

# Listening in on the listening brain

Nina Kraus

feature  
article

The measured response of the auditory brainstem to complex aural stimuli does more than reveal hearing acumen. It also gives insight into how experience molds the perception of sound.

**Nina Kraus** (<http://www.brainvolts.northwestern.edu>) holds the Hugh Knowles Chair at Northwestern University in Evanston, Illinois, where she is a professor of communication sciences, neurobiology and physiology, and otolaryngology.

An article published in the journal *Brain* some 40 years ago forever changed the world of hearing assessment.<sup>1</sup> Don Jewett and John Williston reported that neural firing recorded from the human scalp with electroencephalogram electrodes could determine whether a sound was heard. The measured electrical impulses originated in the midbrain, a part of the auditory brainstem, and auditory brainstem response (ABR) recording entered the clinic as an objective, passive means to determine whether newborn babies could hear. Audiologists no longer needed to wait until children were old enough to raise their hands in response to the beeps and bleeps of an audiometer to determine whether they could hear normally. There soon emerged other uses of ABR—tumor detection, diagnosis of nerve-damaging diseases such as multiple sclerosis, and more.<sup>2</sup>

## From clicks to peaks

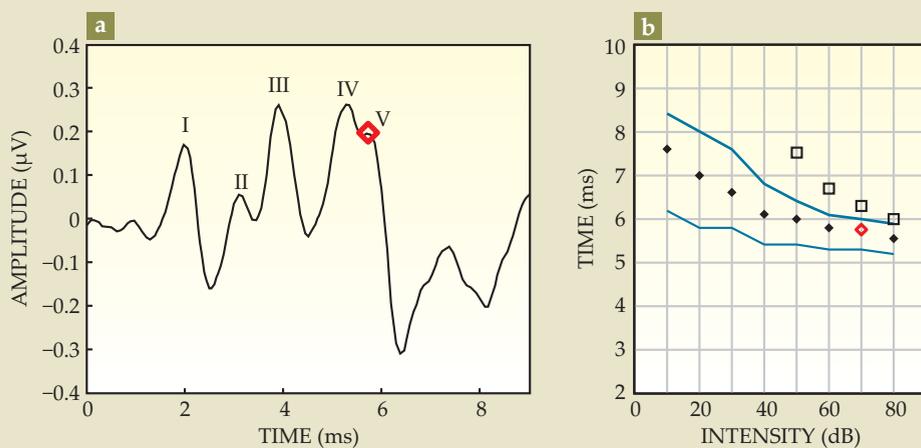
The detection of a sound by the cochlea of the inner ear initiates a volley of neural firing that progresses from inner ear to midbrain to thalamus to primary sensory cortex and beyond. A normal ABR signal indicates a healthy inner ear because if the cochlea fails to react to the sound, then the neural volley will not take off, the brainstem neurons will not fire, and the electrodes have nothing to pick up.

A click, something like a finger snap, had long been the stimulus of choice for an overall probe in ABR hearing assessment. Essentially an impulse with a flat frequency spectrum, the click evokes an ABR with the signature shape shown in

figure 1a. The timing of specific response peaks is related to hearing sensitivity. In a person with normal hearing, a click presented at a reasonably loud level evokes five clear peaks within about 6 milliseconds of the click's initiation. Identified by Roman numerals I through V, those peaks signal neural firing in sequentially higher-level subcortical structures in the auditory pathway.

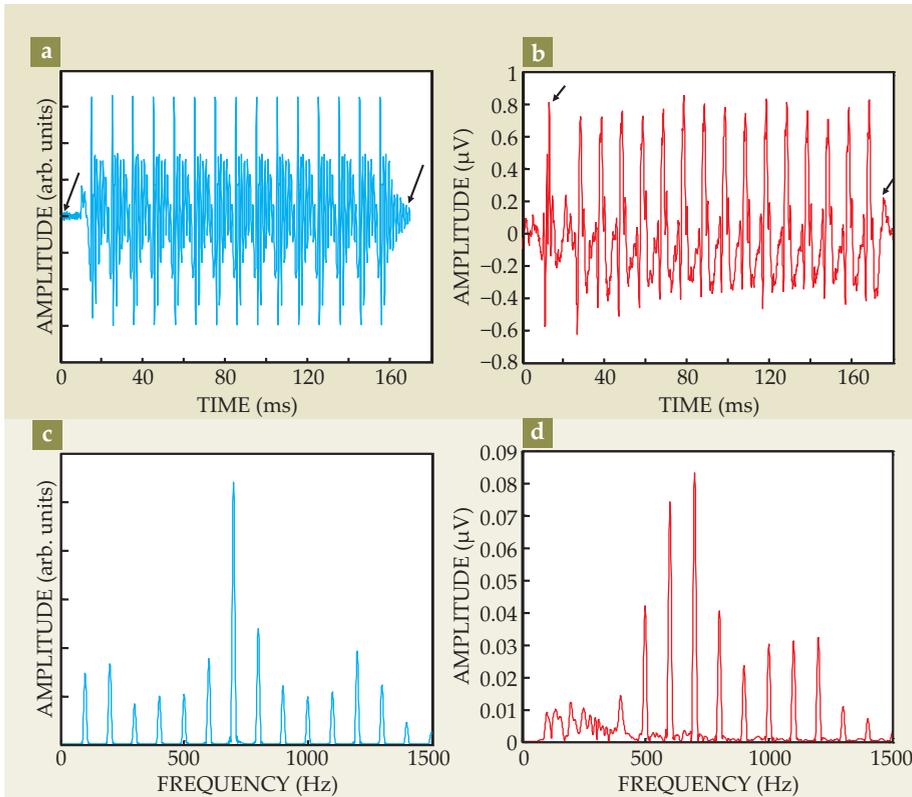
In a thorough ABR audiometric evaluation, a clinician presents clicks in a range of intensities first to one ear, then the other. One useful diagnostic is a plot of the time at which peak V arises versus the intensity of the click. Figure 1b gives simplified plots from a normal-hearing individual and a hearing-impaired one. Peaks that form earlier than V, especially peaks I and III, enable more refined diagnosis. Their presence or absence and the time intervals between peaks I and III, III and V, and I and V help to establish auditory nerve pathology or to distinguish, for example, hearing loss originating in the middle ear from that originating in the inner ear.

An ABR recording electrode cannot be placed directly in the brain. The best one can do is to attach the electrode to the scalp with conducting paste. The response to a single click is thus swamped by the ongoing electrical activity from all parts of the brain, not to mention artifactual muscle activity and the buzz from nearby electrical appliances. A key to brainstem response testing, therefore, is signal averaging. The peaks shown in figure 1a are invariant in that every time a click is detected, the precise voltage pattern of the figure emanates from the brainstem. On the other hand, the larger



**Figure 1. The brainstem** responds to clicks. **(a)** For a normal-hearing person, a square-wave click presented to the right ear at an intensity of 70 decibels (dB) reproducibly evokes a response that includes five clinically useful peaks, usually identified by roman numerals. The red diamond identifies peak V, which occurs 5.75 ms after the sound is presented. **(b)** The simplified plot here shows the time at which peak V arises, as a function of the intensity of a click presented at time zero. The red diamond corresponds to panel a. The blue

lines delimit the time range for normal-hearing individuals. Open squares illustrate a typical time–intensity relationship for someone with hearing loss. That person, whose hearing threshold is 50 dB, evinced no response below the threshold loudness.



**Figure 2.** The complex auditory brainstem response (cABR) resembles the complex stimulus that evokes it. **(a)** A plot of time versus amplitude (in arbitrary units) for the stimulus “ah.” A 10-ms, low-amplitude, aperiodic transient precedes the high-amplitude periodic vowel. Transients arise from events such as the burst of air as it erupts from between pursed lips in a “p” sound or the striking of a hammer in a piano’s soundboard. The left arrow indicates the initiation of the transient, the right arrow the termination of the vowel sound. Other speech sounds have additional timing cues, such as consonant–vowel transitions, bursts, and other aperiodicities that occur on millisecond or even shorter time scales. **(b)** The cABR is dominated by periodic peaks corresponding to the periodic vowel sounding. The response is a mix of low- and high-frequency neural activity reflected, respectively, in the period and shape of the repeated peaks. Also visible in the cABR are peaks, indicated by arrows, that correspond to the vowel’s onset and termination. In

all cases the cABR peaks have a precise timing relation to the evoking stimulus. **(c, d)** Fourier transforms of the post-transient time spectra reveal similar frequency spectra.

voltages arising from nonauditory parts of the nervous system, muscle activity, and electrical noise are random. Thus, on averaging over hundreds of clicks, the background noise destructively sums to zero and a response showing the invariant peaks is all that remains. Moreover, because the duration of the recorded brain activity is so short—for a click ABR, 10 ms or so—one can rapidly repeat the stimulus.

The clean, interpretable response shown in figure 1a is an average over 3000 clicks. With the clicks presented at a rate of 30 per second, the response averaging required less than two minutes. When the presentation rate is further increased or the intensity lowered, peaks I–IV tend to disappear, though peak V and the negative trough following it remain.

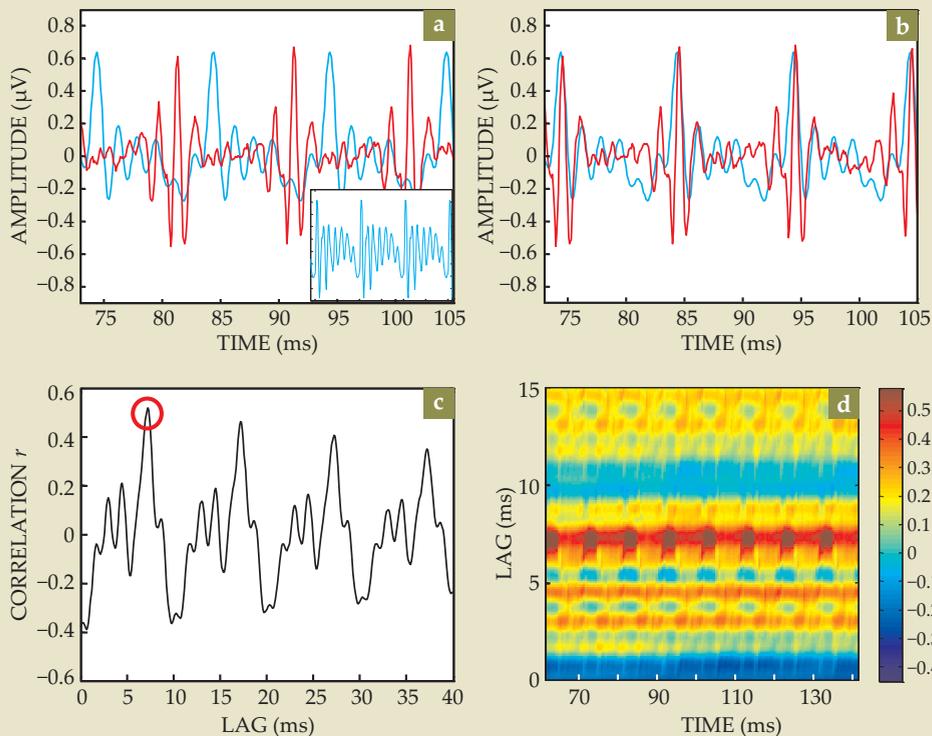
A second ABR stimulus of choice is a short sine-wave tone burst. Whereas clicks serve nicely as an overall probe of hearing, tone bursts provide finer-grained pitch assessment. So, for example, the sine waves may reveal that low-frequency hearing is normal but high-frequency hearing is abnormal. The 21st century has seen a growing interest in complex ABR (cABR) testing—that is, probing the auditory brainstem’s response to complex sounds. Complex ABR provides a wealth of information unobtainable from click- or tone-evoked ABR about sound processing in the auditory pathway—including information about experience with language and music. The reasons for the increasing use of cABR are threefold. First, a reasonably transparent mapping connects the evoking stimulus and the response in cABR. Second, cABR provides information about the efferent auditory system—the downward connections that begin in the cerebral cortex and end in the inner ear. Third, the data can be easily and reliably obtained in individuals.

## Stimulus–response transparency

As figure 1a makes clear, the brainstem response to a click in no way resembles the square wave that stimulated the response. The click is nearly instantaneous—about 0.1 ms long in most ABR systems—but the response evolves over the course of about 7 ms. In contrast, cABRs are elicited by complex stimuli that last several orders of magnitude longer than a click. For most cABR work, stimuli persist for 0.1–0.5 s or so, but some research uses stimuli that continue for several seconds.<sup>3</sup> The response to a periodic stimulus such as a musical note or a speech utterance is essentially the 7-ms-duration response of figure 1a repeated over and over again for the duration of the sound. That repetitive neural response to periodic sounds is known as phase locking; it’s the property that imparts a striking similarity between stimulus and cABR. In fact, if you take a cABR and play it through a loudspeaker, you can often figure out what sound induced the response. (The online version of this article includes three sound files as examples.)

Figures 2a and 2b nicely illustrate the similarity between stimulus and response. The response peaks match up with the stimulus peaks, albeit with a short time delay due to neural conduction and synaptic delays between the cochlea and the auditory brainstem. In the sounding of the vowel sound “ah” that generated the figure, a 10-ms low-amplitude aperiodic transient precedes the high-amplitude periodic vowel. Corresponding peaks in the cABR are precisely associated with both transient and steady-state auditory phenomena. Different stimuli evoke different cABR transients; which peaks are important depend on the stimulus, the population being studied, and the questions that the researcher is asking.

Other ways of looking at the stimulus and response



**Figure 3.** The fidelity of a complex auditory brainstem response (cABR) to its evoking sound and the similarity of one cABR to another, given similar stimuli, permit correlational analyses. **(a)** The inset shows the stimulus (in arbitrary amplitude units) that induced the response (red) in the main plot. Once the stimulus is filtered (blue, main plot) to approximately match the brainstem's attenuation of high frequencies, the formal similarity between stimulus and response is evident. Nonetheless, the response lags, reflecting neural propagation time. **(b)** Shifting the response earlier in time by 6.8 ms highlights the similarity between filtered stimulus and cABR. As described in the text, a standard statistical analysis yields an  $r$  value that describes how well stimulus and response are correlated; for the data plotted here, the correlation is 0.59. **(c)** This

correlogram shows the correlation as a function of time lag. The red circle indicates the 6.8-ms lag that yields the largest  $r$  value. The periodicity in the peaks of the correlation function reflects the periodicity of the stimulus. **(d)** To obtain a finer-grained analysis, investigators can determine  $r$  values by correlating stimulus and response over short time windows of their choosing. In the correlogram shown here, the time window is 20 ms long and is centered at the value given by the horizontal coordinate; colors indicate how  $r$  values fluctuate over time.

reveal transparency as well. Any chosen segment of the stimulus and response may be submitted to a Fourier transformation to obtain frequency spectra. For example, to obtain figures 2c and 2d from figures 2a and 2b, the chosen segment of the “ah” stimulus and cABR is the period from 20 to 170 ms. The two frequency spectra are clearly similar, though the auditory structures involved in the response tend to cut off frequencies at the high end of the cABR.

The cABR research that has proliferated during the past decade has involved speech, nonspeech vocalizations such as a baby’s cry, multitone complexes, music, environmental sounds, and other stimuli. Disparate studies by scientists around the world share two things. First, they employ a variety of digital signal-processing techniques beyond those required for click ABRs. I’ll review a few of those in this article. Second, as I will emphasize below, they have revealed that the brainstem is not just a passive relay station for auditory information. Rather, it is a hub of ascending (ear to brain) sound processing and descending (brain to ear) modulation of the incoming signal.

### Correlation techniques

In large part, cABR work has employed periodic rather than stochastic stimuli. The resulting phase locking lends itself to some familiar digital signal-processing techniques. A simple Fourier transform can reveal information about the spectral energy present in the neural firing. The absolute and relative sizes of the resulting frequency peaks reflect the fidelity with which the sound is being processed. Other methods of processing data yield representations that permit a glimpse of how the spectrum of the response unfolds over time.

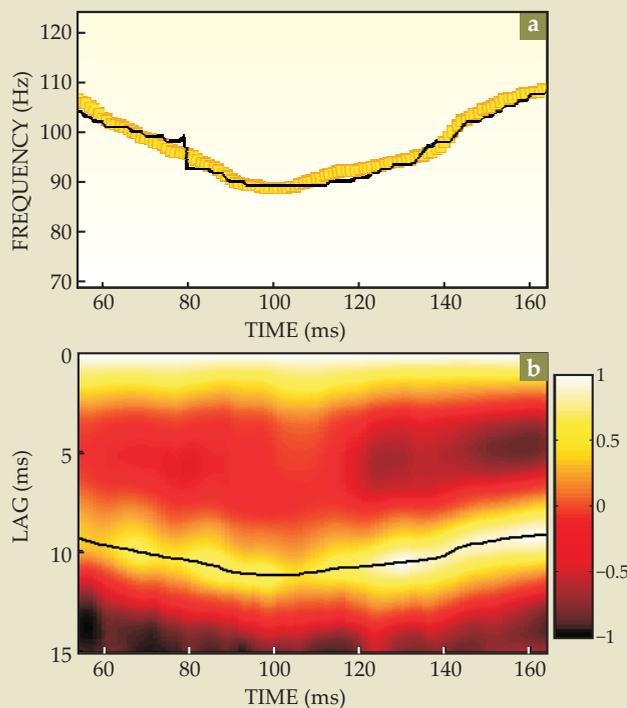
Linear correlation techniques result in measures that

demonstrate impressive similarity between stimulus and response. One routine method of statistically cross-correlating stimulus and response leads to so-called  $r$  values that range from  $r = 1$  for a response that’s an exact copy of the stimulus to  $r = 0$  for a response with absolutely no correspondence to the stimulus.

For two reasons that I have already noted, however, a direct time-domain comparison of stimulus and cABR will not fully expose the similarity of the spectra. First, because the cABR reflects activity in midbrain structures several synapses away from the auditory periphery, high-frequency response tends to be attenuated. Thus the stimulus must be appropriately filtered if one is to achieve a meaningful correlation with the response. Second, finite neural propagation speed means that peaks in the response are delayed in relation to their counterparts in the stimulus. That time lag is evident in figure 3a. Shifting the response data by 6.8 ms, as in figure 3b, results in a correlation ( $r$  value) of nearly 0.6. In general, the lag time that maximizes the correlation depends on stimulus, intensity of presentation, and the individual, but 6.8 ms is a representative value.

The remaining panels of figure 3 give the correlation as a function of time lag, both over the entire period of the recording (panel c) and for 20-ms segments of the recording (panel d). Such cross-correlation techniques may likewise be used to compare one response with another. In that manner, a researcher can assess the timing delay introduced in the response by, for example, a faster stimulus presentation rate, a softer intensity of presentation, or background noise.

The synthesized syllables that evoked the responses in figures 2 and 3 maintain a steady pitch. That monotone quality is in contrast with normal speech, which commonly in-



**Figure 4. The intonation of your voice** as you ask a question is a common example of a pitch sweep that is linguistically important. Intonation also conveys emotion and a sense of how you mean something. Autocorrelations and other techniques amenable to complex auditory brainstem response (cABR) analysis can track pitch changes in speech. **(a)** The black curve shows the pitch trajectory—how the pitch changes over time—of a stimulus similar to a specific tone in Mandarin Chinese. The yellow curve gives the frequency of a cABR as it phase locks to the stimulus’s fundamental frequency. The yellow and black nearly overlap, which demonstrates the accuracy of the brainstem’s pitch tracking. **(b)** In this autocorrelogram, colors denote the correlation of the cABR signal and a lagging version of the same signal. The data represent 40-ms windows, with the original signal centered at the value given by the horizontal coordinate. The bright yellow band near the bottom signals the high correlations that occur when the lag nearly matches the fundamental period. The black line corresponds to the stimulus’s fundamental period. Autocorrelation techniques can also track overtones of a fundamental (harmonics). The accuracy with which cABR tracks a stimulus’s pitch depends on experience and pathologies.

cludes pitch glides. Some of those are incidental, but some—such as English interrogatives or Mandarin Chinese tones—are mandatory. Figure 4a illustrates a syllable with a high-low-high pitch contour, not unlike one of the Mandarin tones, and the tracking of the pitch in the brainstem. Short-time Fourier analysis or short-time autocorrelations such as shown in figure 4b can help expose the neural underpinnings of such pitch tracking.

The assessment of timing details in an ABR can be particularly useful in comparative studies that, for example, compare two individuals’ cABRs with the same sound or the responses of the same individual with two different sounds. Unfortunately, intricate and complex stimuli often evoke cABRs that do not exhibit the easy-to-identify peaks of an ABR click response. Fortunately, there are techniques that are more

sophisticated than visual identification of response peaks. One of those is cross-phase analysis, a technique applicable in particular to cABRs resulting from stop consonants.

As its name suggests, a stop consonant is one formed by a stoppage of air flow through the vocal tract. The interruption is accomplished by, for example, briefly closing the lips for the consonant “b” and by tapping the tongue against the palate for the consonant “d.” Despite the obvious mechanical differences in the production of those two sounds, acoustically they are quite similar. With the help of a speech synthesizer, an investigator can strictly control the acoustic dissimilarities between “bah” and “dah,” reducing them to subtle differences in overtones caused by the resonance properties of the mouth as speech articulators such as the tongue and lips shift from one consonant to another. A listener would readily distinguish the resulting synthetic syllables as “bah” and “dah,” but they actually differ only in their high-frequency content.

In fact, the frequencies at which the differences occur are greater than the maximum frequency for which the brainstem can achieve phase locking. Nonetheless, cABR timing features encode the high-frequency differences; figure 5 shows how. Cross-phase analysis such as that in the figure uncovers timing differences that are both too subtle and too widespread in frequency to manifest themselves as simple and discrete peak-timing differences. The technique may also be applied to investigate the masking of an auditory signal by background noise.

The analysis techniques I have discussed are but a sampling of those currently used. And new techniques are continuing to emerge as cABR research becomes more widespread.<sup>4</sup> But why go through all that trouble if a cABR is simply a fancy hearing test? We do so because a cABR relates to real-life skills such as literacy and the ability to pick out a message in a noisy environment, and it reflects life experiences with language and music.

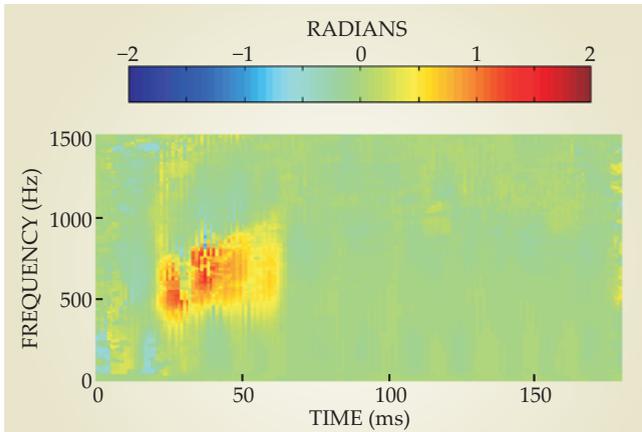
## A cABR depends on experience

The neural routes that connect the sensory organs and the brain run both ways. As figure 6 illustrates, the afferent pathway sends information toward the brain and the efferent pathway sends information toward the sensory organs. Just as the brain “tells” a pianist’s fingers how to move, it exerts an influence all along the auditory pathway.

Until recently, people assumed that passively evoked ABRs reflected one-way processing—that of the afferent, ear-to-brain path. To a first approximation, that assumption is true for a stimulus such as a click. Because of the signature shape of the click-evoked response and the small amount of variability among individuals, a click ABR is a nearly infallible indicator of hearing sensitivity. A response present and at the right timing indicates normal hearing. A response with delayed timing suggests hearing loss. Is there no response at all? That indicates no hearing or perhaps a neuropathological condition.

Afferent processing in the auditory system is only half of the equation. The downward-projecting efferent auditory system has a profound effect. Even activity in the hair cells of the cochlea—the peripheral extreme of the chain—is modulated by higher-level processing.<sup>5</sup> A cABR thus represents a snapshot of both afferent and efferent processing; while still a faithful representation of afferent processing, it is modulated by the total of an individual’s experience with the evoking sound.

What types of experience manifest themselves in cABR? Language background, for one. A classic example concerns the pitch tracking of Mandarin syllables. In Mandarin, unlike



**Figure 5.** The sounds “bah” and “dah” are acoustically quite similar, differing only in their high-frequency components. Even though the auditory brainstem cannot phase lock to such high frequencies, its subtly different timing response to the two syllables can be observed, as revealed in the figure. Most of the field is green, corresponding to a common time response. That coherence is to be expected because for most of their duration the two syllables share the acoustically identical vowel “ah.” However, the two stimuli are different from 20 to 60 ms; in that portion of the complex auditory brainstem response, warm colors signify that the response to “dah” is slightly earlier than the response to “bah.” For any given frequency  $f$ , 1 radian corresponds to a timing difference of  $1/2\pi f$ .

English and other Western languages, the tone of voice helps determine the meaning of a word. A number of studies, notably by A. Ravi Krishnan’s group at Purdue University, have demonstrated that native Mandarin speakers track the pitch changes in Mandarin syllables more accurately than people with no experience with Mandarin or other tonal languages.<sup>6</sup> Krishnan and Purdue colleague Jackson Gandour argue that experience fine-tunes brainstem structures via an efferent sharpening that originates in the cortex. Their idea is in line with the “reverse hierarchy” principle that higher-level cortical structures can sharpen lower-level structures based on a determination of biological relevance. In general, the scientific community is moving away from hierarchical, domain-specific understanding of speech processing toward a picture of an interactive processing system that merges lower and higher structures.<sup>7</sup>

A good deal of research demonstrates that musical experience positively affects broader skills, including those related to motor function, verbal facility, attention, and memory.<sup>8</sup> I am particularly fascinated by how cABR can be used to examine the effects of musical experience on brainstem processing. My colleagues and I at Northwestern University, among others, have discovered that highly trained musicians display enhancements in their cABRs not only to musical sounds but also to speech sounds and nonspeech vocalizations.<sup>9</sup> That result is not wholly unexpected, because scientists have long known that musical experience rewires the auditory system<sup>10</sup> and, as discussed above, the cortex influences the brainstem. The enhancement of a musician’s cABR to speech sounds, though, is intriguing.

Taken together, cABR studies on musicians provide evidence of basic brain rewiring as a result of music training and also furnish an objective means to track music’s effect on the sensory and cognitive systems. One should not get the

impression, however, that every aspect of a cABR is enhanced in individuals with music or language experience. Rather, only certain sounds induce improved response in experienced listeners, and the enhancements may be evident only in particular features of the cABR—for example, timing but not pitch tracking might be affected.

## Applications

Features of a cABR can indicate proficiency at real-life skills, including literacy and the ability to listen to speech in a noisy background. To follow running speech, particularly in a noisy environment, the listener needs to organize relevant sounds—a companion’s voice—into a coherent object or stream while ignoring the rest. That task is accomplished in part with timing and pitch cues. In addition to subjective reporting, various objective tests can assess an individual’s ability to comprehend speech in a noisy environment. My colleagues and I have undertaken several studies in which we’ve analyzed cABRs with respect to their timing and pitch representation. The quality of those objective representations tracks well with subjective and other objective measures of ability to follow speech amidst noise. The consistency of cABRs with other measures cuts across different ages and even occurs in hearing-impaired individuals.<sup>11</sup>

It’s not difficult to imagine a correlation between the ability to follow a conversation and brainstem processing. But it may surprise you to learn that literacy, as measured by standardized paper-and-pencil evaluations, also has a relationship with cABR—one that is especially striking in children. In particular, response timings are faster, and high-frequency components of the response spectrum are stronger, in better readers.<sup>12</sup> My coworkers and I postulate that in some cases poor reading is due to an inadequate efferent sharpening of the lower sensory pathways and that the deficiency, discernable in a cABR, prohibits the poor reader from establishing the sound-to-meaning relationships required for efficient reading.

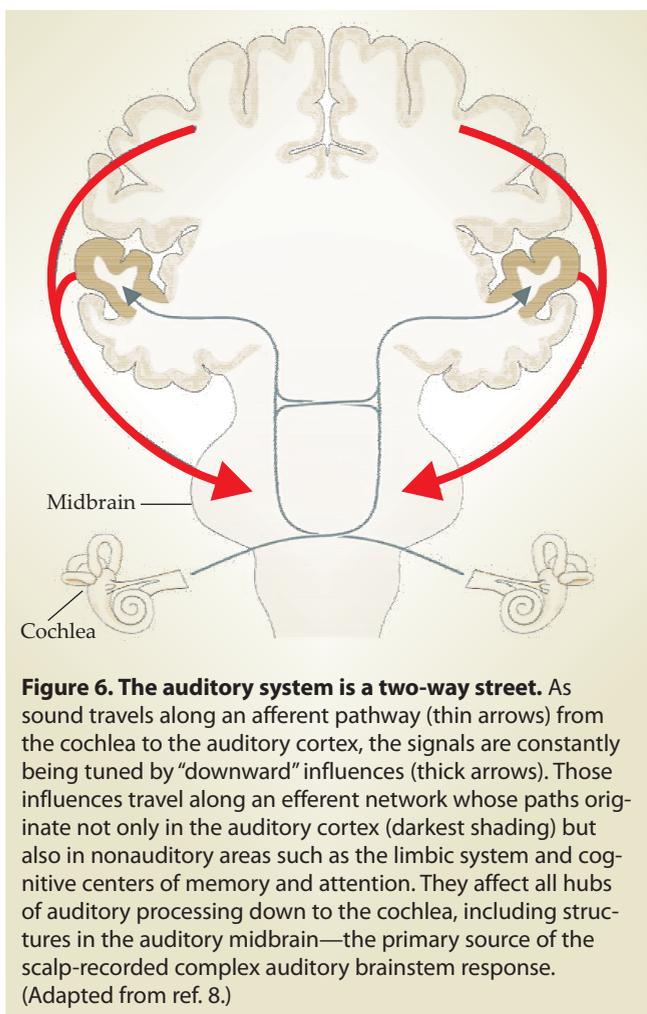
As individuals learn to listen better in noise or as they improve their reading skills, their cABRs will reflect those changes. Several studies have shown that weeks-long training can affect cABR.<sup>13</sup> In some cases even a single session can lead to measurable change, in accord with the nervous system’s known sensitivity to patterns in spoken language.<sup>14</sup>

In sum, the auditory brainstem is far from being an inert relay station between cochlea and cerebral cortex. Rather, it incorporates a rich conjunction of ascending and descending neural processes, and cABRs are able to tap into the wealth of information that is found within. Since cABRs reflect not only expertise and experience but also deficiencies in speech perception and reading, clinical applications should be able both to assess auditory function and to track the neural changes that accompany exposure or training.

## The future of cABR technology

Although interest in cABR is increasing, at present the technique is confined to a few labs that have adapted suboptimal equipment and developed from scratch their own analysis routines. It is my hope that as its utility becomes more widely known, cABR recording and analysis technology will enter the marketplace in a user-friendly form—much as click ABR itself did decades ago.

In addition to the applications emphasized in this article, in particular studying the influences of neuroeducational outcomes on auditory function, cABR may prove useful for investigating the auditory effects of such factors as nutrition, exercise, and hormones. As a gauge of biological processes, cABR might assist in the development and fine-tuning of



such devices as cochlear implants, hearing aids and other hearing devices, microphones, amplifiers, and speakers. Nor is the application of cABR limited to humans; it can provide a probe into auditory physiology in animals. Once momentum is established in the lab, it should carry cABR into the clinic and, one might hope, into schools.

## References

1. D. L. Jewett, J. S. Williston, *Brain* **94**, 681 (1971).
2. T. W. Picton, *Human Auditory Evoked Potentials*, Plural, San Diego, CA (2010).
3. G. C. Galbraith et al., *NeuroReport* **15**, 2057 (2004).
4. E. Skoe, N. Kraus, *Ear Hear.* **31**, 302 (2010).
5. X. Perrot et al., *Cereb. Cortex* **16**, 941 (2006).
6. For a review, see A. Krishnan, J. Gandour, *Brain Lang.* **110**, 135 (2009).
7. M. Ahissar, S. Hochstein, *Trends Cogn. Sci.* **8**, 457 (2004); R. J. Zatorre, J. T. Gandour, *Philos. Trans. R. Soc. London B. Biol. Sci.* **363**, 1087 (2008).
8. For a review, see N. Kraus, B. Chandrasekaran, *Nat. Rev. Neurosci.* **11**, 599 (2010).
9. D. Strait et al., *Eur. J. Neurosci.* **29**, 661 (2009); A. Parbery-Clark, E. Skoe, N. Kraus, *J. Neurosci.* **29**, 14100 (2009).
10. G. Schlaug, *Ann. NY Acad. Sci.* **930**, 281 (2001); T. F. Münte, E. Altenmüller, L. Jäncke, *Nat. Rev. Neurosci.* **3**, 473 (2002).
11. S. Anderson, N. Kraus, *Trends Amplif.* **14**, 73 (2010).
12. K. Banai et al., *Cereb. Cortex* **19**, 2699 (2009).
13. J. Song et al., *J. Cogn. Neurosci.* **20**, 1892 (2008); N. M. Russo et al., *Behav. Brain Funct.* **6**, 60 (2010); S. Carcagno, C. J. Plack, *J. Assoc. Res. Otolaryngol.* **12**, 89 (2011).
14. B. Chandrasekaran et al., *Neuron* **64**, 311 (2009). ■

VACUUM & ABATEMENT INNOVATIVE TECHNOLOGY GLOBAL STRENGTH LOCAL SUPPORT

**Think Turbopump, Think Edwards...**

Did you know Edwards is the leading manufacturer and supplier of turbopumps worldwide?

**nEXT Turbopumps** for pumping speeds up to 400l/s

- Optimal pumping speeds and compression ratios
- True end user serviceability without complex tooling

**STP Maglev Turbopumps** for pumping speeds from 300 to 4500l/s

- Very low noise and vibration
- Zero maintenance in most R&D applications

Whatever your turbopump requirement, whether discreet pump or ready made pumping stations, be assured Edwards has the solution and application expertise to meet your needs.

Think Turbopump, Think Edwards:  
[www.edwardsvacuum.com](http://www.edwardsvacuum.com)  
[turbo@edwardsvacuum.com](mailto:turbo@edwardsvacuum.com)

**EDWARDS**  
[WWW.EDWARDSVACUUM.COM](http://WWW.EDWARDSVACUUM.COM)

© Edwards Limited 2011. All Rights Reserved