

Hearing impairments hidden in normal listeners

Shihab A. Shamma¹

University of Maryland, College Park, MD 20742

Early and accurate diagnosis of sensory deficits is a public health concern with major scientific and financial implications. Several fundamental requirements conspire to make this goal a difficult one to achieve; among them is the necessity for well-designed quick tests that can pinpoint the nature of the deficits. Another is the need for a deep understanding of the functional organization of the sensory system and the cues mediating perception.

To appreciate how these requirements shape our current diagnostic paradigms, consider the striking difference between the procedures routinely used to screen children and young adults for visual and auditory impairments. In vision, even the simplest eye examinations test the ability to identify the shapes of letters or the direction of arrows, and hence the acuity of the visual system in encoding the meaningful structure of the visual scene. By contrast, hearing tests are often limited to measuring audiometric thresholds that reveal the just-audible tone intensity as a function of frequency. The relevance of these hearing tests is analogous in vision to asking subjects to indicate the presence or absence of a spot of light falling on different retinal locations. It is therefore understandable why the utility of basic auditory screening procedures is limited to indications of “hard of hearing,” failing in most cases to discover a heterogeneity of hearing disabilities in a population of listeners or to provide an indication of hearing impairments that can benefit from early intervention.

Auditory scientists are of course well aware of these limitations and have devised a variety of alternative tests to probe deeper into auditory function. These efforts, however, are hampered by the lack of an agreed-upon set of basic auditory physiological and perceptual cues whose measurements reliably reflect hearing ability and fidelity. Thus, despite enormous progress, we remain largely uncertain about the neurobiological mechanisms underlying perception of basic attributes of sound, such as pitch (1, 2), localization (3, 4), and timbre (5, 6). One important source of this difficulty stems from a fundamental duality of spectral and temporal cues in early auditory processing that has bedeviled auditory research for more than a century (7). The essence of this controversy is illustrated in Fig. 1. It stems from the duality of cochlear representations (Fig. 1A) and the way these cues

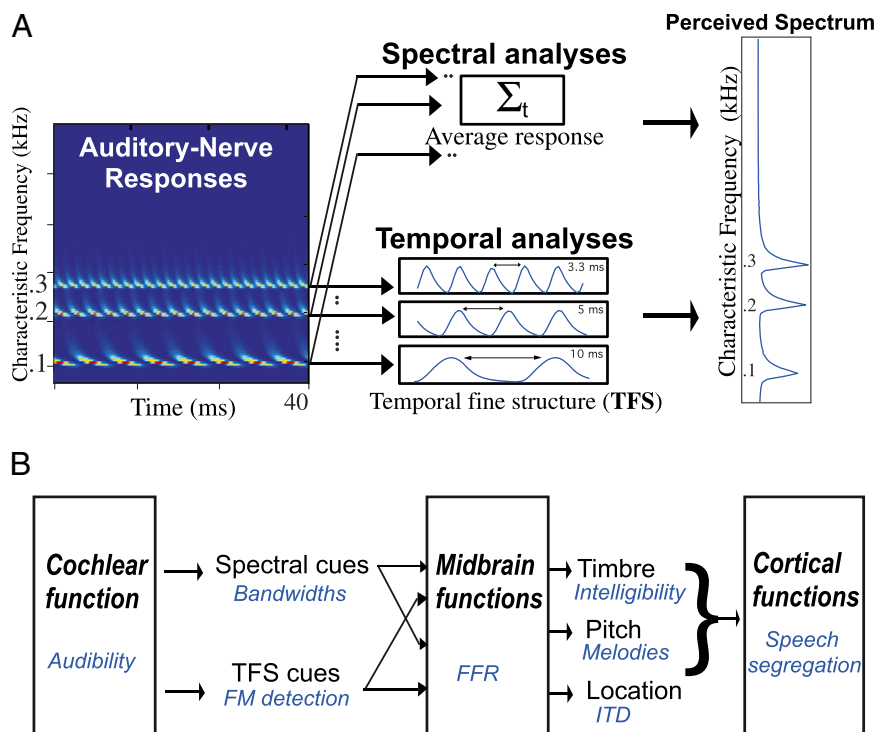


Fig. 1. Auditory functions and cues. (A) Spectral and temporal cues in cochlear processing. Cochlear analysis of three harmonic tones (100, 200, and 300 Hz) results in spatially well-separated clusters of responses along the tonotopic axis of the auditory nerve (spectral analysis). Each tone also evokes a periodic response pattern time-locked to the frequency of the tone and referred to as the temporal fine structure (TFS). The perceived spectral pattern can be estimated from either the spectral cues (by simple averaging of each channel) and/or from the temporal cues by analysis of the response periodicities. (B) Auditory functions and diagnostics. A conceptual sequence of auditory transformations and the common procedures (in blue) to probe the cues at each stage. Audibility threshold measurements often reflect general cochlear function. More specific insights into its spectral analysis are gained from “bandwidth” estimates (16). Fidelity of the TFS is assessed by detecting thresholds of frequency modulation (FM) (10). Spectral and temporal cues become intimately intermingled as they jointly contribute to generate the perceptual attributes (timbre, pitch, and location) in the midbrain and toward the cortex. Pitch percepts are often evaluated using “melodic judgments and resolution estimates” (5). Timbre is a more ambiguous percept that is broadly reflected by measurements of speech intelligibility and musical instrument identification (9). Localization of azimuthal sources near the midline is thought to reflect well the fidelity of the interaural time delays (ITDs) derived from the TFS (17). Finally, auditory cortical function engages numerous cognitive abilities, such as attention and memory, as well as bottom-up spectral and temporal cues through the perceptual attributes. “Speech segregation tests” can be designed to diagnose the contribution of each specific subset of these cues to perception, as demonstrated by Ruggles et al. (8).

become intermingled in the midbrain as they contribute jointly in mediating various auditory perceptual attributes (Fig. 1B).

The study by Ruggles et al. in PNAS (8) elegantly demonstrates how much can be gained from a careful consideration of auditory principles in diagnosing hearing impairments, both at a scientific level and as a public health goal. Having selected a group of “normal” subjects on the basis of widely used audiometric threshold criteria, the authors design a series of tests with suprathreshold sounds targeted to detect evidence of “temporal dysfunction,”

specifically the inability to encode properly the temporal cues thought to be critical for performing well on these tests.

The first test is a speech segregation task, typical of scenarios commonly referred to as the “cocktail party problem”

Author contributions: S.A.S. wrote the paper.

The author declares no conflict of interest.

See companion article on page 15516 of issue 37 in volume 108.

¹E-mail: sas@umd.edu.

(9), in which the subject has to attend to one of several simultaneous speakers. To do so successfully, the listener has to accurately perceive the location of a target speaker and comprehend her speech while resisting interference from other nearby speakers. This complex process requires not only the encoding of precise temporal information to localize the desired source but also a myriad of other spectral and temporal cues to decode the speech. The authors astutely observe that badly performing subjects comprehended well the received speech, though not that of the intended target speaker. One sensible interpretation of this result is that subjects could not localize well, an auditory function that is strongly mediated by interaural time delay cues that are neurally encoded by the fine-time-structure responses on the auditory nerve [the TFS (temporal fine structure) cues in Fig. 1A].

The second test explores the integrity of these temporal cues more directly by measuring detection thresholds of frequency modulations on a single tone (Fig. 1B), a task that is widely assumed to require precise temporal cues but is entirely unrelated to binaural hearing (10). Again, badly performing subjects on the first test displayed a similar difficulty in this task, suggesting that their deficits had more to do specifically with temporal encoding mechanisms and less to do with the binaural processing of these cues.

Both of the above tests are psychoacoustic, requiring careful controls, extensive measurements, and willing subjects who follow instructions. By contrast, the

third test the authors administer is a noninvasive physiological measurement of the auditory brainstem response, the summed neural activity from the cochlear

Much can be gained from a careful consideration of auditory principles in diagnosing hearing impairments.

nucleus and nearby midbrain structures midway up the auditory pathway. Voiced speech and other stimuli containing periodic components evoke a frequency-following response (FFR) that reflects stimulus-locked neural responses, which may be interpreted as an approximate, and admittedly ill-defined, global correlate of the accuracy of temporal encoding (11). Be that as it may, the FFR of the subjects that performed well on the previous two psychoacoustic tests displayed strong responses to the fundamental periodicities in the speech signal. Interestingly, this finding is consistent with many previous reports that have associated enhanced FFR with learned auditory skills, such as in musicians and speakers of tonal languages (12), and weaker brainstem responses (ABR) after exposure to loud noise (13). ABR measurements do not involve a psychoacoustic task requiring comprehension of instructions or speech samples. As

such, the test decouples higher cognitive abilities and functions (e.g., deficits in attention or linguistic abilities) from the expression of the basic temporal cues at the earlier auditory stages, thus pointing to the possible origin of the observed impairments. Finally, an attractive aspect of the ABR test is its suitability for infants and other noncooperative subjects, because it requires no feedback or even conscious listening (14).

The implications of the Ruggles et al. study (8) are numerous and important. On the scientific level, this study confirms the critical role of the TFS in sound perception and analysis of complex auditory scenes (15). On the clinical and therapeutic levels, it demonstrates the feasibility and benefits of more sophisticated auditory screening tests. It also points to the urgent need for a larger battery of tests to probe a wider set of cues at different levels of the auditory system (10). Such information will pave the way for the design of more effective hearing aids that are tailored to ameliorate the exact diagnosed impairments. For instance, one may incorporate finer spectral analysis to counter abnormally broad cochlear filters or a crisper representation of onsets and offsets to enhance the temporal modulations. Finally, as the authors point out, the societal implications of better auditory screening are substantial because it makes apparent the dangers of pervasive habits such as listening to blaring music and living with loud environmental sounds, all the while hiding behind the false security of normal audiometric thresholds!

- Oxenham AJ, Bernstein JG, Penagos H (2004) Correct tonotopic representation is necessary for complex pitch perception. *Proc Natl Acad Sci USA* 101:1421–1425.
- Shamma S, Klein D (2000) The case of the missing pitch templates: How harmonic templates emerge in the early auditory system. *J Acoust Soc Am* 107:2631–2644.
- McAlpine D, Jiang D, Palmer A (2001) A neural code for low-frequency sound localization in mammals. *Nat Neurosci* 4:396–401.
- Popper AN, Fay RR, eds (2005) *Sound Source Localization* (Springer, New York).
- Moore BCJ (2004) *An Introduction to the Psychology of Hearing* (Elsevier Academic Press, London).
- Shamma S (1999) Physiological bases of timbre perception. *The New Cognitive Neurosciences*, ed Michael Gazzaniga (MIT Press, Cambridge, MA).
- Pickles JO (2003) *An Introduction to the Physiology of Hearing* (Emerald Group, Bingley, UK).
- Ruggles D, Bharadwaj H, Shinn-Cunningham BG (2011) Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proc Natl Acad Sci USA* 108:15516–15521.
- Bronkhorst AW (2000) The cocktail party phenomenon: A review on speech intelligibility in multiple-talker conditions. *Acust Acta Acust* 86:117–128.
- Strelcyk O, Dau T (2009) Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. *J Acoust Soc Am* 125:3328–3345.
- Smith JC, Marsh JT, Brown WS (1975) Far-field recorded frequency-following responses: Evidence for the locus of brainstem sources. *Electroencephalogr Clin Neurophysiol* 39:465–472.
- Wong PC, Skoe E, Russo NM, Dees T, Kraus N (2007) Musical experience shapes human brainstem encoding of linguistic pitch patterns. *Nat Neurosci* 10:420–422.
- Kujawa SG, Liberman MC (2009) Adding insult to injury: Cochlear nerve degeneration after “temporary” noise-induced hearing loss. *J Neurosci* 29:14077–14085.
- Russo NM, et al. (2008) Deficient brainstem encoding of pitch in children with Autism Spectrum Disorders. *Clin Neurophysiol* 119:1720–1731.
- Lorenzi C, Moore BCJ (2008) Role of temporal envelope and fine structure cues in speech perception: A review. *Auditory Signal Processing in Hearing-Impaired Listeners, First International Symposium on Auditory and Audiological Research (ISAAR 2007)*, eds Dau T, Buchholz JM, Harte JM, Christiansen TU (Center-Tryk, Holbaek, Denmark).
- Glasberg BR, Moore BCJ (1990) Derivation of auditory filter shapes from notched-noise data. *Hear Res* 47:103–138.
- Joris PX, Smith PH, Yin TC (1998) Coincidence detection in the auditory system: 50 years after Jeffress. *Neuron* 21:1235–1238.