Title
Perceptual consequences of disrupted auditory nerve activity

Permalink
https://escholarship.org/uc/item/76p335v2

Journal
Journal of Neurophysiology, 93(6)

ISSN
0022-3077

Authors
Zeng, FG
Kong, YY
Michalewski, HJ
et al.

Publication Date
2005-06-01

DOI
10.1152/jn.00985.2004

License
CC BY 4.0

Peer reviewed
**INTRODUCTION**

Perception is a delicate chain of events including conversion of a sensory stimulus into electrical signals at the receptor level, transmission of the electrical signals via the peripheral nervous system, and processing and interpretation of the electrical signal in the CNS. Any breakdown in the process can have significant consequences in perception. In audition, perceptual consequences of both peripheral and central auditory disorders have been studied extensively. For example, peripheral damage in the inner ear and the auditory nerve leads to threshold elevation, abnormal loudness, pitch, and temporal processing (Buss et al. 1998; Formby 1986; Moore 1996; Moore and Oxenham 1998; Nienhuys and Clark 1978; Oxenham and Bacon 2003; Prosen et al. 1981; Ryan and Dallos 1975); central disorders and degeneration produce complex processing deficits in speech and sound object recognition (Cacace and McFarland 1998; Gordon-Salant and Fitzgibbons 1999; Levine et al. 1993; Wright et al. 1997); and electric stimulation of the auditory nerve via a cochlear implant in deaf persons results in fundamental changes in the brain affecting behaviors ranging from basic psychophysics to language development (Giraud et al. 2001; Simmons et al. 1965; Svirsky et al. 2000; Zeng and Shannon 1994). Examination of the above-mentioned studies shows that detailed documentation of behavioral changes coupled with a clearly defined pathology can make significant contributions in two important ways. Clinically, these studies often lead to better diagnosis and management of a particular disease. Theoretically, these studies shed light on the relative contribution of different neural codes to perception at a mechanistic level.

Here we focus on perceptual consequences of a recently defined hearing disorder that preserves the outer hair cell function but apparently disrupts auditory nerve activity. This hearing disorder was first described in one single subject and considered to involve a dysfunction of the auditory nerve (Starr et al. 1991). Subsequently, 10 subjects with similar symptoms were identified. Since eight of them had accompanying peripheral neuropathy, the term “auditory neuropathy” (AN) was coined (Starr et al. 1991, 1996). The clinical diagnosis of AN has been typically characterized by the presence of otoacoustic emission and/or cochlear microphonics and the concurrent absence of the averaged auditory brain stem responses. The presence of otoacoustic emission and/or cochlear microphonics is indicative of normal cochlear outer hair cell activity, whereas the absence of the auditory brain stem responses is indicative of disrupted auditory nerve activity. Despite absent auditory brain stem responses, it is evident that sound information must have been transmitted, because individuals with AN can hear sound, have normal brain imaging, and identifiable, although usually delayed, cortical potentials (Rance et al. 2002; Starr et al. 2003; Zeng et al. 1999).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

**Department of Anatomy and Neurobiology, Biomedical Engineering, Cognitive Sciences, Neurology, and Otolaryngology–Head and Neck Surgery, University of California, Irvine, California**

Submitted 21 September 2004; accepted in final form 17 December 2004

Fan-Gang Zeng, Ying-Yee Kong, Henry J. Michalewski, and Arnold Starr. Perceptual consequences of disrupted auditory nerve activity. *J Neurophysiol* 93: 3050–3063, 2005. First published December 22, 2004; doi:10.1152/jn.00985.2004. Perceptual consequences of disrupted auditory nerve activity were systematically studied in 21 subjects who had been clinically diagnosed with auditory neuropathy (AN), a recently defined disorder characterized by normal outer hair cell function but disrupted auditory nerve function. Neurological and electrophysiological evidence suggests that disrupted auditory nerve activity is due to desynchronized or reduced neural activity or both. Psychophysical measures showed that the disrupted neural activity has minimal effects on intensity-related perception, such as loudness discrimination, pitch discrimination at high frequencies, and sound localization using interaural level differences. In contrast, the disrupted neural activity significantly impairs timing related perception, such as pitch discrimination at low frequencies, temporal integration, gap detection, temporal modulation detection, backward and forward masking, signal detection in noise, binaural beats, and sound localization using interaural time differences. These perceptual consequences are the opposite of what is typically observed in cochlear-impaired subjects who have impaired intensity perception but relatively normal temporal processing after taking their impaired intensity perception into account. These differences in perceptual consequences between auditory neuropathy and cochlear damage suggest the use of different neural codes in auditory perception: a suboptimal spike count code for intensity processing, a synchronized spike code for temporal processing, and a duplex code for frequency processing. We also proposed two underlying physiological models based on desynchronized and reduced discharge in the auditory nerve to successfully account for the observed neurological and behavioral data. These methods and measures cannot differentiate between these two AN models, but future studies using electric stimulation of the auditory nerve via a cochlear implant might. These results not only show the unique contribution of neural synchrony to sensory perception but also provide guidance for translational research in terms of better diagnosis and management of human communication disorders.
The distorted auditory nerve activity has been suggested in the form of either desynchronized or reduced discharge in the auditory nerve because both could lead to absent or abnormal brain stem responses and present but delayed cortical responses (Berliner et al. 2003b; Starr et al. 1996, 2003). Desynchronized neural discharge can occur due to demyelination and ion channel dysfunction in the auditory nerve (Starr et al. 1998; Waxman 1977) and/or dysfunctional synaptic transmission between the inner hair cells and the auditory nerve (Fuchs et al. 2003; Glowatzki and Fuchs 2002). Loss of the neural input to the brain can occur due to inner hair cell loss (Harrison 1998; Salvi et al. 1999; Sawada et al. 2001) and/or auditory nerve loss (Hallpike et al. 1980; Spoendlin 1974; Starr et al. 2003). Although the term AN has been widely accepted clinically as a diagnosis, alternative terms such as “auditory dys-synchrony” have been suggested to reflect the common phenomenon that likely has several underlying pathologies (Berlin et al. 2003a; Rapin and Gravel 2003). We note that the auditory nerve loss model received strong support as a recent temporal bone study in a 77-year-old AN subject showed a 95% reduction of the ganglion cells but essentially normal number and morphology of the inner and outer hair cells except for a 30% reduction of the outer hair cells in just the apical turn (Starr et al. 2003).

The etiologies for AN are multiple, ranging from genetic, including mutations in several genes (MPZ, NDRG1, and PMP22) that are critical for peripheral nerve myelination and axonal survival (Chapon et al. 1999; De Jonghe et al. 1999; Kalaydjieva et al. 2000; Maier et al. 2003), to infectious (measles, mumps), metabolic (diabetes, hypoxia), congenital (atriesia), and neoplastic (tumors) (Starr et al. 1996). The prevalence of AN has been estimated to be as high as 10% of the children identified as having hearing loss (Berlin et al. 2003a; Rance et al. 1999). Because AN patients typically do not derive benefits from conventional hearing aids, treatment options are limited, with only recent success being reported with cochlear implantation (Berlin et al. 1996; Buss et al. 2002; Mason et al. 2003; Miyamoto et al. 1999; Peterson et al. 2003; Shallop et al. 2001; Trautwein et al. 2000).

Limited behavioral data have been reported on the perceptual consequences of auditory neuropathy (Kraus et al. 2000; Rance et al. 2004; Starr et al. 1991; Zeng et al. 1999). Starr et al. (1991) and Kraus et al. (2000) each presented a case study in which the subject’s audiological, behavioral, and electrophysiological performance was systematically documented. Kraus et al. noted in their case study that the AN subject had a nearly normal audiogram but significant difficulty in speech perception in noise. Zeng et al. (1999) found significant temporal processing impairment in eight AN subjects and were able to correlate the degrees of their temporal processing impairment with the degrees of their speech perception deficits. Rance et al. (2004) found additional frequency discrimination deficits in some of their 14 AN subjects, who also had concurrent temporal processing impairment.

This study systematically examined perceptual consequences of disrupted auditory nerve activity in 21 subjects who have been clinically diagnosed with AN. Psychophysical data will be reported in basic auditory processing of intensity, frequency, and time, as well as complex processing of simultaneous masking, simultaneous masking, and binaural hearing. Temporal processing from 8 of these 21 subjects was previously reported (Zeng et al. 1999) and will be combined into the present data. Intensity discrimination from five AN subjects, frequency discrimination from three subjects, and detection of a tone in noise from two subjects were also previously reported in non–peer-reviewed book chapters or conference proceedings (Zeng 2000, 2001a,b). Although the group difference between the AN and normal-hearing control subjects is emphasized here, interesting individual cases will be presented to highlight the perceptual consequences of disrupted auditory nerve activity.

METHODS

Subjects

A total of 21 previously diagnosed AN subjects participated in this study. Table 1 summarizes their personal and audiological information, identified only by a code name. Subjects AN1-8 were identical to those who participated in an earlier study on temporal and speech perception (Zeng et al. 1999). The subjects included 13 females and 8 males, and were from 6 to 53 yr old, with a mean age of 21 yr. Their degree of hearing loss varied from nearly normal hearing with a pure-tone average threshold (PTA at all tested frequencies from 125 to 8,000 Hz) in the 20 dB HL range to severe hearing loss, with a PTA in the 70 dB HL range. Figure 1 shows the mean audiogram from the 21 AN subjects. Clinical diagnosis of AN relied on the presence of otoacoustic emission (16 of 19 tested subjects) and/or cochlear microphonics (19 of 20 tested subjects), as well as concurrent absence or severe abnormalities of auditory brain stem responses beyond that expected for the degree of pure-tone hearing loss (all 21 subjects) and absence of middle ear reflex (all 17 tested subjects). Cortical potentials were generally present but in a distorted and/or delayed form (15 of 17 tested subjects). Brain imaging was also performed in 13 of the 21 tested subjects with high-resolution CT, PET, or MRI, showing no discernible sign of an abnormal CNS. Finally, neurological exams were conducted to identify peripheral neuropathy in 7 of the 21 tested subjects, including 5 of 7 subjects having biopsies of sural nerve confirming the presence of peripheral neuropathy.

A total of 34 normal-hearing subjects also participated in this study to serve as the control. To the extent possible, age-matched subjects or unaffected family members served as the control. Tests were usually performed on the same day for both the AN and control subjects. This strategy was particularly useful when the test involved child subjects who apparently liked the alternative test sessions during which the unaffected siblings took turns performing the test. Informed consent was obtained for adult subjects; a special child’s assent form with parental consent was obtained for child subjects whose age was <17 yr. Local Institutional Review Board approval was obtained for these assent and consent forms, as well as the experimental procedures. No adverse incidents occurred during and/or after the test.

Stimuli

All stimuli were generated digitally using TDT System II equipment (Tucker-Davis Technologies, Gainesville, FL). A 16-bit D/A converter was used with a 44,100-Hz sampling rate. A 2.5-ms ramp was applied to all stimuli to avoid spectral splitter. The full digital range was used to generate a 1,000-Hz calibration tone that reached a maximal level of 100 dB SPL in Sennheiser HDA200 headphones (Wedemark, Germany), measured by a B&K Type 2260 sound level meter in a Zwischenloki real-ear simulator (Bruel and Kjær Sound and Vibration Measurement, Nærum, Denmark). The corresponding voltage (132 mV) was measured daily, whereas the sound pressure level was measured periodically for equipment calibration and maintenance. The subject performed all tasks in a double-walled, sound attenuating booth (Industrial Acoustics, Bronx, NY).
Three sets of experiments were conducted to characterize fundamental detection and discrimination abilities in AN subjects. First, intensity discrimination was measured as a function of level from near threshold to the maximal comfortable loudness for a 200-ms, 1,000-Hz tone. Second, frequency discrimination was measured as a function of frequency from 250 to 8,000 Hz in octave steps. These tones were presented at the maximal comfortable loudness and all had a 200-ms duration. Third, temporal processing measures were obtained for temporal integration, gap detection, and temporal modulation detection with a broadband (20–14,000 Hz) white noise. Temporal integration was measured as a function of duration from 5 to 500 ms. Gap detection was measured as a function of presentation level from 5 to 50 dB sensation levels (SL). A temporal gap was produced by a 100-ms, 1,000-Hz tone, whereas the signal was either a 200-ms tone that was temporally centered on the noise or a 9-ms tone that was presented with either a 1- or 300-ms delay from the onset of the masker in backward masking and the reverse in forward masking, was varied from 1 to 500 ms. To account for the individual difference in absolute threshold, the masked threshold was normalized by the following equation

\[ y = -10 \log \left( \frac{x_0}{1 + \left( \frac{f_f}{f_c} \right)^2} \right) \]  

where \( y \) is modulation index (in dB (–10log(m)), \( f \) is the modulation frequency in hertz, –10log(\( x_0 \)) is the peak sensitivity or gain in dB, and \( f_c \) is the 3-dB cut-off frequency or bandwidth in hertz.

Three additional sets of experiments were conducted to measure complex processing in nonsimultaneous masking, simultaneous masking, and binaural hearing. The total number of subjects who participated in these experiments ranged from 3 to 10 because they were required to have either normal thresholds or mild hearing loss. Four AN subjects (AN4, 9, 10, and 14) participated in the backward and forward masking experiment. Backward masking refers to a condition where the detection of a signal is affected by a masker occurring after the signal, whereas forward masking refers to a condition where the detection of a signal is affected by a masker occurring before the signal. The masker was a 100-ms, 1,000-Hz tone, whereas the signal was a 9-ms, 1,000-Hz tone. Signal delay, defined as the offset of the signal and the onset of the masker in backward masking and the reverse in forward masking, was varied from 1 to 500 ms. To account for the individual difference in absolute threshold, the masked threshold was normalized by the following equation

\[ \text{Normalized threshold} = 100 \times \left( \frac{TH - TH_0}{TH_{1ms} - TH_0} \right) \]  

where \( TH \) is the threshold in dB SPL at any delay, \( TH_{1ms} \) is the threshold at the 1-ms delay, and \( TH_0 \) is the threshold in quiet.

Ten AN subjects participated in the simultaneous masking experiment (AN2, 7, 10, 11, 12, 13, 15, 16, 18, and 19). The noise masker was white noise that had a 20- to 14,000-Hz bandwidth and a 500-ms duration. The signal was either a 200-ms tone that was temporally centered on the noise or a 9-ms tone that was presented with either a 3- or 300-ms delay from the onset of the noise. Signal frequency varied from 250, 500, 1,000, 2,000, to 4,000 Hz depending on the subject’s threshold (<30 dB HL). Detection of a long tone in noise has been traditionally used to measure spectral resolution (Fletcher 1938) and pathology related to “dead regions” in the cochlea (Moore...
2004). On the other hand, threshold difference detection of a brief tone in noise between onset and steady-state conditions has been termed “overshoot” and used to measure pathology related to outer hair cell and olivocochlear bundle functions (Bacon and Takahashi 1992; Carlyon and Sloan 1987; McFadden and Champlin 1990; Zeng et al. 2000).

Three AN subjects (AN10, 17, and 18) with mild-to-moderate hearing loss participated in the following binaural hearing tests, including sound localization using the interaural level time difference (ILD) and interaural time difference (ITD) cues, as well as monaural and binaural beats. In the ILD experiment, 500-ms pure tones were presented with either 0-dB ILD (90 dB SPL at both left and right ears) or 10-dB ILD (95 dB SPL at left and 85 dB SPL at right ear). The tone frequency varied from 250 to 8,000 Hz in octave steps. In the ITD experiment, a pair of 500-ms, 500-Hz tones was presented to two ears with ITD varying 0, 30, 60, and 90° in phase, corresponding to 0-, 167-, 333-, and 500-μs lead time in the right ear. Monaural and binaural beats were measured by presenting two 1,000-ms tones with a 3-Hz difference to either the same ear (monaural beats) or the first tone to one ear and the second tone to the other ear (binaural beats). In the monaural beat experiment, the frequency was 500, 1,000, 2,000, 4,000, or 8,000 Hz for the first tone and 3 Hz higher, respectively, for the second tone. In the binaural beat experiment, the frequency was 500 Hz for the first tone and 503 Hz for the second tone.

**Psychophysical procedure**

An adaptive, three-interval, three-alternative, forced-choice, two-down and one-up procedure was employed to track the 70.7% percent correct response criterion (Levitt 1971). This procedure was used in all objective measures, including temporal integration, intensity, and frequency discrimination, backward and forward masking, simultaneous masking, and monaural beat detection. During each trial, the subject heard three sounds that were visually marked by three intervals on a computer screen. One of the three intervals contained the signal, whereas the other two contained the standard. The order of the signal and standard sounds was randomized (3-alternative). The subject had to choose the signal (forced-choice) and was given a visual feedback regarding the correct response. The initial difference between the signal and the standard was large so it was easy for the subject to tell which interval contained the signal. The difference was reduced after two consecutive correct responses and increased after one incorrect response (2-down, 1-up). A reversal was recorded when the subject made an incorrect response from two or more consecutive correct responses or vice versa. A large step size was used for the first three or four reversals, and a small step size was used for the remaining reversals (adaptive step size). Each run had a total of 12 reversals. The reported data were averaged from the last eight reversals.

An additional seven-alternative, forced-choice procedure was employed for sound localization measurement. The seven alternatives are represented by integers from −3 to 3, corresponding to the leftmost and the rightmost position (with the center position being represented by 0). The seven numbers were labeled with seven boxes that were geometrically arranged to form a semicircle on a computer screen. The subject had to choose a number corresponding to his or her perceived sound position. If the sounds from two ears could not be fused, two separate images might be perceived. In this case, the subject was instructed to report the separate images without having to choose any alternative. However, this situation was not encountered in this study.

Finally, a subjective report and counting procedure was employed to measure the presence of monaural and binaural beats. The subject was instructed to pay attention to loudness fluctuation and report the number of beats heard. In the monaural beat experiment, the modulation threshold was measured using the same three-interval, two-up and one-down, adaptive procedure as described above.

**Statistical analysis**

A between-subjects ANOVA with uneven sample sizes was used to test whether there was a significant difference in performance between the AN and normal-hearing subjects. Should the overall difference reach the significance level (P < 0.05), posthoc tests were usually conducted to examine the conditions under which the difference was significant. ANOVA was conducted with SPSS for Windows (version 11.0 2002, SPSS, Chicago, IL).

**RESULTS**

**Intensity discrimination**

Figure 2 shows intensity discrimination as a function of sensation level (dB above the individual subject’s threshold) in eight AN subjects (●) and eight normal controls (○). Both groups showed the typical pattern that had been termed as “the near-miss to Weber’s law,” in which the difference limen decreased slightly as a function of the stimulus level (McGill and Goldberg 1968). Although the AN subjects showed slightly larger difference limens at low levels than the normal controls, no significant main effect was observed between groups [F(1,69) = 3.17; P > 0.05]. This result indicates that AN subjects encounter no significant difficulty in performing pure-tone intensity discrimination.

**Frequency discrimination**

Figure 3 shows frequency discrimination as a function of standard frequency in 12 AN subjects (●) and 4 normal controls (○). A significant difference in performance was observed between the AN subjects and the normal controls [F(1,86) = 5.53; P < 0.05]. The normal controls required <10 Hz to discriminate a pitch difference for frequencies ≤ 1,000 Hz, but the AN subjects required a difference that was about two orders of magnitude higher than the normal difference limen. Interestingly, the difference between the two groups reduced with frequency and was not significant at 8,000 Hz (P > 0.10). This result suggests that AN subjects have pro-

---

**Psychophysical procedure**

An adaptive, three-interval, three-alternative, forced-choice, two-down and one-up procedure was employed to track the 70.7% percent correct response criterion (Levitt 1971). This procedure was used in all objective measures, including temporal integration, intensity, and frequency discrimination, backward and forward masking, simultaneous masking, and monaural beat detection. During each trial, the subject heard three sounds that were visually marked by three intervals on a computer screen. One of the three intervals contained the signal, whereas the other two contained the standard. The order of the signal and standard sounds was randomized (3-alternative). The subject had to choose the signal (forced-choice) and was given a visual feedback regarding the correct response. The initial difference between the signal and the standard was large so it was easy for the subject to tell which interval contained the signal. The difference was reduced after two consecutive correct responses and increased after one incorrect response (2-down, 1-up). A reversal was recorded when the subject made an incorrect response from two or more consecutive correct responses or vice versa. A large step size was used for the first three or four reversals, and a small step size was used for the remaining reversals (adaptive step size). Each run had a total of 12 reversals. The reported data were averaged from the last eight reversals.

An additional seven-alternative, forced-choice procedure was employed for sound localization measurement. The seven alternatives are represented by integers from −3 to 3, corresponding to the leftmost and the rightmost position (with the center position being represented by 0). The seven numbers were labeled with seven boxes that were geometrically arranged to form a semicircle on a computer screen. The subject had to choose a number corresponding to his or her perceived sound position. If the sounds from two ears could not be fused, two separate images might be perceived. In this case, the subject was instructed to report the separate images without having to choose any alternative. However, this situation was not encountered in this study.

Finally, a subjective report and counting procedure was employed to measure the presence of monaural and binaural beats. The subject was instructed to pay attention to loudness fluctuation and report the number of beats heard. In the monaural beat experiment, the modulation threshold was measured using the same three-interval, two-up and one-down, adaptive procedure as described above.

**Statistical analysis**

A between-subjects ANOVA with uneven sample sizes was used to test whether there was a significant difference in performance between the AN and normal-hearing subjects. Should the overall difference reach the significance level (P < 0.05), posthoc tests were usually conducted to examine the conditions under which the difference was significant. ANOVA was conducted with SPSS for Windows (version 11.0 2002, SPSS, Chicago, IL).

**RESULTS**

**Intensity discrimination**

Figure 2 shows intensity discrimination as a function of sensation level (dB above the individual subject’s threshold) in eight AN subjects (●) and eight normal controls (○). Both groups showed the typical pattern that had been termed as “the near-miss to Weber’s law,” in which the difference limen decreased slightly as a function of the stimulus level (McGill and Goldberg 1968). Although the AN subjects showed slightly larger difference limens at low levels than the normal controls, no significant main effect was observed between groups [F(1,69) = 3.17; P > 0.05]. This result indicates that AN subjects encounter no significant difficulty in performing pure-tone intensity discrimination.

**Frequency discrimination**

Figure 3 shows frequency discrimination as a function of standard frequency in 12 AN subjects (●) and 4 normal controls (○). A significant difference in performance was observed between the AN subjects and the normal controls [F(1,86) = 5.53; P < 0.05]. The normal controls required <10 Hz to discriminate a pitch difference for frequencies ≤ 1,000 Hz, but the AN subjects required a difference that was about two orders of magnitude higher than the normal difference limen. Interestingly, the difference between the two groups reduced with frequency and was not significant at 8,000 Hz (P > 0.10). This result suggests that AN subjects have pro-
found impairment in pitch discrimination at low frequencies (<4,000 Hz) but not at high frequencies (>4,000 Hz).

Temporal processing

Figure 4 shows temporal integration as a function of duration in 16 AN subjects (●) and 4 normal controls (▼). Both AN and normal-hearing subjects showed a 100- to 200-ms course of temporal integration, but the slope of the temporal integration function was slightly elevated in the AN subjects (−3.9 dB per doubling duration) compared with the normal-hearing subjects (−3.0 dB per doubling duration). A between-subjects ANOVA revealed a significant group effect between the AN and control subjects \( F(1,99) = 5.81; P < 0.05 \). The normal controls required about a 50-ms silent interval to detect a gap at 5-dB low sensation (very soft sound) but improved to 3 ms at high sensation levels (40 and 50 dB). The AN subjects performed similarly to the normal controls at low sensation levels (5 and 10 dB) but required significantly longer gaps (15–20 ms) than the normal-hearing subjects at higher sensation levels (postdoc \( t \)-test, \( P < 0.05 \)). This result suggests that the AN subjects have difficulty in gap detection even at comfortable loudness levels. Figure 6 shows modulation detection as a function of modulation frequency in 16 AN subjects (●) and 4 normal controls (▼). There was a significant group effect between the AN and control subjects \( F(1,107) = 111.10; P < 0.001 \). The normal controls showed a typical low-pass pattern, with peak sensitivity of −19.9 dB (10% modulation) and 3-dB cut-off frequency of 258.1 Hz (\( r = 0.96 \)). The AN subjects showed a lower peak
sensitivity of $-8.7$ dB (37% modulation) and a lower cut-off frequency of 17.0 Hz ($r = 0.81$). The relatively poor fit in AN subjects was due to the band-pass characteristic in the data. This result suggests that AN subjects have difficulty in detecting both slow and fast temporal modulations.

**Backward and forward masking**

Figure 7 shows backward (left) and forward (right) masking as a function of signal delay in four AN subjects (●) and four normal controls (○). Both masking conditions produced a significant group effect between the AN and normal-hearing subjects [$F(1,33) = 52.39, P < 0.001$ for backward masking and $F(1,36) = 10.43, P < 0.05$ for forward masking]. In backward masking, the normal controls showed $\leq 15\%$ masking when the masker and the signal was separated by 20 ms or longer, whereas the AN subjects still showed 60% masking even at 100-ms signal delay. In forward masking, both the normal and AN subjects showed a 100- to 200-ms recovery time, with the AN subjects having significantly more masking than the normal controls between 5- and 50-ms delays. Note also the clear asymmetrical pattern between backward and forward masking in the normal controls and the relatively symmetrical masking patterns in the AN subjects. This result suggests that AN subjects cannot effectively separate sounds occurring successively.

**Simultaneous masking: long tones**

Figure 8 shows detection of a 200-ms, 1,000-Hz pure-tone signal in noise as a function of noise level in seven AN subjects (●) and five normal controls (○). The dotted line represents the predicted detection threshold using the equivalent rectangular bandwidth (ERB) of the presumed auditory filter at 1,000 Hz (Moore and Glasberg 1987). The normal controls showed essentially the same amount of masking as predicted by the ERB model, whereas the AN subjects showed significant excessive masking of about 20 dB [$F(1,38) = 71.10, P < 0.001$]. This result shows that AN subjects have difficulty in detecting signals in noise.

**Simultaneous masking: short tones**

Figure 9 presents six individual cases to show both unique and common features in simultaneous masking by AN subjects. First, excessive masking was also observed at frequencies other than 1,000 Hz, including 250 Hz (AN10), 2 kHz (AN12 and AN14), and 4 kHz (AN7, AN10, and AN16). Second, excessive masking was observed independently of the threshold at the test frequency, which could be normal (AN7 and AN10 at 4 kHz) or elevated (all the remaining conditions). Third, the slope of the masking function varied between subjects, including 1) a relatively normal slope, namely, a 1-dB increase in noise level produced a 1-dB increase in the signal detection threshold (AN7 at 4 kHz, AN10 at 4 kHz, AN14, and AN16), 2) an abnormally steeper slope, in which a 1-dB increase in noise level produced a 2-dB increase in signal detection threshold (AN12 and AN13), and 3) an abnormally shallower slope, in which a 2- or 3-dB increase in noise level produced a 1-dB increase in signal detection threshold (AN7 at 1 kHz and AN10 at 250 Hz). Fourth, excessive masking could be produced by extremely low-level noise (AN10 at 4 kHz with $-20$ dB noise level, AN12 with 10 dB noise level, and AN16 with $-20$ dB noise level). Finally, excessive masking may be observed in both adult and children subjects. At the time of testing, AN12 was 7 yr old, with her 10-yr-old sister serving as a control; AN14 was 13 yr old, with his 14-yr-old brother serving as a control; and AN16 was 17 yr old, with her 14-yr-old sister serving as a control.

**Simultaneous masking: short tones**

Figure 10 presents individual cases to examine the masking effect on short-duration signals from six AN subjects who had $\geq 30$ dB HL hearing at tested frequencies and five normal controls who were either age-matched (AN10) or unaffected family members of the tested AN subject. First, recall that the AN subjects required 6- to 9-dB higher levels than the normal control to detect short signals in quiet (Fig. 4). This difficulty of detecting short-duration signals was clearly increased in noise, requiring an average 22.6 ± 9.2 (SD) dB higher signal...
level than the normal control to detect the signal in noise. The average difference was 21.4 dB for the 3-ms delay condition and 23.5 dB for the 300-ms delay condition, with no significant difference between the conditions ($t$-test, $P = 0.58$).

Second, note the apparent overshoot effect, the greater threshold at 3-ms delay than 300-ms delay in the four normal controls (AN7, AN10, AN13, and AN16), particularly at intermediate noise levels between 10 and 30 dB SPL. The overshoot effect was also present in the six AN subjects, with the largest effect of 25.3 dB for AN13 at 10-dB noise level and the smallest effect of 2.0 dB for AN18. Taking all noise levels into account, the AN and control groups have an average overshoot effect of 3.2 ± 6.8 and 5.6 ± 6.5 dB, respectively. Taking only the noise level at which the largest overshoot was observed into account, the AN and control groups have an average overshoot effect of 11.0 ± 8.5 and 12.6 ± 6.7 dB, respectively. There were no group differences ($P > 0.30$) for the overshoot effect (averaged or maximal measures). These results suggest that AN subjects generally have significant difficulty in detection of brief signals in noise in general, but they do not seem to have any additional difficulty in detecting a signal at the onset of the noise than in the steady state of the noise.

**Binaural processing: interaural level difference**

Figure 11 shows sound localization using interaural level differences (ILD) as a function of signal frequency from three AN subjects (circles and squares) and three normal controls (inverted and regular triangles). Except for producing a sound image that was lateralized to the right (about 2), a 0-dB ILD produced essentially a center position (0) for both the AN and control groups [$F(1,23) = 0.74; P = 0.40$]. On the other hand, a -10-dB ILD produced a sound image that was essentially lateralized to the left for both AN and control groups [$F(1,23) = 1.70; P = 0.21$]. This result indicates that AN subjects can use the interaural level difference cue to localize sound.

**Binaural processing: interaural time difference**

Figure 12 shows sound localization using ITDs as a function of phase difference for a 500-Hz sinusoidal sound in three AN subjects (○) and three normal controls (△). As the phase difference (with right ear leading) was increased, the normal control group reported a sound image from the center position (0) to the rightmost position (3), whereas the AN subjects reported no changes in the perceived sound position.
FIG. 10. Simultaneous masking of short-duration tones in noise in 6 individual AN subjects (filled symbols) and their corresponding normal-hearing controls (open symbols). Triangles represent the onset condition in which the tone is presented only 3 ms after the onset of the noise, whereas squares represent the steady-state condition in which the tone is presented about 300 ms after the onset of the noise. Difference in detection threshold between the onset and the steady-state condition is referred to as the overshoot.

FIG. 11. Sound localization using the interaural level difference (ILD) cue in 3 AN (filled symbols) and 3 normal-hearing (open symbols) subjects. Lateralization (−3 representing the left position, 0 the center position, and 3 the right position) is plotted as a function of test frequency (Hz). The 0-dB ILD condition is represented by either circles in AN subjects or inverted triangles in normal controls. The −10-dB ILD condition is represented by either squares in AN subjects or triangles in normal controls. Error bars, ±SE.

FIG. 12. Sound localization using the interaural time difference (ITD) cue in 3 AN (●) and 3 normal-hearing (○) subjects. Lateralization (0 representing the center position and 3 the right position) is plotted as a function of interaural phase difference between a pair of 500-Hz pure tones (0, 30, 60, and 90° corresponding to 0-, 167-, 333-, and 500-μs lead time in the right ear). Error bars, ±SE.
[\(F(1,16) = 114.8; P < 0.0001\)]. This result indicates that AN subjects cannot use the interaural time difference cue to localize sound.

**Binaural processing: beats and fusion**

Both the AN and normal control subjects could perceive the monaural beats and reported the correct number of beats in the stimulus. Figure 13 shows the modulation threshold for monaural beat detection from three AN subjects (●) and three normal controls (○). The AN subjects could detect about 10% modulation (~20 dB), whereas the normal control subjects could detect about 4% modulation (~28 dB). However, this difference was not significant \([F(1,18) = 3.0; P = 0.10]\) due to both the relatively small sample size \((n = 3)\) and the large individual variability in this task.

After performing the monaural beat experiment, verifying that the AN subjects could perceive reliably monaural beat sensation, the binaural beat experiment was performed with a 500-Hz tone being presented to the left ear and a 503-Hz tone being presented to the right ear. Both the AN and normal control reported a fused auditory image, with no beat sensation reported by the AN subjects, but a clearly audible and reliable beat sensation reported by the normal control. Because detection of monaural beats requires only spike synchrony to the 3-Hz modulation in the waveform envelope, whereas detection of binaural beats requires spike synchrony to the rapidly varying carrier frequency (500 Hz in 1 ear and 503 Hz in the other), these results suggest that AN subjects can follow slow temporal fluctuation (3 Hz) to perceive monaural beats but cannot follow fast fluctuation (500 and 503 Hz) to perceive binaural beats.

**Comparison with other hearing disorders**

The most common type of hearing disorder is of the cochlear origin, which is usually caused by noise trauma and/or ototoxic drugs, resulting in damaged inner and outer hair cells (Liberman and Dodds 1984). Perceptual consequences of cochlear damage have been extensively studied (Moore 1996). Here we highlight the differences as well as the similarities in perceptual consequences between AN and cochlear damage.

First, cochlear damage changes intensity related perception. Cochlear damage produces steeper than normal loudness growth (loudness recruitment) and better than normal intensity discrimination at equal sensation levels (Fowler 1936; Turner et al. 1989). AN produces no significant effect on intensity discrimination, and if anything at all, slightly worsens performance at low sensation levels (Fig. 2). We also measured loudness growth function in two AN subjects and found no sign of loudness recruitment (Zeng et al. 2001a).

Second, cochlear damage impairs frequency discrimination uniformly across all frequencies but generally does not increase the difference limen by more than one order of magnitude of the normal values (Freyman and Nelson 1991). In contrast, AN impairs frequency discrimination selectively, increasing the difference limen by almost two orders of magnitude at low frequencies but producing no significant effect on the difference limen at high frequencies (Fig. 3). The only similarity is the lack of significant correlation between the value of difference limen and the amount of hearing loss at the test frequency in both AN and cochlear damage.

Third, after taking reduced audibility and nonlinear compression into account (Oxenham and Bacon 2003), cochlear damage usually does not impair temporal processing, such as temporal integration function (Florentine et al. 1988), gap detection (Florentine and Buus 1984; Nelson and Thomas 1997), temporal modulation detection (Bacon and Gleitman 1984; Nelson and Thomas 1997), gap detection (Fig. 5), temporal modulation detection (Fig. 6), and forward and backward masking (Nelson and Freyman 1987). However, these data show significantly impaired temporal processing in AN as measured by the same tasks, including detection of short sounds (Fig. 5), gap detection (Fig. 5), temporal modulation detection (Fig. 6), and forward and backward masking (Fig. 7). We would like to reaffirm that neither audibility nor cochlear compression is likely a confounding factor, because many of the AN subjects tested here have only mild-to-moderate hearing and all of them have preserved outer hair cell functions as evidenced by the presence of otoacoustic emission and/or cochlear microphonics. We also note that the impaired temporal processing in AN is similar to that in other hearing disorders that have
confirmed neurological origins such as the presence of tumors in the auditory nerve (Formby 1986).

Although cochlear damage widens the auditory filter, it generally will not increase the detection threshold for long-duration tones in noise beyond several decibels if the damage involves only the outer hair cells (Moore et al. 1993). As a matter of fact, one would expect to only have a 3-dB increase in the threshold even if the auditory filter’s bandwidth is doubled. The slope of the growth of the masking function is typically 1 in both normal-hearing and cochlear-impaired subjects, implicating that a 1-dB increase in noise level will produce a 1-dB masking effect on detection of the tone (Stelmachowicz et al. 1987). However, the average detection threshold of tones in noise is increased by 20 dB (Fig. 8) or more for short tones (Fig. 10), and the slope of the growth of masking is often >1 (Fig. 9) in AN. This excessive masking has also been observed in listeners with dead regions or inner hair cell loss but rarely in listeners with outer hair cell loss (Moore 2004). The presence of the overshoot effect in AN (Fig. 10) is also in contrast to the greatly reduced overshoot associated with cochlear damage (Bacon and Takahashi 1992), even when the damage is temporary and reversible, such as that induced by taking aspirin (McFadden and Champlin 1990).

Finally, after taking audibility and asymmetric hearing loss into account, cochlear damage typically has little or no effect on binaural tasks, such as sound localization using interaural level and timing differences (Hall et al. 1984; Hauser et al. 1983; Hawkins and Wightman 1980; Smoski and Trahiotis 1986). Similar to cochlear damage, AN subjects can use the interaural level difference to localize sound (Fig. 11). Different from cochlear damage, AN subjects cannot use the interaural timing difference at all (Fig. 12).

In summary, these findings show that perceptual consequences of disrupted auditory nerve activity are significantly different from those of hearing loss of cochlear origin. On the one hand, perception related to intensity is usually normal in AN but severely affected with cochlear damage. On the other hand, perception related to timing is significantly affected in AN but relatively normal with cochlear damage. These differences suggest that different physiological mechanisms are likely to be involved in AN and cochlear damage.

**Physiological mechanisms**

The term AN was coined because many of the subjects initially studied had some form of peripheral neuropathy based on neurological examination (Starr et al. 1996). Many more patients with AN have since been identified, but the underlying physiological mechanisms remain unclear and have been a major source of debate (Berlin et al. 2003a; Rapin and Gravel 1991; Zeng and Turner 1991). Consistent with this finding, neither the desynchronized nor the reduced neural activity should affect intensity related perception. On the other hand, animal lesion studies and computational models have shown that frequency discrimination, particularly at low frequencies, is critically dependent on the presence of inner hair cells (Nienhuys and Clark 1978) and requires both phase locking and combinatorial information from many nerve fibers (Heinz et al. 2001a, b).

While it is apparent that desynchronized neural activity impairs temporal processing, it is less apparent how reduced
spike count can also impair temporal processing. Referring back to the gap detection data (Fig. 5), we note clearly that both the normal-hearing and the AN subjects performed poorly at low sensation levels, producing about a 50-ms gap detection threshold at 5 dB SL. At low levels, a sound would likely activate a small number of nerve fibers in both normal and neuropathy cases. However, as the level is increased, the sound would recruit and activate more nerve fibers in normal-hearing subjects but not in AN subjects due to the reduced number of receptors, nerve fibers, or both. Therefore the gap detection improves with level in normal-hearing subjects but not in AN subjects.

Finally, both the desynchronized spike model and the receptor/neuron loss model can explain the increased threshold for detection of tones in noise. In the desynchronized spike model, the detection threshold can be significantly increased because the more sensitive phase-locking cue is absent; the less sensitive overall rate cue has to be used (Colburn et al. 2003). In the receptor/neuron loss model, the detection threshold can be significantly increased because there is essentially a hole or dead region in the cochlea, forcing the subject to use a less sensitive “off-frequency” cue (Moore 2004).

At present, no neurophysiological or psychoacoustic tests can reliably differentiate between these two models. For example, statistical analysis found no significant differences in intensity, frequency, and temporal processing between the 7 AN subjects who had confirmed peripheral neuropathy and the remaining 14 subjects who did not. Since many of the AN subjects are expected to receive a cochlear implant, future studies using electric stimulation of the auditory nerve might help differentiate the site of lesion in AN. Should the damage be mostly restricted to the inner hair cell loss, cochlear implants would be able to effectively stimulate the residual neurons, restoring evoked brain stem potentials. Indeed, electric stimulation has been reported to restore electrically evoked brain stem potentials in several AN subjects who have received a cochlear implant (Shallop et al. 2001; Starr et al. 2004).

Translational research

In addition to its apparent theoretical implications on the relationship between neural coding and sensory perception, this study sheds light into better diagnosis and treatment options for persons with AN. An effective and efficient means of screening for AN is the gap detection task at high sensation levels. We have developed a web-based gap detection program that can be used (http://www.ucihs.uci.edu/hesp/webtest/gap-detection/ie_gap.html). We have successfully used this JAVA-based program to study a large family with inherited deafness likely affecting the inner hair cells and the peripheral process of the auditory nerve fibers (Starr et al. 2004). An enhanced off-line version of this program is available for nonprofit use by writing to the corresponding author of this study.

At present, conventional hearing aids provide no or minimal benefits to alleviate the unique difficulty associated with AN, particularly in speech recognition with noise (Berlin et al. 2003a,b). These perceptual data suggest several innovative signal processing algorithms that may lead to improved performance in AN (Zeng 2000; Zeng and Liu 2005). One idea is to eliminate low-frequency components but to preserve, or perhaps emphasize, high-frequency components in speech sounds. This idea is based on the present observation that many AN subjects have extremely poor frequency discrimination at low frequencies but nearly normal discrimination at high frequencies. The low-frequency sound could cause undesirable masking and might be better off filtered out. Another idea is to accentuate the temporal waveform modulation in speech sounds to compensate for the impaired temporal processing in AN. These features are not available in today’s hearing aid devices but may provide significant benefits to persons with AN.

Yet another treatment option is cochlear implant action. Some affected persons have already received a cochlear im-

---

**FIG. 14.** Phenomenological models of AN. A: normal auditory pathway converting the “gap” stimuli (bottom trace) via 3 synchronized nerve fibers (<0.5 ms; top 3 traces) into an undistorted central representation of the gap at the output (4th trace). B: 1 AN model with desynchronized nerve conduction, in which the central representation of the gap is distorted due to different delays (~1.5 ms). C: another AN model with reduced nerve conduction, in which the central representation of the gap is also difficult to detect because of its similarity to the background spontaneous activity.
plant or will receive an auditory brain stem implant should the brain stem implant performance justify the risk (Brackmann et al. 1993). The present data and models can potentially provide guidance to these treatment options. For example, a presurgical promontory stimulation may determine the extent to which electric stimulation may restore neural synchrony in cochlear implant candidates (Kileny et al. 1994). While the absence of electrically evoked potentials may not be a good indicator of the postsurgical performance (Nikolopoulos et al. 2000), its presence would certainly increase the likelihood of success in restoring neural synchrony via a cochlear implant in persons with AN. Another example is appropriate programming of the cochlear implant in persons with AN. Fast rate of stimulation to improve temporal encoding is a current trend in cochlear implant programming (Rubinstein et al. 1999). Should demyelination and axonal loss be a significant factor in AN, high-rate electric stimulation may produce adverse effects such as neural fatigue or even nonresponse (Stephanova and Daskalova 2004). Given this consideration, high-rate stimulation may provide less benefit than low-rate stimulation in persons with AN.

In summary, a systematic and comprehensive report is presented on perceptual consequences of disrupted auditory nerve activity in 21 persons who were clinically diagnosed with AN. The result shows that, in AN subjects, intensity-related perception is nearly normal, but temporal processing is impaired. A particularly interesting finding was that frequency discrimination is severely impaired at low frequencies but normal at high frequencies, reinforcing the duplex encoding of pitch using the phase-locking cue at low frequencies and the place cue at high frequencies. These perceptual consequences are significantly different from those caused by the more commonly observed hearing loss of cochlear damage, which has typically impaired intensity perception but relatively normal frequency and temporal processing once the impaired intensity perception is taken into account. These results strongly suggest the use of different neural codes in auditory perception: a suboptimal spike count code for intensity perception, a synchronized spike code for temporal processing, and a duplex code for frequency processing. Two AN models based on desynchronized discharge and neuronal loss are proposed to account for the presently observed perceptual consequences. Although current psychoacoustic and electrophysiological methods cannot differentiate between these two AN models, electric stimulation of the auditory nerve via a cochlear implant may allow us to differentiate between them. Nevertheless, these results suggest several translational research ideas that can potentially improve the diagnosis and treatment of AN.

REFERENCES


