Department of Veterans Affairs

JRRD Journal of Rehabilitation Research & Development
Volume 47, Number 3, 2010
Pages 243–260

Measuring consonant identification in nonsense syllables, words, and sentences

David L. Woods, PhD;[1–4*] E. William Yund, PhD;[1] T. J. Herron, MS[1]
[1]Human Cognitive Neurophysiology Laboratory, Department of Veterans Affairs Northern California Health Care System, Martinez, CA; [2]Department of Neurology, University of California (UC) Davis, Davis, CA; [3]Center for Neurosciences, UC Davis, Davis, CA; [4]Center for Mind and Brain, UC Davis, Davis, CA

Abstract—Sensorineural hearing loss (SNHL) produces deficits in speech comprehension in noise that primarily are due to impairments in identifying consonants. Here, we describe the California Syllable Test (CaST) that quantifies the identification of common American English consonants. In experiment I, 16 young subjects with normal hearing identified 720 consonant-vowel-consonant (CVC) syllables in three test sessions. Consonants were identified slightly more accurately in words than nonsense syllables, and small interactions were found between the processing of initial and final consonants. Consonant-identification performance correlated strongly with sentence reception thresholds (SeRTs) measured with both the Hearing in Noise Test and QuickSIN (Etymotic Research; Elk Grove Village, Illinois). At SeRTs, subjects with normal hearing could identify 32.5% of consonants in isolated CVCs. In experiment II, a patient with moderate SNHL showed large elevations in consonant-identification thresholds and smaller elevations in SeRTs. At SeRT levels, the patient could identify only 12.5% of consonants in isolated CVCs, indicating that sentence comprehension relied disproportionately on vowel cues and semantic constraints. Consonant-profile analysis revealed disproportional impairments in identifying consonants dependent on high-frequency acoustic cues. Consonant confusion analysis revealed a reorganization of consonant perception. The CaST is a promising tool for evaluating consonant-specific processing deficits in patients with hearing impairment.

Key words: audiological rehabilitation, CaST, consonants, CVC, hearing impairment, Hearing in Noise Test, hearing loss, QuickSIN, sensorineural hearing loss, sentence reception threshold, signal-to-noise ratio, vowels.

INTRODUCTION

The most common complaint of patients with hearing loss is difficulty understanding conversational speech in the presence of noise [1]. Patients with mild to moderate sensorineural hearing loss (SNHL) have significantly higher sentence reception thresholds (SeRTs), even with hearing aids [2–5], than subjects with normal hearing. This finding reflects a degradation of the acoustic cues that have their greatest impact on consonant discrimination [6]. In contrast, vowel identification is relatively well preserved [7]. For example, Ferguson and Kewley-Port found that patients with mild SNHL accurately identified 73 percent of vowels at speech-to-babble ratios of –3 dB, i.e., well below typical SeRTs [8]. As a result, improving consonant comprehension is a major focus of

Abbreviations: ANOVA = analysis of variance, B = baseline, CaST = California Syllable Test, CV = consonant-vowel, CVC = consonant-vowel-consonant, HINT = Hearing in Noise Test, PTA = pure tone average, SeRT = sentence reception threshold, SIN = Speech-in-Noise, SNHL = sensorineural hearing loss, SNR = signal-to-noise ratio, SPL = sound pressure level, VA = Department of Veterans Affairs.

***Address all correspondence to David L. Woods, PhD; Neurology Service (127E), VA Northern California Health Care System, 150 Muir Road, Martinez, CA 95553; 925-372-2571; fax: 925-229-2315. Email: dlwoods@ucdavis.edu**
DOI:10.1682/JRRD.2009.04.0040

current research on hearing aid design [9–11] and adaptive perceptual training [12–15].

How should consonant processing be assessed? One common approach is to measure SeRTs with sentence tests such as the Hearing in Noise Test (HINT) [16] and QuickSIN (Speech-in-Noise test) (Etymotic Research; Elk Grove Village, Illinois) [17]. However, SeRTs are relatively insensitive to consonant processing deficits for several reasons. First, sentence context greatly enhances the accuracy of word report even when individual words cannot be clearly heard. For example, Boothroyd and Nittrouer studied the relationship between phoneme-identification and SeRTs and found that the SeRTs for high-probability sentences occurred at signal-to-noise ratios (SNRs) where <50 percent of phonemes (including vowels) could be identified in isolated words [18]. Other groups have obtained similar results [16,19]. Second, the phonemes that are actually heard, rather than deduced from semantic context, come disproportionally from words that occur early in sentences. In declarative sentences of the sort used in most SeRT tests, syllable intensity declines over the course of the sentence (see "Methods" section). Since masking noise amplitudes are constant, SNRs are much higher for consonants that occur early in sentences than for consonants that occur later. Thus, sentence context is particularly necessary to perceive words that occur later in the sentence. Because of the critical role of semantic context, sentence tests may underestimate phonological impairments in patients with exceptional semantic skills and overestimate deficits in patients with cognitive impairments or impaired semantic processing of standard American English because of bilingualism or ethnic speech patterns.

Which consonants are normally audible at SeRTs? We recently found that the SNRs needed to equate the identification of 21 common American English consonants varied by >40 dB in young subjects with normal hearing [20]. Consonants could be divided into three groups on the basis of their SNR thresholds. Group A consonants (/ʃ/, /ʧ/, /t/, /s/, /z/, /ʤ/, and /r/) were accurately identified in isolated syllables at SNRs below typical SeRTs. Group B consonants (/d/, /g/, /l/, /m/, /n/, /f/, and /k/) were identified at SNRs that were 3 to 5 dB above SeRTs. Group C (/b/, /v/, /ð/, /ŋ/, /θ/, /p/, and /h/) consonants could only be identified at SNRs that were >10 dB above typical SeRTs. Thus, at the SNRs that characterize SeRTs, almost all group A consonants will be identified, including many that will be identifiable even in words that occur later in sentences. In addition, some consonants from group B will be identifiable in well-articulated words that receive strong emphasis early in the sentence. However, group C consonants will almost never be presented at SNRs that would permit their identification in the absence of strong contextual cues. This finding suggests that SeRTs primarily reflect a subject's ability to identify vowels and consonants in group A as well as some consonants in group B. Consequently, even in subjects with normal hearing, SeRT testing will fail to evaluate the phonological processing of >50 percent of common American English consonants.

How is consonant-identification performance affected by SNHL? Since SNHL typically produces greater audiometric deficits at high frequencies, disproportionate impairments in consonant identification would be expected for those consonants whose discrimination depends disproportionately on low-intensity, high-frequency acoustic cues [6,21–22]. These consonants are primarily plosives and nonsibilant fricatives that generally require SNRs that are well above SeRTs for their identification. Paradoxically, sentence tests would thus appear to be largely insensitive to the identification of deficits in consonants most severely impaired by SNHL. This result leads to the prediction that many patients with SNHL may show relatively small elevations in SeRTs despite significant difficulties in identifying many consonants that occur frequently in American English. Although impairments in the processing of these consonants may not increase SeRTs in the simple declarative sentences used in most sentence tests, they will impair comprehension and memory for less predictable spoken materials [23] and contribute to patient fatigue.

Currently, no widely accepted tests of consonant-specific processing deficits exist. Testing with word lists can reveal overall deficits in phonological processing [24–25], but word-recognition scores confound the identification of different consonants and are also influenced by word familiarity. The accurate measurement of consonant-identification performance for a complete set of American English consonants is a particularly challenging task that has resisted easy solution for several reasons. First, a large consonant set must be used so that the full range of possible consonant confusions can be evaluated. Second, because of systematic differences in the processing of initial and final consonants in syllables [20,26], consonant processing should be assessed in both the initial and final syllable positions. Finally, testing consonant identification over a broad range of SNRs is necessary for assessing consonant processing over SNR ranges that produce hit

rates (percentage of correct responses) of 40 to 70 percent for each consonant. Since the SNRs needed to identify different consonants differ by >40 dB [20], consonant identification must be tested at many different SNRs. For example, if testing is done at 6 dB intervals, at least seven different SNRs are needed for testing the full range of consonant-identification performance.

Only a few investigators have undertaken such lengthy experiments. In their original study, Miller and Nicely evaluated consonant confusions in five subjects using consonant-vowels (CVs) presented at seven different SNRs spanning a 35 dB range [27]. Each subject was tested over several months to provide 50 responses to each syllable at each SNR. Wang and Bilger characterized consonant identification in subjects who underwent three successive 2.5 h test sessions following an orientation session [28]. Four different groups of four subjects each were tested with different sets of 16 consonants. Two groups identified CVs and two groups identified vowel-consonants, using consonants randomly paired with three vowels. Each subject produced 72 responses at six different SNRs spanning a 25 dB range. Phatak and Allen investigated the processing of 16 consonants and 4 vowels in CVs presented at six different SNRs from –22 dB to quiet [29]. Subjects (n = 14) required approximately 15 h to produce 56 responses to each consonant at each SNR.

Thus, previous studies that have characterized consonant identification in noise have used methods that are too time-consuming for routine application. Alternative approaches, such as testing subsets of confusable consonants over more limited SNR ranges [30], are less time-consuming. However, because of the limited number of consonants and response alternatives, this approach may underestimate consonant confusions that complicate conversational listening. Moreover, small consonant sets give subjects the opportunity to use strategies (e.g., reporting the most hard-to-identify consonant during trials where no consonant was clearly heard) that differ from those available in more natural listening conditions.

Identifying consonant-vowel-consonant (CVC) tokens offers a more time-efficient method of assessing consonant-identification performance than testing with single-consonant syllables because each CVC can elicit two consonant-identification responses. However, CVC tests are complicated by difficulties in creating the token corpora, by potential token-learning effects, and by potential differences in the identifiability of consonants in words and nonsense syllables. Boothroyd and Nittrouer developed phonetically matched sets of 120 CVC words and 120 nonsense syllables by combining 10 different initial consonants, 10 different vowels, and 10 different final consonants [18]. Boothroyd and Nittrouer found significant differences in consonant-identification performance for words and nonsense syllables presented in separate lists. Other investigators have obtained similar results [6,19]. However, because only 240 of 1,000 possible CVC combinations were used in these experiments, they did not analyze the processing of initial and final consonants independent of each other and the accompanying vowel. In addition, because of the relatively small number of CVC tokens, they had to test separate subject groups at each SNR to avoid token repetition. A number of other investigators have also used lists of CVC words to evaluate consonant processing in patients with hearing loss. These tests require between 20 [31] and 54 min [32] to administer to subjects with hearing impairment to identify global phoneme deficits that are significantly correlated with SeRTs, but these tests fail to permit the isolation of consonant-specific identification deficits.

This article describes the California Syllable Test (CaST), a 48 min test designed to assess a subject's ability to identify a large set of common American English consonants in noise. CaST CVC syllables were constructed by the exhaustive combination of 20 initial consonants, 3 vowels, and 20 final consonants and include both nonsense syllables and words. The CaST uses an extremely large corpus, since two recordings of each of the 1,200 syllables were obtained from each of four talkers to create a total of 9,600 tokens. During the administration of the CaST, 720 tokens are pseudorandomly selected from the corpus for measuring the identification of each of 20 initial and final consonants across a range of SNRs needed for defining their psychometric functions. In group studies, the CaST provides information about consonant-identification thresholds, confusion patterns, and vowel- and syllable-position influences on consonant-feature processing [20].

In the first experiment of the current article, we described the use of the CaST as a test of consonant identification in individual subjects, focusing on the test-retest reliability and the relationship of CaST scores to SeRTs and audiometric thresholds. In a second experiment, we demonstrated the application of the CaST to understanding consonant-processing impairments in a subject with SNHL. We also investigated two factors that might influence audiological applications. Since CaST

tokens were derived from the exhaustive combination of 20 initial and final consonants and 3 vowels, they included a mixture of nonsense syllables (66.42%) as well as syllables recognizable as close approximations to words in standard American English (33.58%). The mixture of words and nonsense syllables containing the same consonants enabled us to examine the role of syllable type on consonant identification. Previous syllable tests have shown that subjects can more accurately identify consonants in words than in nonsense syllables when words and nonsense syllables are presented in separate blocks [18]. Similarly, previous investigators have also found interactions between the processing of initial and final consonants in words but not in nonsense syllables when both were presented in separate lists [18–19]. The CaST presents word and nonsense-syllable tokens in random order, thus permitting the examination of word superiority effects and interactions in the processing of initial and final consonants in conditions where category-report bias was minimized.

METHODS: EXPERIMENT I

Subjects

Sixteen young subjects (eight females and eight males, aged 18–30 yr) with normal hearing (thresholds ≤ 20 dB hearing level at 250–4,000 Hz) each participated in three sessions over a period of 3 to 11 days. Each session included CaST, HINT, and QuickSIN assessment.

Syllable Tokens

The CaST includes 1,200 CVC syllables constructed from the exhaustive combination of 20 initial consonants, 20 final consonants, and 3 different vowels (/ɑ/, /i/, /u/). Nineteen consonants (/p/, /t/, /k/, /f/, /θ/, /s/, /ʃ/, /tʃ/, /b/, /d/, /g/, /v/, /ð/, /z/, /dʒ/, /m/, /n/, /l/, and /r/) occurred in both initial and final consonant positions, while /h/ occurred only in the initial position and /ŋ/ only in the final position. CaST tokens were obtained from four syllable sets (4,800 syllables each) that had been recorded from each of four phonetically trained talkers (two males and two females). The four talkers had been raised in different parts of the United States (two from the Midwest and two from California) and had slightly different American English speech patterns. Syllables were digitized (16-bit resolution and 44.1 kHz sampling rate) under MATLAB (The MathWorks Inc; Natick, Massachusetts) control. We reviewed the complete syllable sets and selected the two best exemplars of each syllable from each talker's corpus. Then two listeners with normal hearing independently reviewed each of the 9,600 syllables in the absence of masking noise to assure the intelligibility of all tokens. Whenever this intelligibility test failed, a new exemplar from the same talker was substituted and further testing was performed among laboratory staff to assure the intelligibility of the substituted tokens. Syllable durations ranged from 350 to 890 ms (mean = 636 ms). For each token, the central 100 ms of each vowel were identified by manual review.

Speech-Spectrum Noise Adjustment

Talker-specific speech-spectrum noise was used to mask CVC tokens. We first obtained the average spectrum for each talker by averaging the spectra of all CVC tokens spoken by that talker. We then used this spectrum to create a finite impulse response function for filtering broadband white noise. Each filtered-noise file was trimmed of the first 0.5 s and cut into 100 different noise segments of 1,200 ms duration. Then we randomly sampled the 100 different noise segments during the testing sessions to mask CVCs spoken by that talker.

Stimuli and Procedures

Testing was performed in a 2.44 × 2.44 m single-walled, sound-attenuating testing room. The interior walls were covered by 2.5 cm acoustic foam, resulting in ambient third-octave noise levels <20 dB sound pressure level (SPL) from 250 to 4,000 Hz. In anticipation of future studies of subjects with hearing impairment wearing hearing aids, we presented stimuli through loudspeakers (M-Audio Studiophile AV 40; Irwindale, California). Immediately before the first CaST session, subjects were briefed with written and oral instructions and received ~5 min of training in identifying CVCs presented without masking noise.

During each CaST session, the CVCs were grouped by a talker into 30 trial blocks. Presentation software (NeuroBehavioral Systems, version 12.0; Albany, California) was used for stimulus delivery, noise level adjustment masking, response monitoring, and d′ calculations. Each trial began with a tone-burst cue (100 ms 1.0 kHz tone, 70 dB SPL) 1 s before the start of the noise (**Figure 1**). Talker-specific noise bursts of 1,200 ms duration were then presented independently from the left and right loudspeakers along with a single CVC presented from

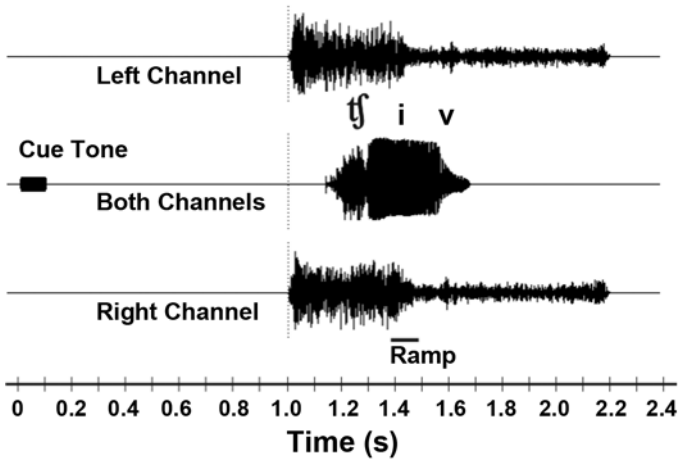

Figure 1.
Trial structure. Trials were cued by 1.0 kHz tone. After 1.0 s, two independent 1,200 ms noise bursts were presented through left and right loudspeakers. Consonant-vowel-consonants (CVCs) were presented simultaneously through both loudspeakers at random intervals after noise-burst onset. Noise amplitudes were linearly adjusted over 100 ms interval during midvowel segment of CVC for appropriate masking levels for different initial and final consonants.

both loudspeakers. Syllable onset time was randomized with the constraint that each CVC began at least 100 ms after noise-burst onset and ended at least 100 ms before noise-burst offset. After familiarizing themselves with a list of acceptable initial and final consonants and vowels posted in the testing room and practicing for 15 min to ensure understanding and accuracy, listeners attempted to repeat the CVC token correctly on each trial. Responses were spoken in quiet into a microphone and phonetically transcribed by an investigator listening through headphones in an adjacent room. Subjects were queried by way of an intercom when responses were invalid or poorly enunciated.* Subjects were given the option of repeating trials in cases of attentional lapse or noise interference (e.g., coughing). Repeated trials occurred on 1.15 percent of trial presentations. Each intertrial interval (approximately 2 s) included the time needed for syllable transcription plus a small delay (0.5 s) before the delivery of the warning tone signaling the next trial. Trials occurred at a rate of approximately 15/min so that each 720-syllable test required about 48 min, excluding rest breaks that occurred at each subject's discretion.

Syllable intensity was randomly roved from 70 to 75 dB SPL in 1 dB steps. Psychometric functions were measured for each initial and final consonant at three different SNRs: B (Baseline), B – 6, and B + 6 dB relative to the baseline SNR that was specific to each initial and final consonant. Consonant-specific baseline levels were established in preliminary experiments. The SNR level (i.e., B – 6, B, or B + 6) varied randomly from trial to trial.

During each test session, 720 tokens were randomly selected without repetition from the syllable corpus of 9,600 tokens. Selection was constrained so that each initial and final consonant was presented 12 times at each SNR. These 12 tokens included syllables containing each of the three vowels (/ɑ/, /i/, /u/) spoken by each of the four talkers. Syllables were selected based on the random combination of the initial consonant, vowel, and final consonant so that each token in the corpus had an equal probability of being presented. Following talker and syllable selection, one token was randomly selected from the two token exemplars for that talker. This procedure resulted in the presentation of 240 tokens (60 from each talker), at each of the three SNR levels (B – 6, B, and B + 6 dB) on each day of testing. Because of the low rate of vowel errors, only consonant identification was scored.

Quantifying Consonant Identification

Consonants were presented at a consonant-specific SNR designed to equate the identifiability of different consonants. Because of the variation in response criteria for different consonants, consonant-identification thresholds were quantified with a modified, multiresponse d′ measure derived from signal detection theory [33]. We adjusted SNRs to minimize variations in the identifiabilities of different consonants and set to produce a mean d′ of 2.20 (approximately 65% correct). We used additional adjustments to equate performance for syllables spoken by different talkers (syllables spoken by female talkers were reduced by 1.8 dB) and for syllables containing different vowels (syllables with /i/ were reduced by 3.0 dB, and those containing /ɑ/ were reduced by 1.2 dB, relative to those containing /u/). Mean SNRs averaged 6.6 dB for initial consonants and 9.9 dB for independently adjusted final consonants. Further methodological details can be found in Woods et al. [20].

*Experimenter response transcription was used in preference to subject transcription for maintaining the naturalness of the listening task, minimizing procedural learning effects, and avoiding scoring biases that might be introduced by listeners untrained in the use of the phonetic alphabet.

Sentence Tests

On each day of testing, we measured sentence comprehension using the HINT [16] and the QuickSIN [17]. HINT sentences were delivered through the loudspeakers at 70 dB SPL, with varying levels of speech-spectrum noise. A total of 80 HINT sentences were presented in four blocks of 20 on each day of testing. We measured thresholds for each of the sentence blocks by initially decreasing SNRs in 4 dB steps until the first incorrect report. Thereafter, we increased SNRs by 2 dB following each incorrect report and decreased by 2 dB following each correct report. We then estimated thresholds by averaging the SNRs over the final 16 sentences in each block, with mean daily thresholds averaged over the four blocks. The QuickSIN involved the delivery of six blocks, each containing six sentences in four-talker babble. Speech-to-babble ratios were reduced by 5 dB on each sentence presentation within a block, and the number of words correctly reported was used for calculating thresholds. We reported QuickSIN thresholds on the standard QuickSIN SNR loss scale, where 0 dB SNR represents normal-hearing performance on the test. Thresholds were averaged over the six blocks presented on each day. The order of presentation of the three different sets of HINT and QuickSIN sentences was randomized across subjects. No HINT and QuickSIN sentences were repeated across testing days.

We also quantified the SNRs of each syllable presented during HINT by measuring the intensities of the vowel segment of each syllable and then quantifying the SNR relative to masking noise at each SeRT. **Figure 2** shows the results for 0 dB SeRTs. On average, SNRs ranged from +2.00 dB for syllable position 2 to –4.83 dB for syllable 6, a range of 6.83 dB. Further SNR declines were evident for syllable 7 in those longer sentences that included a seventh syllable. Measurements of the QuickSIN sentences showed a similar pattern: intensities were for syllable 2 and declined by 6.22 dB by syllable 12. These measurements establish that the clarity of acoustic cues was greatest for syllables occurring early in sentences in both tests and declined substantially for words occurring later in the sentences.

Statistical Analysis

We analyzed the data with analysis of variance (ANOVA) for repeated measures using the open-source CLEAVE program (T. J. Herron, www.ebire.org/hcnlab/).

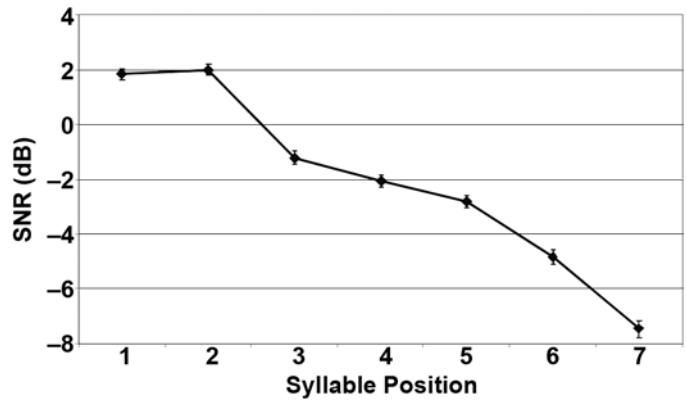

Figure 2.
Signal-to-noise ratios (SNRs) for successive syllables presented in sentences at sentence reception thresholds of 0 dB. Error bars show standard errors of the mean.

The original degrees of freedom are reported for each test with the significance levels adjusted with use of the Box-Greenhouse-Geisser correction for inhomogeneity of variance when appropriate [34]. In these cases, the original degrees of freedom are reported along with corrected significance levels.

RESULTS: EXPERIMENT I

Mean d′ scores were 2.18 for initial and 2.19 for final consonants, very close to the target d′ (2.20). The SNR levels required to equate the identifiability of different consonants varied by >40 dB. Mean d′ thresholds of 1.6 (generating an average hit rate of approximately 50%) were estimated from psychometric functions and are shown in **Table 1**, along with their associated variance measures (standard error of the mean). Consonant thresholds differed systematically in different consonant groups. Mean thresholds for consonants in group A (/ʃ/, /tʃ/, /t/, /s/, /z/, /dʒ/, and /r/) were –4.0 dB (range for different subjects, –11.1 to –1.6 dB), mean thresholds for consonants in group B (/d/, /g/, /l/, /m/, /n/, /f/, and /k/) were 5.6 dB (range 2.3 to 8.2 dB), and mean thresholds for consonants in group C (/b/, /v/, /ð/, /ŋ/, /θ/, /p/, and /h/) were 11.6 dB (range 7.0 to 18.9 dB). Across all test sessions, thresholds for consonants in group A were highly correlated with thresholds for consonants in groups B and C (r = 0.85 and r = 0.78, respectively), and thresholds for consonants in group B were highly correlated with thresholds for consonants in group C (r = 0.90).

Table 1.
Estimated signal-to-noise ratios (in dB) needed to produce identical consonant-identification performance (mean d′ score = 1.6) in the California Syllable Test of all consonants in experiment I.

Consonant	b	d	g	r	l	ŋ	n	m	v	ð	z	dʒ	tʃ	ʃ	s	θ	f	p	t	k	h
Initial																					
Mean	7.4	1.4	4.5	–0.3	6.2	—	1.5	3.6	7.6	10.8	–5.1*	–1.8*	–2.6*	–2.9*	–10.9*	8.1	3.8	8.1	–3.5*	3.5	11.8
SEM	0.4	0.4	0.3	0.4	0.4	—	0.3	0.4	0.5	0.7	0.3	0.4	0.3	0.5	0.6	0.6	0.5	0.5	0.4	0.4	0.4
Final																					
Mean	10.5	9.2	12.0	–4.1*	–1.9*	16.4	12.4	12.0	16.6	27.2	–4.7*	–3.1*	–3.3*	–3.2*	–11.3*	8.1	3.8	8.6	0.3	4.5	—
SEM	0.4	0.5	0.5	0.4	0.4	0.5	0.3	0.3	0.6	0.7	0.5	0.4	0.4	0.5	0.5	0.7	0.5	0.4	0.3	0.2	—

*13 consonants (32.5%) below mean Hearing in Noise Test sentence reception thresholds (–1.8 dB).
SEM = standard error of the mean.

The confusion matrixes obtained in experiment I (averaged over SNRs, subjects, and syllables) are presented in Woods et al. [20]. The patterns of confusion resemble those reported in previous studies [27–29]. A high incidence of place, place + manner, manner, and voicing errors was found, along with a relatively low incidence of multifeature errors that declined rapidly over the B – 6 to B + 6 dB SNR range. These results are presented in detail elsewhere [20].

Consonant Identification in Words and Nonsense Syllables

Since syllables were randomly selected from the syllable corpus, the percentage of words among the 34,560 syllables actually delivered (33.54%) was very similar to the percentage of words in the corpus (33.58%). **Figure 3** shows consonant-identification performance for words and nonsense syllables at the three SNRs. Consonants were identified more accurately in words than nonsense syllables (by an average of 4.7%) as reflected in a significant main effect of syllable type ($F_{1, 15} = 16.83$, $p < 0.001$). Specific comparisons revealed that the percentage of consonants correctly identified in words exceeded the percentage correctly identified in nonwords at baseline SNRs ($F_{1, 15} = 8.82$, $p < 0.01$) and B + 6 dB ($F_{1, 15} = 29.58$, $p < 0.001$), but not at B – 6 dB ($F_{1, 15} = 2.19$, $p < 0.16$). In addition, consonants had steeper psychometric functions in words than nonsense syllables as reflected in a significant SNR × syllable-type interaction ($F_{2, 30} = 15.48$, $p < 0.001$).

We also analyzed the frequency of word and nonword responses. This analysis showed that the overall percentage of word responses (35.98%) exceeded (by 2.4%) the percentage of word stimuli actually delivered ($F_{1, 15} = 6.01$, $p < 0.03$). The small word-response bias did not change significantly with SNR ($F_{2, 30} = 0.79$). We performed subsequent ANOVA to examine the incorrect

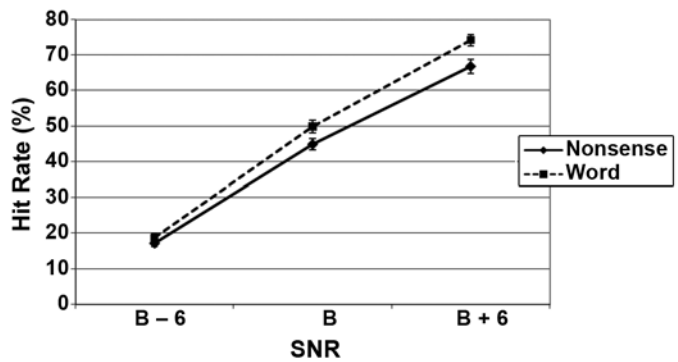

Figure 3.
Hit rates for words and nonsense syllables at different SNRs. Error bars show standard errors of the mean. B = baseline.

responses elicited by word and nonsense-syllable tokens. This analysis revealed a highly significant interaction between the category of the stimulus and the category of the incorrect response ($F_{1, 15} = 214.95$, $p < 0.001$). Incorrect responses to words were more likely to be words than expected by chance (44.56% vs 33.54%), and incorrect responses to nonsense syllables were more likely to be nonsense syllables than predicted by chance (68.55% vs 66.46%). The magnitude of this category bias increased with SNR, as reflected in a significant category-bias × SNR interaction ($F_{2, 30} = 3.87$, $p < 0.05$).

To explore further the nature of this category-specific response bias, we examined the probability of occurrence of different consonants in the word and nonsense-syllable tokens of the corpus. This analysis revealed that the frequency of occurrence of some consonants in words deviated significantly from the aggregate probability of word and nonsense-syllable tokens, as shown in **Table 2**. Some consonants occurred much less frequently in words than would be expected by chance (e.g., /ð/, 7%), whereas others occurred more frequently (e.g., /t/ = 53%). In particular,

Table 2.
Percentage of occurrence of each initial and final consonant in word stimuli. Overall, word stimuli constituted 33.54% of corpus.

Consonant	b	d	g	r	l	ŋ	n	m	V	ð	z	dʒ	tʃ	ʃ	s	θ	f	p	t	k	h
Initial	61	40	31	49	55	—	28	44	6	5	9	24	25	26	47	9	24	46	51	44	47
Final	22	53	17	51	46	34	49	35	17	9	37	7	32	19	30	28	32	58	55	41	—

most fricatives (/f/, /θ/, /v/, /ð/, and /ʃ/) and the affricate /tʃ/ occurred infrequently in words, while plosives (e.g., /b/, /d/, /t/, /k/, and /p/) and liquids (/r/ and /l/) occurred disproportionately in words. Thus, words and nonwords were derived from partially distinct consonant pools. Because single-feature place of articulation errors was the most common confusion observed [20], in correct word or nonsense-syllable reports remained in the same syllable word or nonsense-syllable category as the stimulus (e.g., /bid/ misreported as /did/).

Differences in consonant occurrence in words and nonsense syllables may also help to account for the increased identifiability of consonants in words at B + 6 SNRs. Plosives and liquids occurred disproportionately in words and had steeper performance/SNR functions than the nonsibilant fricatives that occurred disproportionately in nonsense syllables [20]. Correlation analysis showed that the probability that a consonant occurred in words correlated positively with the slope of its psychometric function ($r = 0.52$, $t(18) = 3.02$, $p < 0.01$). Thus, consonants occurring in words are expected to be perceived more accurately at B + 6 dB SNRs than those occurring in nonsense syllables.

Interactions in Processing of Initial and Final Consonants

We examined interactions between the processing of initial and final consonants in words and nonsense syllables. Positive interactions between the processing of initial and final consonants would be reflected in a relative increase in the percentage of trials, where both consonants were identified either correctly or incorrectly, whereas negative interactions would be reflected in a relative decrease in concordant responses. To quantify such interactions, we estimated the predicted probability of concordant responses (both correct or both incorrect) from the observed probabilities of individual initial and final consonant identification for each subject at each SNR. Then, the observed probabilities of concordant responses (both correct + both incorrect) were compared with the probabilities that would be expected by chance.

These results were analyzed with repeated-measures ANOVA with subjects, SNR, syllable type (word or nonsense syllable), and concordance (observed vs predicted probabilities) as factors. The concordance factor was significant ($F_{1, 15} = 27.68$, $p < 0.001$), reflecting the fact that concordant responses occurred more frequently (by +1.6%) than predicted by chance. Further analyses showed that concordant responses significantly exceeded predicted concordant responses at each SNR (B – 6 dB, +2.3%, $F_{1, 18} = 18.47$, $p < 0.006$; B decibel, +1.7%, $F_{1, 18} = 13.25$, $p < 0.003$; B + 6 dB, +0.6%, $F_{1, 18} = 5.10$, $p < 0.05$). A significant interaction was also found between concordance and SNR ($F_{2, 30} = 3.94$, $p < 0.04$) because at B – 6 and B, concordant responses exceeded predicted concordant responses by a greater degree than at B + 6 dB. Finally, a significant interaction was found between concordance and syllable type ($F_{1, 15} = 5.25$, $p < 0.04$), because of the greater concordant responses for words (+2.2%) than for nonsense syllables (+1.1%). However, when nonsense syllables were analyzed in isolation, the concordance factor remained significant ($F_{1, 15} = 16.47$, $p < 0.001$).

Learning Effects

Mean d′ scores (averaged over initial and final consonants) increased over the 3 successive days of testing. Repeated-measures ANOVA with SNR, days, and position as factors showed a significant effect of days ($F_{2, 28} = 13.51$, $p < 0.001$), reflecting a mean improvement of 0.10 d′ units (2.47% hit rate) over the 3 days of testing that was equivalent to an SNR improvement of 0.65 dB. The learning effects neither differed significantly between initial and final consonants ($F_{2, 28} = 1.59$) nor differed in the magnitude of improvement at different SNRs ($F_{4, 56} = 1.35$).

Intersubject Differences and Test-Retest Reliability

Figure 4 shows mean d′ scores (averaged over SNRs) for each of the 16 subjects on each of the 3 days of CaST assessment. Highly significant differences were found between subjects ($F_{15, 30} = 20.68$, $p < 0.001$) with mean d′ scores ranging from 1.81 to 2.33. Test results from individual subjects showed good test-retest reliability: the

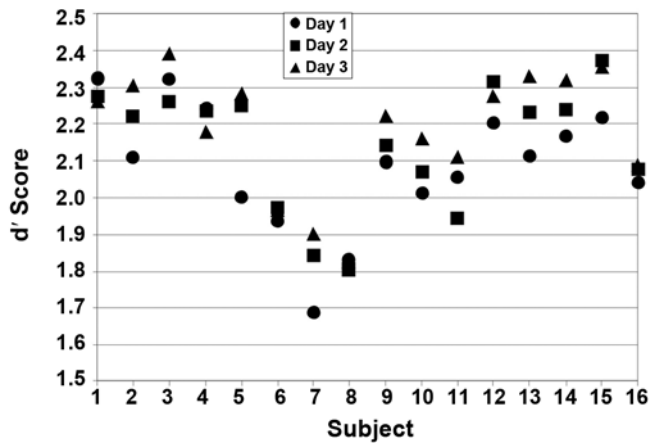

Figure 4.
Mean d′ scores for each subject on each of 3 days of testing.

average within-subject variance was 0.07 d′ units after factoring out mean learning effects. Estimates based on the average psychometric slope of 0.16 d′/dB suggested that average consonant-identification thresholds were measured precisely to approximate 0.7 dB on each testing day.

Correlations Between Audiometric and CaST Thresholds

Overall, d′ scores of different subjects showed weak negative correlations, with variations in their audiometric thresholds at all frequencies ($r = -0.35$, $t(14) = 1.50$, $p < 0.1$) and for the pure tone average (PTA) (PTA = 500–2,000 Hz) ($r = -0.33$, $t(14) = 1.40$, $p < 0.1$). However, separate analyses of consonant groups showed that mean audiometric thresholds did not correlate significantly with group A or group B consonant thresholds but did correlate significantly with group C consonant thresholds ($r = -0.42$, $t(14) = 1.88$, $p < 0.05$). This result was due largely to high-frequency hearing thresholds (3,000–8,000 Hz) that correlated more highly with group C thresholds ($r = -0.34$, $t(14) = 1.45$, $p < 0.1$) than with group A thresholds ($r = -0.14$) or group B ($r = -0.16$) consonants.

Correlations Between SeRT Testing and CaST

HINT SeRTs averaged –1.79 dB (range –1.17 to –2.25). QuickSIN SeRTs averaged +0.35 dB (range – 0.50 to 1.50). HINT and QuickSIN thresholds varied significantly across subjects ($F_{(15, 30)} = 3.26$, $p < 0.003$, and $F_{15, 30} = 6.21$, $p < 0.0001$, respectively). HINT thresholds were not significantly correlated with audiometric thresholds ($r = -0.06$, not significant), but a positive correlation was found between QuickSIN and audiometric thresholds ($r = 0.45$, $t(14) = 2.13$, $p < 0.05$). Performance showed a trend toward improvement over successive days of testing on the HINT ($F_{2, 30} = 2.86$, $p < 0.09$) but not on the QuickSIN ($F_{2, 30} = 0.04$).

Mean CaST thresholds for each subject correlated significantly with SeRTs measured with both the HINT ($r = 0.62$, $t(14) = 3.70$, $p < 0.005$) and the QuickSIN ($r = 0.54$, $t(14) = 2.86$, $p < 0.02$). Indeed, correlations between CaST scores and SeRTs were slightly greater than the correlations between the two SeRT measures ($r = 0.45$, $t(14) = 2.14$, $p < 0.05$). However, average CaST thresholds were significantly higher than mean SeRTs. In fact, an examination of **Table 1** shows that only 32.5 percent of consonants had thresholds below average HINT thresholds. Finally, significant correlations were also found between CaST thresholds and SeRTs measured separately for each of the three consonant groups. For the HINT, the correlations were slightly higher with group B consonant thresholds ($r = 0.71$) than with group A thresholds ($r = 0.56$) or C ($r = 0.58$). A similar pattern was seen for the QuickSIN: thresholds were more strongly correlated with group B consonant thresholds ($r = 0.61$) than with group A ($r = 0.42$) or group C ($r = 0.50$) thresholds.

DISCUSSION: EXPERIMENT I

The confusion patterns obtained from 1 h of CaST assessment closely resembled those reported in lengthier previous studies [27–29]. Previous studies suggest that SNR levels must be adjusted between 18 [27] and 24 dB [29] to produce comparable hit rates across all consonants in a 16-consonant set. We found that even larger SNR ranges (22.7 dB for initial consonants and 38.5 dB for final consonants) were needed to equate consonant identifiability in 20 consonant sets. The increased range of SNRs needed to equate consonant identifiability in the larger consonant sets likely reflects the increased number of possible consonant confusions. The addition of the consonants /l/, /r/, /tʃ/, /h/, and /ŋ/ to the 16 consonant sets used by others increased potential confusions for many consonants (particularly /ð/), reducing their discriminability, and increasing their required baseline SNR levels.

Word and Nonsense-Syllable Identification

In the current experiment, consonants were slightly more accurately identified in words than in nonsense syllables, particularly at high SNRs. These differences occurred even though words and nonsense syllables were delivered in mixed random order. Two factors appeared to account for the increased accuracy of consonant identification in words. First, a small overall response bias toward words was found: the probability of word responses was 2.5 percent higher than the probability of words in the corpus. Second, a large category-specific response bias was found: incorrect responses on word trials were likely to be words, and incorrect responses on nonsense-syllable trials were likely to be nonsense syllables. These effects were primarily due to phonological factors that reflected differences in the consonant pools of word and nonsense-syllable tokens. Some consonants (e.g., unvoiced plosives) occurred disproportionately in words, while others (e.g., voiced fricatives) occurred disproportionately in nonsense syllables. Thus, the common phonological confusions (i.e., single-feature place confusions) resulted in syllable reports that remained in the same category as the syllable presented. Finally, the psychometric functions for the consonants that occurred disproportionately in words were steeper than for consonants that occurred disproportionately in nonsense syllables. Thus, as SNRs increased, the accuracy of word report would be expected to increase more than the accuracy of nonsense-syllable report.

Interactions Between Processing of Initial and Final Consonants

Positive interactions were observed between the identification of initial and final consonants: subjects were more likely to produce concordant responses (either both correct or both incorrect) than predicted by chance. Such facilitatory interactions might be expected for several reasons. First, the subject's level of attention may have varied from trial to trial. On trials during which attention was well focused, the probability of detecting both consonants would be expected to increase. Conversely, if the subject was not attending to the stimuli, the probability of detecting either consonant would be expected to decrease. Second, the rapid identification of the initial consonant might have facilitated formant tracking in the vowel and hence improve the identification of the final consonant. Alternatively, the rapid identification of the initial consonant might have freed phonetic processing resources for final consonant analysis. Although positive interactions were significant at all SNRs, they increased as SNRs were reduced. These results are consistent with models in which syllable elements are processed in an interactive, holistic manner. They argue against models hypothesizing a competition between processing resources devoted to analyzing the initial consonant and those devoted to analyzing the final consonant. Such models would predict concordance below chance levels, particularly at low SNRs.

Although interactions were observed for consonants in nonsense syllables, larger interactions were found in words as previously reported by Boothroyd and Nittrouer [18]. Therefore, we performed further analysis to characterize Boothroyd and Nittrouer's *k*-factor (related to context, $k = 1.0$ for no context) and the *j*-factor (indicating the number of units of information, i.e., $j = 2.0$ for two independent consonants). In comparing words with nonsense syllables, we found *k*-factor = 1.15. Thus, even in conditions in which the majority of stimuli and responses were nonsense syllables, subjects still adopted an implicit word context. Not surprisingly, this benefit was reduced with respect to Boothroyd and Nittrouer's experiment, in which words and nonsense syllables were presented in separate blocks ($k = 1.32$). An analysis of the number of independent units of information revealed *j*-values of 1.70 for words and 1.93 for nonsense syllables. The fact that both values were <2.0 indicated interdependence of initial and final consonant processing for both syllable types, while the greater *j* reduction for words as opposed to nonsense syllables was consistent with a greater interaction of initial and final consonant processing in words that was revealed by ANOVA.

Learning Effects

Most subjects improved their performance over the 3 days of the CaST, with an average improvement of 0.10 d′ units (0.63 dB). Learning effects likely reflected increased familiarity with the talkers' voices, improved estimation of syllable timing during the noise-masking interval, and/or greater familiarity with the permissible stimulus and response alternatives. Similar small procedural-learning effects occur on repeated administration of sentence tests such as the HINT and QuickSIN [42].

Intersubject Variation in Consonant Identification and Sentence Processing

The CaST revealed significant individual differences in consonant-identification ability among young native

American English speakers with normal hearing. The d′ scores across different subjects spanned a range of 0.52 d′ units, corresponding to an SNR difference of 3.25 dB. Intersubject differences in overall performance on the CaST were not significantly correlated with audiometric thresholds, but thresholds for the hardest-to-identify group C consonants did correlate with hearing thresholds, particularly at high frequencies. CaST thresholds accurately predicted SeRTs measured with both the HINT and the QuickSIN. Thus, the CaST measurements of a subject's basic ability to identify consonants in noise provided an accurate estimate of the bottom-up phonological information that subjects could extract when listening to coherent sentences at low SNRs and hence correlated with SeRTs.

We found that only 32.5 percent of consonants could be accurately identified in isolated syllables at the SNRs that characterize SeRTs. This result agrees well with Boothroyd and Nittrouer [18], who reported that about 45 percent of phonemes (vowels included) could be identified in nonsense syllables at SeRTs of predictable sentences. Further analysis showed that consonants fell into three categories:

1. Group A consonants (/ʃ/, /tʃ/, /t/, /s/, /z/, /dʒ/, and /r/) had average thresholds that were 2.2 dB below HINT SeRTs.
2. Group B consonants (/d/, /g/, /l/, /m/, /n/, /f/, and /k/)had average thresholds that were 7.4 dB above HINT SeRTs.
3. Group C consonants (e.g., /b/, /v/, /ð/, /ŋ/, /θ/, /p/, and /h/) had average thresholds that were 13.3 dB above HINT SeRTs.

The strongest correlations between SeRTs and CaST thresholds were observed for consonants in group B. One possible explanation for this result is that the consonants in group A were accurately identified by almost all subjects during sentence testing, while the consonants in group C contributed little to sentence understanding regardless of consonant-identification ability. In contrast, some consonants in group B could be identified during SeRT testing, particularly in subjects with low thresholds for group B consonants. Thus, additional phonetic information from the accurate perception of group B consonants would differentially contribute to lowering SeRT thresholds.

METHODS: EXPERIMENT II

A Case Study

We performed a second experiment to evaluate the capability of the CaST to reveal consonant-processing deficits in a subject with significant bilateral SNHL. A number of studies have reported high correlations between deficits in phoneme processing measures and elevations in SeRTs [21,31,35]. Olsen et al. studied both subjects with normal hearing and those with hearing impairment using word and sentence tests. The population with hearing impairment demonstrated impairments on both tests. However, they also demonstrated increased benefits of sentence context among subjects with hearing impairment [6].

What pattern of phonological impairment would be expected in patients with high-frequency SNHL? Patients with mild to moderate SNHL typically retain low-frequency hearing and show relatively well-preserved discrimination of vowels, syllable durations, and intonation cues compared with consonants. In sentence testing, these cues plus efficient semantic and syntactic processing can mask much larger deficits in consonant perception. Because the phonological discrimination of some consonant manners (e.g., fricatives) depends disproportionally on high-frequency acoustic cues [36], phonological impairments would be expected to vary with consonant manner of articulation.

In addition, the pattern of consonant confusions might be altered in SNHL of gradual onset because of the progressive degradation of the acoustic cues normally used to discriminate consonants. As a result, patients with hearing impairment may use different acoustic cues in consonant discrimination [37] and hence might show altered patterns of consonant confusions compared with subjects with normal hearing.

Elevated SeRTs are also typically reported in patients with hearing loss [6]. However, SeRTs would be expected to show less elevation than CaST thresholds for two reasons. First, subjects with hearing impairment process semantic and syntactic cues as well or better than subjects with normal hearing [6,38]. Second, vowel and intonation processing is better preserved in patients with SNHL than is consonant processing [7,39–40]. Thus, subjects with hearing impairment may compensate for impairments in consonant-identification performance by increasing their reliance on nonconsonant phonological cues (e.g., vowels, intonation, syllable duration) and syntactic and semantic processing.

Subject

The subject was a 65-year-old female patient with mild to moderate SNHL of gradual onset who underwent three test sessions over a 1-week period. Each session included CaST, HINT, and QuickSIN assessment. Audiograms for the control group and the subject with hearing impairment are shown in **Figure 5**. Test procedures were identical for the subject with hearing impairment and the control group, except that the subject with hearing impairment underwent CaST with SNRs increased by 6 dB for all consonants with respect to the SNR levels used in the young control population.

RESULTS: EXPERIMENT II

Despite the fact that SNRs had been increased by 6 dB, the patient's mean d′ scores were slightly reduced compared with those of the control subjects for both initial (2.08 vs 2.18) and final (2.04 vs 2.19) consonants. Based on mean psychometric function slopes from the control population, the subject with hearing impairment required a mean SNR increase of 6.8 dB to achieve identification performance equivalent to the mean performance seen in the control group. Sentence testing also revealed elevated SeRTs on both the HINT (–0.6 dB, +1.2 dB compared with controls) and the QuickSIN (2.2 dB, +1.8 dB compared with controls) that reached significance for both tests (HINT, z score = 3.95, $p < 0.001$; QuickSIN z score = 3.67, $p < 0.001$).

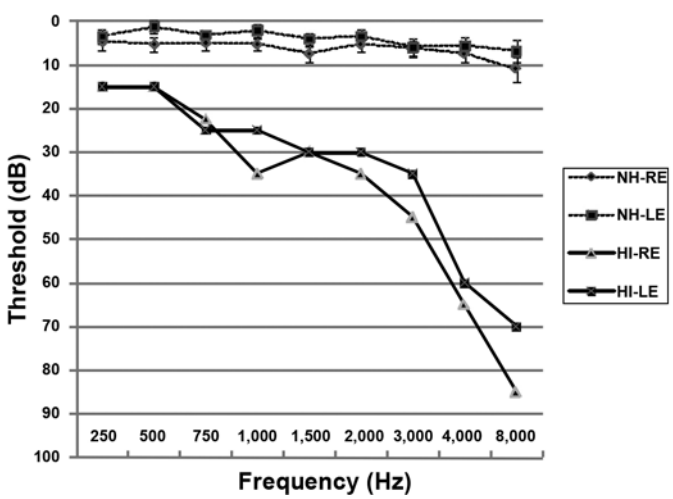

Figure 5.
Audiometric test results showing mean thresholds and standard errors of the mean for subjects in control group with normal hearing (NH) and audiometric results for subject with hearing impairment (HI). LE = left ear, RE = right ear.

Estimated CaST identification thresholds are shown in **Table 3**. Average CaST threshold elevations were significantly greater than SeRT elevations so that only 12.5 percent of consonants had SNR thresholds below average HINT SeRTs (see values with asterisks in **Table 3**). The magnitude of SNR elevation varied substantially for different consonants as shown in **Figure 6**. Small elevations were seen for affricates, liquids, and nasals; intermediate elevations were seen for plosives; and large elevations were observed for most fricatives. Overall, consonant-identification thresholds were significantly elevated for 14 of the 19 consonants that occurred in both initial and final syllable position (z score range 3.1 to 12.4).

The patient's confusion matrixes for initial and final consonant processing are presented in **Tables 4** and **5,** respectively. These confusion matrixes reveal that consonants fall into confusable clusters of varying sizes as in the control population. However, the subject with hearing impairment showed frequent confusions that were rarely seen in the control population. For example, in initial syllable position, the subject with hearing impairment frequently confused /b/ with /f/ and also confused both /ð/ and /v/ with many other consonants.

The pattern of consonant confusions can be visualized in individual subjects using cluster-analysis techniques [20,41] as shown in **Figure 7**. Observed consonant confusions of the patient are shown as colored x's, with the magnitude of displacements from the initial consonant locations (dotted lines) reflecting the type of

Table 3.
Estimated signal-to-noise ratios (in dB) needed to produce consonant-identification performance (mean d′ score = 1.6) in the California Syllable Test for subject with hearing impairment in experiment II.

Consonant	b	d	g	r	l	ŋ	n	m	v	ð	z	ʤ	ʧ	ʃ	s	θ	f	p	t	k	h
Initial	16.3	6.8	8.4	2.2	10.2	—	4.4	6.9	15.7	38.6	2.2	2.0	–0.5	1.0	–1.1*	32.5	14.9	15.5	4.2	7.6	19.6
Final	19.1	15.1	16.4	–1.3*	0.7	23.1	22.8	17.4	28.9	52.0	1.8	–1.7*	–1.9*	–0.2	–1.6*	26.6	18.2	17.2	6.9	11.9	—

*Five consonants (12.5%) at or below subject's Hearing in Noise Test threshold (–0.6 dB).

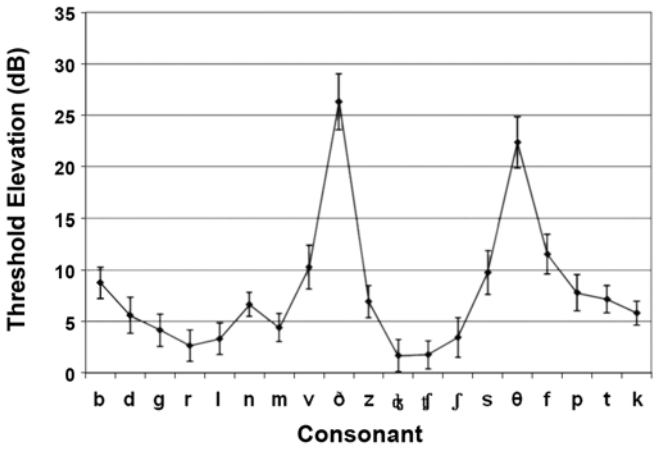

Figure 6.
Mean threshold elevations needed to obtain d′ scores of 2.2 for each consonant for subject with hearing impairment relative to control subjects. Error bars show standard deviations estimated from 3 days of testing.

confusions. The distance between consonant "x" pairs for the patient reflects their discriminability. The polygons in **Figure 7** show the range of consonant confusions observed for initial and final consonants (**Figure 7(a)** and **(b)**, respectively) in each of the 16 control subjects, with polygon color-coding consonant identity. For example, in control subjects, initial unvoiced plosives and /h/ (cyan) are clustered together in the upper left of the confusion circle, with /k/ occurring at the intersection of /h/ and /p/ spaces and /t/ well discriminated both from the other unvoiced plosives and /h/. In the patient, /h/ falls within a normal location and /k/ is rarely confused with other consonants and so remains near the circle periphery. However, /t/ falls within the normal /p/ cluster and is poorly discriminated from /p/.

Confusion abnormalities were more striking for other consonants. For example, for initial consonants (**Figure 7(a)**), the /b/ confusion cluster for control subjects (dark green, center right) is located close to the confusion clusters of other voiced plosives. In contrast, the initial /b/ (**Figure 7(a)**) for the subject with hearing impairment was located near the center of the normal /t/ cluster, close to the unvoiced plosives. This result reflects the fact that the subject with hearing impairment frequently confused /b/-/f/ initial syllable position. As a result, /b/ and, to a lesser extent, /f/ were displaced toward a location intermediate between their initial locations. The subject with hearing impairment also showed abnormal clustering of /ð/-/v/ in initial-syllable position. In subjects with normal hearing, /v/ and /ð/ clusters are located in the lower right portion of the confusion circle because their confusions are largely restricted to each other, liquids, and nasals. In the subject with hearing impairment, both /v/ and /ð/ confusions were displaced to a point near the confusion circle center (red dot), indicating that they were frequently confused with many other consonants, including voiced and unvoiced plosives. In addition, the locations of the /v/ and /ð/ of the subject with hearing impairment were virtually superimposed, reflecting near-chance discrimination between these two consonants. In contrast, the /ʃ/ and /ʧ/ of the subject with hearing impairment are located near the circle circumference (lower right). This result reflects the fact that the subject with hearing impairment benefited from the 6 dB increase in SNRs to accurately discriminate these consonants both from each other and from other consonants.

Among final consonants (**Figure 7(b)**), the subject with hearing impairment showed relatively normal locations of voiced and unvoiced plosives, although impaired /p/-/t/ discrimination was again found. Confusions for /s/ with voiced plosives exceeded similar confusions in control subjects, so /s/ was displaced toward the circle center. As in the initial consonant position, the subject with hearing impairment showed very poor discrimination of /v/-/ð/. However, /v/ and /ð/ confusions with plosives were reduced in the final consonant position so that both consonants remained in a location similar to that of subjects with normal hearing. As in initial consonant position, the subject with hearing impairment effectively discriminated affricates and /ʃ/ both from each other and from other consonants.

DISCUSSION: EXPERIMENT II

The audiogram of the subject with hearing impairment had a sloping contour with mild bilateral losses (mean 30 dB) at 750 to 2,000 Hz that increased to losses of 60 dB at 4,000 Hz and nearly 90 dB at 8,000 Hz. Consonant-identification mean thresholds increased by 6.8 dB. Small threshold elevations were seen for affricates, liquids, and nasals; intermediate elevations were seen for plosives; and large elevations were observed for fricatives. Sentence testing revealed the patient's SeRT elevations (1.2 dB in the HINT and 1.8 dB in the QuickSIN) were much smaller than the increases in

Table 4.
Initial consonant confusions for subject with hearing impairment.

Consonant	b	d	g	r	l	n	m	v	ð	z	ʤ	ʧ	ʃ	s	θ	f	p	t	k	h
b	50	2	2	2	2	0	0	3	0	1	0	0	0	0	6	26	0	0	0	14
d	3	78	3	0	2	0	2	0	1	0	1	0	0	6	1	1	2	0	1	7
g	1	9	76	2	0	0	0	1	1	0	2	1	0	1	0	3	1	0	3	7
r	1	0	1	89	6	0	0	3	0	0	0	0	0	0	1	1	1	0	0	5
l	0	0	0	4	92	2	2	3	1	1	0	0	0	0	0	1	0	0	1	1
n	0	1	1	5	20	74	4	0	0	0	0	0	0	0	0	0	0	0	0	3
m	5	1	1	2	11	7	74	5	0	0	1	0	0	0	0	0	0	0	0	1
v	10	0	0	5	5	0	2	66	1	2	0	0	0	0	3	8	1	0	1	4
ð	1	9	1	0	22	0	1	38	10	9	0	0	0	3	8	6	0	0	0	0
z	2	7	2	0	7	1	1	4	2	53	5	2	1	11	2	1	0	3	1	3
ʤ	0	2	5	2	1	1	0	0	0	1	78	11	0	1	0	0	1	4	0	1
ʧ	0	1	0	0	0	0	0	0	0	0	4	92	3	1	0	0	0	6	1	0
ʃ	0	0	0	0	0	0	0	0	0	0	3	28	66	2	0	1	0	6	1	1
s	1	2	3	1	2	0	1	0	0	7	2	2	3	58	1	6	2	11	2	4
f	7	0	0	1	0	0	0	1	0	0	0	0	1	3	10	79	0	0	1	5
θ	1	0	0	0	0	0	0	0	1	2	0	0	0	16	32	47	0	2	0	7
p	0	0	0	1	0	2	0	0	0	0	0	0	0	0	0	3	75	5	3	19
t	0	0	0	1	1	0	0	0	0	0	0	3	1	5	1	7	7	58	7	17
k	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	8	6	87	5
h	0	0	0	0	1	1	0	0	0	0	0	0	0	1	1	1	10	3	4	86

Table 5.
Final consonant confusions for subject with hearing impairment.

Consonant	b	d	g	r	l	ŋ	n	m	v	ð	z	ʤ	ʧ	ʃ	s	θ	f	p	t	k
b	61	7	6	0	0	1	1	3	20	5	0	0	0	0	0	1	0	3	0	0
d	3	76	4	0	0	0	0	0	6	10	2	0	0	0	1	1	3	0	2	0
g	7	3	80	0	0	0	0	1	10	5	0	0	0	0	0	1	0	0	1	0
r	1	2	3	73	4	1	3	1	9	0	4	1	1	0	1	0	0	2	1	1
l	2	1	4	4	77	2	2	4	9	0	2	0	0	0	0	0	0	0	0	1
ŋ	0	0	0	0	0	57	18	25	5	2	1	0	0	0	0	0	0	0	0	0
n	0	0	0	0	0	9	55	34	6	2	0	1	0	0	1	0	0	0	0	0
m	1	0	0	0	0	3	4	95	4	1	0	0	0	0	0	0	0	0	0	0
v	7	1	3	1	2	0	0	1	80	11	2	0	0	0	0	0	0	0	0	0
ð	3	8	5	0	2	0	1	0	57	30	2	0	0	0	0	0	0	0	0	0
z	6	5	4	0	2	0	0	3	9	4	56	10	1	1	2	1	0	0	1	3
ʤ	1	0	2	0	0	0	0	1	1	0	0	97	5	0	0	0	0	0	0	1
ʧ	0	0	0	0	0	0	1	0	0	0	0	5	95	2	2	0	0	0	0	3
ʃ	1	0	0	0	0	0	0	0	0	1	1	11	17	70	1	0	3	0	1	2
s	7	6	2	0	1	2	0	2	4	1	3	6	4	3	46	1	2	4	8	6
θ	0	1	2	1	1	1	0	0	5	0	0	0	0	4	12	45	31	0	2	3
f	4	1	0	0	1	0	2	1	5	2	1	1	0	0	4	17	57	5	1	6
p	9	0	0	0	0	1	1	0	3	0	1	0	0	0	1	4	9	57	10	12
t	2	2	0	0	0	1	0	1	1	0	0	0	3	0	4	4	1	9	71	9
k	2	0	1	0	0	0	1	0	1	0	0	0	0	0	0	6	9	8	8	72

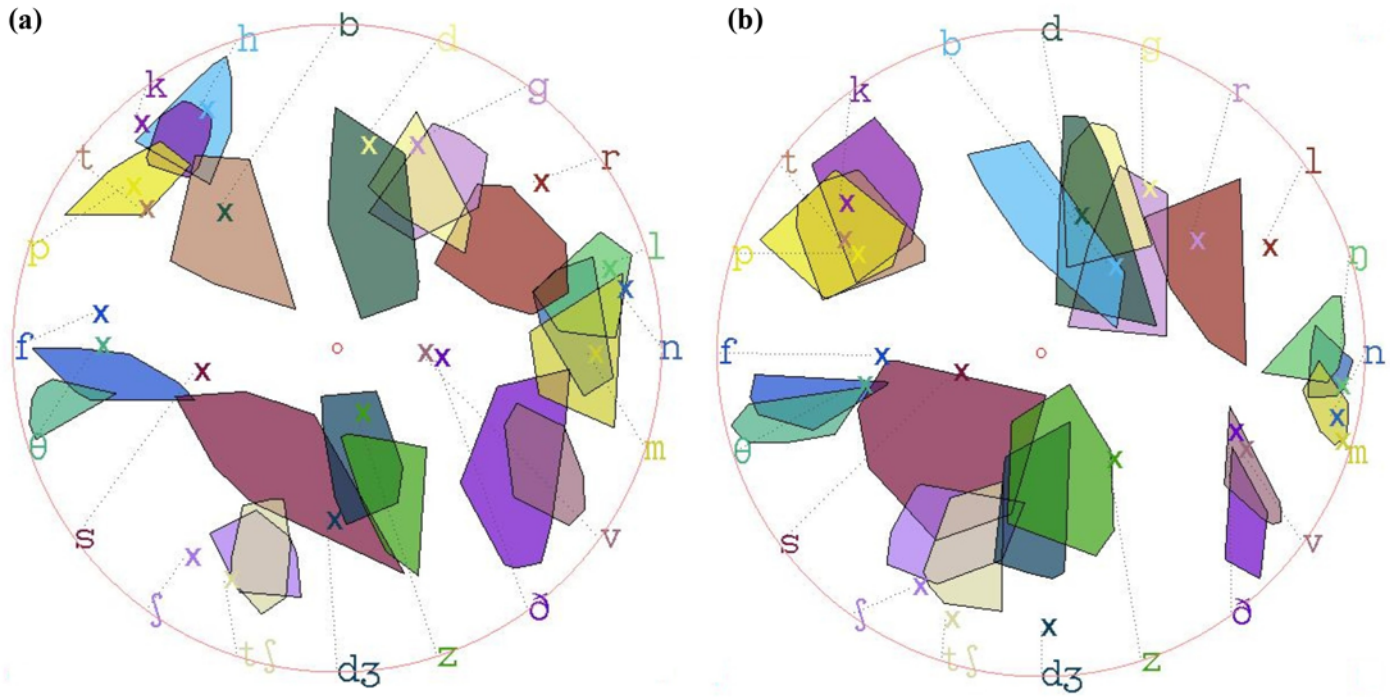

Figure 7.
Abnormalities in consonant-confusion clustering. **(a)** Initial and **(b)** final consonant-confusion clusters for control population and subject with hearing impairment. Location and area of each polygon shows range of consonant confusions in control population with polygon color coding consonant identity. Observed consonant confusions for patient are shown as colored x's, with displacements from starting consonant locations shown as dotted lines. Data have been averaged over three California Syllable Test sessions.

consonant-identification thresholds. This finding is consistent with the results from experiment I, suggesting that only 32.5 percent of consonants are identifiable at SeRTs in subjects with normal hearing. In the patient, only 12.5 percent of consonants could be identified at SeRTs. This finding suggests that the sentence comprehension of the subject with hearing impairment relied more on nonconsonant cues and sentence context than does comprehension in subjects with normal hearing.

Consonant-confusion analysis revealed a number of unusual confusions in the subject with hearing impairment. We found increased confusions among unvoiced plosives /p/ and /t/ and in the higher-than-normal incidence of multifeature errors involving nonsibilant fricatives in initial syllable position. This increase likely reflected that the subject with hearing impairment was unable to use the high-frequency cues that distinguish these phonemes and therefore produced a more random pattern of responses than those seen in subjects with normal hearing. Interestingly, however, some unusual multifeature confusion occurred systematically. For example, the subject with hearing impairment made frequent manner + voicing errors in confusions of initial /b/-/f/. In contrast, few manner + voicing confusions were seen for the similar plosive-fricative pair, /d/-/th/. Further studies of larger groups of subjects with hearing impairment losses are needed to determine if these unusual confusion patterns represent idiosyncratic or systematic adaptations to high-frequency hearing loss.

The CaST provided accurate estimates of the ability of individual subjects to identify a large selection of initial and final consonants in spoken American English. Because CaST tokens were randomly sampled from an extremely large corpus that included both within- and between-talker variation, CaST results likely reflect typical consonant-identification patterns of spoken American English CVCs. Because word and nonsense-syllable tokens are presented in random order, the CaST minimizes the influence of semantic and syntactic processing set. Thus, compared with sentence or word tests, it directly measures the ability of subjects to use the acoustic features of speech to identify consonants.

The CaST also quantifies identification performance for each consonant, including the 67 percent of consonants whose thresholds are normally above SeRTs and thereby contribute little to SeRTs measured with common sentence tests. Many of these consonants occur frequently in American English words, and some remain difficult to identify in everyday listening conditions at moderate noise levels, particularly for subjects with hearing impairments. Effective audiological rehabilitation to improve identification of the more difficult consonants would be expected to improve patient comprehension, reduce patient effort in everyday listening conditions, and enhance patient satisfaction with hearing aids. The CaST can help quantify these improvements.

Cluster analysis permits the visualization of abnormal patterns of consonant confusion in patients with hearing loss. As SNHL develops, patients are deprived of the normal acoustic cues needed to discriminate different consonants and come to use other phonetic cues that remain available. Neuroplastic changes may occur in the phoneme-processing regions of auditory cortex as the patient comes to rely excessively on vowel- and nonoptimal consonant cues [12]. These changes may contribute to abnormal consonant confusions that cannot be explained simply on peripheral hearing loss.

CONCLUSIONS

SNHL produces deficits in consonant identification in noise that cannot be accurately measured with existing sentence comprehension tests. The CaST measures a patient's ability to identify consonants using a large randomly sampled token corpus to measure consonant-identification performance for 21 common American English consonants. CaST consonant-identification thresholds correlated with SeRTs measured with the HINT and QuickSIN. However, consonants could be divided into three groups based on the SNRs needed for their identification. Consonants in group A and some consonants in group B were identifiable at SNRs at or below the SeRT. In contrast, other consonants in group B and all consonants in group C had identification thresholds that were well above SeRTs measured. This finding suggests that SeRTs primarily reflect the contribution of one-third of American English consonants, while the remaining consonants contribute little to SeRTs. Large deficits in consonant processing were seen in a subject with bilateral high-frequency hearing loss along with small elevations in SeRTs. A comparison of SeRTs and consonant-identification thresholds suggested that the patient relied disproportionally on nonconsonant phonological cues. Consonant-identification profile analysis showed that deficits were particularly striking for hard-to-identify consonants, including nonsibilant fricatives. Confusion-cluster analysis revealed abnormal confusion patterns that may have reflected idiosyncratic central nervous system adaptations to peripheral hearing loss. Consonant profile analysis with the CaST well predicts speech comprehension in a variety of noise-masking conditions and provides insight into consonant-identification difficulties that cannot be detected with current sentence testing.

ACKNOWLEDGMENTS

Author Contributions:
Study concept and design: D. L. Woods, E. W. Yund.
Acquisition of data: E. W. Yund.
Analysis and interpretation of data: D. L. Woods, E. W. Yund, T. J. Herron.
Drafting of manuscript: D. L. Woods.
Critical revision of manuscript for important intellectual content: E. W. Yund, T. J. Herron.
Obtained funding: D. L. Woods.
Administrative, technical, or material support: D. L. Woods.
Study supervision: D. L. Woods.
Financial Disclosures: David L. Woods is affiliated with Neurobehavioral Systems, Inc, the creators of Presentation software used for the creation of these experiments.
Funding/Support: This material was based on work supported by the Department of Veterans Affairs (VA) Rehabilitation, Research and Development Service, grant C4739R, to Dr. Woods. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the VA or the Department of Defense.
Additional Contributions: Technical assistance was provided by Matthew Ua Cruadhlaoich and Zoe Doss. Pierre Divenyi provided helpful suggestions on early versions of this article.
Institutional Review: The VA gave institutional human subjects approval.
Participant Follow-Up: The authors do not plan to inform participants of the publication of this study.

REFERENCES

1. Arlinger S. Negative consequences of uncorrected hearing loss—A review. Int J Audiol. 2003;42 Suppl 2:2S17–20. [PMID: 12918624]
2. Kuk F, Keenan D, Lau CC, Dinulescu N, Cortez R, Keogh P. Real-world performance of a reverse-horn vent. J Am

Acad Audiol. 2005;16(9):653–61. [PMID: 16515137] DOI:10.3766/jaaa.16.9.3

3. Walden TC, Walden BE. Predicting success with hearing aids in everyday living. J Am Acad Audiol. 2004;15(5): 342–52. [PMID: 15506496] DOI:10.3766/jaaa.15.5.2
4. Valente M, Mispagel KM. Unaided and aided performance with a directional open-fit hearing aid. Int J Audiol. 2008; 47(6):329–36. [PMID: 18569105] DOI:10.1080/14992020801894832
5. Walden TC, Walden BE. Unilateral versus bilateral amplification for adults with impaired hearing. J Am Acad Audiol. 2005;16(8):574–84. [PMID: 16295244] DOI:10.3766/jaaa.16.8.6
6. Olsen WO, Van Tasell DJ, Speaks CE. The Carhart Memorial Lecture, American Auditory Society, Salt Lake City, Utah, 1996. Phoneme and word recognition for words in isolation and in sentences. Ear Hear. 1997;18(3):175–88. [PMID: 9201453] DOI:10.1097/00003446-199706000-00001
7. Owens E, Talbott CB, Schubert ED. Vowel discrimination in hearing-impaired listeners. J Speech Hear Res. 1968;11(3): 648–55. [PMID: 5722492]
8. Ferguson SH, Kewley-Port D. Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2002;112(1):259–71. [PMID: 12141351] DOI:10.1121/1.1482078
9. Ahlstrom JB, Horwitz AR, Dubno JR. Spatial benefit of bilateral hearing aids. Ear Hear. 2009;30(2):203–18. [PMID: 19194292] DOI:10.1097/AUD.0b013e31819769c1
10. Takahashi G, Martinez CD, Beamer S, Bridges J, Noffsinger D, Sugiura K, Bratt GW, Williams DW. Subjective measures of hearing aid benefit and satisfaction in the NIDCD/VA follow-up study. J Am Acad Audiol. 2007;18(4):323–49. [PMID: 17580727] DOI:10.3766/jaaa.18.4.6
11. Blamey PJ, Fiket HJ, Steele BR. Improving speech intelligibility in background noise with an adaptive directional microphone. J Am Acad Audiol. 2006;17(7):519–30. [PMID: 16927516] DOI:10.3766/jaaa.17.7.7
12. Woods DL, Yund EW. Perceptual training of phoneme identification for hearing loss. Semin Hear. 2007;28(2): 110–19. DOI:10.1055/s-2007-973437
13. Stecker GC, Bowman GA, Yund EW, Herron TJ, Roup CM, Woods DL. Perceptual training improves syllable identification in new and experienced hearing aid users. J Rehabil Res Dev. 2006;43(4):537–52. [PMID: 17123192] DOI:10.1682/JRRD.2005.11.0171
14. Sweetow RW, Sabes JH. The need for and development of an adaptive Listening and Communication Enhancement (LACE) Program. J Am Acad Audiol. 2006;17(8):538–58. [PMID: 16999250] DOI:10.3766/jaaa.17.8.2
15. Sweetow RW, Sabes JH. Technologic advances in aural rehabilitation: Applications and innovative methods of service delivery. Trends Amplif. 2007;11(2):101–11. [PMID: 17494876] DOI:10.1177/1084713807301321
16. Nilsson M, Soli SD, Sullivan JA. Development of the Hearing in Noise Test for the measurement of speech reception thresholds in quiet and in noise. J Acoust Soc Am. 1994;95(2):1085–99. [PMID: 8132902] DOI:10.1121/1.408469
17. Killion MC, Niquette PA, Gudmundsen GI, Revit LJ, Banerjee S. Development of a quick speech-in-noise test for measuring signal-to-noise ratio loss in normal-hearing and hearing-impaired listeners. J Acoust Soc Am. 2004;116(4 Pt 1):2395–2405. [PMID: 15532670] DOI:10.1121/1.1784440
18. Boothroyd A, Nittrouer S. Mathematical treatment of context effects in phoneme and word recognition. J Acoust Soc Am. 1988;84(1):101–14. [PMID: 3411038] DOI:10.1121/1.396976
19. Benki JR. Quantitative evaluation of lexical status, word frequency, and neighborhood density as context effects in spoken word recognition. J Acoust Soc Am. 2003;113(3): 1689–1705. [PMID: 12656401] DOI:10.1121/1.1534102
20. Woods DL, Yund EW, Ua Cruadhlaoich MAI. Consonant identification in consonant-vowel-consonant syllables in speech spectrum noise. J Acoust Soc Am. 2010;127(3): 1609–23. DOI:10.1121/1.3293005
21. Revoile S, Kozma-Spytek L, Holden-Pitt L, Picket tJ, Droge J. Acoustic-phonetic context considerations for speech recognition testing of hearing-impaired listeners. Ear Hear. 1995;16(3):254–62. [PMID: 7672474] DOI:10.1097/00003446-199506000-00002
22. Bilger RC, Wang MD. Consonant confusions in patients with sensorineural hearing loss. J Speech Hear Res. 1976; 19(4):718–48. [PMID: 1003952]
23. Pichora-Fuller MK, Schneider BA, Daneman M. How young and old adults listen to and remember speech in noise. J Acoust Soc Am. 1995;97(1):593–608. [PMID: 7860836] DOI:10.1121/1.412282
24. Dubno JR, Lee FS, Klein AJ, Matthews LJ, Lam CF. Confidence limits for maximum word-recognition scores. J Speech Hear Res. 1995;38(2):490–502. [PMID: 7596114]
25. Hurley RM, Sells JP. An abbreviated word recognition protocol based on item difficulty. Ear Hear. 2003;24(2):111–18.

[PMID: 12677108]
DOI:10.1097/01.AUD.0000058113.56906.0D

26. Redford MA, Diehl RL. The relative perceptual distinctiveness of initial and final consonants in CVC syllables. J Acoust Soc Am. 1999;106(3 Pt 1):1555–65. [PMID: 10489711] DOI:10.1121/1.427152
27. Miller GA, Nicely PE. An analysis of perceptual confusions among some English consonants. J Acoust Soc Am. 1955;27(2):338–52. DOI:10.1121/1.1907526
28. Wang MD, Bilger RC. Consonant confusions in noise: A study of perceptual features. J Acoust Soc Am. 1973;54(5):1248–66. [PMID: 4765809] DOI:10.1121/1.1914417
29. Phatak SA, Allen JB. Consonant and vowel confusions in speech-weighted noise. J Acoust Soc Am. 2007;121(4):2312–26. [PMID: 17471744] DOI:10.1121/1.2642397
30. Dubno JR, Levitt H. Predicting consonant confusions from acoustic analysis. J Acoust Soc Am. 1981;69(1):249–61. [PMID: 7217523] DOI:10.1121/1.385345
31. Bosman AJ, Smoorenburg GF. Intelligibility of Dutch CVC syllables and sentences for listeners with normal hearing and with three types of hearing impairment. Int J Audiology. 1995;34(5):260–84. [PMID: 8837785] DOI:10.3109/00206099509071918
32. Gelfand SA. Optimizing the reliability of speech recognition scores. J Speech Lang Hear Res. 1998;41(5):1088–1102. [PMID: 9771631]
33. Green DM, Swets JA. Signal detection theory and psychophysics. Huntington (NY): Robert E. Krieger; 1974.
34. Greenhouse SW, Geisser S. On methods in the analysis of profile data. Psychometrika. 1959;24(2):95–112. DOI:10.1007/BF02289823
35. Yund EW, Woods DL. Content and procedural learning in repeated sentence tests of speech perception. Ear Hear. Forthcoming 2010.
36. Revoile SG, Kozma-Spytek L, Holden-Pitt L, Pickett JM, Droge J. VCVs vs CVCs for stop/fricative distinctions by hearing-impaired and normal-hearing listeners. J Acoust Soc Am. 1991;89(1):457–60. [PMID: 2002178] DOI:10.1121/1.400483
37. Zeng FG, Turner CW. Recognition of voiceless fricatives by normal and hearing-impaired subjects. J Speech Hear Res. 1990;33(3):440–49. [PMID: 2232762]
38. Hedrick MS, Jesteadt W. Effect of relative amplitude, presentation level, and vowel duration on perception of voiceless stop consonants by normal and hearing-impaired listeners. J Acoust Soc Am. 1996;100(5):3398–3407. [PMID: 8914319] DOI:10.1121/1.416981
39. Abada SH, Baum SR, Titone D. The effects of contextual strength on phonetic identification in younger and older listeners. Exp Aging Res. 2008;34(3):232–50. [PMID: 18568981] DOI:10.1080/03610730802070183
40. Sammeth CA, Dorman MF, Stearns CJ. The role of consonant-vowel amplitude ratio in the recognition of voiceless stop consonants by listeners with hearing impairment. J Speech Lang Hear Res. 1999;42(1):42–55. [PMID: 10025542]
41. Coughlin M, Kewley-Port D, Humes LE. The relation between identification and discrimination of vowels in young and elderly listeners. J Acoust Soc Am. 1998;104(6):3597–3607. [PMID: 9857518] DOI:10.1121/1.423942
42. Cohen J. Graph twiddling in a MapReduce World. Comput Sci Eng. 2009;11(4):29–41. DOI:10.1109/MCSE.2009.120

Submitted for publication April 8, 2009. Accepted in revised form January 6, 2010.