Temporal Sensitivity Measured Shortly After Cochlear Implantation Predicts 6-Month Speech Recognition Outcome

Article *in* Ear and Hearing · April 2018

DOI: 10.1097/AUD.0000000000000588

Alexandra A Ludwig
University of Leipzig

Temporal Sensitivity Measured Shortly After Cochlear Implantation Predicts 6-Month Speech Recognition Outcome

Julia Erb,[1,2,6] Alexandra Annemarie Ludwig,[3,4] Dunja Kunke,[1] Michael Fuchs,[3,5] and Jonas Obleser[1,6]

Objectives: Psychoacoustic tests assessed shortly after cochlear implantation are useful predictors of the rehabilitative speech outcome. While largely independent, both spectral and temporal resolution tests are important to provide an accurate prediction of speech recognition. However, rapid tests of temporal sensitivity are currently lacking. Here, we propose a simple amplitude modulation rate discrimination (AMRD) paradigm that is validated by predicting future speech recognition in adult cochlear implant (CI) patients.

Design: In 34 newly implanted patients, we used an adaptive AMRD paradigm, where broadband noise was modulated at the speech-relevant rate of ~4 Hz. In a longitudinal study, speech recognition in quiet was assessed using the closed-set Freiburger number test shortly after cochlear implantation (t_0) as well as the open-set Freiburger monosyllabic word test 6 months later (t_6).

Results: Both AMRD thresholds at t_0 ($r = -0.51$) and speech recognition scores at t_0 ($r = 0.56$) predicted speech recognition scores at t_6. However, AMRD and speech recognition at t_0 were uncorrelated, suggesting that those measures capture partially distinct perceptual abilities. A multiple regression model predicting 6-month speech recognition outcome with deafness duration and speech recognition at t_0 improved from adjusted $R^2 = 0.30$ to adjusted $R^2 = 0.44$ when AMRD threshold was added as a predictor.

Conclusions: These findings identify AMRD thresholds as a reliable, nonredundant predictor above and beyond established speech tests for CI outcome. This AMRD test could potentially be developed into a rapid clinical temporal-resolution test to be integrated into the postoperative test battery to improve the reliability of speech outcome prognosis.

Key words: Amplitude modulation, Cochlear implant, Deafness, Hearing aid, Hearing diagnostics, Hearing loss, Psychoacoustics, Speech perception, Temporal resolution.

(Ear & Hearing 2018;XX;00–00)

INTRODUCTION

Cochlear implants can restore hearing in deaf patients by means of direct electric stimulation of the auditory nerve. The CI-transduced speech signal is considerably distorted, however, and CI recipients vary largely in how well they adapt to their device. Some learn to comprehend speech signals even under difficult listening conditions while others hardly benefit from their device. The source of this variability is still elusive (Moberly et al., 2016). Clinical factors such as age at implantation, duration and etiology of deafness, the number and position of electrodes, as well as cognitive abilities have all been acknowledged to impact speech recognition in adult CI recipients (Blamey et al., 1992; Lazard et al., 2012; Holden et al., 2013). However, those factors can merely explain approximately 20% of the observed individual variability (Lazard et al., 2012). Currently, diagnostic assessment batteries reliably predicting CI patients' speech recognition success are lacking. Therefore, psychoacoustic tests are needed to aid prediction of speech outcome.

As the CI provides limited spectral resolution, listeners are forced to rely on temporal information to extract speech cues. Various psychoacoustic studies have shown that sensitivity to temporal envelope cues, as characterized by the temporal modulation transfer function (TMTF; Bacon and Viemeister, 1985), is crucial for speech comprehension in CI listeners, both in direct electric and acoustic stimulation experiments. The motivation of the present study was to use low-rate temporal modulation sensitivity to predict 6-month speech outcome.

In direct electrical stimulation experiments bypassing the sound processor, Fu (2002) assessed AM detection thresholds at the rate of 100 Hz in experienced CI users (> 6 years postimplantation). AM sensitivity was significantly correlated with phoneme recognition scores. Characterizing full electric TMTFs across a larger range of modulation rates, Shannon (1992) showed that CI patients' detection performance was best at rates of 80 to 100 Hz and declined above 140 Hz. Further, at 50 to 100 Hz, AM rate discrimination and detection thresholds were found to correlate with lexical tone recognition in Mandarin (Luo et al., 2008). Similarly, speech intonation was recognized more accurately by CI listeners with higher sensitivity to amplitude modulations at 50 to 300 Hz (Chatterjee and Peng, 2008), suggesting that these high-frequency amplitude modulations are important in conveying intonation information. TMTFs typically have low-pass characteristics, in both normal-hearing and CI listeners (Park et al., 2015). The rate of decay of the TMTF at high-frequency modulations is also associated with vowel and consonant recognition with a CI (Cazals et al., 1994).

However, direct electric stimulation bypasses the clinical sound processor. In a more realistic setting where stimuli are played through the clinical sound processor, recent studies evaluated AM detection through sound presentation in

[1]Max Planck Research Group "Auditory Cognition", Max Planck Institute for Cognitive and Brain Sciences, Leipzig, Germany; [2]Department of Cognitive Neuroscience, Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands; [3]Department of Otorhinolaryngology, Section Phoniatrics and Audiology, University Clinic of Leipzig, Leipzig, Germany; [4]Faculty of Biosciences, Pharmacy and Psychology, Institute for Biology, University of Leipzig, Leipzig, Germany; [5]Cochlea-Implant-Center Leipzig, Leipzig, Germany; and [6]Department of Psychology, University of Lübeck, Lübeck, Germany.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and text of this article on the journal's Web site (www.ear-hearing.com).

0196/0202/2018/XXXX-00/0 • Ear & Hearing • Copyright © 2018 Wolters Kluwer Health, Inc. All rights reserved • Printed in the U.S.A.

<zdoi; 10.1097/AUD.0000000000000588>

free field. Gnansia et al. (2014) observed a correlation of CI users' acoustic AM detection thresholds at 8 Hz with phoneme identification in quiet. Two further studies comprehensively characterized TMTFs in pre- and postlingually deafened experienced CI users (Won et al., 2011; De Ruiter et al., 2015). They confirmed a correlation of AM detection thresholds averaged across different rates (Won et al., 2011: 10 to 300 Hz; De Ruiter et al., 2015: 5 to 100 Hz) with word recognition scores. A recent longitudinal study attempted to prospectively predict 1-year speech recognition outcome by assessing TMTFs immediately after CI activation, but failed to do so, possibly due to the relatively high modulation rates tested (10 to 300 Hz; Drennan et al., 2016). Thus, while accumulating evidence confirms that CI user's AM sensitivity correlates with speech comprehension, evidence for its predictive power for future speech recognition abilities and adaptation to the CI-transduced speech signal is lacking.

Most previous modulation detection experiments have concentrated on relatively high modulation frequencies above 10 Hz although lower frequency modulations potentially play a crucial role in CI speech perception. It is known that in normal-hearing listeners the low-frequency temporal-envelope information of 3 to 5 Hz is most important for speech comprehension (Drullman et al., 1994a; Shannon et al., 1995; Elliott and Theunissen, 2009); the modulation spectrum of speech also peaks in this frequency range (Giraud and Poeppel, 2012; Ding et al., 2017). Moreover, in normal-hearing adults, listening to a CI simulation (vocoded speech), AMRD thresholds at 4 Hz are predictive of their ability to adapt to degraded speech (Erb et al., 2012) and exhibit correlations with brain structure (grey-matter density in the pulvinar, a thalamic structure; Erb et al., 2012) as well as functional dynamics (as assessed using functional magnetic resonance imaging; Erb et al., 2013).

Nonlinguistic psychophysical tests have two advantages over the more commonly employed speech tests. First, nonspeech tests can predict performance for individuals with native languages for which validated speech tests are currently unavailable (Drennan et al., 2016). Besides the temporal resolution tests (see above), other nonlinguistic psychoacoustic measures, namely spectral resolution, as assessed by spectral ripple discrimination, are strongly associated with CI speech recognition in quiet and particularly in noise, when spectral cues become critical (Henry and Turner, 2003; Won et al., 2007; Drennan et al., 2016; Holden et al., 2016; Lawler et al., 2017). Crucially, temporal and spectral sensitivities are uncorrelated, indicating that the two measures assess fundamentally different, complementary psychoacoustical abilities (Won et al., 2011). Therefore, early acute assessment of both spectral and temporal resolution is needed to improve the reliability of speech outcome prognosis.

Second, some studies argue that nonlinguistic measures are less sensitive to learning, that is, psychophysical abilities tend to remain relatively constant over the first year after implantation (Drennan et al. 2016; Won et al. 2007). Note, however, the caveat, that stability of psychoacoustic thresholds is debated; a recent study reported good threshold reproducibility in both spectral and temporal resolution tasks for same-day testing, but observed a learning effect over a 2-week period after cochlear implantation (de Jong et al., 2017).

The rationale of the present study was to test whether AMRD performance at slow rates shortly after implantation could prospectively predict 6-month CI speech outcome. We present a novel psychoacoustic approach with respect to the following points: (1) We designed a psychoacoustic test assessing AMRD at speech-relevant 4 Hz that was administered shortly after CI activation; (2) We evaluated patients' speech recognition in a longitudinal approach at CI activation and 6 months later. Given that the current version of this AMRD test was already relatively rapid to administer (~20 minutes), it has the potential to be developed into an even more rapid, clinical temporal-resolution test to improve prediction of CI speech outcome.

MATERIALS AND METHODS

Participants

Thirty-four adult CI patients (aged 23 to 85, median 60 years, 24 female) participated in this study and were recruited from the CI center of the University Clinic, Leipzig. Participants gave informed consent, and procedures were approved by, and in accordance with, the guidelines of the local ethics committee (University of Leipzig). Patients were provided with unilateral implants from either cochlear (13 patients; implant: CI512; processor: CP910) or Med-El (21 patients; implant: Concerto; processor: Opus 2). Patients were tested using their standard clinical sound processor settings (without noise reduction), which encoded signals using the following strategies: cochlear: advanced combination encoder, stimulation with 22 electrodes; Med-El: fine structure 4 (FS4) combined with channel-specific sampling sequence (CSSS), stimulation with 12 electrodes.

All patients were native speakers of German. Etiology and types of severe-to-profound sensorineural hearing loss (HL) were variable, including otitis media, measles, meningitis, sepsis, acute HL, congenital and progressive HL. Duration of profound HL before implantation varied from 0 to 58 years (Table 1): Four patients had early-onset deafness; 27 patients were late, that is, postlingually deafened; for three remaining patients, we were not able to obtain a reliable estimate of onset of deafness (for detailed demographic information and individual subject characteristics see supplemental Table S1, Supplemental Digital Content 1, http://links.lww.com/EANDH/A424).

Amplitude modulation rate discrimination

AMRD thresholds were assessed using an adaptive staircase procedure. Stimuli were sinusoidally amplitude-modulated random broad-band noises. The standard stimulus was modulated at 4 Hz. The deviant stimulus was modulated at different rates of 2 to 6 Hz (for determination of deviant modulation rate on each trial see adaptive tracking procedure below). The onset phase of the sinusoidal modulation varied randomly, and stimulus length varied randomly between 900 and 1100 ms to avoid

TABLE 1. Descriptive Statistics for Collected Measures

Variables	Median	Range	IQR
Speech t_6 (% correct)	45	5–80	30
Speech t_0 (% correct)	85	0–100	30
AMRD threshold (Hz)	0.41	0.13–1.5	0.31
Deafness duration (yrs)	23	0–58	25.5

AM rate was varied around a 4-Hz AM rate standard.
AMRD, amplitude modulation rate discrimination.

the possibility that listeners could discriminate stimuli based on the number of cycles in an interval. Modulation depth was held constant at $m = 100\%$ to ensure that it was well above threshold and could be perceived by all patients to avoid effects of modulation depth on thresholds. Stimuli were presented with an inter-stimulus interval of 800 ms. Stimuli were peak-normalized with respect to each other.

Experimental Procedure • Patients were tested approximately 6 weeks after initial activation, that is, approximately 10 weeks after cochlear implantation. Testing took place in a sound-proof audiometric cabin. Stimuli were played on a PC in Matlab 2010b and were presented in a free-field set-up via an AT900-audiometer (Auritec). The presentation level was set to 30 dB above the patient's absolute hearing threshold for band-limited noise averaged across the center frequencies of 500 to 4000 Hz to ensure audibility of the stimuli for all patients. Thus, presentation levels ranged from 54 to 90 dB sound pressure level (SPL; the median was 69 dB SPL). Note that during mapping, the acoustic dynamic range was different for individual subjects. The most comfortable level as well as threshold level were set individually. Although changing the presentation level could lead to the effect of the automatic gain control being different for each patient, previous research indicates that AM detection thresholds are robust to these factors. Won et al. (2011) observed that modulation detection thresholds were independent of presentation levels (ranging from 50 to 75 dBA) and not affected by the operation of automatic gain control when stimuli were presented at 65 dBA. Note, however, the potential caveat, that in contrast to the psychophysics task where we could adjust the presentation level individually, for the speech task, presentation level was fixed at 65 dB SPL, due to the speech test being part of a standardized clinical routine.

AMRD thresholds were estimated using a three-alternative forced choice (3-AFC) adaptive staircase procedure. On each trial, participants heard two standards and one deviant stimulus; the position of the deviant within the trial was randomly varied from trial to trial. The patients' task was to verbally indicate the position ("one," "two," or "three") at which they heard the deviant sound. The experimenter entered the response and started the next trial. We used a 2-down, 1-up adaptive staircase procedure to measure the AMRD threshold converging on 70.7% correct responses (Levitt & Rabiner, 1971). The initial step size was 0.5 Hz and changed to 0.25 Hz after four reversals. The staircase procedure was terminated after 12 reversals. The AMRD threshold was calculated as the absolute value of the difference between average level at the last eight reversals and the standard modulation rate.

Each participant completed two tracking procedures: In the first descending track, the deviant AM rate starting level was 6 Hz (i.e., deviant stimuli were modulated at a faster rate than standard rate of 4 Hz, starting at 6 Hz and descending to 4 Hz). In the second ascending track, deviant AM rate starting level was 2 Hz (i.e., the deviant rate was slower than the standard AM rate). The initial AM rate difference of 2 Hz was chosen to ensure that standard and deviant modulation rate were initially well discriminable by all patients.

Therefore, we obtained two threshold estimates of which we took the smaller one ("whenever the participant did best"). Note that the ascending and descending AMRD thresholds were significantly correlated (Pearson's $r = 0.47$; $p = 0.01$). The aggregate measures, that is, average threshold and minimal threshold were also highly correlated (Pearson's $r = 0.9$; $p < 0.001$). Importantly, the results of the correlational analyses did not depend on the choice of threshold selection, that is, both average and minimal AMRD threshold correlated significantly with speech recognition at t_6 (supplemental Table S2, Supplemental Digital Content 2, http://links.lww.com/EANDH/A425). Note further that 31 participants completed both tracking histories, but 3 participants finished only one tracking history due to time limitations.

Speech Tests

Speech recognition in quiet was assessed as part of the follow-up program a few days and 6 months postactivation using the "Freiburger Sprachaudiogramm" [speech audiogram (Hahlbrock, 1953, Keller, 1977)], which is the gold standard in clinical speech testing in Germany. In the tests, participants were asked to repeat 20 numbers or monosyllabic words. The Freiburger number test in quiet was applied shortly after CI activation (t_0) and the Freiburger monosyllabic word test in quiet was administered 6 months postactivation (t_6). The signal (recorded female voice) was presented in free field through loudspeakers placed in front of the participant (distance: 1 m) at a presentation level of 65 dB SPL. Speech recognition scores were calculated as percent correctly repeated items. For reference, normal-hearing listeners attain speech recognition scores of at least 90% correct in these tests.

Statistical Analyses

To predict 6-month speech outcome, we computed the correlation coefficient Pearson's r in Matlab 2014a. Because a Kolmogorov-Smirnov test indicated that the variables AMRD threshold, Freiburger number test at t_0 (speech t_0) and Freiburger monosyllabic word test at t_6 (speech t_6) were not normally distributed ($p < 0.001$ for all three variables), we used permutation statistics to determine significance. We obtained the empirical null-distribution of Pearson's r by randomly permuting participants (n = 10,000 permutations; e.g., the AMRD threshold of one subject was randomly paired with the speech t_6 score of a different subject) and computing the correlation coefficient for each permutation. The two-tailed p-value was calculated as the proportion of permutations that yielded a test statistic equal or more extreme than the observed one. Missing data (e.g., deafness duration for three patients) were excluded from the analysis using pairwise deletion.

When correlating two noisy measures, ordinary least square regression will underestimate the slope of the relationship between the two measures. Therefore, we used Model II Standard Major Axis Regression in R to estimate regression slopes (Legendre and Legendre, 1998; Legendre, 2008).

We ran multiple linear regression models to test the contribution of single variables to the prediction of 6-month speech outcome. We then compared goodness of fit of different linear regression models by determining the Bayesian information criterion (BIC), which penalizes for the number of fitted parameters and thereby allows for a fair comparison between models (Schwarz, 1978). To triage the risk of poor speech outcome, we also calculated the odds ratio. All data are available on open science framework (DOI 10.17605/OSF.IO/5YPWC).

RESULTS

Correlations

The Freiburger monosyllabic word test at t_6 correlated negatively with AMRD thresholds (Fig. 1A) and positively with the Freiburger number test at t_0 (Fig. 1B). Both correlations were significant after Bonferroni correction for multiple comparisons. These results indicate that 6-month speech recognition was predictable already shortly after CI activation from (1) higher sensitivity to temporal modulations and (2) better speech recognition scores. The correlation between speech recognition at t_6 and AMRD thresholds remained significant after removing the outlier AMRD threshold of 1.5 Hz that was > 3 standard deviations away from the mean (Fig. 1A).

In contrast, the correlation between speech recognition at t_0 and AMRD thresholds had a relatively low r-value and did not reach significance ($r = -0.29$; $p = 0.088$, permutation test), indicating that both measures capture partially distinct aspects of auditory sensitivity. Note that the number test at t_0 is less demanding than the monosyllabic speech test at t_6; 13 participants performed at ceiling in the number test at the beginning of rehabilitation.

Deafness duration correlated negatively with speech recognition at t_6 (Fig. 1C) and with speech recognition at t_0 ($r = -0.36$; $p = 0.048$, permutation test), suggesting that longer periods of deafness before cochlear implantation tend to be associated with poor future speech recognition. Note, however, that both correlations did not survive correction for multiple comparisons. Deafness duration was not related to AMRD thresholds ($r = 0.2$; $p = 0.276$), indicating that the period of auditory deprivation did not affect AMRD ability substantially. Descriptive statistics for all variables are provided in Table 1.

De Ruiter et al. (2015) demonstrated that modulation detection thresholds differ for pre- and postlingually deafened listeners. Here, we did not find a difference in AMRD thresholds between pre- and postlingually deafened patients, according to a two-tailed Wilcoxon rank-sum test (median prelingual: 0.34 Hz; median postlingual: 0.41 Hz; $W = 48$, $z = -0.92$, $p = 0.36$), nor in speech recognition at t_6 (median prelingual: 50% correct; median postlingual: 45% correct; $W = 65.5$, $z = 0.06$, $p = 0.95$), possibly due to the small number of prelingually (n = 4) compared with postlingually deafened listeners (n = 27). Consistent with De Ruiter et al. (2015), all correlations remained significant when ran exclusively on postlingually deafened patients (supplemental Figure S1A, Supplemental Digital Content 3, http://links.lww.com/EANDH/A426).

In contrast, we found significant differences of CI type, that is, between users of cochlear (n = 13) and Med-El (n = 21) devices. According to a two-tailed Wilcoxon rank-sum test, cochlear users (median: 0.289 Hz) had significantly smaller (better) AMRD thresholds than Med-El users (median: 0.531 Hz; $W = 142$, $z = -3.02$, $p = 0.003$). Cochlear users also performed significantly better in speech recognition at t_6 (median: 60% correct) than Med-El users (median: 40% correct; $W = 302$, z = 2.65, $p = 0.008$, Wilcoxon rank-sum test). We followed this up calculating separate correlations for cochlear and Med-El users (supplemental Figure S1B, Supplemental Digital Content 3, http://links.lww.com/EANDH/A426). Note, however, that we had not hypothesized about differences between users of different CI types, as the respective subgroup sizes are too small to yield powered results for all but very large effect sizes. Nevertheless, as exploratory analyses we include supplemental Figure S1B showing correlations separately for cochlear and Med-El users between the following variables: (1) speech t_6 and AMRD thresholds; (2) speech t_6 and speech t_0; (3) speech t_6 and deafness duration.

Multiple Regression

We ran multiple linear regression analyses with the dependent variable speech recognition at t_6 and the independent variables speech recognition at t_0, AMRD threshold, deafness duration, and CI type. The results are shown in Table 2.

Linear regression on speech recognition at t_6 with the sole predictor speech recognition at t_0 resulted in R^2 = 32% of explained variance (Table 2). Adding deafness duration to the linear regression model did not improve the fit according to a likelihood ratio test ($\chi^2(1) = 1.73$, $p = 0.188$). Notably, however, when adding AMRD thresholds to the model, AMRD ability accounted for an additional 12–14% of the variance (after adjustment) in speech recognition at t_6 (Table 2). The likelihood ratio test was significant ($\chi^2(1) = 7.47$, $p = 0.006$), confirming

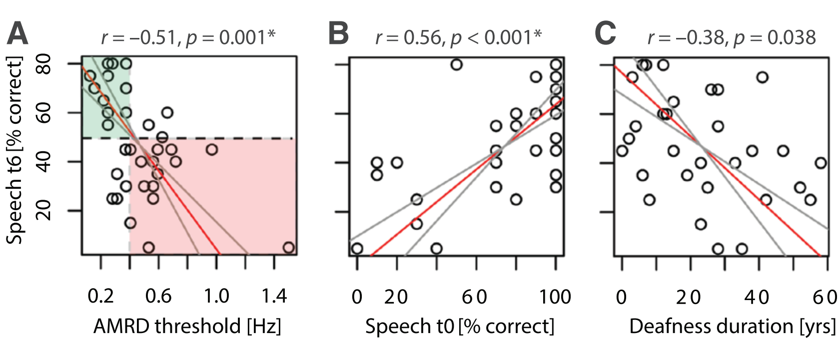

Figure 1. Correlations of 6-month speech recognition with all other measures. Scatter diagrams show the standard major axis regression line (red) and its 95% confidence region (grey). Pearson's correlation coefficients are indicated above scatter plots; p values were obtained using exact permutation test (n = 10,000 permutations). A, Speech recognition scores 6 months postactivation (speech t_6) significantly correlate with AMRD thresholds. The correlation remains significant when removing the outlier AMRD threshold of 1.5 Hz ($r = 0.4$; $p = 0.02$). The median for AMRD thresholds and speech t_6 are shown with the dashed line. According to this sample (n = 34), the odds for an adverse speech-comprehension outcome after 6 months (i.e., comprehension < 50%) are 17.6 higher in CI recipients who presented with elevated AMRD thresholds briefly after CI implantation and activation (lower-right quadrant, red). B, Correlation of speech t_6 with speech recognition scores shortly after CI activation (speech t_0), and deafness duration (C). *Significant after Bonferroni correction for multiple comparisons.

TABLE 2. Multiple Linear Regression Models Predicting Speech Recognition at t_6

Predictors	R^2	Adjusted R^2	F-statistic	p Value
Speech t_0	0.32	0.30	$F(1,32) = 14.44$	0.001
Speech t_0 + deafness duration	0.35	0.30	$F(2,27) = 7.53$	0.002
AMRD threshold	0.26	0.24	$F(1,32) = 11.51$	0.002
Speech t_0 + **AMRD threshold**	0.45	0.42	$F(2,31) = 12.53$	< 0.001
Speech t_0 + deafness duration + **AMRD threshold**	0.49	0.44	$F(3,27) = 8.74$	< 0.001
Speech t_0 + **AMRD threshold** + CI type	0.55	0.51	$F(3,30) = 12.47$	< 0.001

AMRD, amplitude modulation rate discrimination.

a better goodness of fit for the model including AMRD thresholds. Thus, AMRD constitutes a promising novel predictor for future speech recognition in CI recipients. This model also had a better fit than the model using AMRD thresholds as only regressor ($\chi^2(1) = 9.69$, $p = 0.002$). Due to the effect of CI type on speech recognition at t_6 (see above), we included CI type as an additional predictor in the multiple regression analyses. This model accounted for adjusted R^2 = 51% of the variance (Table 2) and had a significantly better fit than the nested model including speech recognition at t_0 and AMRD thresholds ($\chi^2(1) = 7.38$, $p = 0.006$).

We then formally compared the model including solely speech recognition at t_0, to the model additionally including AMRD thresholds: The model including both speech recognition at t_0 and AMRD thresholds had a smaller Bayes information criterion (BIC = 85.9) than the nested model including speech recognition at t_0 only (BIC = 89.9) and thus a better fit even when penalizing for the higher number of fitted parameters. This can also be expressed as a Bayes Factor of 7.2, indicating more than sevenfold higher probability to observe these data under the combined model including AMRD compared with the nested model including speech recognition at t_0 only.

What are the potential clinical implications of our findings for the outcome prognosis after cochlear implantation? To find a tentative answer, we defined the clinically important cutoff in 6-month speech recognition as 50% correct (median was 45%, Table 1 and Fig. 1A). AMRD performance was split into poor (> 0.4 Hz; note that the median threshold was 0.4 Hz) and good discriminators (≤ 0.4 Hz; Fig. 1A). To quantify how strongly the occurrence of poor AMRD was associated with poor speech outcome, we calculated the odds ratio and relative risk. Poor AMRD performers were significantly more likely to have a poor 6-month speech outcome than good AMRD performers (odds ratio = 17.6; 95% CI, 2.9–107.6; relative risk = 2.8 [1.6–31.2]; Fig. 1A). Thus, poor AMRD shortly after CI activation increases the odds for poor 6-month speech recognition by a factor of 17.6 relative to good AMRD.

DISCUSSION

The main objective of the present study was to investigate whether auditory temporal sensitivity shortly after implantation could prospectively predict speech recognition abilities and adaptation to the CI-transduced speech signal. Our results show that AMRD thresholds constitute a promising novel predictor for 6-month CI outcome. Thus, for the prediction of 6-month speech recognition, adjusted R^2 improved from 0.3 to 0.44 when AMRD thresholds were added to speech recognition and deafness duration in a multiple regression. An odds ratio of 17.6 confirmed that poor AMRD performance shortly after CI activation constitutes a risk factor for poor CI outcome 6 months later. A shorter clinical test of AMRD ability could be developed as predictor of CI speech outcome in addition to the more commonly assessed word recognition tests, as they independently assess distinct auditory sensitivities.

Temporal Sensitivity and Speech Recognition

Our finding of a correlation between AMRD and speech recognition scores (both at the time of CI activation and 6 months later) is consistent with previous observations of temporal sensitivity being crucial for speech comprehension with a CI. A large body of experiments assessing electric (Cazals et al., 1994; Fu, 2002; Luo et al., 2008; Chatterjee and Oberzut, 2011) and acoustic modulation thresholds (Won et al., 2011; Gnansia et al., 2014; De Ruiter et al., 2015) demonstrate a relationship between AM sensitivity and speech perception in experienced CI users.

Above and beyond these previous studies, we show that sensitivity to slow AM rates is also helpful in prospectively predicting CI users' speech comprehension over a time scale of half a year after implantation. In normal-hearing participants, listening to a CI simulation we observed that AMRD thresholds centered at 4 Hz are predictive of short-term adaptation: Improvement during 20 minutes of listening to noise-vocoded speech correlated with AM rate sensitivity (Erb et al., 2012, 2013). Yet, in CI patients, Drennan et al. (2016) failed to prospectively predict 1-year outcome based on TMTFs, possibly due to the focus on higher AM rates: The authors did not measure sensitivities to temporal rates below 10 Hz although those are known to be crucial for speech recognition (Drullman et al., 1994a, b; Elliott and Theunissen, 2009).

Clinical Implications

Our results suggest that assessment of AMRD abilities could be a valuable addition to the clinical postoperative test battery and has the potential to improve the reliability of speech outcome prognosis. AMRD thresholds as predictor for prospective speech recognition abilities have several practical advantages over the more commonly employed speech tests (e.g., Freiburger speech audiogram). First, nonlinguistic psychoacoustic tests are independent of native language and can thus be employed in speakers of languages for which no validated speech tests are currently available (Drennan et al., 2016). Second, AMRD assessment was relatively rapid and efficient. In contrast to previous studies characterizing complete TMTFs that involved extensive testing of ~2 hours (Won et al., 2011; De Ruiter et al., 2015; Drennan et al., 2016), our psychoacoustic test took ~20 minutes per patient. An even shorter clinical screening test could be developed based on our results. Instead of using an adaptive procedure, a shorter clinical version could be based on the method of constant stimuli (cf. Drennan et al. (2014) for a clinical spectral ripple test), with only few trials presented at a fixed modulation rate at a critical level (e.g., the median = 0.4 Hz; Fig. 1A). Patients who fail this test (i.e., perform at chance level) would be at risk of poor speech comprehension 6 months later.

Consequently, additional training measures would be in order for those poorly performing patients to improve their perception of temporal speech cues. One of the major challenges for future CI research is to develop treatments and interventions for patients who respond poorly to their CI (Pisoni et al., 2018). While the currently available computer-based training programs (e.g., Angel Sound, LACE, Sound and Way Beyond; for review see Olson, 2015) are generic programs that often only benefit a subgroup of patients, novel individualized interventions for poorly performing listeners are needed. Comprehensive assessment of both auditory and cognitive profiles is critical for the development of evidence-based treatments for individual patients. Some success has been observed with the training of discrimination of specific acoustic cues. For example, pitch discrimination training improved sensitivity to spectral cues in speech signals (Fu and Galvin, 2008; Ingvalson and Wong, 2013).

Moreover, regarding sound encoding strategies of the clinical sound processor, our findings imply that accurate transmission of temporal modulation information is crucial for speech recognition (Won et al., 2011). Thus, temporal resolution tests can also help choosing the most suitable stimulation strategies during the first 6 months after activation, to accurately deliver temporal cues. Actually, the observed effect of CI type on AMRD thresholds and speech recognition at t_6 may relate to differences in stimulation strategy, in particular, the difference in how low AM rate is encoded by FS4 (Med-El) and advanced combination encoder (cochlear).

A caveat of the Freiburger number test at t_0 was that 13 participants performed at ceiling (Fig. 1B), most likely due to the number test being relatively easy (in contrast to the more challenging monosyllabic speech test at t_6). Ceiling effects may artificially have reduced the predictive power of the easier speech task. Our adaptive psychoacoustic AMRD task has the benefit that it inherently avoids the problem of ceiling effects and can thus better capture variability in auditory performance.

Our approach was limited to predicting perception of CI-transduced speech in quiet. Hence, whether our results generalize to more natural listening situations, where usually some levels of background noise are present, needs further investigation. Speech perception in noise relies more heavily on spectral cues. Spectral discrimination abilities also become crucial in different listening situations, such as music perception with a CI (Drennan and Rubinstein, 2008).

CONCLUSIONS

The present results demonstrate that future speech recognition in CI listeners can be predicted with notably high accuracy based on AMRD abilities early after implantation. The current data provide compelling evidence that psychoacoustic assessment 10 weeks postoperatively delivers relevant information about CI patients' speech comprehension abilities scores 6 months later. A shorter clinical test of AMRD ability could potentially be designed to be incorporated into the postoperative rehabilitation program. Based on our results, personalized rehabilitation strategies could be developed to identify and train poor performers in efficiently exploiting temporal cues for speech recognition.

ACKNOWLEDGMENTS

The authors thank all patients for their efforts in taking part in this study.

Research was supported by a Max Planck Research Group grant (to J. O.). We thank the editor and three anonymous reviewers for providing very constructive criticism.

The authors have no conflicts of interest to disclose.

Address for correspondence: Julia Erb, University of Lübeck, Department of Psychology, Maria-Goeppert-Straße 9a, 23562 Lübeck, Germany. E-mail: erbjulia@gmail.com

Received July 3, 2017; accepted February 24, 2018.

REFERENCES

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

Blamey, P. J., Pyman, B. C., Gordon, M., et al. (1992). Factors predicting postoperative sentence scores in postlinguistically deaf adult cochlear implant patients. *Ann Otol Rhinol Laryngol*, *101*, 342–348.

Cazals, Y., Pelizzone, M., Saudan, O., et al. (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.

Chatterjee, M., & Oberzut, C. (2011). Detection and rate discrimination of amplitude modulation in electrical hearing. *J Acoust Soc Am*, *130*, 1567–1580.

de Jong M. A. M., Briaire J. J., Frijns J. H. M. (2017). Learning effects in psychophysical tests of spectral and temporal resolution. *Ear Hear*. Sept. 26. [Epub ahead of print]

De Ruiter, A. M., Debruyne, J. A., Chenault, M. N., et al. (2015). Amplitude modulation detection and speech recognition in late-implanted prelingually and postlingually deafened cochlear implant users. *Ear Hear*, *36*, 557–566.

Ding, N., Patel, A. D., Chen, L., et al. (2017). Temporal modulations in speech and music. Neuroscience and biobehavioral reviews*., 81(Pt B)*, 181–187.

Drennan, W. R., Anderson, E. S., Won, J. H., et al. (2014). Validation of a clinical assessment of spectral-ripple resolution for cochlear implant users. *Ear Hear*, *35*, e92–e98.

Drennan, W. R., & Rubinstein, J. T. (2008). Music perception in cochlear implant users and its relationship with psychophysical capabilities. *J Rehabil Res Dev*, *45*, 779–789.

Drennan, W. R., Won, J. H., Timme, A. O., et al. (2016). Nonlinguistic outcome measures in adult cochlear implant users over the first year of implantation. *Ear Hear*, *37*, 354–364.

Drullman, R., Festen, J. M., Plomp, R. (1994a). Effect of reducing slow temporal modulations on speech reception. *J Acoust Soc Am*, *95*, 2670–2680.

Drullman, R., Festen, J. M., Plomp, R. (1994b). Effect of temporal envelope smearing on speech reception. *J Acoust Soc Am*, *95*, 1053–1064.

Elliott, T. M., & Theunissen, F. E. (2009). The modulation transfer function for speech intelligibility. *PLoS Comput Biol*, *5*, e1000302.

Erb, J., Henry, M. J., Eisner, F., et al. (2012). Auditory skills and brain morphology predict individual differences in adaptation to degraded speech. *Neuropsychologia*, *50*, 2154–2164.

Erb, J., Henry, M. J., Eisner, F., et al. (2013). The brain dynamics of rapid perceptual adaptation to adverse listening conditions. *J Neurosci*, *33*, 10688–10697.

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

Fu, Q. J., & Galvin, J. J. 3rd. (2008). Maximizing cochlear implant patients' performance with advanced speech training procedures. *Hear Res*, *242*, 198–208.

Giraud, A. L., & Poeppel, D. (2012). Cortical oscillations and speech processing: emerging computational principles and operations. *Nat Neurosci*, *15*, 511–517.

Gnansia, D., Lazard, D. S., Léger, A. C., et al. (2014). Role of slow temporal modulations in speech identification for cochlear implant users. *Int J Audiol*, *53*, 48–54.

Hahlbrock, K. H. (1953). [Speech audiometry and new word-tests]. *Arch Ohren Nasen Kehlkopfheilkd*, *162*, 394–431.

Henry, B. A., & Turner, C. W. (2003). The resolution of complex spectral patterns by cochlear implant and normal-hearing listeners. *J Acoust Soc Am*, *113*, 2861–2873.

Holden, L. K., Finley, C. C., Firszt, J. B., et al. (2013). Factors affecting open-set word recognition in adults with cochlear implants. *Ear Hear*, *34*, 342–360.

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

Ingvalson, E. M., & Wong, P. C. (2013). Training to improve language outcomes in cochlear implant recipients. *Front Psychol*, *4*, 263.

Keller F. (1977). Verschiedene Aussprachen des Sprachverstehenstests nach DIN 45621 (Freiburger Test). *Biomed Technol*, *22*, 292–298.

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.

Lazard, D. S., Vincent, C., Venail, F., et al. (2012). Pre-, per- and postoperative factors affecting performance of postlinguistically deaf adults using cochlear implants: a new conceptual model over time. *PLoS One*, *7*, e48739.

Legendre, P. (2008). Model II regression user's guide, R Edition.

Legendre, P., Legendre, L. (1998). *Numerical Ecology*. Amsterdam, The Netherlands: Elsevier Science BV.

Levitt, H., & Rabiner, L. R. (1971). Analysis of fundamental frequency contours in speech. *J Acoust Soc Am*, *49*, Suppl 2:569+.

Luo, X., Fu, Q. J., Wei, C. G., et al. (2008). Speech recognition and temporal amplitude modulation processing by Mandarin-speaking cochlear implant users. *Ear Hear*, *29*, 957–970.

Moberly, A. C., Bates, C., Harris, M. S., et al. (2016). The enigma of poor performance by adults with cochlear implants. *Otol Neurotol*, *37*, 1522–1528.

Olson, A. D. (2015). Options for auditory training for adults with hearing loss. *Semin Hear*, *36*, 284–295.

Park, M. H., Won, J. H., Horn, D. L., et al. (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.

Pisoni, D. B., Kronenberger, W. G., Harris, M. S., et al. (2018). Three challenges for future research on cochlear implants. *World J Otorhinolaryngol Head Neck Surg*. Jan. 2. [Epub ahead of print]

Schwarz, G. (1978). Estimating the dimension of a model. *Ann Stat*, *6*, 461–464.

Shannon, R. V., Zeng, F. G., Kamath, V., et al. (1995). Speech recognition with primarily temporal cues. *Science*, *270*, 303–304.

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

Won, J. H., Drennan, W. R., Rubinstein, J. T. (2007). Spectral-ripple resolution correlates with speech reception in noise in cochlear implant users. *J Assoc Res Otolaryngol*, *8*, 384–392.