long GR, Tubis A. Multiple internal reflections in the cochlea and their effect on DPOAE fine structure. J Acoust Soc Am 2002;112:2882–97.

Elsisy H, Krishnan A. Comparison of the acoustic and neural distortion product at 2f1 − f2 in normal-hearing adults. Int J Audiol 2008;47:431–8.

Guinan Jr JJ. Olivocochlear efferents: anatomy, physiology, function, and the measurement of efferent effects in humans. Ear Hear 2006;27:589–607.

Guinan Jr JJ, Backus BC, Lilaonitkul W, Aharonson V. Medial olivocochlear efferent reflex in humans: otoacoustic emission (OAE) measurement issues and the advantages of stimulus frequency OAEs. J Assoc Res Otolaryngol 2003;l4:521–40.

Hall JW. Handbook of auditory evoked responses. Boston: Allyn & Bacon; 1992. p. xi, 871p.

Hoormann J, Falkenstein M, Hohnsbein J, Blanke L. The human frequency-following response (FFR): normal variability and relation to the click-evoked brainstem response. Hear Res 1992;59:179–88.

Hornickel J, Skoe E, Kraus N. Subcortical laterality of speech processing. Audiol Neurootol 2009a;14:198–207.

Hornickel J, Skoe E, Nicol T, Zecker S, Kraus N. Subcortical differentiation of voiced stop consonants: relationships to reading and speech in noise perception. In: Association for Research in Otolaryngology Midwinter Meeting, Baltimore, MD; 2009b.

Johnson KL, Nicol TG, Kraus N. Brain stem response to speech: a biological marker of auditory processing. Ear Hear 2005;26:424–34.

Johnson KL, Nicol TG, Zecker SG, Bradlow A, Skoe E, Kraus N. Brainstem encoding of voiced consonant–vowel stop syllables. Clin Neurophysiol 2008;119:2623–35.

Kemp DT. Stimulated acoustic emissions from within the human auditory system. J Acoust Soc Am 1978;64:1386–91.

Kemp DT. Evidence of mechanical nonlinearity and frequency selective wave amplification in the cochlea. Arch Otorhinolaryngol 1979;224:37–45.

King C, Warrier CM, Hayes E, Kraus N. Deficits in auditory brainstem pathway encoding of speech sounds in children with learning problems. Neurosci Lett 2002;319:111–5.

Klatt DH. Software for a cascade/parallel formant synthesizer. J Acoust Soc Am 1980;67:971–95.

Kraus N, Nicol T. Brainstem origins for cortical 'what' and 'where' pathways in the auditory system. Trends Neurosci 2005;28:176–81.

Kraus NSE, Parbery-Clark A, Ashley R. Experience-induced malleability in neural encoding of pitch, timbre and timing: implications for language and music. Ann N Y Acad Sci: Neurosciences and Music III, in press.

Krishnan A, Xu Y, Gandour J, Cariani P. Encoding of pitch in the human brainstem is sensitive to language experience. Brain Res Cogn Brain Res 2005;25:161–8.

Marsh JT, Worden FG, Smith JC. Auditory frequency-following response: neural or artifact? Science 1970;169:1222–3.

Mauermann M, Uppenkamp S, van Hengel PW, Kollmeier B. Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. I. Fine structure and higher-order DPOAE as a function of the frequency ratio f2/f1. J Acoust Soc Am 1999a;106:3473–83.

Mauermann M, Uppenkamp S, van Hengel PW, Kollmeier B. Evidence for the distortion product frequency place as a source of distortion product otoacoustic emission (DPOAE) fine structure in humans. II. Fine structure for different shapes of cochlear hearing loss. J Acoust Soc Am 1999b;106:3484–91.

Moushegian G, Rupert AL, Stillman RD. Laboratory note, scalp-recorded early responses in man to frequencies in the speech range. Electroencephalogr Clin Neurophysiol 1973;l35:665–7.

Musacchia G, Sams M, Skoe E, Kraus N. Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. Proc Natl Acad Sci USA 2007;104:15894–8.

Oswald JA, Rosner T, Janssen T. Hybrid measurement of auditory steady-state responses and distortion product otoacoustic emissions using an amplitude-modulated primary tone. J Acoust Soc Am 2006;119:3886–95.

Purcell DW, Van Roon P, John MS, Picton TW. Simultaneous latency estimations for distortion product otoacoustic emissions and envelope following responses. J Acoust Soc Am 2006;119:2869–80.

Robles L, Ruggero MA. Mechanics of the mammalian cochlea. Physiol Rev 2001;81:1305–52.

Russo N, Nicol T, Musacchia G, Kraus N. Brainstem responses to speech syllables. Clin Neurophysiol 2004;115:2021–30.

Russo NM, Nicol TG, Zecker SG, Hayes EA, Kraus N. Auditory training improves neural timing in the human brainstem. Behav Brain Res 2005;156:95–103.

Shannon RV, Zeng FG, Kamath V, Wygonski J, Ekelid M. Speech recognition with primarily temporal cues. Science 1995;270:303–4.

Shera CA, Guinan Jr JJ. Evoked otoacoustic emissions arise by two fundamentally different mechanisms: a taxonomy for mammalian OAEs. J Acoust Soc Am 1999;105:782–98.

Shera CA, Guinan Jr JJ, Oxenham AJ. Revised estimates of human cochlear tuning from otoacoustic and behavioral measurements. Proc Natl Acad Sci USA 2002;99:3318–23.

Siegel JH, Cerka AJ, Recio-Spinoso A, Temchin AN, van Dijk P, Ruggero MA. Delays of stimulus-frequency otoacoustic emissions and cochlear vibrations contradict the theory of coherent reflection filtering. J Acoust Soc Am 2005;118:2434–43.

Smith JC, Marsh JT, Brown WS. Far-field recorded frequency-following responses: evidence for the locus of brainstem sources. Electroencephalogr Clin Neurophysiol 1975;l39:465–72.

Song JH, Skoe E, Wong PC, Kraus N. Plasticity in the adult human auditory brainstem following short-term linguistic training. J Cogn Neurosci 2008.

Strait DSE, Kraus N, Ashley R. Musical experience and neural efficiency: effects of training on subcortical processing of vocal expressions of emotion. Eur J Neurosci 2009;29:661–8.

Suga N, Ma X. Multiparametric corticofugal modulation and plasticity in the auditory system. Nat Rev Neurosci 2003;4:783–94.

Swaminathan J, Krishnan A, Gandour JT. Pitch encoding in speech and nonspeech contexts in the human auditory brainstem. Neuroreport 2008;19:1163–7.

Tabachnick BG, Fidell LS. Using multivariate statistics. Boston: Allyn & Bacon; 2007.

Talmadge CL, Long GR, Tubis A, Dhar S. Experimental confirmation of the two-source interference model for the fine structure of distortion product otoacoustic emissions. J Acoust Soc Am 1999;105:275–92.

Wagner W, Plinkert PK, Vonthein R, Plontke SK. Fine structure of distortion product otoacoustic emissions: its dependence on age and hearing threshold and clinical implications. Eur Arch Otorhinolaryngol 2008;265:1165–72.

Wible B, Nicol T, Kraus N. Atypical brainstem representation of onset and formant structure of speech sounds in children with language-based learning problems. Biol Psychol 2004;67:299–317.

Wong PC, Skoe E, Russo NM, Dees T, Kraus N. Musical experience shapes human brainstem encoding of linguistic pitch patterns. Nat Neurosci 2007;10:420–2.

Worden FG, Marsh JT. Frequency-following (microphonic-like) neural responses evoked by sound. Electroencephalogr Clin Neurophysiol 1968;25:42–52.