Evaluation of a Tool for Measuring Temporal Modulation Detection

David M. Landsberger and Natalia Stupak

Objectives: A software tool (EasyMDT) that measures temporal modulation detection thresholds of a broadband noise carrier is presented. EasyMDT is designed to be both easy and quick to promote the use in environments where testing time is limited, and testers may not have extensive technical expertise to use typical research software. In addition, by providing a standardized stimulus and protocol, data collected by all groups using the software can be compared directly. Details of EasyMDT, including a description of the protocol, stimuli, interface and how to obtain the software, are provided along with representative sample data from both normal-hearing listeners and cochlear implant (CI) users. Performance with the EasyMDT is compared with speech understanding metrics as well as a metric of spectral-temporal resolution.

Design: A "Full Curve" of modulation detection thresholds is measured using a three-interval forced-choice adaptive task in a single block for 7 modulation frequencies (10, 50, 75, 100, 150, 200, and 300 Hz). Similarly, the modulation detection thresholds were measured for only one modulation frequency in a block (either 100 Hz or 150 Hz). Modulation detection thresholds and block duration were recorded. In addition, performance on speech recognition tasks (CNC words, consonant identification, vowel identification, and AzBio sentences in noise) and a spectral-temporal resolution task (SMRT; Aronoff and Landsberger) were measured. Modulation detection thresholds were measured for both normal-hearing listeners and CI users. Only CI users participated in the speech and spectral-temporal tests.

Results: Modulation detection thresholds measured with EasyMDT were consistent with those previously reported from other laboratories. Modulation detection thresholds at a single modulation frequency (100 Hz or 150 Hz) were predictive of modulation detection thresholds measured as part of the Full Curve consisting of all 7 modulation frequencies. Testing durations for CI users dropped from an average of over 18 minutes for the Full Curve to under 3 minutes for either of the single modulation frequency measures. Modulation detection thresholds at 100 Hz correlated with CNC words, consonant identification, and AzBio sentences in noise, but not vowel identification. No correlations were found between modulation detection and spectral-temporal resolution.

Conclusions: The EasyMDT is designed to be an easy-to-use tool that provides a nonlinguistic measure that can predict speech understanding. The test duration is short enough that it can be incorporated into clinical practice or as part of an experimental battery. The software is available for free download at https://www.ear-lab.org/software-downloads.html. The software is designed to have a minimum barrier of entry as well as provide a standardized protocol allowing direct comparison of modulation detection thresholds across studies and groups.

Key words: Cochlear implant, Temporal modulation detection, Software tools

(Ear & Hearing XXX;XX;00–00)

INTRODUCTION

Temporal processing is an important attribute of auditory processing and a key component of speech understanding (Steeneken and Houtgast 1980; Van Tassel et al. 1987; Freyman et al. 1991). Performance on various temporal tasks has been shown to correlate with speech understanding in both normal hearing (Steeneken and Houtgast 1980; van Tassel et al. 1987) and various hearing-impaired populations (Cazals et al. 1994; Fu 2002). As such, there has been great interest in studying psychoacoustic temporal properties in auditory labs (e.g., Hall and Grosse 1994; Won et al. 2011; Park et al. 2015; Landsberger et al. 2019; Zhou et al. 2020). Furthermore, if a temporal processing task was sufficiently quick and simple, it might also be of interest as a clinical measure of nonlinguistic auditory processing.

One measure of temporal processing is described in Bacon and Viemeister (1985) in which the minimum detectable modulation depth of a broadband noise carrier (i.e., the modulation depth threshold or MDT) is measured. The process is repeated for multiple modulation rates. Using these data, a curve is generated describing modulation detection thresholds as a function of modulation frequency. This curve is called the Temporal Modulation Transfer Function (TMTF) for modulation detection. The Bacon and Viemeister (1985) protocol was replicated by Won et al. (2011) with cochlear implant (CI) users. With the CI users, performance on modulation detection was found to be correlated with CNC words in quiet as well as speech reception thresholds for spondees in steady state noise, suggesting that it may be a nonlinguistic measure capturing an important component of speech perception. Furthermore, the test seems to be devoid of spectral confounds. In addition to no obvious cues available in a visual analysis of spectrograms of the stimuli, Won et al. (2011) failed to find a correlation between modulation detection thresholds and a spectral ripple task. Furthermore, performance on the task was unaffected by the number of electrodes provided to the CI user.

We developed a software tool called EasyMDT to evaluate temporal modulation detection with several goals in mind. First, the software was designed to be easy to use for both the experimenter and for the participant. This encourages the use of the test by people who are less technically oriented or do not have the time to invest in developing their own variant. Second, the software was made freely available to any interested researcher or clinician and can be downloaded from https://www.ear-lab.org/software-downloads.html. This is important in that it reduces the barriers of entry, allowing a researcher or clinician to obtain the tool quickly without limitations of licensing or budgeting. Third, by providing a tool with a fixed protocol and stimulus set, data collected across many studies and groups can be compared directly. As such, each new study using the tool can compare results with previous studies using the same tool as well as provide further reference data for future studies. Fourth, the test should be as efficient as possible. Shorter test durations promote use and reduce testing fatigue. Furthermore, if the test is sufficiently quick, it could become feasible to use in a clinical setting.

Department of Otolaryngology, New York University Grossman School of Medicine, New York, New York, USA.

0196/0202/XXX/XXXX-00/0 • Ear & Hearing • Copyright © 2021 Wolters Kluwer Health, Inc. All rights reserved • Printed in the U.S.A.

<zdoi; 10.1097/AUD.0000000000001107>

The present article will describe the EasyMDT tool, including the protocol, stimuli, and interface as well as how to obtain it. Representative sample data collected with the tool on both normal-hearing (NH) listeners and CI users will be presented. It is expected that the time taken to collect a full TMTF curve would be sufficiently long, preventing use in a clinical environment and discouraging inclusion as part of a battery of tests in a research environment. To address these issues, results and test duration will be compared for measurements of the full TMTF curve and separately measured modulation detection thresholds of individual modulation frequencies. Performance will be measured on CNC word, sentence, vowel, and consonant tasks to determine their relationships with data collected using the EasyMDT tool. In addition, MDTs will be compared with a spectral-temporal task (Spectral-temporally Modulated Ripple Test; SMRT; Aronoff & Landsberger (2013)) that has also been demonstrated to predict speech understanding for CI users (e.g., Holden et al. 2016; Lawler et al. 2017).

MATERIALS AND METHODS

Subjects

Twelve CI users and 10 NH listeners participated in this experiment. The CI users were between 33 and 81 years of age (mean 59 years, standard deviation 15). All CI users were tested using their clinical processors and standard settings. Bilateral CI users were tested only with their self-reported "best ear" (i.e., in a unilateral condition). Participants with residual hearing were tested with an earplug in the corresponding ear. Users of all FDA approved manufacturers (6 Advanced Bionics users, 3 Cochlear Ltd. users, and 3 MED-EL users) participated. NH listeners were between the ages of 23 and 37 years (mean 29 years, standard deviation 5) and required to pass a 25 dB HL screening at 250, 500, 1000, 2000, and 4000 Hz. All participants provided informed consent in accordance with the IRB regulations for the New York University School of Medicine. Specific demographics for the CI users are presented in Table 1.

Modulation Detection (EasyMDT)

The EasyMDT test measures the minimum modulation depth required to detect amplitude modulations in a broadband white noise.

Stimuli • All stimuli consist of 1000 ms of broadband white noise with 10 ms onset and offset ramps sampled at 44,100 Hz and a 16-bit depth. Reference stimuli are unmodulated while target stimuli are sinusoidally amplitude modulated. Modulation rates available in the software (and used in the experiment) were 10, 50, 75, 100, 150, 200, and 300 Hz. The modulation depths varied across trials. Modulations began with a 90° phase. To maintain equal long-term intensity between the modulated and unmodulated stimuli, each sample in the target waveforms was divided by $1+(m^2/2)$, where m is the modulation depth (e.g., Viemeister 1979). Note that the stimuli were generated to replicate the stimuli used in Won et al. (2011), which in turn were based on Bacon and Viemeister (1985).

Procedure • The minimum detectable modulation depth is measured using a three-interval forced-choice (3IFC) adaptive procedure. In a given trial, two of the three stimuli consisted of unmodulated references and the third stimulus consisted of the modulated target. The interval containing the target stimulus was randomly selected for each trial. Each interval was separated by 300 ms of silence. The initial modulation depth of the target stimulus was 100%. The participant was asked to identify which of the three intervals corresponded to the sound that was different than the others. No feedback was given. The modulation depth of the target stimulus in subsequent trials was adjusted in 1-dB steps using a 1-up, 1-down procedure. The procedure was repeated until 12 reversals were completed. The last 6 reversals were averaged as the estimated modulation depth threshold. In a block of trials, modulation depth thresholds could be measured

TABLE 1. Subject demographics for CI users

Code	Gender	Age at testing	Etiology	Onset	Everyday devices	Ear tested	Age at implantation of test ear	Device	Electrode	Strategy
C101	M	72	Unknown	Progressive	Bimodal	RE	66	Advanced Bionics	HR90K HiFocus 1J	HiRes Optima-P
C105	F	55	Unknown	Progressive	Bilateral	LE	42; 48*	Advanced Bionics	HR90K HiFocus 1J	HiRes Optima-S
C106	M	40	Unknown	Congenital	Bimodal	RE	32	Advanced Bionics	HR90K HiFocus 1J	HiRes Optima-S
C107	F	45	Unknown	Progressive	Bimodal	RE	31	Advanced Bionics	CII HiFocus 1J	HiRes Optima-P
C110	M	33	Unknown	Congenital	Bilateral	LE	22	Advanced Bionics	HR90K HiFocus 1J	HiRes Optima-P
C111	M	73	Unknown	Progressive	Bilateral	RE	53; 60*; 62*	Advanced Bionics	HR90K HiFocus 1J	HiRes Optima-S
M102	F	69	Unknown	Unknown	Bilateral	LE	63	MED-EL	Sonata Standard	FSP
M104	F	55	Unknown	Congenital	Bilateral	LE	50	MED-EL	Concert Medium	FSP
M108	F	81	Unknown	Unknown	Bimodal	RE	72	MED-EL	Sonata Medium	FS4
N102	F	64	Unknown	Progressive	Bimodal	RE	61	Cochlear Ltd.	Freedom CI24RE	ACE
N105	M	70	Assumed Genetic	Progressive	Bimodal	RE	68	Cochlear Ltd.	Freedom CI24RE	ACE
N106	F	51	Viral	Sudden Bilateral HL	Unilateral CI	LE	41	Cochlear Ltd.	Freedom CI24RE	ACE

**Indicates surgical revision.*
LE, left ear; RE, right ear.

for any user-selectable subset of the seven modulation frequencies, including all or just one modulation frequency. If multiple modulation frequencies are selected, the modulation frequency tested is randomly selected for each trial until a modulation depth threshold (i.e., 12 reversals) is measured at each modulation frequency. Using the EasyMDT software, the participant responded by clicking on one of three boxes that corresponded to the three presented stimuli using a desktop PC with a choice of a touchscreen computer monitor or a computer mouse.

Interface • The interface for the EasyMDT software is designed to be as simple as possible. The software does not need to be installed and therefore can be used even on managed computers in a clinic for which the experimenter or clinician does not have administrative rights. Running the software consists of clicking on an executable file (EasyMDT.exe) on a windows computer and the EasyMDT main screen appears (top panel of Fig. 1). The options in this menu are intentionally limited. One button will play a calibration noise, and another will begin the experiment. The only other options are to provide a practice run or a button that allows changing of the selected frequencies. In addition, the tester can enter an optional participant code name and comment. Once the experiment started, the EasyMDT Response Screen (bottom panel of Fig. 1) is presented in which the three sounds are presented with three corresponding buttons. The buttons are large and can be pressed using either a mouse or a finger if a touchscreen is used. The buttons light up in red when the corresponding sound is played to simplify the association of the sound with the button.

Spectral-temporal Ripple Task

The Spectral-temporally Modulated Ripple Test (SMRT; Aronoff and Landsberger 2013) measures spectral-temporal ripple discrimination. The SMRT software, like EasyMDT, is a freely available tool that provides information about auditory processing in less than 5 minutes (Landsberger et al. 2019), making it easily implementable in research and clinical settings.

Procedure • SMRT stimuli consist of spectrally rippled broadband noise with phase drifts that change in time at 5 Hz. The peaks and valleys of the signal are modulated in time to avoid producing cues available from attending to a single electrode or a narrow frequency band. An adaptive 1-up/1-down 3-Interval Forced Choice (3IFC) task is implemented. For each trial, two of the intervals contain a reference stimulus and a third interval contains a target stimulus. Reference stimuli are at 20 ripples per octave (RPO). The target is 0.5 RPO initially and is varied adaptively in 0.2 RPO steps. Participants indicate the target stimulus. Thresholds are calculated based on the average of the last 6 of 10 reversals.

Like the EasyMDT software, the participant administered their responses by clicking on one of three boxes which corresponded to the three presented stimuli using a desktop PC with a choice of a touchscreen computer monitor or a computer mouse.

Speech Tests

Participants were evaluated with three lists of AzBio sentences in noise (Spahr et al. 2012), three lists of CNC words (Peterson & Lehiste 1962), as well as single runs of the vowel and consonant identification tasks. Due to time constraints, of the 12 CI users who completed EasyMDT, 8 completed AzBio sentences in noise, 10 completed CNC words presented in quiet, 9 completed Consonant Recognition and Vowel Identification.

AzBio • Three AzBio lists were randomly selected and presented using the standard 10-talker babble noise at +10 dB SNR for each participant. Each list consists of 20 sentences comprised of 5 sentences from 4 different talkers (2 male and 2 female). After each sentence, listeners were asked to verbally report what they had understood. Performance was scored by the number of words correctly reported, and a total percent correct was calculated for each list (Spahr et al. 2012). The test was administered using a custom Matlab script and scored manually.

CNC Words • Testing of the consonant-nucleus-consonant (Peterson & Lehiste 1962) word lists was conducted using the i-Star software (available free for download from istar.emilyfufoundation.org). Each list consisted of 50 monosyllabic words presented in quiet by a male speaker. Participants verbally repeated what was heard, and the responses were manually entered into the i-Star software by an audiologist. Scores were calculated as percent of correctly identified words. The CNC word list used was randomly selected for each participant.

Vowel and Consonant Identification • Vowel and consonant identification were measured separately using the i-Cast software (available for free download from icast.emilyfufoundation.org). The stimuli consisted of digitized natural productions from 5 men and 5 women. Consonant stimuli consisted of 20 consonant sounds (b, d, g, p, t, k, m, n, l, r, y, w, f, s, sh, th, ch, and j) in an /a/-consonant-/a/ format (e.g., “aba”) resulting in a total of 200 trials (20 consonants × 10 talkers). Vowel stimuli consisted of 12 vowel sounds in an /h/-vowel-/d/ format (heed, hid, head, had, who’d, hood, hod, hud, hawed, hoed, heard, and hayed) resulting in a total of 120 trials (12 vowels × 10 talkers). Listeners were required to indicate which vowel or consonant they heard by clicking on a corresponding button on the screen in a closed-set task. Scores were calculated as percent of correctly identified sounds.

Global Procedure

All stimuli for all tests were presented at 60 dB SPL in a sound field from a loudspeaker at 0º azimuth in a sound-treated, double-wall booth. The CI group participated in all testing conditions, while the NH only participated in the EasyMDT testing conditions. Modulation detection thresholds were collected first using the EasyMDT software. In a single run, modulation detection was measured for: only 100 Hz (the “100 Hz Alone” condition), only 150 Hz (the “150 Hz Alone” condition), or at seven different frequencies (10, 50, 75, 100, 150, 200, and 300 Hz; the “Full Curve” condition). In the Full Curve condition, the modulation frequency being evaluated randomly varied between the seven modulation frequencies from trial to trial until modulation detection thresholds were measured for all modulation frequencies. Each subject completed 6 runs of the 100 Hz Alone and 150 Hz Alone conditions and 3 runs in the Full Curve condition. Following MDT collection, each CI participant completed 3 runs of the SMRT. Speech tests were evaluated after the measurement of MDTs and SMRT. Three AzBio lists at +10 dB SNR and 3 CNC lists in quiet, one set of vowels (120 stimuli) and one set of consonants (200 stimuli) were evaluated. For most participants, testing was completed in a single session ranging for 3 to 4 hours. If testing occurred across multiple shorter sessions, it was ensured that no changes had occurred to the participants maps or settings between testing sessions.

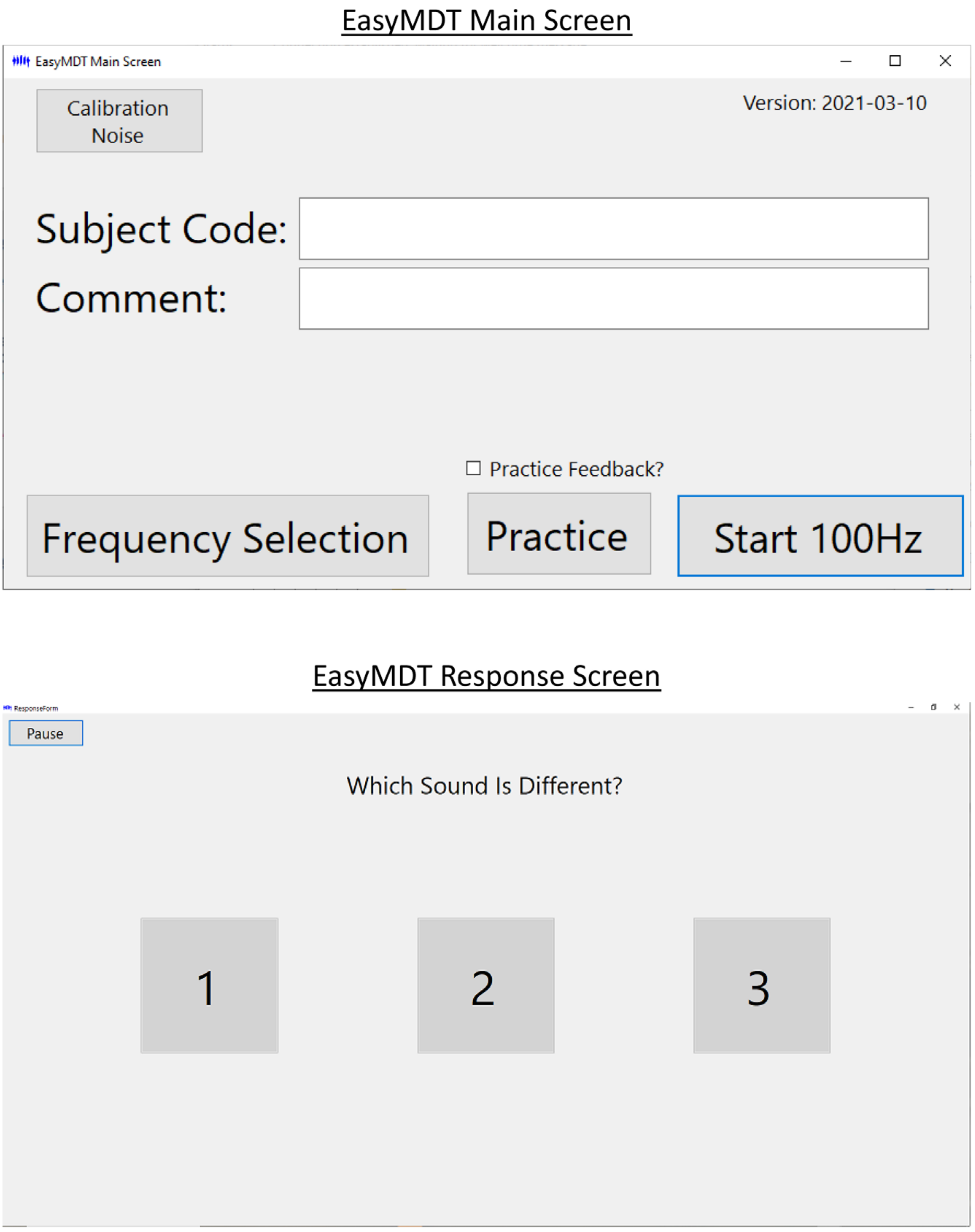

Fig. 1. Interface of the EasyMDT software. The top panel shows the EasyMDT main screen presented to the experimenter. The bottom panel shows the EasyMDT response screen presented to the participant.

RESULTS

Analysis of the Full Curve

The average modulation detection thresholds for CI and NH listeners for each subject were calculated for each modulation frequency evaluated and presented in the left panel of Figure 2. In addition, NH data extracted from Viemeister (1979) and the CI data extracted from Won et al. (2011) are presented. The present NH data are represented by the circles and CI data by triangles. Performance was best at 10 Hz for both CI and NH listeners and decreased as a function of modulation frequency for both groups. Modulation detection thresholds were consistently better for NH than CI users. The present CI data show similar results to that of Won et al. (2011); however, the present NH participants performed slightly worse than those of Viemeister (1979), although maintaining a similar slope across modulation frequencies. A mixed-effects analysis of variance (ANOVA) with modulation frequency as the within-subject factor and listening group as the between-subject factor was used to analyze the data collected in the present experiment. Main effects of modulation frequency ($F[6,120]=201.96$, $p<0.001$) and listening group (NH or CI; $F[1,20]=43.79$, $p<0.001$) were found as was the interaction ($F[6,120]=9.21$, $p<0.001$). The average modulation detection thresholds across modulation frequencies are plotted in the right panel. For both NH and CI curves, the average NH and CI MDTs are similar to the MDTs at 100 Hz.

Comparing 100 Hz Alone and 150 Hz Alone MDTs to Single-frequency MDTs Extracted From the Full Curve

It is important to determine the relationship between MDTs measured at a single modulation frequency (e.g., 100 Hz Alone) and the MDTs at the same modulation frequency measured as part of the Full Curve (e.g., only the 100 Hz data extracted from the Full Curve). Figure 3 represents scatter plots of 100 Hz Alone (left) and 150 Hz Alone (right) plotted against MDTs for the same frequencies collected in the Full Curve for the CI and NH participants. Strong correlations were detected for both modulation frequencies (100 Hz: $r=0.914$, $n=22$, $p<0.001$; 150 Hz: $r=0.932$, $n=22$, $p<0.001$). A linear regression analysis determines that the best fitting line for 100 Hz data is Full Curve (100Hz)=(0.748 * Alone) – 3.989 and for 150 Hz data is Full Curve (150Hz)=(0.916 * Alone) – 0.272. No significant differences between the Full Curve and either the 100 Hz Alone or 150 Hz Alone data were detected. 95% confidence intervals suggest that the true difference between the Full Curve and Alone measures was less than 1.8 dB (100 Hz: –0.706 to 1.415 dB; 150 Hz: –1.778 to 0.379). As expected, the NH listeners performed better than the CI listeners for both modulation frequencies tested with little overlap between the two distributions. A two-way mixed-effect ANOVA with modulation frequency and testing condition (Alone or extracted from the Full Curve) as within-subject factors and listening group (CI or NH) as a between-subject factor detected a main effect of listening group ($F [1,20]=40.050$, $p<0.001$) as MDTs were consistently lower (better) for NH listeners. A main effect of modulation frequency ($F [1,20]=39.718$, $p<0.001$) was detected as MDTs for 100 Hz were consistently lower (better) than MDTs for 150 Hz. The interaction ($F [1,20]=7.230$, $p=0.014$) was also detected, presumably driven by an increased difference between MDTs for NH and CI users at the 150 Hz modulation frequency for stimuli extracted from the Full Curve.

Comparing 100 Hz Alone and 150 Hz Alone MDTs to the Average MDTs Across the Full Curve

Modulation Detection Thresholds for the 100 Hz Alone and 150 Hz Alone conditions were compared with the average scores for the Full Curve to provide an estimate of how

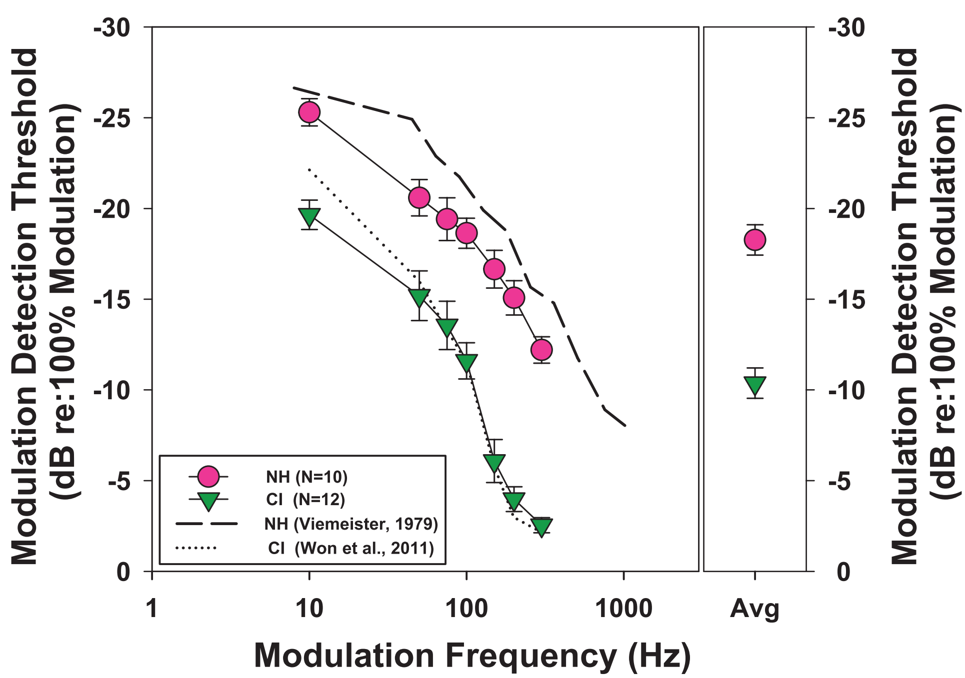

Fig. 2. Average MDT scores across modulation frequencies tested. Circles represent average NH scores, and triangles represent average CI scores. Dashed and dotted lines represent the NH and CI data from Viemeister (1979) and Won et al. (2011), respectively. Error bars indicate ±1 standard error of the mean.

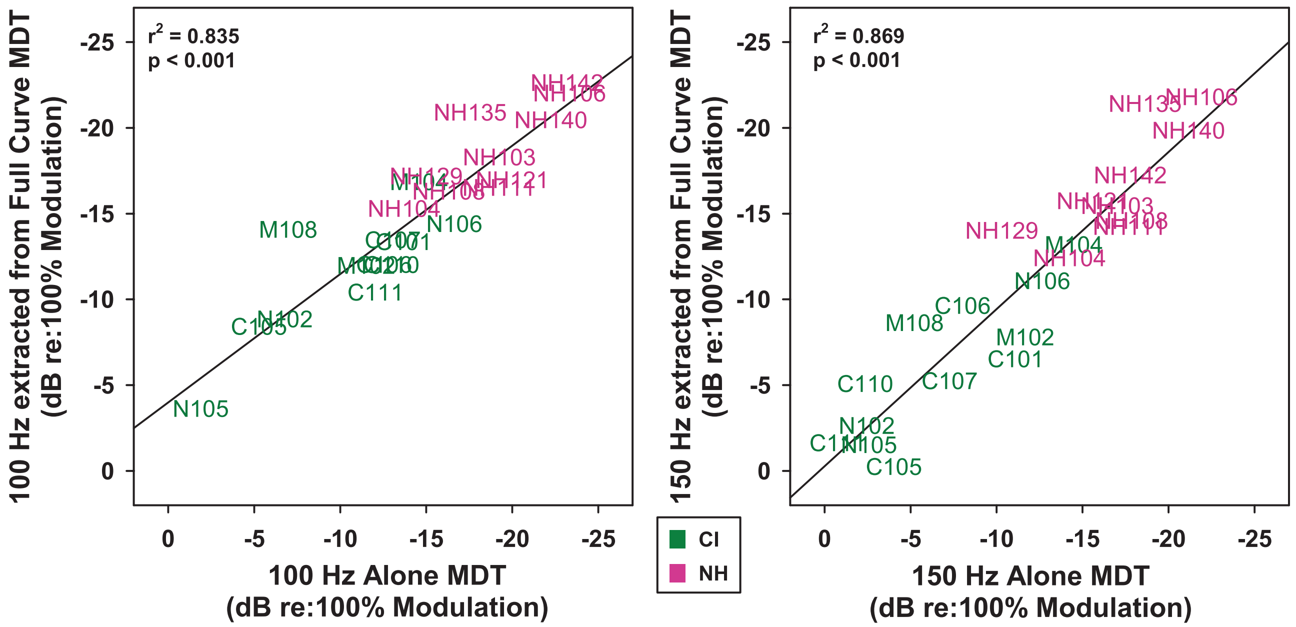

Fig. 3. Scatterplots of 100 Hz (left) and 150 Hz (right) Alone MDTs with 100 Hz and 150 Hz MDTs extracted from the Full Curve for each participant. Green labels indicate participants with CIs. Red labels indicate participants with NH.

well the individual MDTs represented a summary of the full data set. Figure 4 represents scatter plots of 100 Hz Alone (left) and 150 Hz Alone (right) plotted against the MDTs averaged across all modulation frequencies of the Full Curve for the CI and NH groups. The NH listeners performed better than the CI listeners for both modulation frequencies tested with little overlap between the two distributions. Strong correlations were detected for both modulation frequencies (100 Hz: r=0.924, n=22, $p<0.001$; 150 Hz: r=0.894, n=22, $p<0.001$). A linear regression analysis determines that the best fitting line for 100 Hz Alone data is Full Curve average=(0.783 * 100 Hz Alone) – 2.652 and for 150 Hz Alone data is Full Curve average=(0.655 * 150 Hz Alone) – 6.372. A significant difference between the 150 Hz Alone and Full Curve average was detected using a 95% confidence interval (0.975 to 3.773 dB). No significant difference between the 100 Hz Alone and Full Curve average was detected using a 95% confidence interval (–1.474 to 0.505 dB), suggesting that the true difference between these measures was less than 1.5 dB.

Test-retest Comparisons

To evaluate test-retest reliability, correlations were calculated between MDTs measured on the first run with MDTs measured on subsequent runs. The correlations between the first run

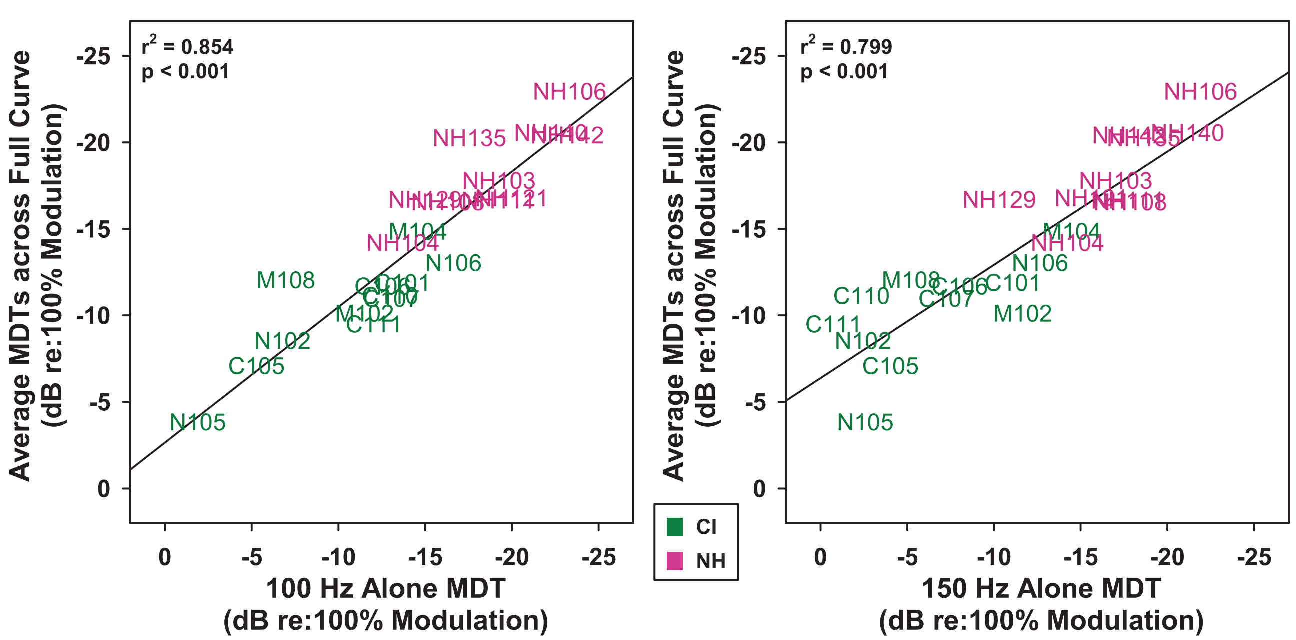

Fig. 4. Scatterplots of 100 Hz (left) and 150 Hz (right) Alone MDTs with average of the Full Curve MDTs for each participant. Green labels indicate participants with CIs. Red labels indicate participants with NH.

of the 100 Hz Alone and 150 Hz Alone and subsequent runs are presented in the right panel of Figure 5. The correlations were strong. On average, the correlation coefficients are 0.85±0.05 and 0.84±0.05 for the 100 Hz Alone and 150 Hz Alone conditions. The correlation between the first run and subsequent runs for individual modulation frequencies extracted from the Full Curve (left panel of Fig. 5) was lower. This suggests that collecting data from a single modulation frequency will provide more reliable data than extracting an MDT from the same modulation frequency from the Full Curve. However, the correlations between the first and subsequent runs of the average of the Full Curve (r = 0.86 and r = 0.88 for runs 2 and 3, respectively) were strong.

In addition, the 95% confidence interval of the differences between the first run and subsequent runs was calculated to determine whether performance on the test changed with run. The differences between the first run and subsequent runs for each individual modulation frequency extracted from the Full Curve as well as the average MDTs from the Full Curve are presented in the left panel of Figure 6. The mean differences between the first and subsequent runs are between +1 and –1 dB for all MDTs and the 95% confidence intervals all include 0. This suggests that the range of expected differences between trials is small and no significant differences were detected. The 95% confidence intervals of the differences between the first and subsequent runs for the 100 Hz Alone and 150 Hz Alone conditions are presented in the right panel of Figure 6. For the 100 Hz Alone condition, the mean differences are small, and the 95% confidence intervals all contain 0. However, the results are somewhat different for the 150 Hz Alone condition in that the performance in runs 2 through 6 was between 1.4 and 2.3 dB better than on the first run. The 95% confidence intervals for differences between the first run and the second, fourth, and fifth runs do not include 0, suggesting that MDTs improved for these runs relative to the first run.

Testing Duration

The duration of testing of all MDT conditions was measured and plotted in Figure 7 for the 100 Hz Alone, 150 Hz Alone, and Full Curve tests for the CI (green) and NH (pink) groups. Average test duration for the Full Curve (CI: 18.1 minutes, NH: 21.1 minutes) was much longer than for the 100 Hz Alone (CI: 2.7 minutes, NH: 3.1 minutes) and 150 Hz Alone (CI: 2.3 minutes, NH: 2.9 minutes). The differences in duration can be explained primarily by the fact that 7 modulation frequencies are measured as part of the Full Curve. As the test is adaptive, the difficulty of the test also affects the testing duration as it affects the number of trials required to reach threshold. As such, less time is required to measure MDTs for poorer performers (such as CI listeners). Similarly, higher modulation frequencies (which are more difficult to detect) provide shorter testing durations than lower modulation frequencies. The duration

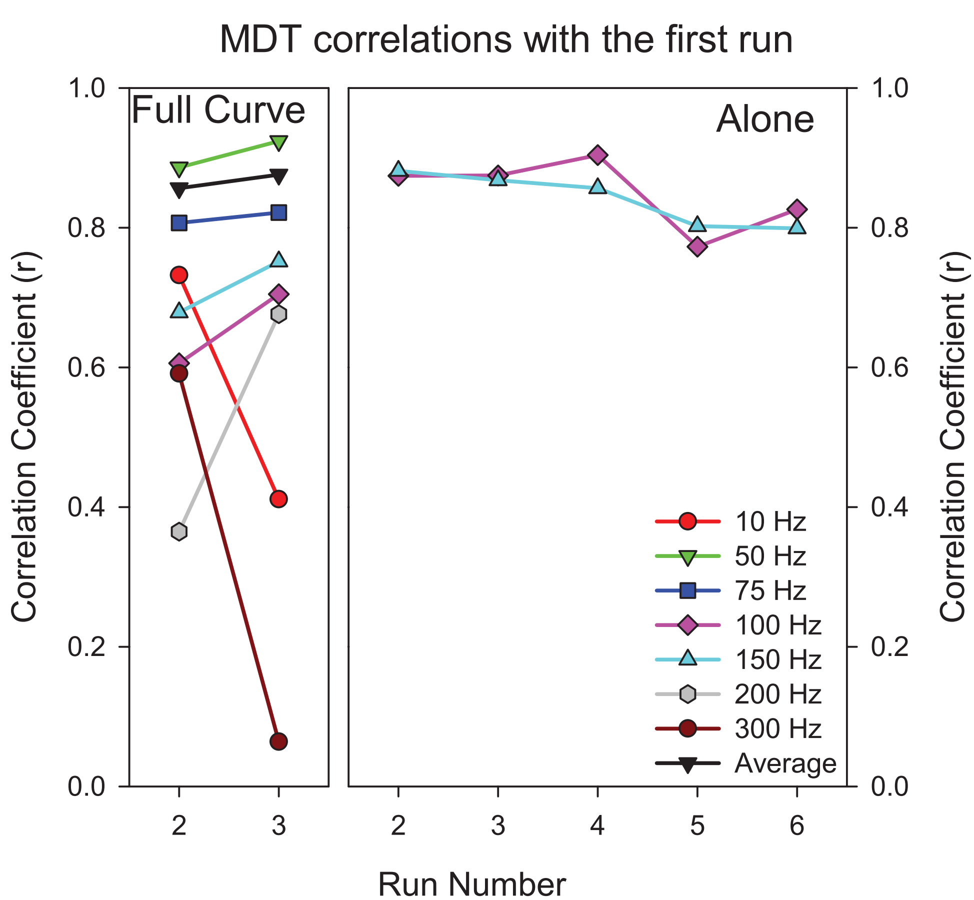

Fig. 5. The correlation coefficients (r) between MDTs measured on the first run with MDTs measured on subsequent runs. The left panel illustrates the correlation coefficients for the average threshold across modulation frequencies (black triangles) as well as each individual modulation frequency extracted from the Full Curve (other symbols). The right panel illustrates the correlation coefficients for the 100 Hz Alone (magenta diamonds) and 150 Hz Alone (cyan triangles) conditions.

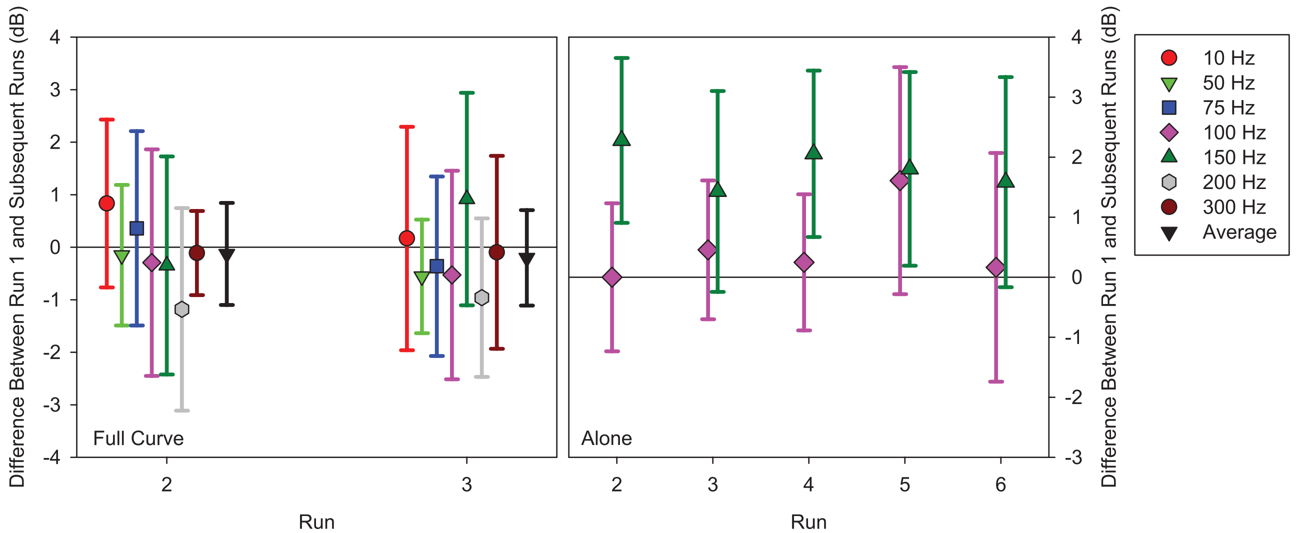

Fig. 6. The mean difference (in dB) between the MDTs measured on the first run with MDTs measured on subsequent runs. Error bars indicate the 95% confidence interval. The left panel illustrates the differences between MDTs for the average threshold across modulation frequencies (black triangles) as well as each individual modulation frequency extracted from the Full Curve (other symbols). The right panel illustrates the differences for the 100 Hz Alone and 150 Hz Alone conditions.

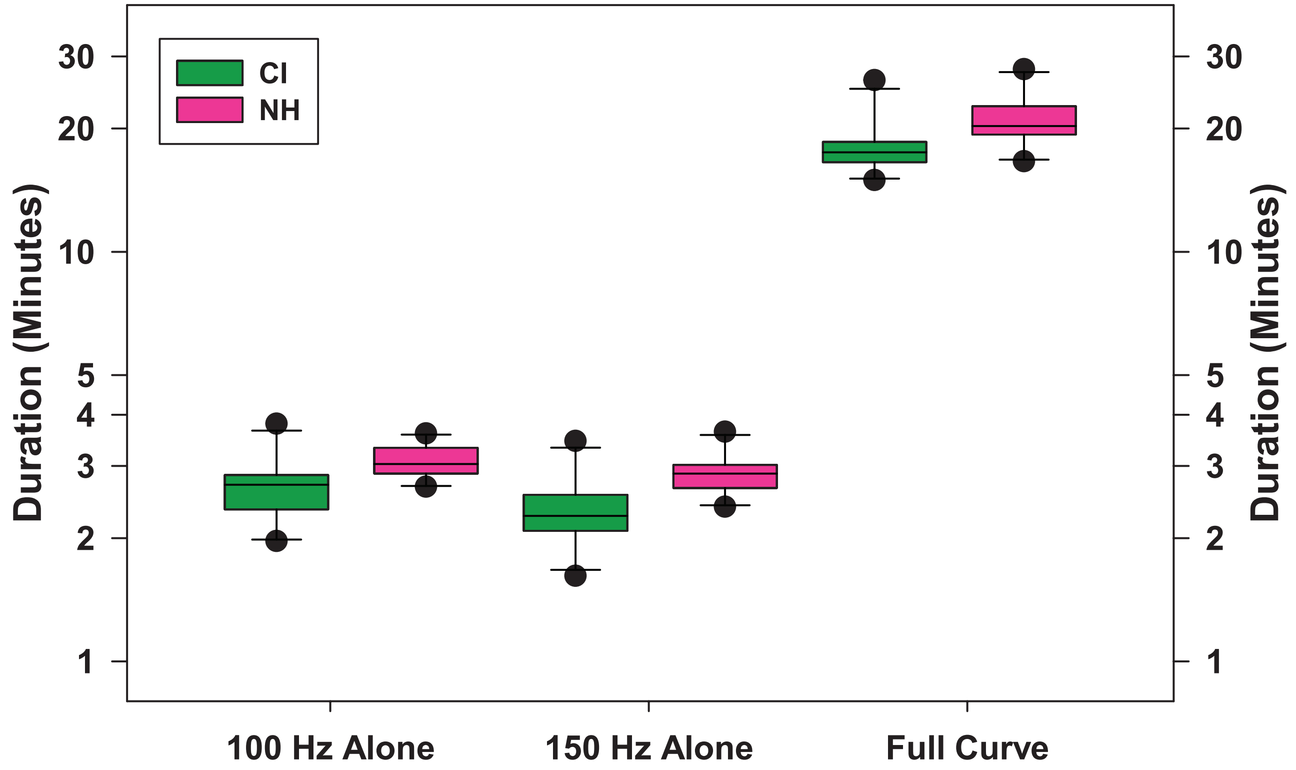

Fig. 7. Average testing durations for 100 Hz Alone, 150 Hz Alone, and the Full Curve. Green boxplots correspond to CI users, whereas pink boxplots correspond to NH users. The top and bottom of each box represents the 25th and 75th percentiles of the corresponding distribution, whereas the horizontal lines through the boxes represent the median. Whiskers represent the 10th and 90th percentiles. Individual outliers (defined as being below the 10th or above the 90th percentile) are plotted as individual points (filled circles).

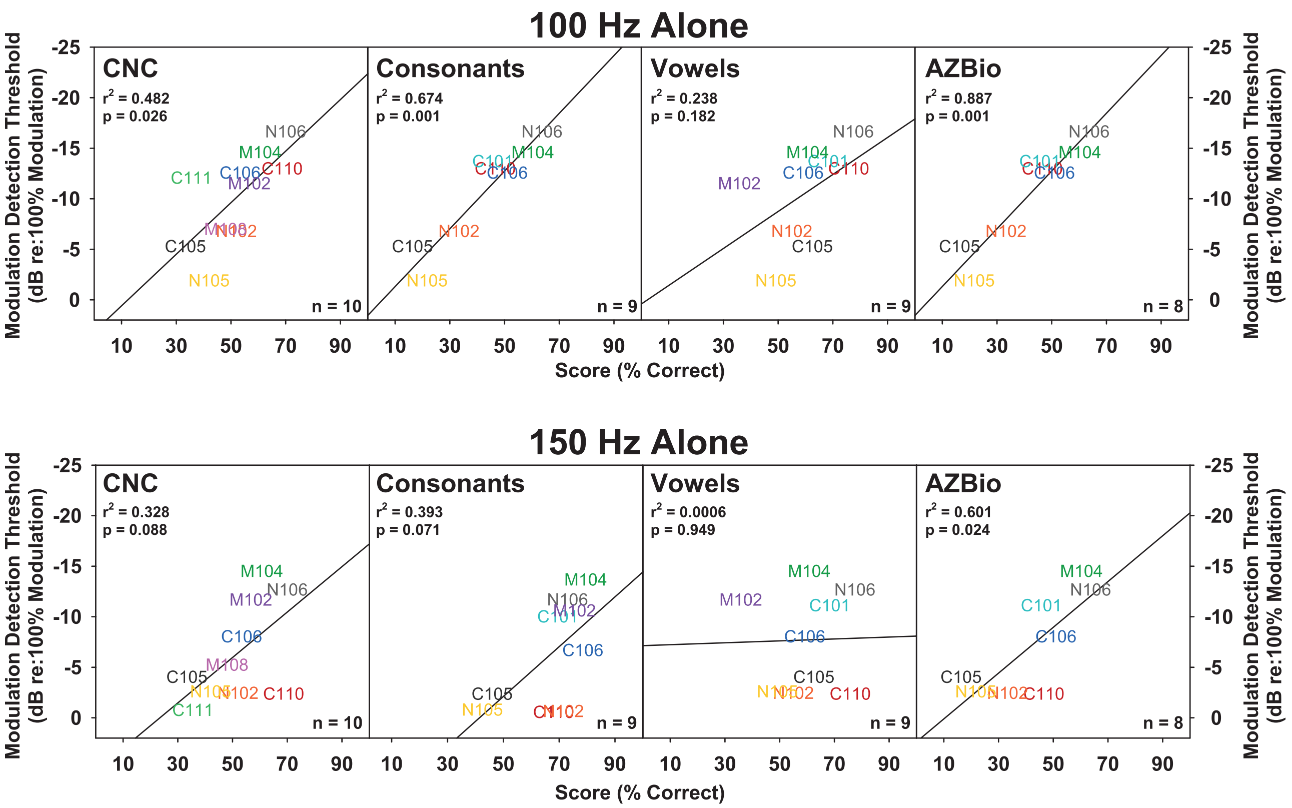

Fig. 8. Scatterplots of speech perception scores and 100 Hz Alone (top row) or 150 Hz Alone (bottom row) MDTs for the CI participants. Columns represent performance on the CNC words, consonant identification, vowel identification, and AzBio sentences for 100 Hz Alone and 150 Hz Alone, respectively.

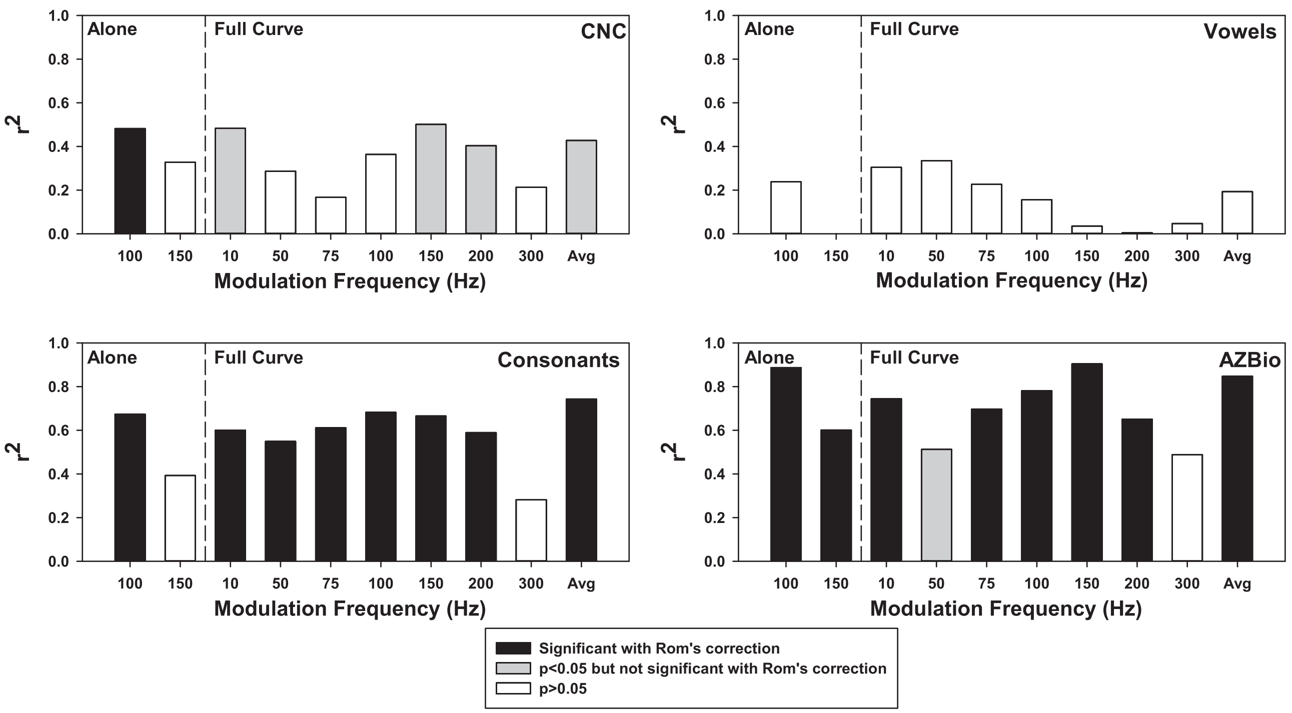

Fig. 9. Strength of correlations (R-square) between MDTs and speech tests. The four panels describe the correlations between MDTs and CNC words (top left), vowel identification (top right), consonant identification (bottom left), and AzBio sentences (bottom right). Within each panel, the strength of the correlation is provided for the 100 Hz Alone and 150 Hz Alone conditions, each individual modulation frequency extracted from the Full Curve, as well as the average MDT across the Full Curve. Correlations that are significant with Rom's correction are represented in black. Correlations where $p<0.05$ but fail to maintain significance after Type I error correction are represented in gray. Correlations with $p>0.05$ are represented in white.

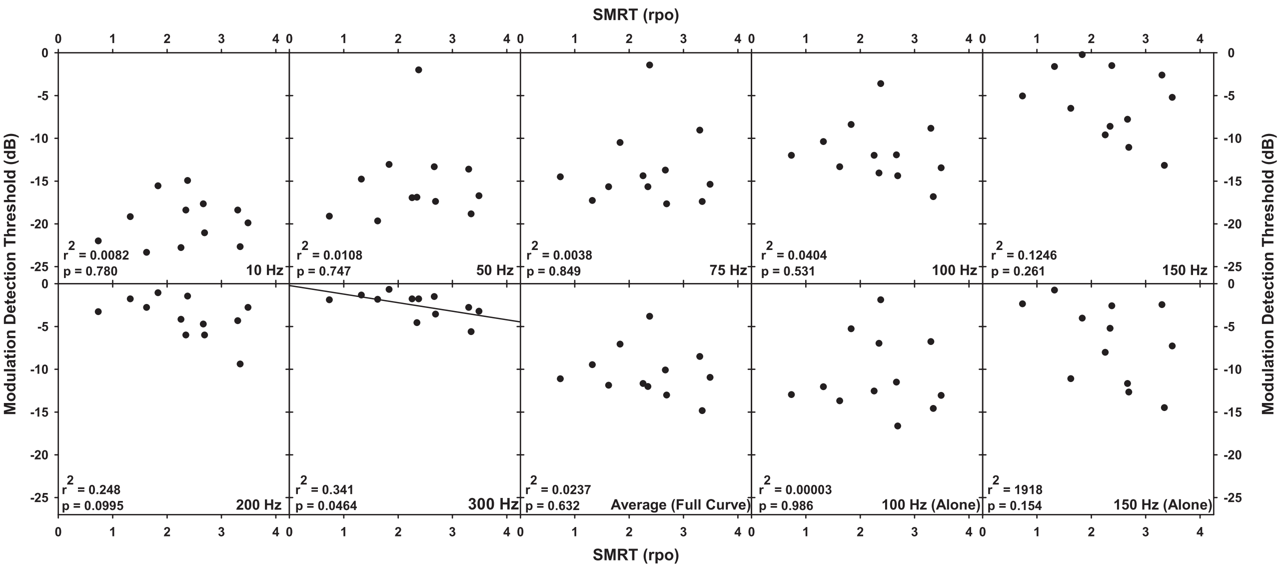

Fig. 10. Scatterplots of SMRT scores with individual MDT modulation frequencies extracted from the Full Curve, the Full Curve Average, and the 100 Hz Alone and 150 Hz Alone measurements.

of the tests were significantly correlated with performance for the 100 Hz Alone (r=–0.724, n=20, p=0.0002), 150 Hz Alone (r=-0.673, n=20, p=0.0008) and the average of the Full Curve (r=–0.538, n=20, p=0.0118) tests.

Correlations Between Modulation Detection Thresholds and Speech understanding

Pearson correlations were calculated between MDTs and the various speech understanding tests (CNC, Vowels, Consonants, and AzBio +10dB Noise). Scatter plots representing these relationships for the 100 Hz Alone (top) and 150 Hz Alone (bottom) MDTs are presented in Figure 8. Significant correlations were found between the 100 Hz Alone MDTs and CNC words (r=-0.694; n=10, p=0.026), consonants (r=-0.821; n=9, p=0.00669), and AzBio sentences (r=-0.942; n=8, p=0.000467). A correlation for the 150 Hz Alone MDTs was only found for the AzBio sentences (r=0.775; n=8, p=0.0238). Additional correlations were calculated between each speech measure and the MDT for each frequency (and average MDT) extracted from the full curve. The strength (r^2) of the correlation and its corresponding significance (or lack thereof) is presented in Figure 9. Black bars indicate significant correlations even after Type I error correction with Rom's method (Rom 1990). Gray bars indicate correlations with p<0.05 but fail to maintain significance after Rom's correction. White bars indicate correlations with p values>0.05.

Modulation Detection Thresholds and SMRT

SMRT scores ranged from 0.7 to 3.5 RPO with an average score of 2.3 RPO. Figure 10 shows scatter plots of the relationship between SMRT and each of the MDT measures (each of the 7 modulation frequencies extracted from the Full Curve, the average MDT across all modulation frequencies in the Full Curve, and the MDTs for 100 Hz Alone and 150 Hz Alone conditions). Only the correlation between SMRT and the 300 Hz MDT provided a p < 0.05 (r=–0.584, n=12, p=0.046). However, after Type I error correction, this correlation did not remain significant. Unlike previous studies (e.g., Holden et al. 2016; Lawler et al. 2017; Zhou 2017; Jeddi et al. 2019; Spitzer et al. 2020), no correlations were found between SMRT and speech metrics (AzBio: r=0.18, n=8, p=0.66; Vowels: r=–0.447, n=9, p=0.29; Consonants: r=0.30, n=9, p=0.43).

DISCUSSION

As expected, the data collected with the EasyMDT software resulted in similar outcomes to previous studies (e.g., Bacon and Viemeister 1985; Won et al. 2011). The stimuli for the EasyMDT software were nearly identical to those used previously as they were generated using the equations provided in Won et al. (2011). However, the protocol for EasyMDT was not consistent with previous work. The most noteworthy difference was that the EasyMDT protocol consists of a 3IFC task without feedback in which the participant is asked to identify the interval with temporal modulations. The Bacon and Viemeister (1985) method uses a 2-interval forced-choice (2IFC) task with feedback, requiring the listener to learn to identify which stimulus has modulations and not just which stimulus is different. The 3IFC protocol was chosen for the EasyMDT software to simplify instructions to the participant. In addition, the two methodologies differed in the adaptive rules and number of reversals making direct comparisons of the data sets difficult. Nevertheless, as shown in Figure 2, the patterns of results are similar for the two protocols, although the thresholds for EasyMDT tend to be higher.

In addition to measuring the Full Curve, MDTs for 100 Hz and 150 Hz were measured in isolation (i.e., 100 Hz Alone and 150 Hz Alone conditions). The purpose of measuring single modulation frequencies was to determine if they could be used in many situations as a substitute for the more time-demanding Full Curve measurement. Indeed, MDTs measured with the 100 Hz Alone and 150 Hz Alone conditions were highly correlated with MDTs for the corresponding frequencies extracted from the Full Curve as well as the average MDT across frequencies for the Full Curve, which was used as a single value representation of the Full Curve by Won et al. (2011). As such, an MDT measure at a single modulation frequency (e.g., 100 Hz Alone) may be a time-efficient substitute for measuring the Full

Curve in many situations. Indeed, for CI users, the average time for MDT measurements drops from 18.1 minutes for the Full Curve to 2.7 minutes for 100 Hz Alone. Therefore, administration of the 100 Hz Alone MDT may be possible in situations where the testing duration of the Full Curve is prohibitively long, such as in clinical environments or as a part of a more extensive research test battery. Although multiple repetitions (two to three) are encouraged to provide a more accurate MDT estimate, the test-retest reliability data presented in Figures 4 and 5 suggest that one trial may be sufficient. It is worth noting that no training or practice was provided to the participants in this study. However, a simple practice mode consisting of the first 4 trials of a 100 Hz MDT with optional feedback has been implemented in the EasyMDT software. Perhaps providing the practice until the listener is comfortable with the task before conducting a complete run would be sufficient.

Performance on the CNC, AzBio, and Consonant Identification speech tests was frequently correlated with MDT measurements. The 100 Hz Alone MDTs were correlated with these three speech tests (Fig. 9), although the correlation was only significant between 100 Hz MDTs extracted from the Full Curve for AzBio sentences and Consonant Identification. This suggests that measuring 100 Hz Alone MDT may be useful for experiments or clinical evaluations in which speech understanding ability needs to be predicted. The 150 Hz Alone MDTs only correlated with AzBio sentence understanding, although the 150 Hz MDTs extracted from the Full Curve correlated with AzBio sentence understanding and consonant identification. Even without Type I error correction, no correlations were found between vowel identification and any of the MDT threshold measurements. Presumably, this is because MDT thresholds are metrics of temporal processing, while vowel identification is dependent on spectral processing (e.g., Kirk et al. 1992; Donaldson et al. 2015; Arenberg et al. 2018). Won et al. (2011) found significant correlations ($p < 0.05$) between MDTs extracted from the Full Curve and CNC word scores for 75 Hz and above. Significant correlations were also found for the average MDT across all frequencies, although it is unclear how many of those correlations would remain significant after Type I error correction. In the present article, correlations were detected using the same criterion as Won et al. ($p < 0.05$) for 10 Hz, 150 Hz, and 200 Hz extracted from the Full Curve as well as the average MDT for the full curve. However, none of these correlations in the present article remained significant after Type I error correction using Rom's correction.

No correlations were detected between MDTs as measured with the EasyMDT software and the SMRT after Type I error correction. This suggests that the SMRT and MDT are measuring different attributes of auditory processing. However, as the SMRT has both temporal and spectral components, it was unknown whether the two tests would provide redundant information. This is important as both the SMRT and EasyMDT are proposed as time-efficient nonlinguistic tests that can be used to predict speech understanding. The only MDT which was significantly correlated with SMRT before Type I error correction was 300 Hz MDT. A visual inspection of the scatterplot representing the 300 Hz MDT reveals a small range of MDTs with all participants performing at or near floor. The correlation appears to be driven by the few participants who can detect any 300 Hz modulations also having relatively high SMRT scores. If the relationship between SMRT and MDT is driven by the 5 Hz temporal component of the SMRT, it would be expected that the correlation would be greatest for modulation frequencies closest to 5 Hz and would be smaller as the modulation frequency increased. However, the data presented provide the opposite results as the correlation increases with increased modulation rate. It therefore may be reasonable in some situations to administer both EasyMDT and SMRT as they seem to provide different information. Given the lack of correlation detected between SMRT and speech scores in this study, the data would suggest only using EasyMDT if the goal is to predict speech outcomes. However, many other studies have demonstrated significant correlations between SMRT and speech in noise recognition with relatively high r^2 values (between 0.60 and 0.84; Holden et al. 2016; Lawler et al. 2017; Zhou 2017). However, an additional study with a greater number of participants may be required to determine whether it is preferable to measure MDTs, SMRT, or both if the purpose is to predict speech understanding.

It is worth noting that there are multiple limitations with this study. Most importantly, the study consists of a relatively small number of participants (12 CI users and 10 NH listeners), with only a subset of the CI users evaluated on the speech tests. The data should be sufficient to provide representative sample data. However, it may be desirable to conduct a larger study to provide more precise normative data. Fortunately, as the software tool is publicly available providing a standardized protocol and stimulus set, this can be done at any lab or clinic around the world or even by collapsing data across clinics and studies. Furthermore, the NH listeners were all relatively young adults (between 23 and 37 years), and it is unclear how the NH data generalize to NH populations such as children or older adults. Another limitation is that, while the software was designed to be easy to use, no formal attempt to evaluate ease of use or otherwise study human factors related to testing was conducted. However, as EasyMDT shares a nearly identical interface and distribution method as the SMRT (Aronoff & Landsberger 2013), we expect it to be similarly easy to use. The SMRT has been used successfully in dozens of published manuscripts and conference presentations (see https://www.ear-lab.org/smrt.html for an incomplete list) without consultation, instruction, or request for technical support from the SMRT developers, suggesting that the software is indeed easy to use.

While the EasyMDT may be a quick and efficient measure of temporal processing abilities, there are important limitations with the EasyMDT (and any MDT test) to consider. As modulation detection depends on detected differences in signal amplitude or intensity, results may be affected by intensity discrimination ability. Additionally, if there is compression in the auditory system such as imposed by a damaged auditory system or the signal processing associated with hearing aids or CIs, the representation of small amplitude changes may be modified. This would result in differing MDTs without necessarily reflecting temporal processing. Similarly, the settings of a CI such as the electric dynamic range will affect modulation detection. The limitations of MDT measurement for CI measurements are described in greater detail in Galvin and Fu (2009) and Fraser and McKay (2012). An alternative test, such as one that involves modulation frequency discrimination instead of modulation detection may avoid these issues. Nevertheless, this task may potentially provide useful information in clinical situations such as testing nonnative speakers (as the test is language independent) or difficult to test populations such as nonverbal patients and those with developmental disabilities. It may also be helpful when testing children, allowing for quick results without testing fatigue.

CONCLUSION

Data are presented from a tool that was designed to allow researchers and clinicians to measure MDTs simply. The tool allows for time-efficient measurements using a standardized protocol across multiple clinical and laboratory environments. The measurements may be useful in that they not only provide information about temporal processing, but they also can be used to predict many speech outcomes. It may therefore be a tool appropriate for evaluating patients for whom speech testing is difficult or inappropriate for many reasons, including developmental issues or nonfluency in the testing language. The software is distributed for free at https://www.ear-lab.org/software-downloads.html.

ACKNOWLEDGMENTS

The authors would like to thank the participants for their time and dedication. Further thanks are given to Qian-Jie Fu for providing server space for the software and demonstrating to our group the importance of releasing useful audiological tools to the public. Support for this research was provided by the National Institute on Deafness and Other Communication Disorders (NIDCD) R01 DC012152 (Landsberger).

Address for correspondence: David M. Landsberger, Department of Otolaryngology, New York University School of Medicine, 550 1st Avenue, STE NBV 5E5, New York, NY 10016, USA. E-mail: David.Landsberger@nyumc.org.

Received December 13, 2020; accepted June 11, 2021

REFERENCES

Arenberg, J. G., Parkinson, W. S., Litvak, L., Chen, C., Kreft, H. A., Oxenham, A. J. (2018). A dynamically focusing cochlear implant strategy can improve vowel identification in noise. *Ear Hear, 39*, 1136–1145.

Aronoff, J. M., & Landsberger, D. M. (2013). The development of a modified spectral ripple test. *J Acoust Soc Am, 134*, EL217–EL222.

Bacon, S. P., & Viemeister, N. F. (1985). Temporal modulation transfer functions in normal-hearing and hearing-impaired listeners. *Audiology, 24*, 117–134.

Cazals, Y., Pelizzone, M., Saudan, O., Boex, C. (1994). Low-pass filtering in amplitude modulation detection associated with vowel and consonant identification in subjects with cochlear implants. *J Acoust Soc Am, 96*, 2048–2054.

Donaldson, G. S., Rogers, C. L., Johnson, L. B., Oh, S. H. (2015). Vowel identification by cochlear implant users: Contributions of duration cues and dynamic spectral cues. *J Acoust Soc Am, 138*, 65–73.

Freyman, R. L., Nerbonne, G. P., Cote, H. A. (1991). Effect of consonant-vowel ratio modification on amplitude envelope cues for consonant recognition. *J Speech Hear Res, 34*, 415–426.

Fraser, M., & McKay, C. M. (2012). Temporal modulation transfer functions in cochlear implantees using a method that limits overall loudness cues. *Hear Res, 283*, 59–69.

Fu, Q. J. (2002). Temporal processing and speech recognition in cochlear implant users. *Neuroreport, 13*, 1635–1639.

Galvin, J. J. III, & Fu, Q. J. (2009). Influence of stimulation rate and loudness growth on modulation detection and intensity discrimination in cochlear implant users. *Hear Res, 250*, 46–54.

Hall, J. W. III, & Grose, J. H. (1994). Development of temporal resolution in children as measured by the temporal modulation transfer function. *J Acoust Soc Am, 96*, 150–154.

Holden, L. K., Firszt, J. B., Reeder, R. M., Uchanski, R. M., Dwyer, N. Y., Holden, T. A. (2016). Factors affecting outcomes in cochlear implant recipients implanted with a perimodiolar electrode array located in Scala Tympani. *Otol Neurotol, 37*, 1662–1668.

Jeddi, Z., Lotfi, Y., Moossavi, A., Bakhshi, E., Hashemi, S. B. (2019). Correlation between auditory spectral resolution and speech perception in children with cochlear implants. *Iran J Med Sci, 44*, 382–389.

Kirk, K. I., Tye-Murray, N., Hurtig, R. R. (1992). The use of static and dynamic vowel cues by multichannel cochlear implant users. *J Acoust Soc Am, 91*, 3487–3498.

Landsberger, D. M., Stupak, N., Green, J., Tona, K., Padilla, M., Martinez, A. S., Eisenberg, L. S., Waltzman, S. (2019). Temporal modulation detection in children and adults with cochlear implants: Initial results. *Otol Neurotol, 40*, e311–e315.

Lawler, M., Yu, J., Aronoff, J. M. (2017). Comparison of the spectral-temporally modulated ripple test with the Arizona biomedical institute sentence test in cochlear implant users. *Ear Hear, 38*, 760–766.

Park, M. H., Won, J. H., Horn, D. L., Rubinstein, J. T. (2015). Acoustic temporal modulation detection in normal-hearing and cochlear implanted listeners: Effects of hearing mechanism and development. *J Assoc Res Otolaryngol, 16*, 389–399.

Peterson, G. E., & Lehiste, I. (1962). Revised CNC lists for auditory tests. *J Speech Hear Disord, 27*, 62–70.

Rom, D. M. (1990). A sequentially rejective test procedure based on a modified Bonferroni inequality. *Biometrika*, *77*, 663–665.

Spahr, A. J., Dorman, M. F., Litvak, L. M., Van Wie, S., Gifford, R. H., Loizou, P. C., Loiselle, L. M., Oakes, T., Cook, S. (2012). Development and validation of the AzBio sentence lists. *Ear Hear, 33*, 112–117.

Spitzer, E. R., Landsberger, D. M., Friedmann, D. R. (2020). Assessing the quality of low-frequency acoustic hearing: Implications for combined electroacoustic stimulation with cochlear implants. *Ear Hear, 42*, 475–486.

Steeneken, H. J., & Houtgast, T. (1980). A physical method for measuring speech-transmission quality. *J Acoust Soc Am, 67*, 318–326.

Van Tasell, D. J., Soli, S. D., Kirby, V. M., Widin, G. P. (1987). Speech waveform envelope cues for consonant recognition. *J Acoust Soc Am, 82*, 1152–1161.

Viemeister, N. F. (1979). Temporal modulation transfer functions based upon modulation thresholds. *J Acoust Soc Am, 66*, 1364–1380.

Won, J. H., Drennan, W. R., Nie, K., Jameyson, E. M., Rubinstein, J. T. (2011). Acoustic temporal modulation detection and speech perception in cochlear implant listeners. *J Acoust Soc Am, 130*, 376–388.

Zhou, N. (2017). Deactivating stimulation sites based on low-rate thresholds improves spectral ripple and speech reception thresholds in cochlear implant users. *J Acoust Soc Am, 141*, EL243.

Zhou, N., Dixon, S., Zhu, Z., Dong, L., Weiner, M. (2020). Spectrotemporal modulation sensitivity in cochlear-implant and Normal-Hearing listeners: Is the performance driven by temporal or spectral modulation sensitivity? *Trends Hear, 24*, 2331216520948385.