Cutting Through the Noise: Noise-induced Cochlear Synaptopathy and Individual Differences in Speech Understanding Among Listeners with Normal Audiograms

Mishaela DiNino, Lori L. Holt, and Barbara G. Shinn-Cunningham

Carnegie Mellon University Department of Psychology and Neuroscience Institute, 5000 Forbes Ave., Pittsburgh, PA 15213

Financial disclosures/conflicts of interest: This study was funded by NIH NIDCD F32 DC018460 (DiNino), NIH NIDCD T32 DC011499 (DiNino), NIH NIDCD R01 DC017734 (Holt), and NIH NIDCD R01 DC013825 (Shinn-Cunningham). There are no conflicts of interest, financial, or otherwise.

All correspondence should be addressed to: Mishaela DiNino, Carnegie Mellon University Department of Psychology, 5000 Forbes Ave. Pittsburgh PA 15213. E-mail: mdinino@andrew.cmu.edu

Abstract

Following a conversation in a crowded restaurant or at a lively party poses immense perceptual challenges for some individuals with normal hearing thresholds. A number of studies have investigated whether noise-induced cochlear synaptopathy (CS; damage to the synapses between cochlear hair cells and the auditory nerve following noise exposure that does not permanently elevate hearing thresholds) contributes to this difficulty. A few studies have observed correlations between proxies of noise-induced CS and speech perception in difficult listening conditions, but many have found no evidence of a relationship. To understand these mixed results, we reviewed previous studies that have examined noise-induced CS and performance on speech perception tasks in adverse listening conditions in adults with normal or near-normal hearing thresholds. Our review suggests that previous investigations, which used superficially similar speech perception paradigms, actually placed very different demands on sensory, perceptual, and cognitive processing. Only speech perception tests that use low signal-to-noise ratios and maximize the importance of fine sensory details— specifically by using test stimuli for which lexical, syntactic, and semantic cues do not contribute to performance— show any relationship to estimated CS levels. Thus, the current controversy as to whether or not noise-induced CS contributes to individual differences in speech perception under challenging listening conditions may be due in part to the fact that many of the speech perception tasks used in past studies are relatively insensitive to CS-induced deficits.

Keywords: cochlear synaptopathy, obscure auditory dysfunction, speech perception, speech in noise

Introduction

A number of animal models demonstrate cochlear synaptopathy, a loss of the synapses between inner hair cells and the auditory nerve, following exposure to high-intensity noise, even if the damage does not result in a permanent increase in hearing thresholds (Furman, Kujawa, & Liberman, 2013; Kujawa & Liberman, 2009; Valero et al., 2017). Less clear is the extent to which noise-induced CS occurs in humans and, if it does, whether it precipitates any perceptually relevant deficits. A large number of carefully controlled studies in humans with normal hearing thresholds (NHTs) have failed to find relationships between performance on perceptual tasks and proxies of noise-induced CS, such as noise exposure history or auditory nerve (AN) integrity metrics. These negative results have called into question the link between CS and clinically relevant perceptual impairments, and even the very existence of noise-induced CS in the human population (e.g., Johannesen, Buzo, & Lopez-Poveda, 2019; Le Prell, Siburt, Lobarinas, Griffiths, & Spankovich, 2018; Prendergast et al., 2017).

Yet, interest in noise-induced CS persists because evidence in animal models suggests that it may contribute to a particularly distressing auditory perceptual deficit: impaired speech-in-noise perception in adults with NHTs. Since these individuals have normal audiograms, they are not diagnosed as having traditional hearing loss; instead, they are labelled as having auditory processing disorder (American-Speech-Language-Hearing Association, 2005), King-Kopetzky syndrome (Hinchcliffe, 1992), or obscure auditory dysfunction (Saunders & Haggard, 1989). Such symptoms have been linked to various deficits in peripheral (Badri, Siegel, & Wright, 2011; Shaw, Jardine, & Fridjhon, 1996; Zhao & Stephens, 2000; Zhao & Stephens, 2006) and central (Jerger et al., 1991; Saunders & Haggard, 1992; Zhao & Stephens, 2000) processing. CS may be an additional candidate to explain this constellation of symptoms: the synaptic loss reduces the number of available AN fibers (particularly fibers with relatively greater importance for encoding loud sound; Furman et al., 2013) and is thus thought to impair perception of speech in the presence of competing auditory signals much more than it affects speech perception in quiet (e.g., Lopez-Poveda, 2014).

In animal models, CS also occurs with aging (e.g., Sergeyenko, Lall, Liberman, & Kujawa, 2013) and noise-induced CS accelerates natural age-related CS (Fernandez, Jeffers, Lall, Liberman, & Kujawa, 2015; Liberman & Kujawa, 2017). Whether CS is caused by noise exposure, the aging process, or both, its perceptual consequences are similar: the common denominator is damage to the synapse. As temporal bone studies suggest that age-related CS does occur in humans (Makary, Shin, Kujawa, Liberman, & Merchant, 2011; Viana et al., 2015; Wu et al., 2019), it is important to explore not only whether noise-induced CS exists in humans, but how it may exacerbate effects of age-related CS.

There are no direct assessments of CS in living humans, complicating attempts to link this synaptic damage to auditory perceptual impairments. Previous investigations have instead relied on indirect proxies, including self-reported noise exposure history and physiological measures that correlate with CS in animal models. Most studies of human CS, as well as several reviews (Bharadwaj et al, 2019; Bramhall et al., 2019; Le Prell, 2019), have acknowledged and discussed the limitations of the metrics used to assess risk of noise-induced CS among humans. As these authors point out, inconsistent results from previous studies may be due in part to the fact that existing (indirect) methods to quantify CS in humans are unreliable—a point to keep in mind as we consider past work.

Here, we review results from 23 studies that asked whether individual differences in the ability to understand speech amongst listeners with NHTS are related to any proxy of noise-induced CS. Across the studies, these proxies include noise exposure history metrics as well as electrophysiological measures of peripheral auditory function (the auditory brainstem response [ABR] wave I amplitude, ABR Wave I/Wave V ratio, summating potential [SP]/action potential [AP] ratio, ABR Wave I growth function in response to increasing sound intensity, envelope following response [EFR], and middle ear muscle reflex [MEMR]). Together, these studies included 41 separate experiments (See Table 1, which summarizes methods and the factors that contribute to good performance for each experiment).

Of the 41 experiments reviewed, only 13 (32%) observed a significant relationship between speech perception performance and a proxy of noise-induced CS (highlighted in gray, Table 1). With less than a third of the literature observing a significant relationship between speech understanding performance and CS proxies, one might justifiably question whether noise-induced CS even occurs in humans. Still, as we describe, the speech perception tasks used in the studies reviewed here placed very different demands on the listener. Some emphasized sensory processing, while others used tasks in which other perceptual and cognitive processes contribute to performance, which may have obscured subtle perceptual deficits caused by CS.

Experiment #	Study #	Reference	Participant Ages	Task	Target Speech Stimuli	Competing Stimuli	Response Set	Presentation Mode	Contributing Factors
Studies using high context speech									
	Studies using unintelligible competing sound								
1	1	Grose et al., 2017	18-35	Sentence identification (Modified BKB-SIN Test)	High-context sentences	Speech-shaped noise	Open	Monaural headphone	Context effects, lexical knowledge, energetic masking
2	2	Johannesen et al., 2019	12-68	Sentence identification (HINT)	High-context sentences	Speech-shaped noise	Open	Monaural headphone	Context effects, lexical knowledge, energetic masking
3	2	Johannesen et al., 2019	12-68	Sentence identification (HINT)	High-context sentences	Speech-like fluctuating signal (International Female Fluctuating Masker, IFFM; Holube et al., 2011)	Open	Monaural headphone	Context effects, lexical knowledge, segregation / selection
	Studies using intelligible competing speech								
4	3	Valderrama et al., 2018	18-55	Sentence identification (LiSN-S Test)	High-context sentences (0˚)	Two streams of ongoing stories from different talkers (+90˚ and -90˚)	Open	Binaural headphone, HRTF-separated speech and noise	Context effects, lexical knowledge, non-spatial and spatial segregation / selection
5	4	Yeend et al., 2017	30-60	Sentence identification (LiSN-S Test)	High-context sentences (0˚)	Two streams of ongoing stories from different talkers (+90˚ and -90˚)	Open	Binaural headphone, HRTF-separated target and noise	Context effects, lexical knowledge, non-spatial and spatial segregation / selection

6	4	Yeend et al., 2017	30-60	Speech comprehension (NAL Dynamic Conversations Test)	Four min, high-context speech monologues (0˚)	Conversational noise (distinct talkers, taking turns, at various locations)	Open	Multi-speaker soundfield simulation in anechoic chamber	Context effects, lexical knowledge, non-spatial and spatial segregation / selection
Studies using low-context sentences									
7	5	Bramhall et al., 2015a	19-90	Sentence identification (QuickSIN)	Low-context sentences	Multitalker babble	Open	Monaural headphone	Context effects, lexical knowledge, energetic masking, non-spatial segregation / selection
8	6	Skoe et al., 2019	18-24	Sentence identification (QuickSIN)	Low-context sentences	Four-talker babble	Open	Diotic headphone	Context effects, energetic masking, non-spatial segregation / selection
9	7	Smith et al., 2019	18-30	Sentence identification (QuickSIN)	Low-context sentences	Four-talker babble	Open	Diotic headphone	Context effects, energetic masking, non-spatial segregation / selection
10	8	Grant et al., 2020	18-63	Sentence identification (Modified QuickSIN)	Low-context sentences	Four-talker babble	Open	Monaural headphone	Context effects, lexical knowledge, energetic masking, non-spatial segregation / selection
11	9	Mepani et al., 2020	18-63	Sentence identification (Modified QuickSIN)	Low-context sentences	Four-talker babble	Open	Monaural headphone	Context effects, lexical knowledge, energetic

									masking, non-spatial segregation / selection
Studies using speech without semantic or syntactic context									
	Studies with no competing sound								
12	10	Liberman et al., 2016	18-41	Word identification	NU-6 words, 45% time compression, 0.3s reverb	None	Open	Monaural headphone	Lexical knowledge, speech feature coding
13	10	Liberman et al., 2016	18-41	Word identification	NU-6 words, 65% time compression, 0.3s reverb	None	Open	Monaural headphone	Lexical knowledge, speech feature coding
14	9	Mepani et al., 2020	18-63	Word identification	NU-6 words, 45% time compression, 0.3s reverb	None	Open	Monaural headphone	Lexical knowledge, speech feature coding
15	9	Mepani et al., 2020	18-63	Word identification	NU-6 words, 65% time compression, 0.3s reverb	None	Open	Monaural headphone	Lexical knowledge, speech feature coding
16	8	Grant et al., 2020	18-63	Word identification	NU-6 words, 45% time compression, 0.3s reverb	None	Open	Monaural headphone	Lexical knowledge, speech feature coding
17	8	Grant et al., 2020	18-63	Word identification	NU-6 words, 65% time compression, 0.3s reverb	None	Open	Monaural headphone	Lexical knowledge, speech feature coding
18	11	Kamerer et al., 2019	20-86	Word identification	NU-6 words, 45% time compression	None	Open	Monaural headphone	Lexical knowledge, speech feature coding
19	11	Kamerer et al., 2019	20-86	Word identification	NU-6 words, 45% time compression, 0.3s reverb	None	Open	Monaural headphone	Lexical knowledge, speech feature coding
	Studies using unintelligible competing sound								

20	10	Liberman et al., 2016	18-41	Word identification	NU-6 words	White noise at 0 dB SNR	Open	Monaural headphone	Lexical knowledge, energetic masking
21	10	Liberman et al., 2016	18-41	Word identification	NU-6 words	White noise at +5 dB SNR	Open	Monaural headphone	Lexical knowledge, energetic masking
22	9	Mepani et al., 2020	18-63	Word identification	NU-6 words	Speech-shaped noise	Open	Monaural headphone	Lexical knowledge, energetic masking
23	12	Shehorn et al., 2020	21-54	Word identification (Modified MD CNC Test)	Words with reverb	Speech-shaped noise	Open	Diotic headphone	Energetic masking
24	8	Grant et al., 2020	18-63	Word identification	NU-6 words	Speech-shaped noise	Open	Monaural headphone	Lexical knowledge, energetic masking
25	11	Kamerer et al., 2019	20-86	Word identification	NU-6 words	Noise (type not reported)	Open	Monaural headphone	Lexical knowledge, energetic masking
26	2	Johannesen et al., 2019	12-68	Word identification	Disyllabic words	Speech-shaped noise	Open	Monaural headphone	Lexical knowledge, energetic masking
27	2	Johannesen et al., 2019	12-68	Word identification	Disyllabic words	Speech-like fluctuating signal (IFFM)	Open	Monaural headphone	Lexical knowledge, non-spatial segregation / selection
28	13	Fulbright et al., 2017	18-30	Word identification (The WIN Test)	NU-6 words	Multitalker babble	Open	Monaural headphone	Lexical knowledge, energetic masking, non-spatial segregation / selection

29	13	Fulbright et al., 2017	18-30	Word identification (The Words in Broadband Noise Test)	NU-6 words	Broadband noise	Open	Monaural headphone	Lexical knowledge, energetic masking
30	14	Grinn et al., 2017	21-27	Word identification (The WIN Test)	NU-6 words	Multitalker babble	Open	Monaural headphone	Lexical knowledge, energetic masking non-spatial segregation / selection
31	15	Le Prell et al., 2018	18-27	Word identification (The WIN Test)	Words	Multitalker babble	Open	Monaural headphone	Lexical knowledge, energetic masking, non-spatial segregation / selection
32	16	Hope et al., 2013	24-39	Syllable identification	VCV syllables	ICRA noise	Closed	Diotic headphone	Energetic masking
33	17	Prendergast et al., 2017b	18-36	Digit stream identification (Digit Triplet Test)	Digit streams	Speech-shaped noise	Closed	Diotic headphone	Energetic masking
34	18	Prendergast et al., 2019	18-60	Digit identification (Digit Triplet Test)	Digit streams	Speech-shaped noise	Closed	Diotic headphone	Energetic masking
	Studies using intelligible competing speech								
35	19	Ruggles et al., 2011	18-55	Digit stream identification	Monotonized digit streams (0˚), varying reverb	Two digit streams identical to target, from -15˚ and +15˚	Closed	Binaural headphone, HRTF-separated speech and noise	Spatial segregation / selection
36	20	Bharadwaj et al., 2015	21-39	Digit stream identification	Monotonized digit streams (ITD 50-400 µs)	Digit stream identical to target, but with	Closed	Binaural headphone, ITD-separated	Spatial segregation / selection

						ITD of opposite sign (symmetric)		speech and noise	
37	17	Prendergast et al., 2017b	18-36	Keyword identification (CRM)	Carrier phrases with callsign, color, and number keywords	Two streams identical to target, but with different keywords and talkers	Closed	Diotic headphone	Non-spatial segregation / selection
38	18	Prendergast et al., 2019	18-60	Keyword identification (CRM)	Carrier phrases with callsign, color, and number keywords	Two streams identical to target, but with different keywords and talkers	Closed	Diotic headphone	Non-spatial segregation / selection
39	21	Guest et al., 2018	18-40	Keyword identification (CRM)	Carrier phrases with callsign, color, and number keywords (0˚)	Two streams identical to target, but with different keywords, talkers, and locations (-60˚ and +60˚)	Closed	Binaural headphone, HRTF-separated speech and noise	Non-spatial and spatial segregation / selection
40	22	Couth et al., 2020	18-27	Keyword identification (CRM)	Carrier phrases with callsign, color, and number keywords (0˚)	Two streams identical to target, but with different keywords, talkers, and locations (-60˚ and +60˚)	Closed	Binaural headphone, HRTF-separated speech and noise	Non-spatial and spatial segregation / selection
41	23	Parthasarathy et al., 2020	Not reported (mean 28.3 +/- 0.9)	Digit identification	Digit streams	Two streams identical to target, but with different talkers	Closed	Binaural headphone	Non-spatial and spatial segregation / selection

Table 1. Summary of 23 previous studies, encompassing 41 separate experiments, investigating the relationship between human noise-induced cochlear synaptopathy and speech perception performance. Studies are grouped first by speech

materials (high-context sentences, low-context sentences, open-set words, closed-set words or syllables) and then, within these, by masker type (none, unintelligible maskers like noise and babble, intelligible speech). The summary shows the total age range of participants in each study (from both control and experimental groups, if applicable), the name of speech-in-noise perception tasks used (if applicable), the noise type employed by each task, the nature of the speech stimuli (e.g., open or closed set), the presentation mode, and the factors expected to contribute to performance on each task. Most descriptions of task parameters in the table are exactly as worded by the study authors. Of the 41 experiments summarized, only 13 (32%) found a positive relationship between speech perception performance and proxies of noise-induced CS (highlighted in gray). Results from the two experiments in light gray were influenced by the effects of traditional hearing loss, whereas those in dark gray were not. Note that some studies reported fewer task details than others; some details were obtained by contacting the authors. BKB-SIN: Bamford-Kowal-Bench Speech-in-Noise. CRM: Coordinate Response Measure. dB: decibels. HINT: Hearing in Noise Test. HRTF: head related transfer function. ITD: interaural time difference. LISN-S: Listening in Spatialized Noise-Sentences. MD CNC: Maryland consonant-nucleus-consonant words. NAL: National Acoustic Laboratories. NU-6: Northwestern University Auditory Test No. 6. QuickSIN: Quick Speech-in-Noise Test. SNR: signal-to-noise ratio. WIN: Words in Noise. ICRA: International Collegium for Rehabilitative Audiology.

Stimulus and Task Differences Impact Whether Relationships with CS are Observed

Table 1 highlights how previous studies of the relationship between noise-induced CS and speech perception in challenging listening contexts have used various combinations of speech stimuli, noise types, presentation modes, and response sets (open- or closed-set) in the speech perception task. Every one of these variables *on its own* can influence the specific demands of the task. As the table shows, it should not be surprising that the human CS literature has yielded inconsistent results. Experiments with seemingly similar objectives engage very different perceptual processes, depending upon the kind of target *speech* they present, whether they present that speech in *noise*—and if so, what the “noise” characteristics are, and how they measure the joint interaction of speech and noise. Therefore, although each of the studies listed in Table 1 quantifies speech understanding, the paradigms differ in substantive ways that may affect whether or not perceptual performance is observed to relate to measures of CS.

For instance, consider two hypothetical “speech-in-noise perception” experiments: one in which a participant listens diotically to a meaningful story masked by simultaneous steady-state noise (without any envelope modulation), and one in which the participant identifies an isolated, closed-set digit presented against a competing digit spoken by the same talker, but coming from a different location in space. Each task uses “speech” presented against a competing “noise.” Yet, these tasks differ fundamentally in the demands placed on the system, the information a listener can use to understand the target speech, and the response used to measure speech comprehension. Given this, the experiments should be expected to interact differently with various auditory pathologies – including CS.

By considering the processes that impact perception of speech and differences in experimental procedures across studies, our review of the literature identifies some factors that may help explain disparate findings across these studies. The following sections describe specific issues related to previously used speech perception tasks that we believe complicate interpretation of the larger literature on the impact of noise-induced CS on speech perception under adverse listening contexts:

1) CS does not affect auditory detection thresholds and thus produces much more subtle deficits than does traditional hearing loss. This may impede attempts to link CS to performance on clinically validated speech perception tasks used to quantify traditional hearing impairments.

2) Speech perception paradigms with high ecological validity involve cognitive processes that may obscure any relationship between CS and task performance.

3) Previously used speech perception tasks vary in the degree to which they emphasize perception of temporal features. Work in animal models suggests such features are particularly susceptible to CS-induced deficits (Parthasarathy & Kujawa, 2018; Shaheen, Valero, & Liberman, 2015); therefore, tasks that most strongly emphasize temporal information may be more likely to show relationships between CS and performance.

In total, our review reveals characteristics of speech perception tasks that are likely to be sensitive to deficits caused by CS. Future studies directed at determining whether CS accounts for difficulties processing speech in challenging listening contexts may benefit from considering these issues when designing the tasks that they use.

Cochlear Synaptopathy Likely Causes Deficits Too Subtle to Influence Scores on Many Clinical Speech Perception Tests

The sensory deficits that CS may cause are likely to be rather subtle compared to those of “traditional” hearing loss (i.e., spectral loss that affects auditory detection thresholds, and thus speech perception performance in quiet). CS predominately targets synapses connected to low- and medium-spontaneous rate (SR) AN fibers, likely because of their cochlear location and relatively high susceptibility to glutamate excitotoxicity (see Figure 1; Furman et al., 2013; Liberman, Suzuki, & Liberman, 2015). CS may challenge speech perception in noisy environments, which are often loud, because individual low-SR fibers make a relatively larger contribution to auditory signal encoding as sound levels increase. For instance, at sound intensities 35-40 dB above threshold, firing rates of the low-threshold, high-SR fibers (which are not as vulnerable to CS) begin to saturate. However, low- and medium-SR fibers continue to increase their firing rates as sound level intensifies (Costalupes, 1985; Costalupes, Young, & Gibson, 1984; Liberman, 1978; Winter, Robertson, & Yates, 1990; Young & Barta, 1986). A reduced low-SR fiber population response (as demonstrated in animal models of CS) likely affects the encoding fidelity of high-intensity auditory stimuli, including sounds in loud, noisy environments, more than encoding of low-intensity sounds in quiet. This may help explain why CS could impair speech perception in noisy listening situations, but preserve speech understanding in quiet.

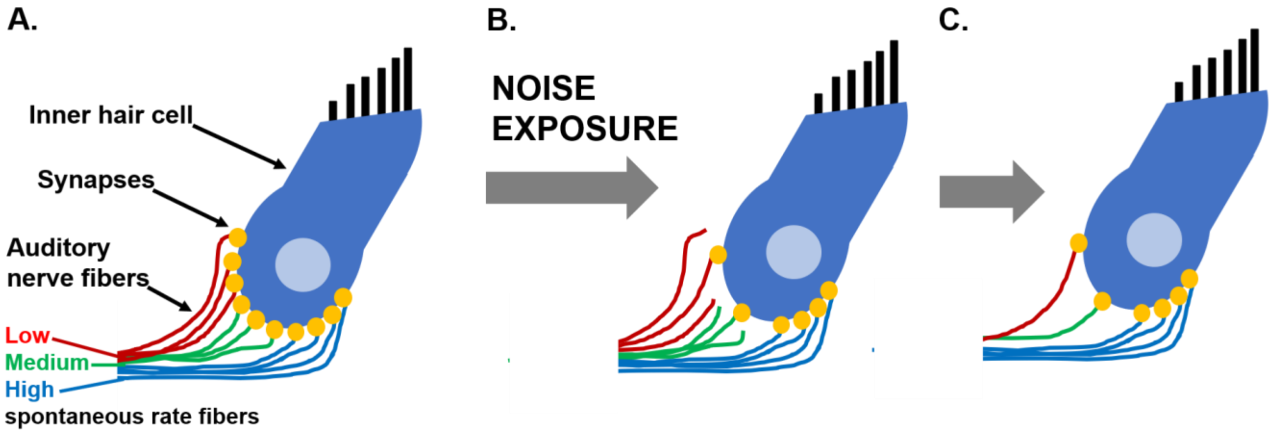

Figure 1. **Illustration of auditory nerve fiber degeneration following noise exposure.** (A) Prior to noise exposure, synapses between the pictured inner hair cell and the auditory nerve are

intact, as are auditory nerve fibers. (B) Noise exposure results in synaptic damage. (C). Auditory nerve fibers degenerate following synaptic loss. Note that the low- and medium-spontaneous rate fibers, located on the modiolar side of the inner hair cell, are particularly affected.

Low-SR fibers likely aid listening in the presence of background noise in other ways, such as their role in the descending auditory pathways, the auditory efferent system (Liberman, 1988; Ryugo, 2008; Ye, Machado, & Kim, 2000; see Carney, 2018 for full discussion). Evidence from studies in animals (Kawase, Delgutte, & Liberman, 1993; Pang & Guinan, 1997), in humans (e.g., Giraud et al., 1997; Kumar & Vanaja, 2004), and from computational modelling (Brown, Ferry, & Meddis, 2010) suggests that efferent pathways enhance sound perception in the presence of competing auditory signals by adapting to ongoing noise. Disruption of the efferent system in animal models also decreases detection and discrimination of sounds in noise (Dewson, 1968; May & McQuone, 1995). CS may thus reduce the effectiveness of the auditory efferent pathway, which is likely to be especially detrimental to understanding sound sources when levels are relatively high.

While CS preferentially affects low-SR fibers, loss of *any* type of AN fiber will have consequences on sound coding. Firing rates, and also firing synchrony, of individual AN fibers increase with sound intensity, as does the number of fibers responding to sound. However, CS-related AN degeneration reduces phase-locking of neural firing to auditory signals (Parthasarathy & Kujawa, 2018; Shaheen et al., 2015), reducing the faithfulness of auditory signal encoding and increasing noise in the auditory representation (Lopez-Poveda, 2014). Accordingly, a model of AN under-sampling (such as would occur with CS-related AN deafferentation) predicts poor sentence identification performance in noise, but not in quiet (Lopez-Poveda & Barrios, 2013).

Cochlear neuropathy from CS, then, is likely to impair perception of auditory signals by degrading temporal coding, especially of suprathreshold sound for which low-SR fibers contribute relatively more to neural coding. Thus, instead of affecting whether a listener can *detect* a sound (like traditional hearing loss), CS-related AN degeneration likely alters the fidelity of the coding of a sound's *content* (Carney, 2018; Lopez-Poveda, 2014; Lopez-Poveda and Barrios, 2013; Plack, Barker, & Prendergast, 2014).

Most clinical tests of speech perception were designed to distinguish listeners with healthy cochlear function from those with traditional hearing loss. Specifically, they have been optimized to quantify damage to the cochlear amplifier, which results in inaudibility and poor frequency selectivity. Given the sensory differences between traditional hearing loss and CS as described above, it should not be surprising that these tests are not well-suited to assessing CS-induced deficits in adults with NHTs. Such tasks utilize SNRs that may be difficult for listeners with traditional hearing loss, but inappropriate for listeners with NHTs, even those with a sensory coding deficit from CS. For example, the Words in Broadband Noise Test (used by Fulbright, Le Prell, Griffiths, & Lobarinas, 2017) and the Words in Noise Test (Wilson & Burks, 2005; used by Fulbright et al., 2017; Grinn, Wiseman, Baker, & Le Prell, 2017; and Le Prell et al., 2018) utilize SNRs ranging from +30 to +20 dB SNR and +24 dB to 0 dB SNR, respectively. In these prior studies, scores on the Words in Broadband Noise Test were not reported, but most participants with NHTs performed at ceiling on the Words in Noise Test until the SNR decreased to +8 dB SNR, leaving only 15 words (five from the three most difficult SNRs) on which participants' identification scores varied.

Similarly, the clinical version of the QuickSIN test (Killion, Niquette, Gudmundsen, Revit, & Banerjee, 2004; used by Bramhall, Ong, Ko, & Parker, 2015; Smith et al., 2019; Skoe, Camera, & Tufts, 2019) presents sentences in noise ranging from +25 dB to 0 dB SNR. This assessment is scored clinically as "SNR loss," the total number of keywords correctly repeated (out of 30) subtracted from 25.5. A recent study of young adults with NHTs confirms that many individuals without traditional hearing loss have little trouble identifying key words even at the most difficult SNR levels: participants' SNR loss fell into the limited range of -1.25 to 2.25 (with lower SNR loss representing better performance) out of the possible range of -4.50 to 25.50 (Skoe, et al., 2019). Listeners with NHTs perform very well, and very similarly, on clinical speech tests that use SNRs designed to be challenging for listeners with traditional hearing loss. It is thus unsurprising that the small variation in task performance observed across listeners with NHTs does not correlate with estimates of CS severity.

Indeed, seven of the 41 experiments we reviewed (Table 1) used one of these clinical tests, but only one found any relationship to proxies of CS (Bramhall et al., 2015). The one observing a relationship included participants with traditional hearing loss—which makes it difficult to attribute any observed relationship to CS, rather than damage to the cochlear amplifier. Thus, although existing clinical speech tests and speech corpora are useful for assessing how overt hearing loss affects speech perception, those that use high SNRs are likely to be insensitive to the more subtle differences in speech perception abilities that CS may cause. As shown in Figure 2A, after excluding the study that was influenced by the effects of traditional hearing loss, no experiments using such clinical speech-in-noise perception tests demonstrated a relationship between a proxy of noise-induced CS and speech perception scores.

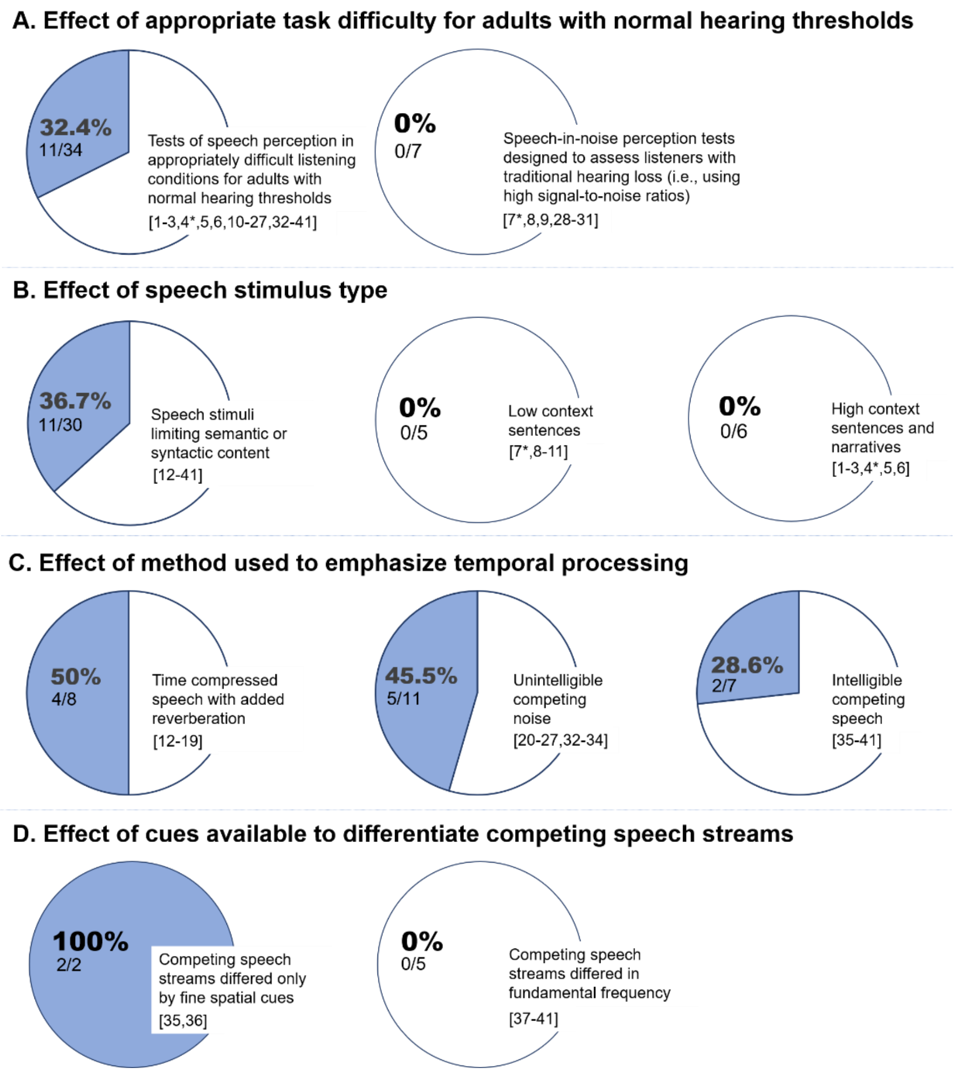

Figure 2. **Pie charts illustrating the percent of studies utilizing particular task parameters that found a relationship between a proxy of noise-induced synaptopathy and speech perception in challenging listening conditions.** The experiments that contributed to data in each pie chart are listed in brackets under each – experiment numbers refer to those from Table 1. Experiments with asterisks indicate one that yielded a positive result, but was influenced by the effects of traditional hearing loss. Experiments are classified by: (A) Suitability for assessing listeners with normal hearing thresholds, (B) Speech stimulus, (C^) Method used to emphasize temporal processing, and (D^) Cues available for differentiating between speech streams.

^Note that (C) and (D) exclude experiments that used relatively high SNRs (those in panel A, left) or sentence stimuli/narratives (those in panel B, middle and right).

High-Context Speech Materials Engage Non-Sensory Factors

Natural speech perception involves a host of cognitive processes, some of which may obscure observation of any potential relationship between impaired speech perception and *subtle* degradations in the peripheral coding of sound, such as those that CS would cause. For instance, speech perception can be guided by syntactic and semantic context that provide top-down constraints that “fill in” phonemes, syllables, or even whole words that are otherwise degraded in the input (e.g., Samuel, 1981). Tasks presenting sentences or narratives (e.g., the Dynamic Conversations Test; Best, Keidser, Freeston, & Buchholz; used by Yeend, Beach, Sharma, & Dillon, 2017) thus provide linguistic context that individuals can leverage to fill in words they did not hear clearly (although context can hinder speech identification at very low SNRs; see Marrufo-Pérez, Eustaquio-Martín, & Lopez-Poveda, 2019). The demands of natural speech processing may also reveal individual differences unrelated to CS that may confound discovery of a relationship between CS and speech perception under challenging conditions. For example, comprehension of sentences or passages requires participants to hold speech in memory before responding and captures individual differences in working memory. Thus, when a task uses meaningful sentences or stories, listeners may lean on top-down perceptual restoration to compensate for subtle sensory deficits, and/or individual differences in the cognitive processes engaged by speech perception (but unrelated to the sensory deficits of CS) may conceal possibly subtle influences of CS on speech perception performance.

Six of the 41 experiments reviewed in Table 1 utilized meaningful sentences embedded in different kinds of competing sound (e.g., HINT sentences, Listening in Spatialized Noise Sentences, Bamford-Kowal-Bench sentences, and sentences from the Dynamic Conversations Test). While one of these experiments reported a relationship between speech perception and estimated CS levels, that study did not rule out differences in individuals' hearing thresholds and also reported marginally significant effects that would not survive correction for multiple comparisons (Valderrama et al., 2018). None of the other five experiments using meaningful sentence materials found a relationship to proxies of CS (see Figure 2B; Grose, Buss, & Hall, 2017; two experiments in Johannesen et al., 2019; two experiments in Yeend et al., 2017).

Some tests reduce the influence of linguistic context effects by employing low-predictability sentences for which context provides little or no information about target words. Still, individual differences in vocabulary and access to linguistic knowledge can affect performance on even simple tasks using low-context sentences under adverse listening conditions (e.g., Banks, Gowen, Munro, & Adank, 2015; Kaandorp, Groot, Festen, Smits, & Goverts, 2015; Carroll, Warzybok, Kollmeier, & Ruigendijk, 2016). These confounds are a source of individual variation unrelated to sensory deficits, again reducing sensitivity to effects of CS.

Of the 41 experiments we reviewed, five presented low-context sentences. One reported a relationship to CS proxies; however, this experiment did not rule out effects due to elevated hearing thresholds (Bramhall et al., 2015). The remaining four reported no relationship to CS (see Figure 2B; Skoe et al., 2019; Smith et al., 2019; one experiment in Grant et al., 2020; and one in Mepani et al., 2020).

The remaining 30 prior experiments listed in Table 1 used either open-set, isolated word recognition tests or closed-set speech identification tasks. These tasks place modest demands on working memory and remove the semantic and syntactic information that could help listeners compensate for subtle sensory deficits. Importantly, as described below and as shown in Figure 2B, the only experiments that *did* find significant relationships between speech perception and CS proxies used such tasks.

Twenty of the 41 experiments we reviewed used open-set word identification tests, in which the presented word can be any possible word; participants are not limited by a set of response options. Although 12 of the 20 experiments reported null results (three in Grant et al., 2020; three experiments in Kamerer et al., 2019; two in Johannesen et al., 2019; two in Fulbright et al., 2017; Grinn et al., 2017; Le Prell et al., 2018), eight experiments did find a relationship to estimated CS levels (four experiments in Liberman, Epstein, Cleveland, Wang, & Maison, 2016; three in Mepani et al., 2020; and Shehorn, Strelcyk, & Zahorik, 2020).

Relative to open-set tasks, closed-set speech identification tests provide participants with a small number of response alternatives and thus further limit the effects of individual differences in lexical knowledge and lexical access on test performance. For instance, the Digit Triplet Test (used by Prendergast et al., 2017; Prendergast et al., 2019) requires participants to identify three digits between one and nine presented in noise. In the Coordinate Response Measure (used by Guest, Munro, Prendergast, Millman, & Plack et al., 2018; Prendergast et al, 2017; Prendergast et al., 2019), participants listen to competing streams of the form "Ready <call sign> go to <color> <number>" and are asked to report back the color (out of four options) and number (between one and four) of the stream that contains a target call sign, such as "Baron." Because of the structure of these stimuli and limited response options, all of these studies reduce reliance on cognitive factors that influence speech intelligibility in daily life. Such tests are clearly less natural than tests using sentences, or even open-set isolated word recognition tests, but are more likely to be sensitive to the impact of a subtle sensory deficit on speech intelligibility.

Ten of the experiments we reviewed used closed-set speech identification tasks. Three found that performance on the speech-in-noise task was related to proxies of CS (Hope, Luxon, & Bamiou, 2013; Ruggles, Bharadwaj, & Shinn-Cunningham, 2011; Bharadwaj, Masud, Mehraei, Verhulst, & Shinn-Cunningham, 2015), but the seven other experiments found no such relationship (two experiments in Prendergast et al., 2017; two in Prendergast et al., 2019; Guest et al., 2018; Couth et al., 2020; and Parthasarathy, Hancock, Bennett, DeGruttola, & Polley, 2020).

While each of these experiments compared perception to different CS proxies that may themselves have influenced study results, overall, this analysis suggests that studies are only likely to reveal a relationship between estimated CS levels and speech understanding if they use speech materials and tasks that minimize context effects and other non-sensory factors (see Figure 2B). This is a tradeoff: closed-set tasks do not have the ecological validity of more natural speech tasks, but cognitive factors may need to be minimized in order to observe the putative relationships between a subtle sensory deficit and speech perception. Ecological validity must be put aside, at least for the moment, in favor of accumulating a body of evidence regarding whether tasks that draw upon processes impacted by CS, in fact, influence human speech perception.

Speech Perception Tasks Vary in Emphasis on Temporal Acoustic Features

In animal models, CS-related AN degeneration degrades encoding of auditory signal timing (Parthasarathy & Kujawa, 2018; Shaheen et al., 2015). Thus, speech perception tasks requiring a listener to rely on the temporal processing important for identifying speech in noisy listening environments might be expected to correlate with measures of CS. Yet, as shown in Figure 2C, even the 26 experiments that used both 1) speech perception tasks with appropriate SNR levels for listeners with NHTs and 2) stimuli that limited non-sensory factors still varied in the methods they utilized to emphasize temporal processing.

Eight of the studies that we reviewed presented isolated words in which temporal features were degraded, thus stressing sensory coding (particularly of temporal representations) more than typical speech. Specifically, to degrade sensory features, these studies time-compressed the words, then added simulated reverberation. Of these eight studies, half found a relationship to CS (see Figure 2C; two experiments in Liberman et al., 2016; two experiments in Mepani et al., 2020) and half did not (two experiments in Kamerer et al., 2019; two experiments in Grant et al., 2020).

Presenting speech with simultaneous, competing sounds introduces greater demands on temporal processing than presenting speech in quiet. The main effects of steady-state or fluctuating noise that is dissimilar from the target is to degrade the representation of target speech features, an effect often known as energetic masking (Durlach et al., 2003b). Specifically, noise renders portions of the speech signal inaudible and reduces the prominence of amplitude modulations important for conveying speech content. Because it is not spectrotemporally sparse, multi-talker speech babble causes a fair amount of energetic masking; its effects are more similar to that of competing noise than to competing speech with the same total energy (Lu & Cooke, 2008). Thus, the primary effects of steady-state noise, fluctuating noise, and speech babble on intelligibility of speech are to reduce sensitivity to temporal features through energetic masking.

The similarity of target speech and any competing sound also influences the factors that limit speech intelligibility (Durlach et al., 2003a). If target speech is presented simultaneously with other intelligible speech, the temporal precision of the auditory representation must be good enough to support both *segregation* of the speech from the noise and *selection* of the target speech from the mixture (e.g., Shinn-Cunningham & Best, 2008). Only then can a listener successfully deploy selective attention to the target and analyze its acoustic content.

Importantly, the acoustic features that are important for source segregation and selection require temporal precision orders of magnitude more precise than those supporting speech perception in quiet or even in the presence of dissimilar noise. For instance, use of fundamental frequency differences between talkers requires temporal coding precision on the order of a few milliseconds. Differences of even a few semitones in the fundamental frequencies of competing talkers are sufficient to support segregation and selection (Binns and Culling, 2007; Madsen, Dau, & Oxenham, 2019). Although this is of substantially greater precision than that necessary for recognition in quiet (owing to the multiple redundant cues that are typically available in speech), source location provides an important cue to support segregation and selection when listeners must focus on target speech and ignore a competing, similar sound (Hawley, Litovsky, & Culling, 2004; Kidd, Arbogast, Mason, & Gallun, 2005). Coding of interaural timing difference (ITDs), the dominant perceptual cue for sound source location (Wightman & Kistler, 1992), requires even greater temporal precision than does pitch coding, on a scale of tens to hundreds of *microseconds*. Tasks that require reliance on spatial cues for segregating speech streams are thus especially likely to be sensitive to CS.

It is worth noting that listeners may not rely on spatial cues to segregate target speech in paradigms in which competing speech sources are spatially separated (as in the Coordinate Response Measure and the Listening in Spatialized Noise – Sentences test; see Table 1). For instance, fundamental frequency differences alone can provide sufficient differentiation of target and masker to support selective attention, rendering spatial cues irrelevant (Brungart, 2001). Further, even if two otherwise identical speech streams are presented from different directions, *forcing* a listener to rely on spatial cues, the task may not be sensitive to subtle differences in temporal coding precision. Figure 3 illustrates this point. If competing streams are presented with a large spatial separation (for instance, as in some past studies; 60°: Couth et al., 2020; Prendergast et al. 2017; Guest et al., 2018; 90°: Yeend et al., 2017; Valderrama et al, 2018), even a listener with poor resolution nonetheless may be able to resolve the streams based on spatial cues. Only if the sources are close enough that listeners with "good" resolution must focus to perform the task are listeners with subtle sensory deficit like CS likely to show impaired performance.

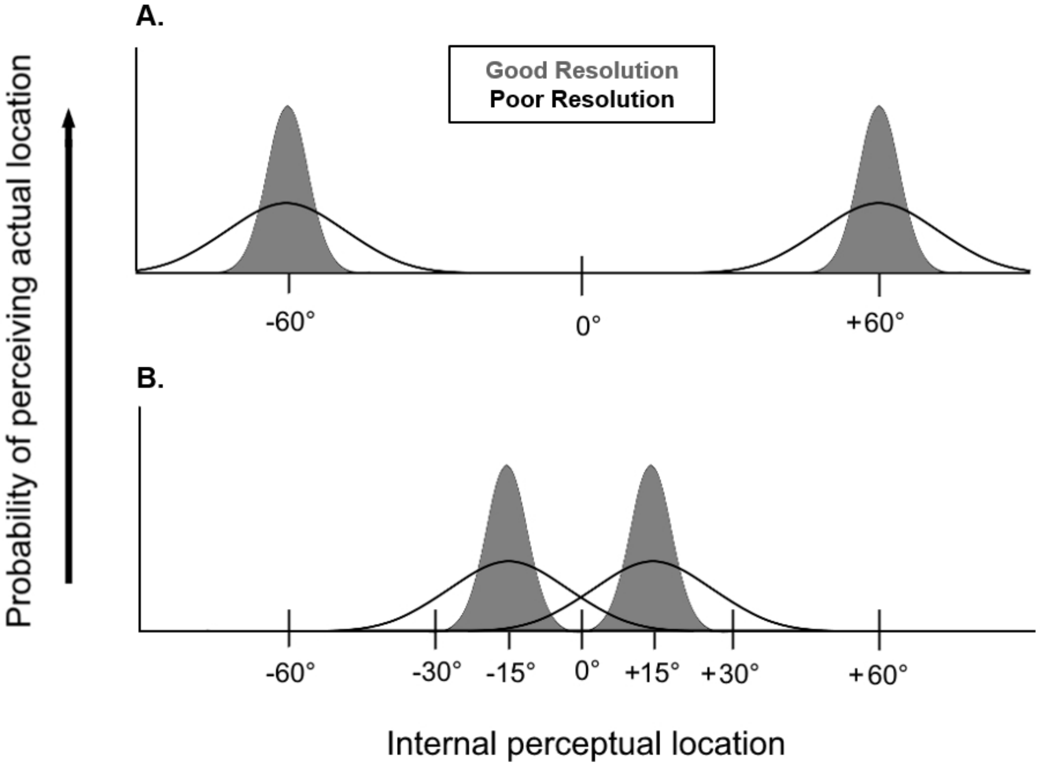

Figure 3. **Cartoon depicting the importance of using small spatial separations between speech and noise to reveal subtle temporal coding deficits.** Each panel shows probability density functions representing the perceived spatial locations of two competing sources symmetrically positioned the left and right, either with a large spatial separation (A) or a spacing that is just resolvable for a listener with good temporal resolution (B). A) For large spatial separations, listeners with good temporal coding (gray, narrow distributions) and poor temporal encoding as might arise with CS (black line, board distributions), would both be able to resolve the spatial locations to perform a spatial selective listening task. Many spatial listening tasks fall into this category. B) For a small spatial separation, listeners with good temporal resolution are more likely to perform well relative to listeners with poorer temporal encoding. This design may thus be more sensitive to CS-related perceptual deficits.

Of the 26 experiments reviewed that used speech perception tests with appropriate SNRs for listeners with NHTs and speech stimuli that limited non-sensory contributions, seven asked listeners to report target speech played with competing, intelligible speech streams and thus emphasized acoustic cues supporting segregation and selection. Five of these found no relationship between speech intelligibility and CS proxies (one experiment in Prendergast et al., 2017; one in Prendergast et al., 2019; Guest et al., 2018; Couth et al, 2020; Parthasarathy et al., 2020). Only two of the experiments reported a positive result (see Figure 2C; Ruggles et al., 2011; Bharadwaj et al., 2015).

Importantly, spatial cues were critical for those two experiments. While some of the studies presenting target speech with competing, intelligible speech played the competing streams from different directions, the talkers also differed across streams, allowing a listener to rely on fundamental frequency cues and rendering spatial cues unnecessary (Guest et al., 2018; Couth et al., 2020; Parthasarathy et al., 2020). In the two experiments that found a relationship, the target speech and the competing speech were from the same talker and differed only because of a small spatial separation, stressing the ability of listeners to utilize fine spatial cues to direct attention (see Figure 3). Thus, as illustrated in Figure 2D, the influence of CS on speech perception in experiments that require listeners to segregate and select target speech may be most pronounced when the task relies upon precise spatial selective attention, which is a critical contributor to understanding speech in noisy listening environments, and which places extreme demands on temporal coding.

Summary and Implications

In this review, we have described several factors that may help explain the mixed results among previous studies of the relationship between noise-induced CS and speech in difficult listening situations. Of the 41 experiments we reviewed, 13 reported a relationship between proxies for noise-induced CS and speech perception performance (see grayed entries). Of these, two (light gray fill in Table 1) did not rule out confounds due to traditional hearing loss (Bramhall et al., 2015; Valderrama et al., 2018). Importantly, each of the remaining 11 experiments that reported a relationship of estimated CS levels to speech perception performance, summarized below, employed speech tasks able to tap into the subtle temporal sensory deficits most associated with CS while also minimizing the higher-order perceptual and cognitive processes that can be drawn into play in speech perception.

- Eight experiments found that isolated word recognition correlated with physiological CS proxies: four experiments presenting words either in steady-state noise or that were sped up and presented with reverberation found correlations with ABR measures (Liberman et al., 2016); three using open-set monosyllabic words that were either presented in broadband noise or sped up and presented with reverberation found correlations with MEMR strength (Mepani et al., 2020); and one using closed-set identification of words in noise with reverberation found a correlation with MEMR thresholds (Shehorn, et al., 2020).
- One experiment found that closed-set syllable identification correlated with occupational noise-exposure history (Hope et al., 2013).

- Two experiments observed significant relationships between EFRs and performance on a speech-against-speech task requiring fine spatial attention (Ruggles et al., 2011, Bharadwaj et al., 2015).

Although differences in CS proxies or other experimental methods may have also limited the sensitivity of the studies examined in this review, it is noteworthy that *all prior studies* using speech-in-noise perception tests with relatively high SNRs and/or speech stimuli with high levels of context did not find a robust relationship between estimated levels of noise-induced CS and speech-in-noise understanding scores. Viewed from the perspective of the subtle sensory challenges introduced by CS, this pattern of results highlights that there are specific characteristics of speech tasks that are most appropriate for investigating the putative influences of CS on speech perception in challenging listening conditions. In particular, tasks that utilize appropriately difficult SNRs for listeners with NHTs and maximize the importance of the sensory representation of temporal acoustic features, while minimizing other perceptual and cognitive factors that could influence an individual's performance, will be best suited to quantify the relationship of CS to speech perception performance. Such tasks can be sensitive to subtle sensory deficits while maintaining at least some ecological validity to the challenges of everyday speech perception.

Resolving the question of whether CS impacts speech perception in human listeners is essential to the future of the field, and there are important clinical implications if CS can explain otherwise puzzling perceptual deficits. A link between auditory perceptual impairments in humans and moderate- to high-intensity sound exposure that does *not* permanently alter hearing thresholds could motivate systemic efforts to improve hearing protection education and guidelines. Compelling evidence that CS contributes to difficulties perceiving speech under adverse listening conditions could change how clinicians diagnose and treat this type of hearing impairment. Even apart from whether CS plays a significant role in human auditory perception, this area of study has incited widespread interest that may lead to the discovery of other neural and perceptual factors that impair speech-in-noise understanding in adults with NHTs. While previous reviews have focused on the need to develop precise assessments of CS levels in humans, our review highlights the importance of using speech perception tasks that tap into the specific deficits that CS may cause.

Acknowledgements

This work was supported by NIH NIDCD F32 DC018460 (DiNino), NIH NIDCD T32 DC011499 (DiNino), NIH NIDCD R01 DC017734 (Holt), and NIH NIDCD R01 DC013825 (Shinn-Cunningham).

References

American Speech-Language-Hearing Association. (2005). *(Central) auditory processing disorders.* (Technical Report). Retrieved from https://www.asha.org/content.aspx?id=10737450473.

Badri, R., Siegel, J. H., & Wright, B. A. (2011). Auditory filter shapes and high-frequency hearing in adults who have impaired speech in noise performance despite clinically normal audiograms. *The Journal of the Acoustical Society of America*, *129*(2), 852–863. https://doi.org/10.1121/1.3523476

Banks, B., Gowen, E., Munro, K. J., & Adank, P. (2015). Cognitive predictors of perceptual adaptation to accented speech. *The Journal of the Acoustical Society of America*, *137*(4), 2015–2024. https://doi.org/10.1121/1.4916265

Best, V., Keidser, G., Freeston, K., & Buchholz, J. M. (2016). A Dynamic Speech Comprehension Test for Assessing Real-World Listening Ability. *Journal of the American Academy of Audiology*, *27*(7), 515–526. https://doi.org/10.3766/jaaa.15089

Bharadwaj, H. M., Mai, A. R., Simpson, J. M., Choi, I., Heinz, M. G., & Shinn-Cunningham, B. G. (2019). Non-invasive assays of cochlear synaptopathy - Candidates and considerations. *Neuroscience*, *407*, 53–66. https://doi.org/10.1016/j.neuroscience.2019.02.031

Bharadwaj, H. M., Masud, S., Mehraei, G., Verhulst, S., & Shinn-Cunningham, B. G. (2015). Individual differences reveal correlates of hidden hearing deficits. *The Journal of Neuroscience*, *35*(5), 2161–2172. https://doi.org/10.1523/JNEUROSCI.3915-14.2015

Binns, C., & Culling, J. F. (2007). The role of fundamental frequency contours in the perception of speech against interfering speech. *The Journal of the Acoustical Society of America*, *122*(3), 1765. https://doi.org/10.1121/1.2751394

Bramhall, N., Beach, E. F., Epp, B., Le Prell, C. G., Lopez-Poveda, E. A., Plack, C. J., … Canlon, B. (2019). The search for noise-induced cochlear synaptopathy in humans: Mission impossible? *Hearing Research*, *377*, 88–103. https://doi.org/10.1016/j.heares.2019.02.016

Bramhall, N., Ong, B., Ko, J., & Parker, M. (2015). Speech perception ability in noise is correlated with Auditory Brainstem Response Wave I amplitude. *Journal of the American Academy of Audiology*, *26*(5), 509–517. https://doi.org/10.3766/jaaa.14100

Brown, G. J., Ferry, R. T., & Meddis, R. (2010). A computer model of auditory efferent suppression: Implications for the recognition of speech in noise. *The Journal of the Acoustical Society of America*, *127*(2), 943–954. https://doi.org/10.1121/1.3273893

Brungart, D. S. (2001). Informational and energetic masking effects in the perception of two simultaneous talkers. *The Journal of the Acoustical Society of America*, *109*(3), 1101–1109. https://doi.org/10.1121/1.1345696

Carney, L. H. (2018). Supra-threshold hearing and fluctuation profiles: Implications for sensorineural and hidden hearing loss. *Journal of the Association for Research in Otolaryngology*, *19*(4), 331–352. https://doi.org/10.1007/s10162-018-0669-5

Carroll, R., Warzybok, A., Kollmeier, B., & Ruigendijk, E. (2016). Age-related differences in lexical access relate to speech recognition in noise. *Frontiers in Psychology*, *7*, 990. https://doi.org/10.3389/fpsyg.2016.00990

Costalupes, J. A. (1985). Representation of tones in noise in the responses of auditory nerve fibers in cats. I. Comparison with detection thresholds. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, *5*(12), 3261–3269.

Costalupes, J. A., Young, E. D., & Gibson, D. J. (1984). Effects of continuous noise backgrounds on rate response of auditory nerve fibers in cat. *Journal of Neurophysiology*, *51*(6), 1326–1344. https://doi.org/10.1152/jn.1984.51.6.1326

Couth, S., Prendergast, G., Guest, H., Munro, K. J., Moore, D. R., Plack, C. J., Ginsborg, J., & Dawes, P. (2020). Investigating the effects of noise exposure on self-report, behavioral and electrophysiological indices of hearing damage in musicians with normal audiometric thresholds. *Hearing Research*, *395*, 108021. https://doi.org/10.1016/j.heares.2020.108021

Dewson, J. H. (1968). Some relationships to stimulus discrimination in noise. *Journal of Neurophysiology*, *31*(1), 122–130. https://doi.org/10.1152/jn.1968.31.1.122

Durlach, N. I., Mason, C. R., Kidd, G., Arbogast, T. L., Colburn, H. S., & Shinn-Cunningham, B. G. (2003a). Note on informational masking (L). *The Journal of the Acoustical Society of America*, *113*(6), 2984–2987. https://doi.org/10.1121/1.1570435

Durlach, N. I., Mason, C. R., Shinn-Cunningham, B. G., Arbogast, T. L., Colburn, H. S., & Kidd, G. (2003b). Informational masking: Counteracting the effects of stimulus uncertainty by decreasing target-masker similarity. *The Journal of the Acoustical Society of America*, *114*(1), 368–379. https://doi.org/10.1121/1.1577562

Fernandez, K. A., Jeffers, P. W. C., Lall, K., Liberman, M. C., & Kujawa, S. G. (2015). Aging after noise exposure: acceleration of cochlear synaptopathy in “recovered” ears. *The Journal of Neuroscience*, *35*(19), 7509–7520. https://doi.org/10.1523/JNEUROSCI.5138-14.2015

Fulbright, A. N. C., Le Prell, C. G., Griffiths, S. K., & Lobarinas, E. (2017). Effects of recreational noise on threshold and suprathreshold measures of auditory function. *Seminars in Hearing*, *38*(4), 298–318. https://doi.org/10.1055/s-0037-1606325

Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2013). Noise-induced cochlear neuropathy is selective for fibers with low spontaneous rates. *Journal of Neurophysiology*, *110*(3), 577–586. https://doi.org/10.1152/jn.00164.2013

Giraud, A. L., Garnier, S., Micheyl, C., Lina, G., Chays, A., & Chéry-Croze, S. (1997). Auditory efferents involved in speech-in-noise intelligibility. *Neuroreport*, *8*(7), 1779–1783. https://doi.org/10.1097/00001756-199705060-00042

Grant, K. J., Mepani, A. M., Wu, P., Hancock, K. E., de Gruttola, V., Liberman, M. C., & Maison, S. F. (2020). Electrophysiological markers of cochlear function correlate with hearing-in-noise performance among audiometrically normal subjects. *Journal of Neurophysiology*, *124*(2), 418–431. https://doi.org/10.1152/jn.00016.2020

Grinn, S. K., Wiseman, K. B., Baker, J. A., & Le Prell, C. G. (2017). Hidden hearing loss? No effect of common recreational noise exposure on cochlear nerve response amplitude in humans. *Frontiers in Neuroscience*, *11*, 465. https://doi.org/10.3389/fnins.2017.00465

Grose, J. H., Buss, E., & Hall, J. W. (2017). Loud music exposure and cochlear synaptopathy in young adults: Isolated auditory brainstem response effects but no perceptual consequences. *Trends in Hearing*, *21*, 2331216517737417. https://doi.org/10.1177/2331216517737417

Guest, H., Munro, K. J., Prendergast, G., Millman, R. E., & Plack, C. J. (2018). Impaired speech perception in noise with a normal audiogram: No evidence for cochlear synaptopathy and no relation to lifetime noise exposure. *Hearing Research*, *364*, 142–151. https://doi.org/10.1016/j.heares.2018.03.008

Guest, H., Munro, K. J., & Plack, C. J. (2019). Acoustic Middle-Ear-Muscle-Reflex Thresholds in Humans with Normal Audiograms: No Relations to Tinnitus, Speech Perception in Noise, or Noise Exposure. *Neuroscience*, *407*, 75–82.

https://doi.org/10.1016/j.neuroscience.2018.12.019

Hawley, M. L., Litovsky, R. Y., & Culling, J. F. (2004). The benefit of binaural hearing in a cocktail party: Effect of location and type of interferer. *The Journal of the Acoustical Society of America*, *115*(2), 833–843. https://doi.org/10.1121/1.1639908

Hienz, R. D., Stiles, P., & May, B. J. (1998). Effects of bilateral olivocochlear lesions on vowel formant discrimination in cats. *Hearing Research*, *116*(1–2), 10–20. https://doi.org/10.1016/s0378-5955(97)00197-4

Henry, K. S., Kale, S., Scheidt, R. E., & Heinz, M. G. (2011). Auditory brainstem responses predict auditory nerve fiber thresholds and frequency selectivity in hearing impaired chinchillas. *Hearing Research*, *280*(1–2), 236–244. https://doi.org/10.1016/j.heares.2011.06.002

Hinchcliffe, R. (1992). King-Kopetzky syndrome: An auditory stress disorder? *Journal of Audiological Medicine*, *1*, 89-98.

Holube, I., Fredelake, S., Vlaming, M., & Kollmeier, B. (2010). Development and analysis of an International Speech Test Signal (ISTS). *International Journal of Audiology*, *49*(12), 891–903. https://doi.org/10.3109/14992027.2010.506889

Hope, A. J., Luxon, L. M., & Bamiou, D.-E. (2013). Effects of chronic noise exposure on speech-in-noise perception in the presence of normal audiometry. *The Journal of Laryngology and Otology*, *127*(3), 233–238. https://doi.org/10.1017/S002221511200299X

Jerger, J., Johnson, K., Jerger, S., Coker, N., Pirozzolo, F., & Gray, L. (1991). Central auditory processing disorder: A case study. *Journal of the American Academy of Audiology*, *2*(1), 36–54.

Johannesen, P. T., Buzo, B. C., & Lopez-Poveda, E. A. (2019). Evidence for age-related cochlear synaptopathy in humans unconnected to speech-in-noise intelligibility deficits. *Hearing Research*, *374*, 35–48. https://doi.org/10.1016/j.heares.2019.01.017

Kamerer, A. M., Kopun, J. G., Fultz, S. E., Allen, C., Neely, S. T., & Rasetshwane, D. M. (2019). Examining physiological and perceptual consequences of noise exposure. *The Journal of the Acoustical Society of America*, *146*(5), 3947. https://doi.org/10.1121/1.5132291

Kaandorp, M. W., Groot, A. M. B. D., Festen, J. M., Smits, C., & Goverts, S. T. (2016). The influence of lexical-access ability and vocabulary knowledge on measures of speech recognition in noise. *International Journal of Audiology*, *55*(3), 157–167. https://doi.org/10.3109/14992027.2015.1104735

Kawase, T., Delgutte, B., & Liberman, M. C. (1993). Antimasking effects of the olivocochlear reflex. II. Enhancement of auditory-nerve response to masked tones. *Journal of Neurophysiology*, *70*(6), 2533–2549. https://doi.org/10.1152/jn.1993.70.6.2533

Kidd, G., Arbogast, T. L., Mason, C. R., & Gallun, F. J. (2005). The advantage of knowing where to listen. *The Journal of the Acoustical Society of America*, *118*(6), 3804–3815. https://doi.org/10.1121/1.2109187

Killion, M.C., Niquette, P.A., Gudmundsen, G.I., Revit, L.J., & Banerjee, S. (2004). Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, *116*(4 pt. 1), 2395-2405. https://doi.org/10.1121/1.1784440

King, K., & Stephens, D. (1992). Auditory and psychological factors in “auditory disability with normal hearing.” *Scandinavian Audiology*, *21*(2), 109–114.

Kujawa, S. G., & Liberman, M. C. (2009). Adding insult to injury: Cochlear nerve degeneration after “temporary” noise-induced hearing loss. *The Journal of Neuroscience* , *29*(45), 14077–14085. https://doi.org/10.1523/JNEUROSCI.2845-09.2009

Kumar, U. A., & Vanaja, C. S. (2004). Functioning of olivocochlear bundle and speech perception in noise. *Ear and Hearing*, *25*(2), 142–146. https://doi.org/10.1097/01.aud.0000120363.56591.e6

Le Prell, C. G. (2019). Effects of noise exposure on auditory brainstem response and speech-in-noise tasks: A review of the literature. *International Journal of Audiology*, *58*(sup1), S3–S32. https://doi.org/10.1080/14992027.2018.1534010

Le Prell, C. G., Siburt, H. W., Lobarinas, E., Griffiths, S. K., & Spankovich, C. (2018). No reliable association between recreational noise exposure and threshold sensitivity, distortion product otoacoustic emission amplitude, or word-in-noise performance in a college student population. *Ear and Hearing*, *39*(6), 1057–1074. https://doi.org/10.1097/AUD.0000000000000575

Liberman, M. C. (1978). Auditory-nerve response from cats raised in a low-noise chamber. *The Journal of the Acoustical Society of America*, *63*(2), 442–455. https://doi.org/10.1121/1.381736

Liberman, M. C. (1988). Physiology of cochlear efferent and afferent neurons: Direct comparisons in the same animal. *Hearing Research*, *34*(2), 179–191. https://doi.org/10.1016/0378-5955(88)90105-0

Liberman, M.C., & Kiang, N.Y. (1984). Single-neuron labeling and chronic cochlear pathology. IV. Stereocilia damage and alterations in rate- and phase-level functions. *Hearing Research, 16*(1), 75-90.

Liberman, L. D., Suzuki, J., & Liberman, M. C. (2015). Dynamics of cochlear synaptopathy after acoustic overexposure. *Journal of the Association for Research in Otolaryngology*, *16*(2), 205–219. https://doi.org/10.1007/s10162-015-0510-3

Liberman, M. C., Epstein, M. J., Cleveland, S. S., Wang, H., & Maison, S. F. (2016). Toward a differential diagnosis of hidden hearing loss in humans. *PloS One*, *11*(9), e0162726. https://doi.org/10.1371/journal.pone.0162726

Liberman, M. C., & Kujawa, S. G. (2017). Cochlear synaptopathy in acquired sensorineural hearing loss: Manifestations and mechanisms. *Hearing Research*, *349*, 138–147. https://doi.org/10.1016/j.heares.2017.01.003

Lin, H. W., Furman, A. C., Kujawa, S. G., & Liberman, M. C. (2011). Primary neural degeneration in the Guinea pig cochlea after reversible noise-induced threshold shift. *Journal of the Association for Research in Otolaryngology*, *12*(5), 605–616. https://doi.org/10.1007/s10162-011-0277-0

Lopez-Poveda, E. A. (2014). Why do I hear but not understand? Stochastic undersampling as a model of degraded neural encoding of speech. *Frontiers in Neuroscience*, *8*, 348. https://doi.org/10.3389/fnins.2014.00348

Lopez-Poveda, E. A., & Barrios, P. (2013). Perception of stochastically undersampled sound waveforms: A model of auditory deafferentation. *Frontiers in Neuroscience*, *7*, 124. https://doi.org/10.3389/fnins.2013.00124

Lu, Y., & Cooke, M. (2008). Speech production modifications produced by competing talkers, babble, and stationary noise. *The Journal of the Acoustical Society of America*, *124*(5), 3261–3275. https://doi.org/10.1121/1.2990705

Madsen, S. M. K., Dau, T., & Oxenham, A. J. (2019). The effect of harmonic number and pitch salience on the ability to understand speech-on-speech based on differences in fundamental frequency. *Proceedings of the International Symposium on Auditory and Audiological Research*, *7*, 45–52.

Makary, C. A., Shin, J., Kujawa, S. G., Liberman, M. C., & Merchant, S. N. (2011). Age-related primary cochlear neuronal degeneration in human temporal bones. *Journal of the Association for Research in Otolaryngology*, *12*(6), 711–717. https://doi.org/10.1007/s10162-011-0283-2

Marmel, F., Rodríguez-Mendoza, M. A., & Lopez-Poveda, E. A. (2015). Stochastic undersampling steepens auditory threshold/duration functions: Implications for understanding auditory deafferentation and aging. *Frontiers in Aging Neuroscience*, *7*, 63. https://doi.org/10.3389/fnagi.2015.00063

Marrufo-Pérez, M. I., Eustaquio-Martín, A., & Lopez-Poveda, E. A. (2019). Speech predictability can hinder communication in difficult listening conditions. *Cognition*, *192*, 103992. https://doi.org/10.1016/j.cognition.2019.06.004

May, B. J., & McQuone, S. J. (1995). Effects of bilateral olivocochlear lesions on pure-tone intensity discrimination in cats. *Auditory Neuroscience*, *1*(4), 385–400.

Mepani, A. M., Kirk, S. A., Hancock, K. E., Bennett, K., de Gruttola, V., Liberman, M. C., & Maison, S. F. (2020). Middle Ear Muscle Reflex and Word Recognition in “Normal-Hearing” Adults: Evidence for Cochlear Synaptopathy? *Ear and Hearing*, *41*(1), 25–38. https://doi.org/10.1097/AUD.0000000000000804

Monson, B. B., Rock, J., Schulz, A., Hoffman, E., & Buss, E. (2019). Ecological cocktail party listening reveals the utility of extended high-frequency hearing. *Hearing Research*, *381*, 107773. https://doi.org/10.1016/j.heares.2019.107773

Pang, X. D., & Guinan, J. J. (1997). Effects of stapedius-muscle contractions on the masking of auditory-nerve responses. *The Journal of the Acoustical Society of America*, *102*(6), 3576–3586. https://doi.org/10.1121/1.420399

Parthasarathy, A., Hancock, K. E., Bennett, K., DeGruttola, V., & Polley, D. B. (2020). Bottom-up and top-down neural signatures of disordered multi-talker speech perception in adults with normal hearing. *ELife*, *9*. https://doi.org/10.7554/eLife.51419

Parthasarathy, A., & Kujawa, S. G. (2018). Synaptopathy in the aging cochlea: Characterizing early-neural deficits in auditory temporal envelope processing. *The Journal of Neuroscience*, *38*(32), 7108–7119. https://doi.org/10.1523/JNEUROSCI.3240-17.2018

Plack, C. J., Barker, D., & Prendergast, G. (2014). Perceptual consequences of “hidden” hearing loss. *Trends in Hearing*, *18*. https://doi.org/10.1177/2331216514550621

Prendergast, G., Couth, S., Millman, R. E., Guest, H., Kluk, K., Munro, K. J., & Plack, C. J. (2019). Effects of age and noise exposure on proxy measures of cochlear synaptopathy. *Trends in Hearing*, *23*, 2331216519877301. https://doi.org/10.1177/2331216519877301

Prendergast, G., Millman, R. E., Guest, H., Munro, K. J., Kluk, K., Dewey, R. S., … Plack, C. J. (2017). Effects of noise exposure on young adults with normal audiograms II: Behavioral measures. *Hearing Research*, *356*, 74–86. https://doi.org/10.1016/j.heares.2017.10.007

Ruggles, D., Bharadwaj, H., & Shinn-Cunningham, B. G. (2011). Normal hearing is not enough to guarantee robust encoding of suprathreshold features important in everyday communication. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(37), 15516–15521. https://doi.org/10.1073/pnas.1108912108

Ryugo, D. K. (2008). Projections of low spontaneous rate, high threshold auditory nerve fibers to the small cell cap of the cochlear nucleus in cats. *Neuroscience*, *154*(1), 114–126.

https://doi.org/10.1016/j.neuroscience.2007.10.052

Samuel, A. G. (1981). Phonemic restoration: Insights from a new methodology. *Journal of Experimental Psychology. General*, *110*(4), 474–494. https://doi.org/10.1037//0096-3445.110.4.474

Saunders, G. H., Field, D. L., & Haggard, M. P. (1992). A clinical test battery for obscure auditory dysfunction (OAD): Development, selection and use of tests. *British Journal of Audiology*, *26*(1), 33–42. https://doi.org/10.3109/03005369209077869

Saunders, G. H., & Haggard, M. P. (1989). The clinical assessment of obscure auditory dysfunction--1. Auditory and psychological factors. *Ear and Hearing*, *10*(3), 200–208.

Saunders, G. H., & Haggard, M. P. (1992). The clinical assessment of "Obscure Auditory Dysfunction" (OAD) 2. Case control analysis of determining factors. *Ear and Hearing*, *13*(4), 241–254. https://doi.org/10.1097/00003446-199208000-00006

Sergeyenko, Y., Lall, K., Liberman, M. C., & Kujawa, S. G. (2013). Age-related cochlear synaptopathy: an early-onset contributor to auditory functional decline. *The Journal of Neuroscience*, *33*(34), 13686–13694. https://doi.org/10.1523/JNEUROSCI.1783-13.2013

Shaheen, L. A., Valero, M. D., & Liberman, M. C. (2015). Towards a diagnosis of cochlear neuropathy with envelope following responses. *Journal of the Association for Research in Otolaryngology*, *16*(6), 727–745. https://doi.org/10.1007/s10162-015-0539-3

Shaw, G. M., Jardine, C. A., & Fridjhon, P. (1996). A pilot investigation of high-frequency audiometry in obscure auditory dysfunction (OAD) patients. *British Journal of Audiology*, *30*(4), 233–237. https://doi.org/10.3109/03005369609076770

Shehorn, J., Strelcyk, O., & Zahorik, P. (2020). Associations between speech recognition at high levels, the middle ear muscle reflex and noise exposure in individuals with normal audiograms. *Hearing Research*, *392*, 107982. https://doi.org/10.1016/j.heares.2020.107982

Shinn-Cunningham, B. G., & Best, V. (2008). Selective attention in normal and impaired hearing. *Trends in Amplification*, *12*(4), 283–299. https://doi.org/10.1177/1084713808325306

Skoe, E., Camera, S., & Tufts, J. (2019). Noise exposure may diminish the musician advantage for perceiving speech in noise. *Ear and Hearing*, *40*(4), 782–793. https://doi.org/10.1097/AUD.0000000000000665

Smith, S. B., Krizman, J., Liu, C., White-Schwoch, T., Nicol, T., & Kraus, N. (2019). Investigating peripheral sources of speech-in-noise variability in listeners with normal audiograms. *Hearing Research*, *371*, 66–74. https://doi.org/10.1016/j.heares.2018.11.008

Valderrama, J. T., Beach, E. F., Yeend, I., Sharma, M., Van Dun, B., & Dillon, H. (2018). Effects of lifetime noise exposure on the middle-age human auditory brainstem response, tinnitus and speech-in-noise intelligibility. *Hearing Research*, *365*, 36–48. https://doi.org/10.1016/j.heares.2018.06.003

Valero, M. D., Burton, J. A., Hauser, S. N., Hackett, T. A., Ramachandran, R., & Liberman, M. C. (2017). Noise-induced cochlear synaptopathy in rhesus monkeys (Macaca mulatta). *Hearing Research*, *353*, 213–223. https://doi.org/10.1016/j.heares.2017.07.003

Viana, L. M., O'Malley, J. T., Burgess, B. J., Jones, D. D., Oliveira, C. A. C. P., Santos, F., … Liberman, M. C. (2015). Cochlear neuropathy in human presbycusis: Confocal analysis of hidden hearing loss in post-mortem tissue. *Hearing Research*, *327*, 78–88. https://doi.org/10.1016/j.heares.2015.04.014

Wightman, F. L., & Kistler, D. J. (1992). The dominant role of low-frequency interaural time differences in sound localization. *The Journal of the Acoustical Society of America*, *91*(3), 1648–1661. https://doi.org/10.1121/1.402445

Wilson, R. H., & Burks, C. A. (2005). Use of 35 words for evaluation of hearing loss in signal-to-babble ratio: A clinic protocol. *Journal of Rehabilitation Research and Development*, *42*(6), 839–852.

Winter, I. M., Robertson, D., & Yates, G. K. (1990). Diversity of characteristic frequency rate-intensity functions in guinea pig auditory nerve fibres. *Hearing Research*, *45*(3), 191–202. https://doi.org/10.1016/0378-5955(90)90120-e

Wu, P. Z., Liberman, L. D., Bennett, K., de Gruttola, V., O'Malley, J. T., & Liberman, M. C. (2019). Primary neural degeneration in the human cochlea: Evidence for hidden hearing loss in the aging ear. *Neuroscience*, *407*, 8–20. https://doi.org/10.1016/j.neuroscience.2018.07.053

Ye, Y., Machado, D. G., & Kim, D. O. (2000). Projection of the marginal shell of the anteroventral cochlear nucleus to olivocochlear neurons in the cat. *The Journal of Comparative Neurology*, *420*(1), 127–138.

Yeend, I., Beach, E. F., Sharma, M., & Dillon, H. (2017). The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. *Hearing Research*, *353*, 224–236. https://doi.org/10.1016/j.heares.2017.07.006

Young, E. D., & Barta, P. E. (1986). Rate responses of auditory nerve fibers to tones in noise near masked threshold. *The Journal of the Acoustical Society of America*, *79*(2), 426–442. https://doi.org/10.1121/1.393530

Zhao, F., & Stephens, D. (2000). Subcategories of patients with King-Kopetzky syndrome. *British Journal of Audiology*, *34*(4), 241–256.

https://doi.org/10.3109/03005364000000134

Zhao, Fei, & Stephens, D. (2006). Distortion product otoacoustic emissions in patients with King-Kopetzky syndrome. *International Journal of Audiology*, *45*(1), 34–39.