Buenos Aires – 5 to 9 September, 2016
Acoustics for the 21st Century…

PROCEEDINGS of the 22nd International Congress on Acoustics

Communication Acoustics: Paper ICA2016-790

Hearing research relevant to communication acoustics

Barbara Shinn-Cunningham[a]

[a] Boston University, United States, shinn@bu.edu

Abstract

Hearing research has been a foundation for understanding communication acoustics. Historically, hearing researchers have viewed individual differences as a nuisance that makes it difficult to interpret how acoustic conditions affect auditory perception. This paper sketches out how we have begun to use individual differences to tease apart the processes that affect perception in young adults who have normal hearing thresholds, with a particular focus on how listeners understand speech when there are competing sound sources. We find that individual subjects show consistent differences in their ability to understand speech in noise, which are correlated with differences in the ability to extract fine temporal details of sounds, as well as with physiological differences in the fidelity with which the brainstem encodes temporal acoustic detail. Growing evidence suggests that cochlear synaptopathy, otherwise known as "hidden hearing loss," may explain these differences in otherwise healthy listeners with normal hearing thresholds. After reviewing this evidence, this talk will consider the implications for developing new technologies that could assist listeners who have difficulty communicating in noisy, reverberant settings.

Keywords: psychoacoustics, hidden hearing loss, cochlear neuropathy

Hearing research relevant to communication acoustics

1 Introduction

Historically, the majority of psychoacoustic studies have explored how variations in perceptual ability depend on acoustic stimulus parameters. Often in such studies, individual differences across listeners confound interpretations; they are a source of noise and interfere with the differences that are the focus of study. However, a growing number of studies have started to exploit repeatable individual differences that are present across listeners with normal audiometric thresholds.

The envelope-following response (EFR) is a measure of electrical activity coming from the subcortical portions of the auditory pathway, measured via electrodes placed on the scalp [1]. The EFR specifically measures the envelope in the electrical response of the brain in response to a periodic sound whose fundamental frequency falls in the range between about 100 – 500 Hz [1]. Importantly, the EFR indexes important differences in temporal coding fidelity in listeners with normal hearing thresholds [2-9]. The finding that individual differences in ability are related to differences in objective physiological measurements supports the idea that sensory coding fidelity differs amongst listeners with normal audiometric thresholds, and that this affects hearing in everyday settings.

2 “Hidden hearing loss” reduces auditory nerve responses

A growing number of animal studies show that noise exposure that causes no permanent damage may nonetheless cause a loss of auditory nerve responses (ANFs) [10, 11]. Noise exposure that does no damage to the cochlear mechanical response can produce a rapid loss of as many as 40-60% of the ANF synapses driven by cochlear inner hair cells, the cells that generate the ascending signal conveying information in the auditory pathway [12, 13]. This loss of synapses subsequently leads to a slow death of the ANF cell bodies (spiral ganglion cells) and central axons [11, 14]. Even in cases where the effects on synapses and spiral ganglion cells are pronounced, the effect on cochlear function can be negligible; cochlear mechanical function (including the tuning of the cochlea) can be normal in animals suffering from this “cochlear neuropathy” [12]. Most hearing screenings reveal losses associated with damage to inner and outer hair cells by looking for: 1) elevated detection thresholds, 2) reduced amplification in the cochlea, 3) wider-than-normal cochlear tuning, and 4) reduced otoacoustic emissions. Yet, with hidden hearing loss, these measures are normal, making the deficit “hidden” to typical hearing screening [15].

3 Hidden hearing loss seems to be present in humans

While there are no data yet to directly support the idea that cochlear neuropathy occurs in humans, a growing number of studies hint that it accounts for some of the individual variability seen in listeners with normal cochlear mechanical function. We have seen repeatedly that listeners with normal hearing thresholds vary significantly in their ability to utilize precise temporal information [4, 16]. This variability correlates with difficulties in using spatial selective attention to focus on and understand speech in a noisy background [5, 16], underscoring the clinical relevance of these differences. Importantly, when we ask the question of "are these individual differences in listeners with "normal hearing" the kinds of differences one might expect based on cochlear neuropathy?," the answer is a resounding "yes."

3.1 Differences in perceptual ability are consistent with cochlear neuropathy

In one such study, young adult subjects were recruited with no special criteria except that they had normal hearing thresholds and no known auditory deficits [16]. Individual differences amongst this cohort were nonetheless large. Perceptual abilities correlated with EFR strength, especially at high sound levels and shallow modulation depths where higher-threshold ANFs are important for coding temporal features. There are consistent relationships between the EFR strength and perceptual thresholds for amplitude modulation detection and for envelope interaural time difference (ITD) discrimination [16]. Both of these perceptual measures rely on fine temporal information, and both are significantly correlated with the strength of the EFR when a shallow modulation drives the brainstem response.

Other studies in humans also support the view that human listeners with normal cochlear function may suffer from different degrees of cochlear neuropathy. For instance, listeners can vary significantly in their ability to discriminate both frequency modulation and interaural time differences (see [17, 18]). The computation of ITDs depends directly on temporal precision in ANF responses and subsequent processing centres (such as neurons in the superior olivary complex). Indeed, sensitivity to ITD cues was one of the perceptual abilities that correlated with EFR strength [16].

On the physiological side, listeners with normal hearing thresholds show large inter-subject variability in the magnitude of ABR wave I (an electrical signal generated in response to a click or short sound by activity of the auditory nerve) [15, 19], again supporting the view that listeners with normal audiograms suffer from cochlear neuropathy to varying degrees. As in animal studies, while ABR wave I amplitude varies significantly across individuals, the magnitude of ABR wave V does not [15, 19], indicating that neuropathy can be accompanied by upstream plasticity that compensates for a loss of response, but not for a loss of temporal precision in responses [20].

Another study has shown that perceptual differences correlate with differences in human ABRs: in young adults with no known hearing deficits, wave I magnitude was related to ITD sensitivity [21]. Consistent with previous animal studies, wave V magnitude was unrelated to wave I magnitude or perceptual ability (although effects of noise on wave V timing were correlated with wave I amplitude). Taken together, these results suggest that cochlear neuropathy is common amongst human listeners who have normal audiograms, many of whom do not even realize that they may have communication difficulties.

4 Hidden hearing loss likely affects acoustic communication

Roughly 5-10% of listeners seeking treatment at audiological clinics have normal hearing thresholds [22, 23]. Typically, these patients are driven to seek help because of difficulty communicating in situations requiring them to focus selective attention. Historically, such listeners were said to have “central auditory processing disorder” [24], a catchall diagnosis that testifies to the fact that underlying causes were not well understood; however, some of these listeners likely are suffering from cochlear neuropathy.

The fact that listeners first notice the effects of cochlear neuropathy when trying to communicate in social settings makes sense, given how neuropathy degrades auditory temporal coding. Spectrotemporal details in a sound mixture are important for grouping of acoustic elements into perceptual objects [25, 26], discrimination of perceptual features like pitch [27] and source location [27, 28], as well as speech perception itself [29]. Importantly, in quiet, subtle hearing deficits may not disrupt speech perception, yet still have a debilitating effect on selective auditory attention.

4.1 Source Segregation

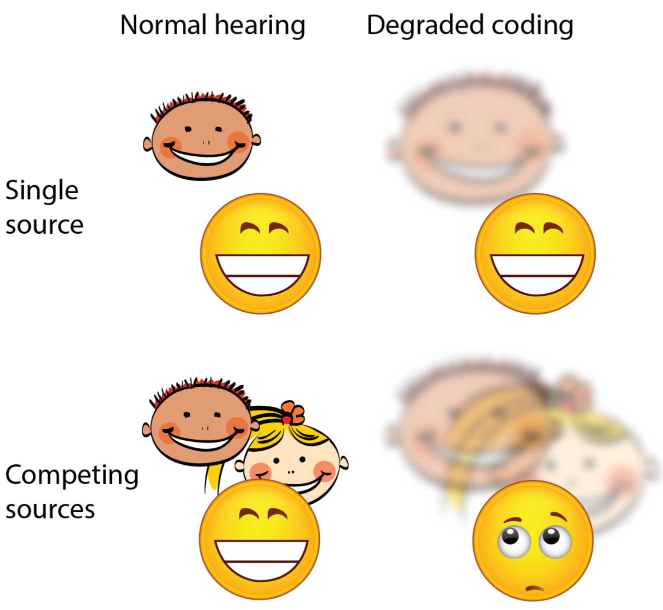
In order to selectively attend, listeners must be able to segregate sounds making up the acoustic mixture entering the ears. Source segregation depends on harmonic structure, ITDs, and other cues computed from acoustic features that are degraded when temporal coding is poor [30, 31]. If temporal features are degraded and the target source cannot properly be segregated from the scene, selective attention will fail [32, 33].

In the auditory domain, when listening to a complex scene, spectrotemporal details (e.g., periodicity, ITD, and amplitude and frequency modulation) are analogous to the edges and colors of a visual scene. These features are less clearly represented when a listener suffers from hidden hearing loss, so that the structural elements critical for parsing the acoustic scene are perceptually indistinct.

4.2 Source Selection

Successfully listening in a complex setting depends on more than simply segregating the sources from one another; it also requires selecting the desired source from the mixture by focusing selective attention. Selective auditory attention enhances the representation of the auditory object with a desired perceptual feature or attribute [34, 35]. The low-level acoustic spectro-temporal structure is what enables a listener to compute perceptual features of objects in a scene that can be used to focus attention. Specifically, low-level features such as periodicity, ITD, and amplitude and frequency modulation support computation of higher-level perceptual quantities such as pitch, location, and timbre. These attributes can be used to “listen to the high-pitched source,” or “the source on the left,” or to “Sally, not Jim.”

One clear example of a high-level feature that is degraded when temporal cues are weak (e.g., due to hidden hearing loss) is spatial location. When temporal cues are weak, the perceived location of a sound source can be smeared. Listeners with a weak temporal code can fail to select the correct source in the scene based on its less-perceptually precise location. For instance, one study found large individual differences in performance on a spatial selective attention task [5]. In this study, when listeners failed, they did not fail to understand speech present in the sound mixture. Instead, they reported the wrong word, coming from the wrong location. That is, perceptual deficits were not severe enough to interfere with understanding the speech that was present in the mixture— the failures happened because listeners could not select the correct talker based on spatial cues. Consistent with the idea that spatial selective attention fails when listeners suffer from hidden hearing loss and poor temporal coding, the individual variations in performance on the selective attention task correlated with differences in EFR strength [3, 5]. Reverberation, which is a natural form of temporal degradation in the signals reaching the ears, exacerbated the selective attention errors. In other words, both external noise in the temporal acoustic features important for conveying location (from reverberation) and internal noise in the computation of ITDs (from differences in temporal coding fidelity in the brainstem) had similar, additive effects in disrupting selective auditory attention.



Source: (Shinn-Cunningham, 2016)

Figure 1: Visual analogy illustrating the effects of a poor peripheral representation on the ability to process sources in a crowded setting.

This idea is illustrated by visual analogy in the cartoons shown in Fig. 1. If there is no competition and the scene is simple, even a poor representation is easy to interpret (top row).

However, when there are competing sources, there is a clear impact of poor sensory coding. In people with good coding fidelity, fine details in the scene ensure that each source is distinct; the listener with a good peripheral representation can focus attention unambiguously to a talker to the left (bottom left). If a listener has a poor peripheral representation, spatial cues are weakly represented and the talker locations can overlap and smear into each other; other features, analogous to edges and colors may also be indistinct, smearing the objects together (bottom right). Even if a listener *were able* to parse the scene into a male and a female talker, they may focus on the wrong talker when trying to focus on "the talker on the left" because of the spatial ambiguity in the scene. Such problems can produce communication problems in settings where there are multiple sources competing for attention that would not show up on a test of speech perception in quiet, or even if there were non-speech sounds present (i.e., in conditions where competing sound objects are so perceptually dissimilar that failures of selection will not occur).

4.3 Everyday communication reveals subtle deficits of hidden hearing loss

These examples demonstrate whey even modest degradations in temporal processing may lead to communication dysfunction in everyday settings [33]. Temporal coding problems interfere with the sound features that support both segregation and selection of the desired source from the mixture. In other words, listening to a talker amidst similar, competing talkers reveals deficits that are may be too subtle to be revealed in other listening situations.

5 Conclusions

Individual differences in the fidelity with which the auditory nerve encodes temporal details in sound likely arise through cochlear neuropathy (death of auditory nerve fibers) due to noise exposure and aging. Everyday communication depends on the ability to segregate a source of interest from a sound mixture and then selectively focus attention on that sound. Both of these operations require exquisite temporal resolution. Because of this, individual differences in auditory temporal coding due to cochlear neuropathy likely influence communication acoustics by interfering with the ability to converse in everyday social settings.

Acknowledgments

This work was funded in part by grants from the National Institute on Deafness and Other Communication Disorders, the National Science Foundation the Oticon Foundation, and the Coulter Foundation.

References

[1] Shinn-Cunningham, B.G., L. Varghese, L. Wang, and H. Bharadwaj, Individual differences in temporal perception, and their implications for everyday listening, in The Frequency-following Response: A Window into Human Communication, N. Kraus, et al., Editors. in press, Springer Nature: New York.

[2] Shinn-Cunningham, B., D.R. Ruggles, and H. Bharadwaj, How Early Aging and Environment Interact in Everyday Listening: From Brainstem to Behavior Through Modeling. Basic Aspects of Hearing: Physiology and Perception, 2013. 787: p. 501-510.

[3] Ruggles, D., H. Bharadwaj, and B.G. Shinn-Cunningham, Why middle-aged listeners have trouble hearing in everyday settings. Curr Biol, 2012. 22(15): p. 1417-22.

[4] Ruggles, D., H. Bharadwaj, and B.G. Shinn-Cunningham, Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. Proc Natl Acad Sci U S A, 2011. 108(37): p. 15516-21.

[5] Ruggles, D. and B. Shinn-Cunningham, Spatial selective auditory attention in the presence of reverberant energy: individual differences in normal-hearing listeners. J Assoc Res Otolaryngol, 2011. 12(3): p. 395-405.

[6] Skoe, E., J. Krizman, E. Spitzer, and N. Kraus, The auditory brainstem is a barometer of rapid auditory learning. Neuroscience, 2013. 243: p. 104-14.

[7] Anderson, S., A. Parbery-Clark, T. White-Schwoch, and N. Kraus, Auditory brainstem response to complex sounds predicts self-reported speech-in-noise performance. J Speech Lang Hear Res, 2013. 56(1): p. 31-43.

[8] Chandrasekaran, B., N. Kraus, and P. Wong, Human inferior colliculus activity relates to individual differences in spoken language learning. Journal of Neurophysiology, 2011.

[9] Anderson, S., E. Skoe, B. Chandrasekaran, and N. Kraus, Neural timing is linked to speech perception in noise. Journal of Neuroscience, 2010. 30(14): p. 4922-6.

[10] Liberman, M.C., Hidden Hearing Loss. Sci Am, 2015. 313(2): p. 48-53.

[11] Kujawa, S.G. and M.C. Liberman, Synaptopathy in the noise-exposed and aging cochlea: Primary neural degeneration in acquired sensorineural hearing loss. Hear Res, 2015. 330(Pt B): p. 191-9.

[12] Kujawa, S.G. and M.C. Liberman, Adding insult to injury: cochlear nerve degeneration after "temporary" noise-induced hearing loss. J Neurosci, 2009. 29(45): p. 14077-85.

[13] Kujawa, S.G. and M.C. Liberman, Acceleration of age-related hearing loss by early noise exposure: evidence of a misspent youth. J Neurosci, 2006. 26(7): p. 2115-23.

[14] Lin, H.W., A.C. Furman, S.G. Kujawa, and M.C. Liberman, Primary neural degeneration in the Guinea pig cochlea after reversible noise-induced threshold shift. J Assoc Res Otolaryngol, 2011. 12(5): p. 605-16.

[15] Schaette, R. and D. McAlpine, Tinnitus with a normal audiogram: physiological evidence for hidden hearing loss and computational model. J Neurosci, 2011. 31(38): p. 13452-7.

[16] Bharadwaj, H.M., S. Masud, G. Mehraei, S. Verhulst, and B.G. Shinn-Cunningham, Individual differences reveal correlates of hidden hearing deficits. Journal of Neuroscience, 2015. 35: p. 2161-2172.

[17] Strelcyk, O. and T. Dau, Relations between frequency selectivity, temporal fine-structure processing, and speech reception in impaired hearing. Journal of the Acoustical Society of America, 2009. 125(5): p. 3328-45.

[18] Grose, J.H. and S.K. Mamo, Processing of temporal fine structure as a function of age. Ear Hear, 2010. 31: p. 755-760.

[19] Stamper, G.C. and T.A. Johnson, Auditory function in normal-hearing, noise-exposed human ears. Ear Hear, 2015. 36(2): p. 172-84.

[20] Chambers, A.R., J. Resnik, Y. Yuan, J.P. Whitton, A.S. Edge, M.C. Liberman, and D.B. Polley, Central Gain Restores Auditory Processing following Near-Complete Cochlear Denervation. Neuron, 2016. 89(4): p. 867-79.

[21] Mehraei, G., A.E. Hickox, H.M. Bharadwaj, H. Goldberg, S. Verhulst, M.C. Liberman, and B.G. Shinn-Cunningham, Auditory Brainstem Response Latency in Noise as a Marker of Cochlear Synaptopathy. J Neurosci, 2016. 36(13): p. 3755-64.

[22] Hind, S.E., R. Haines-Bazrafshan, C.L. Benton, W. Brassington, B. Towle, and D.R. Moore, Prevalence of clinical referrals having hearing thresholds within normal limits. Int J Audiol, 2011. 50(10): p. 708-16.

[23] Kumar, G., F. Amen, and D. Roy, Normal hearing tests: is a further appointment really necessary? J R Soc Med, 2007. 100(2): p. 66.

[24] Rosen, S., M. Cohen, and I. Vanniasegaram, Auditory and cognitive abilities of children suspected of auditory processing disorder (APD). Int J Pediatr Otorhinolaryngol, 2010. 74(6): p. 594-600.

[25] Christiansen, S.K. and A.J. Oxenham, Assessing the effects of temporal coherence on auditory stream formation through comodulation masking release. J Acoust Soc Am, 2014. 135(6): p. 3520-9.

[26] Shamma, S.A., M. Elhilali, and C. Micheyl, Temporal coherence and attention in auditory scene analysis. Trends in Neurosciences, 2011. 34(3): p. 114-23.

[27] Smith, Z.M., B. Delgutte, and A.J. Oxenham, Chimaeric sounds reveal dichotomies in auditory perception. Nature, 2002. 416(6876): p. 87-90.

[28] Blauert, J., Spatial Hearing (2e). 1997, Cambridge, MA: MIT Press.

[29] Zeng, F.G., K. Nie, G.S. Stickney, Y.Y. Kong, M. Vongphoe, A. Bhargave, C. Wei, and K. Cao, Speech recognition with amplitude and frequency modulations. Proc Natl Acad Sci U S A, 2005. 102(7): p. 2293-8.

[30] Bregman, A.S., Auditory Scene Analysis: The Perceptual Organization of Sound. 1990, Cambridge, MA: MIT Press.

[31] Carlyon, R.P., How the brain separates sounds. Trends Cogn Sci, 2004. 8(10): p. 465-71.

[32] Shinn-Cunningham, B.G., Object-based auditory and visual attention. Trends Cogn Sci, 2008. 12(5): p. 182-6.

[33] Shinn-Cunningham, B.G. and V. Best, Selective attention in normal and impaired hearing. Trends in Amplification, 2008. 12(4): p. 283-99.

[34] Lee, A.K.C., E. Larson, and R.K. Maddox, Mapping cortical dynamics using simultaneous MEG/EEG and anatomically-constrained minimum-norm estimates: an auditory attention example. J Vis Exp, 2012(68): p. e4262.

[35] Maddox, R.K. and B.G. Shinn-Cunningham, Influence of task-relevant and task-irrelevant feature continuity on selective auditory attention. J Assoc Res Otolaryngol, 2012. 13(1): p. 119-29.