Article

Comparing the Speech Perception of Cochlear Implant Users with Three Different Finnish Speech Intelligibility Tests in Noise

Tytti Willberg [1,2,*] , Ville Sivonen [3] , Pia Linder [4] and Aarno Dietz [4]

1   Department of Otorhinolaryngology, Turku University Hospital, 20521 Turku, Finland
2   Institute of Clinical Medicine, University of Eastern Finland, 70211 Kuopio, Finland
3   Department of Otorhinolaryngology—Head and Neck Surgery, Head and Neck Center,
    Helsinki University Hospital and University of Helsinki, 00029 Helsinki, Finland; ville.sivonen@hus.fi
4   Department of Otorhinolaryngology, Kuopio University Hospital, 70029 Kuopio, Finland;
    pia.linder@kuh.fi (P.L.); aarno.dietz@kuh.fi (A.D.)
*   Correspondence: tytti.willberg@tyks.fi

Abstract: Background: A large number of different speech-in-noise (SIN) tests are available for testing cochlear implant (CI) recipients, but few studies have compared the different tests in the same patient population to assess how well their results correlate. Methods: A clinically representative group of 80 CI users conducted the Finnish versions of the matrix sentence test, the simplified matrix sentence test, and the digit triplet test. The results were analyzed for correlations between the different tests and for differences among the participants, including age and device modality. Results: Strong and statistically significant correlations were observed between all of the tests. No floor or ceiling effects were observed with any of the tests when using the adaptive test procedure. Age or the length of device use showed no correlation to SIN perception, but bilateral CI users showed slightly better results in comparison to unilateral or bimodal users. Conclusions: Three SIN tests that differ in length and complexity of the test material provided comparable results in a diverse CI user group.

Keywords: speech perception; speech perception in noise; cochlear implant; matrix sentence test; digit triplet test; outcomes; adult; hearing loss; audiology

1. Introduction

Cochlear implantation (CI) has been a standard treatment for severe-to-profound hearing loss (HL) for more than two decades. CIs are safe and effective intervention for HL [1–4], even in elderly patients [5]. Reaching open-set speech perception is currently a realistic and attainable goal for the majority of CI users [6–8]. However, while CIs provide most patients adequate speech perception in quiet listening conditions, speech perception typically deteriorates drastically in noise [6,9], and noisy listening environments remain a challenge for CI users.

Testing CI users' speech perception both in quiet and in noise is essential for obtaining a comprehensive view of hearing in everyday listening conditions. Many current national and international guidelines recommend including some type of a speech-in-noise (SIN) test in the test batteries used for evaluating CI candidacy and rehabilitation outcomes [10–12]. However, whereas audiometry is a standardized procedure that provides reliable and comparable results across languages and diverse patient populations, SIN tests are language specific and subject to multiple confounding factors, which complicates the interpretation of their results [13].

A plethora of different types of SIN tests are available for many languages to accommodate diverse patient and testing needs [14–17]. For example, complex sentences better reflect everyday conversations than isolated words, and sentence-level tests are therefore usually preferred when testing in the clinic. On the other hand, children and

adults with limited language knowledge benefit from shorter speech materials with familiar words [18,19]. The COVID-19 situation has increased the need and interest for tele-audiology and remote hearing testing [20,21]. Luckily, simple, closed-set test materials (i.e., digit strings) are well suited for remote testing [15,22], and remote testing is feasible and reliable with current CI technology [23,24].

One of the confounding factors for interpreting SIN test results is the background noise that can vary from a steady state [17,25,26] to a fluctuating noise or emerge as speech babble [16,27]. Speaker selection also introduces variability, as some tests only use one speaker with clear articulation [25,26,28], while others have multiple speakers using different dialects [27]. These differences in SIN test content and construction significantly affect the test scores. Therefore, SIN test scores that are typically expressed either as a percentage of correctly recognized words or as speech reception threshold (SRT) are test-specific.

Only a few studies have compared two or more different sentence-level SIN tests in the same hearing-impaired patient population [29–31]. Data from a study by Dillon et al. [31] demonstrate how differences between test-specific characteristics lead to different results in the same patient population. When a group of CI users were tested with the Hearing in Noise Test (HINT) [25], a SIN test that uses simple, short everyday sentences presented in a steady background noise, the speech recognition scores in noise were, on average, between 80–100%. However, when the same CI users conducted the AzBio sentence test [16], which uses more complex sentence material presented in babble-noise, the scores were 20–60%.

Even the same SIN test can produce different results if different test protocols are used. For example, at +10 dB signal-to-noise ratio (SNR) a group of CI users reached a mean score of 46.8% for the AzBio sentence test, while at a more difficult SNR of +5 dB, the mean result was only 23.3% [32]. Some SIN tests have a strictly defined, standardized protocol that clearly states which SNRs to use [17,25]. Clinical guidelines also provide recommendations for a standardized use of the tests [10]. However, a recent survey revealed that the recommended procedures are not always followed in the clinic [33], which can lead to biased results and further complicate their interpretation.

Comparing CI users' speech perception results from two different SIN tests is difficult without intricate knowledge of both tests, testing conditions, and the typical CI user performance range for both tests. Often, when a new SIN test is developed, the only reported reference value is the normative reference value for normal-hearing (NH) listeners [17,25,27]. However, NH listeners typically reach 100% speech intelligibility even with the more complex SIN tests [34]. Since restoring normal speech perception in noise is not possible with the current CI systems, CI users usually fail to reach the normative range for most SIN tests [2,32,35,36]. In addition to reference data for NH listeners, reference data for CI users would be valuable in the clinical practice. The variability in CI rehabilitation results is considerable [1,3,37], and the concept of CI user reference data is complicated, as it is inevitable that the average performance of a large CI patient population may be an unrealistic goal for some CI users and an underachievement for others. Nevertheless, the need for CI user reference data that is collected from a large, unselected patient population under standardized, well-defined testing conditions is recognized in multiple recent publications [12,33,38].

Ideally, all speech perception in noise assessments could be carried out with one, internationally comparable SIN test. However, due to differences in patient-related factors (i.e., age, memory span, language knowledge), test environments (clinic vs. home), and testing requirements (accurate clinical diagnostic vs screening), no single test can currently accommodate all of these needs. To meet the requirements of a comprehensive SIN test battery, three SIN tests have been developed in Finnish. The Finnish matrix sentence test (FMST) [26] is the standard for SIN testing in Finnish, with reference data for CI users published in 2015 [39]. The two more recently developed SIN tests, the Finnish simplified matrix sentence test (FINSIMAT) [40] and the Finnish digit triplet test (FDTT) [41] have not yet been validated for CI users. Our first aim was to validate these tests in a clinically

representative CI patient population by assessing their correlation with the FMST. The second aim of our study was to provide reference data for the new tests in CI users.

2. Materials and Methods

2.1. Participants

The participants in this study were 80 consecutive adult and adolescent CI users from Kuopio University Central Hospital and Helsinki University Central Hospital, who volunteered to participate in the study during their routine follow-up visit. Data were collected between December 2016 and February 2019. Altogether 83 CI users met the inclusion criteria (native Finnish speakers, no cognitive impairments, at least four months of CI experience). However, three CI users had to be excluded since their speech recognition score for the FMST presented at +10 dB SNR was less than 70%. The reason for exclusion was the adaptive test procedure the FMST uses; the adaptive test procedure converges to a SNR where 50% of the test items are correctly recognized, and it only functions reliably if the majority of the test items are recognized at +10 dB SNR [7]. The study was approved by the Research Ethics Committees of the Hospital District of Northern Savo and HUS, and informed consent was obtained from all of the participants.

2.2. Speech Intelligibility Tests

The differences between the three tests are summarized in Table 1. The FMST [26] is based on the Oldenburg Sentence Test (OLSA), and the general principles of this type of matrix test are described in detail elsewhere [17,42–44]. The test material consists of 50 common everyday words combined into five-word sentences. All of the sentences have the same structure, but the content is semantically unpredictable (e.g., "Johanna buys seven large cups." or "Harri sees four old cars."). The FINSIMAT [40] is based on the simplified version of the OLSA, the Oldenburg Kinder Satztest [45]. Simplified matrix sentence tests use a shortened version of the original word matrix to form three-word pseudo-sentences (i.e., "seven large cups" or "four old cars"). The third SIN test in our study, the FDTT [41], uses digit triplets (i.e., 2-7-1) as speech material. It is based on a test originally developed by Smits et al. [15] for large-scale hearing screening.

Table 1. The characteristics of the speech intelligibility in noise tests used in the study.

	Finnish Matrix Sentence Test (FMST)	Finnish Simplified Matrix Sentence Test (FINSIMAT)	Finnish Digit Triplet Test (FDTT)
Test material	Five-word sentences	Three-word pseudo-sentences	Digit triplets (*two-six-nine*)
Content of a test list	20 sentences	14 pseudo-sentences	30 triplets
Step size for the adaptive procedure	Decreases progressively	Decreases progressively	Fixed at 2 dB SNR
Scoring method	Word scoring	Word scoring	Triplet scoring
Method for determining the SRT	Maximum likelihood method	Maximum likelihood method	The mean SNR of the last 27 triplets
Reference value (SD) for normal-hearing listeners [1]	−9.7 (0.7) dB SNR	−11.2 (1.0) dB SNR	−10.8 (0.5) dB SNR

[1] Reference values were obtained monaurally with headphones: Finnish matrix sentence test [26], Finnish simplified matrix sentence test [40], and Finnish digit triplet test [41]; SNR: signal-to-noise ratio.

All of the tests use speech-shaped background noise that is created by superimposing the test material 30 times. All of the tests can be conducted at a fixed SNR, but the preferred clinical practice is to use an adaptive test procedure, where the presentation level of the next speech item is based on the responses for the previous items. The FDTT uses a fixed step size of 2 dB, and the whole triplet needs to be recognized for the answer to be considered correct. The FMST and the FINSIMAT use word scoring, and the step size changes based on how many of the words in the (pseudo)sentence were correctly recognized [46]. When using the adaptive test procedure, the results are expressed as a SRT estimate, which is the SNR that is needed to correctly recognize 50% of the test material. For the FMST and the FINSIMAT, the test software obtains the SRT estimates by fitting the test data to the test-specific psychometric function. For the FDTT the SRT estimates are calculated

by averaging the SNRs of the last 27 triplets. In our clinical procedure, the noise is held constant at 65 dB SPL, and the starting SNR is 0 dB for all of the tests.

2.3. Test Procedure

The participants conducted all of the tests during a single visit with their preferred device configuration (CI, CI + HA or CI + CI) and instructions to use their preferred device settings. All of the measurements were conducted in a sound field with the participants seated at 1 m distance in front of the loudspeaker that was used to present the signal and the noise (S0N0). Testing began with two practice measurements with the FMST: the first FMST test list was presented at +10 dB SNR, and the second test list using the adaptive test procedure. The practice phase was followed by the actual test measurements: one test list was conducted for each test using the adaptive test procedure. The test material was presented in a counterbalanced order across participants. The participants received no feedback during the testing, but they were told in advance which test would be conducted next. Before the practice phase and the test measurements with the FMST and the FINSIMAT, the participants were shown the word matrix in writing for each test in order to (re)familiarize them with the words, but the written word matrices were not available during the actual test measurements.

2.4. Statistical Analysis

The baseline characteristics and the final results were described with standard descriptive statistics. Frequency and relative frequency were used to describe categorical variables. Normal distribution was tested using the Shapiro-Wilk test. When the data were normally distributed, means, standard deviations (SD), and ranges were used describe the continuous variables. Medians and ranges were used for variables with non-normal distribution.

Comparisons between the different device configurations were performed using one-way analysis of variance (ANOVA) with Sidak correction for multiple comparisons for post-hoc tests. Based on the normality of the results, the Pearson correlation coefficient (r), or Spearman's rank-order correlation coefficient (r_S) was used to assess the strength of association among the continuous variables. Data were analyzed for all of the participants together and, for additional analysis, the participants were divided in to two age groups: (1) age $\leq$ 65 and (2) age > 65 years. IBM SPSS Statistics (Version 27) was used for all statistical analysis with a significance criterion of $p \leq 0.05$.

3. Results

The demographics of the participants are summarized in Table 2. The mean age of the participants was 56 years (SD = 16.7). Of all the participants, 55 (69%) were 65 years of age or younger. Detailed hearing history was not recorded as our aim was not to assess explanatory factors behind the results, but to compare the tests to each other. The range for CI experience was wide, as the most recent CI recipient had only five months of CI experience while the most experienced user had used a CI for more than 18 years. Med-El was the most commonly used CI manufacturer (49% of all participants). All of the participants were fitted with current CI and HA technology at the time of the data collection. The CI processors used were Naída CI Q70/90 (Advanced Bionics); Opus 2, Sonnet, Rondo (Med-El); CP810, CP910, Kanso (Cochlear). All other data were normally distributed except for CI experience, recognition scores at +10dB SNR, and age in the two age groups.

For the FMST, the SRT estimates showed a large variance in the participants' speech perception, with results ranging from +0.8 to −7.7 dB SNR. A similar variance was observed for the FINSIMAT (+0.1 to −9.1 dB SNR) and the FDTT (−0.4 to −8.2 dB SNR). The mean SRT estimates for the tests are summarized in Table 2.

Table 2. Summary of the demographics and test results of the participants.

	Age ≤65	Age >65	All
Number of participants	55	25	80
Kuopio	33 (61%)	21 (39%)	54
Helsinki	22 (85%)	4 (15%)	26
Age, median (range), y	53 (17–65)	73 (66–85)	60 (17–85)
Length of CI experience, median (range), mo	32 (5–229)	14 (6–87)	26 (5–229)
Device configuration			
CI	29 (66%)	15 (34%)	44
CI + HA	8 (57%)	6 (43%)	14
CI + CI	18 (82%)	4 (18%)	22
Cochlear implant manufacturer			
Advanced Bionics	5 (45%)	6 (55%)	11
Cochlear	20 (67%)	10 (33%)	30
Med-El	29 (74%)	10 (26%)	39
Speech perception scores			
FMST, +10 dB SNR median (range), %	98 (78–100)	96 (76–100)	97 (76–100)
FMST, adaptive procedure mean (SD), dB SNR	−4.2 (2.0)	−4.0 (2.1)	−4.2 (2.1)
FINSIMAT, adaptive procedure mean (SD), dB SNR	−5.0 (2.0)	−5.1 (1.8)	−5.0 (1.9)
FDTT, adaptive procedure mean (SD), dB SNR	−4.7 (2.0)	−4.4 (1.6)	−4.6 (1.9)

FMST, Finnish matrix sentence test; FINSIMAT, Finnish simplified matrix sentence test; FDTT, Finnish digit triplet test; SNR, signal-to-noise ratio.

Figure 1 shows all the SRT estimates plotted against each other. All of the test scores were strongly correlated, and the correlations were statistically significant ($p < 0.001$ for all). The strongest correlation was between the FMST and the FINSIMAT (r = 0.819), and it was only slightly weaker between the FMST and the FDTT (r = 0.779). The FINSIMAT and the FDTT had an equally good correlation (r = 0.769). No correlation was observed between the SRT estimates and age, or between the SRT estimates and the CI experience (Figure 2). The differences in the SRT estimates between the younger and older CI users were small (Figure 3a), and statistically insignificant for any of the tests (t_{FMST} (78) = −0.503, $p = 0.617$; $t_{FINSIMAT}$ (78) = 0.384, $p = 0.702$; t_{FDTT} (78) = −0.679, $p = 0.499$).
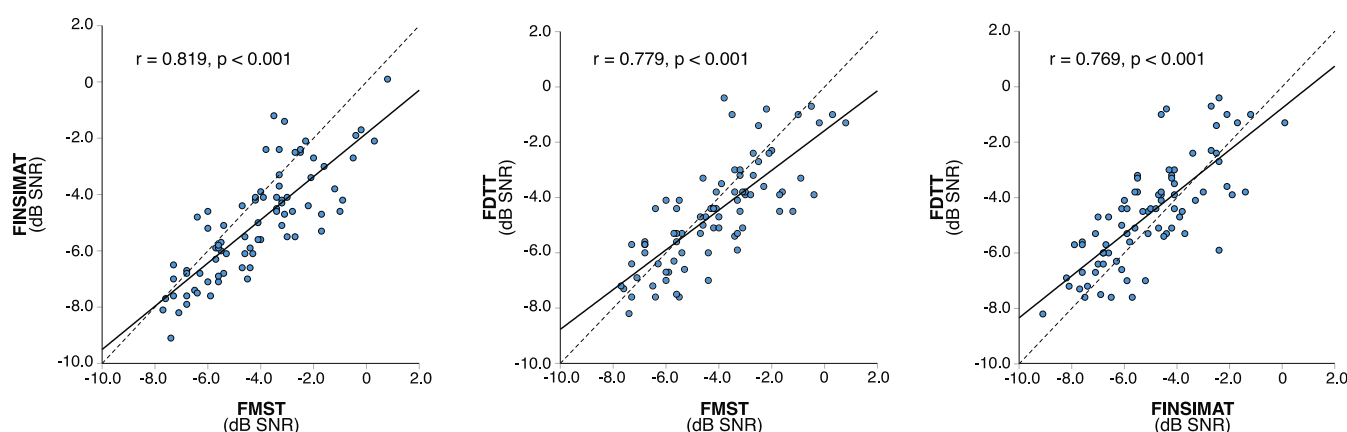

Figure 1. Speech reception threshold (SRT) estimates in dB signal-to-noise ratio (SNR) plotted against each other for the three-speech intelligibility in noise tests with a linear regression plotted as a solid line. Pearson correlation coefficients (r) and *p*-values are included for all regression plots. FMST, Finnish matrix sentence test; FINSIMAT, Finnish simplified matrix sentence test; FDTT, Finnish digit triplet test.
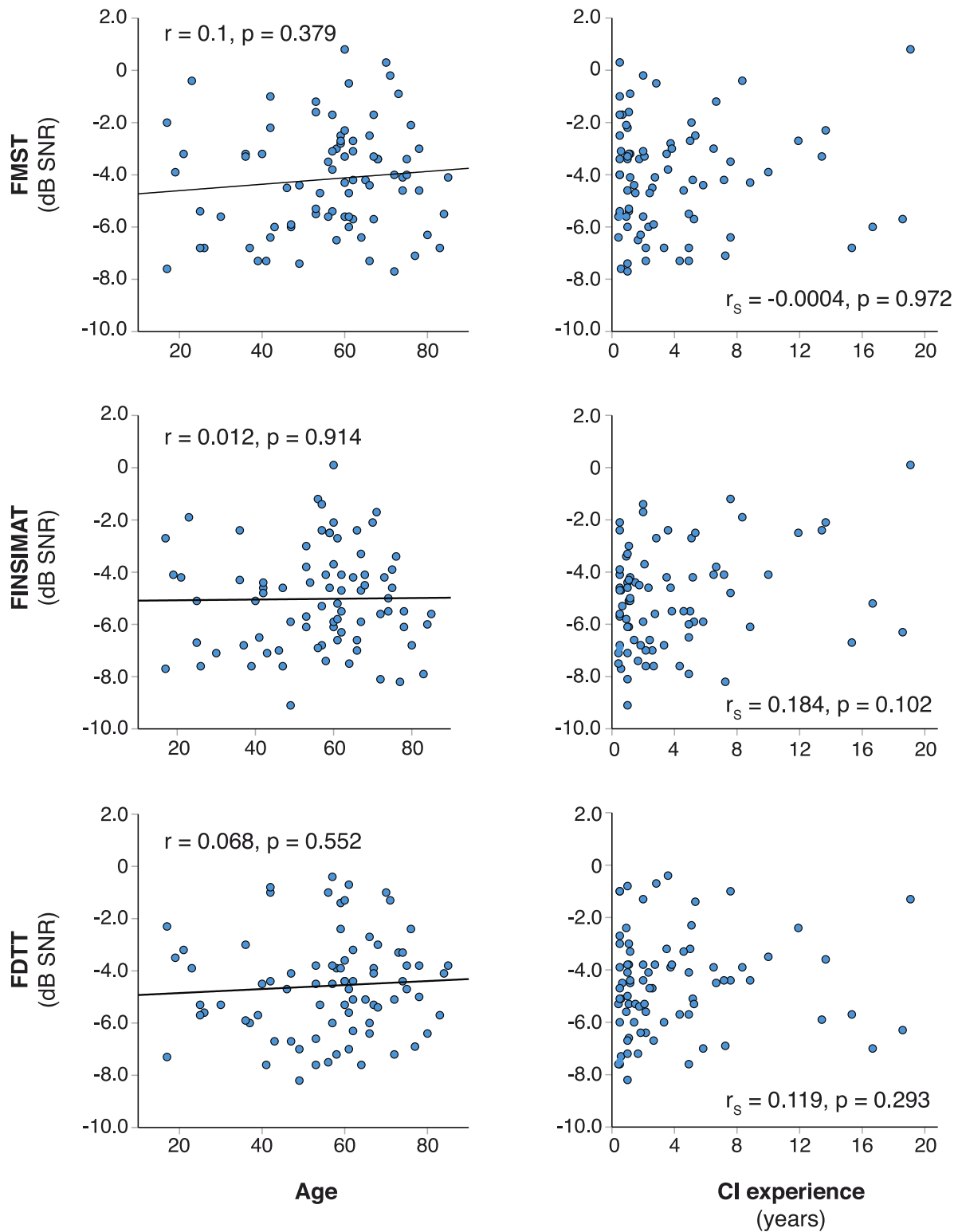
Figure 2. Age and cochlear implant (CI) experience plotted against the speech reception threshold (SRT) estimates for all three tests. For correlations to age, a linear regression is plotted as a solid line, and Pearson correlation coefficients (r) with corresponding *p*-values are included. Spearman's rank-order correlation coefficient (r_S) was used for CI experience data.
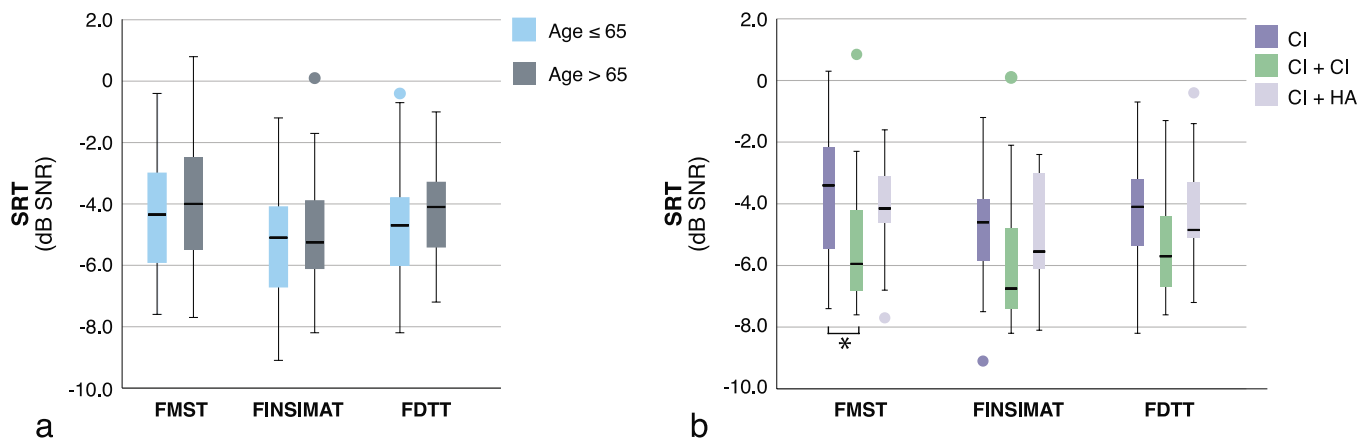
Figure 3. Comparisons between the different age groups (**a**), and device modalities (**b**). The asterisk notes the only statistically significant difference between the results. FMST, Finnish matrix sentence test, FINSIMAT, Finnish simplified matrix sentence test; FDTT, Finnish digit triplet test; CI, cochlear implant; HA, hearing aid.

Participants with bilateral CIs had better SRT estimates for all of the tests (Figure 3b). However, the differences were only statistically significant for the FMST (F(277) = 6.311, *p* = 0.003), and the post-hoc comparisons only showed a statistically significant difference between unilateral and bilateral CI users (*p* = 0.002). A ceiling effect was observed for the FMST presented at +10 dB SNR, but not when using the adaptive test procedure (Figure 4).
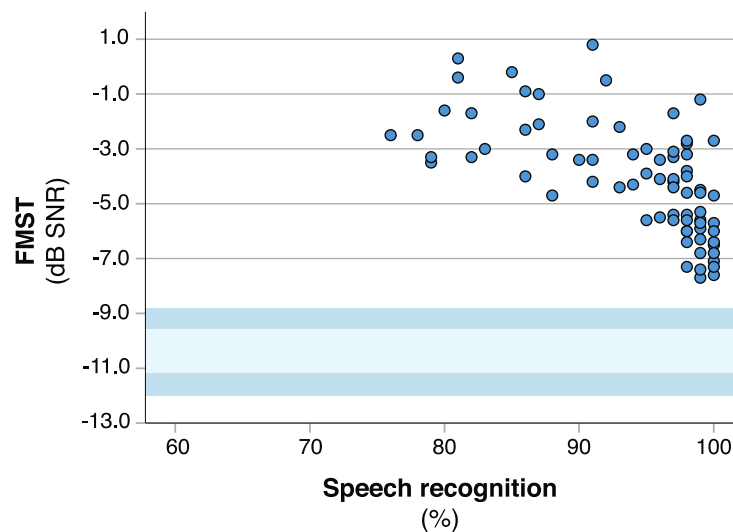

Figure 4. Speech recognition scores at +10 dB plotted against the speech recognition threshold estimates obtained with the adaptive test procedure for the Finnish matrix sentence test (FMST). The light blue represents the normative range ±1 standard deviation (SD) for the FMST, and the dark blue areas denote the ±2 SD range.

4. Discussion

Our aims were to validate the FINSIMAT and the FDTT in CI users by comparing the results to the FMST, and to collect reference data for the two tests in an unselected and representative CI patient population. Despite differences in speech material and difficulty level, we observed concise and correlating results in our heterogenous patient population for all three of the Finnish SIN tests.

The mean SRT estimates for all three tests in our CI patient population showed a similar pattern to the normative reference values (see Table 1). For NH listeners, the reference value for the FMST is the highest, whereas the reference value for FINSIMAT is the lowest (i.e.,

most negative), and the reference value for the FDTT is situated in-between. Figure 3b demonstrates the differences between the three SIN tests by showing how the mean SRT estimates for the FMST were always the highest, and the mean SRT estimates for the FINSIMAT were the lowest, despite the differences in the overall performance level between unilateral, bilateral, and bimodal CI users.

The differences in the SRT estimates between the three tests are likely due to the differences in speech materials and test procedures, since all of the tests used the same female speaker. The FMST is the most complex of these tests due to longer sentences and relatively long test-lists (20 sentences per test-list). The test is precise and suited for an accurate diagnostic assessment in patients with hearing loss, since the test-retest value was within 1 dB for CI users [7,39]. The results we obtained with the FMST are similar to previously published results for CI users. In the validation study for the FMST [39], the mean SRT estimate for the participants was −3.5 dB SNR (SD = 1.7). Our participants reached a slightly better mean SRT estimate of −4.2 dB SNR (SD = 2.1). However, during the 2015 validation study [39], all of the participants were tested in the unilateral CI condition, whereas in the present study all of the participants were measured with their preferred device configuration, and thus only 55% of our participants were measured in the unilateral CI condition. The mean SRT estimate for the FMST for our unilateral users (−3.6 dB SNR (SD = 1.9)) did not differ from the original validation data. Sivonen et al. [47] reported a slightly better mean SRT estimate of −4.6 dB SNR for a selected group of unilateral CI users. For a group of CI users with good functional low-frequency hearing preoperatively, Iso-Mustajärvi et al. [48] reported a mean post-operative SRT estimate of −5.2 dB SNR for the FMST, reflecting the significance of residual hearing and hearing preservation surgery in the CI performance.

The FINSIMAT has the lowest SRT estimate, both in NH listeners as well as in our patient population (see Tables 1 and 2). The speech material for the FINSIMAT is simple and restricted. The short pseudo-sentences are easier to retain in memory during the listening task than the longer FMST sentences. These qualities likely contribute to the lower reference value for NH listeners, and also to the lower SRT estimate in our study population. The shortness and simplicity of the test is essential for testing children, since their speech perception in noise is still maturing [19,49], and children with hearing loss may have a shorter auditory memory span [50,51]. Likewise, these qualities might render the test the most reliable SIN measure in the elderly, since ageing and age-related cognitive decline affect the speech perception in noise independently of the hearing level [52–54]. In the clinical protocol for the FMST, two practice lists are conducted before the actual test measurements, and sometimes after the practice, some elderly patients are too tired to focus on the listening task during the actual test measurements, which may affect the reliability of the results. Therefore, we are currently assessing whether the shorter FINSIMAT would provide more reliable results in this patient population [40].

Plenty of data have been published on the use of the standard matrix sentence tests in CI users [48,55–57], but the present study is, to our knowledge, the first on the simplified matrix sentence tests in adult CI users. The results for CI users are similar to the results from hearing impaired adults, who also showed slightly better SRT estimates for the FINSIMAT than for the FMST [40]. In that study, the FINSIMAT was presented monoaurally with headphones, and the mean SRT estimate for the unilateral test condition was −4.1 dB SNR (SD = 2.0). The result is slightly worse than the mean SRT estimate for our participants (−5.0 dB SNR (SD = 1.9)).

The FDTT was originally developed for hearing screening [15], but it has been increasingly adopted into hearing diagnostics in recent years [23,24]. The DTTs are especially well suited for remote testing since responses can easily be recorded using a picture of a number pad on a touch-screen device. Even though the speech material for DTTs is extremely simple and restricted, previous studies have reported good correlation between DTTs and other SIN tests in CI users [23,58]. To our knowledge, ours are the first published data to compare a DTT to a matrix sentence test in CI users.

Even though all of our tests were designed for different purposes and patient populations, and despite producing SRT estimates that are not directly comparable, they were all designed to measure the same thing: the ability to perceive speech in noise. All of the test results should therefore correlate with each other, and we did observe strong and statistically significant correlations between all of the tests. Since a previous validation study has already confirmed that the FMST is well-suited to assessing speech perception in CI users [39], the FMST was used as a reference point for data from the other tests. For the FINSIMAT, the correlation with the FMST was 0.82 ($p < 0.001$), which is similar to previously published results from elderly listeners with mild-to-moderate hearing impairment (r = 0.89, $p < 0.001$) [40]. For the FDTT, the correlation with the FMST was equally good (r = 0.78, $p < 0.001$), even though the speech materials differ significantly in length and complexity. A previous study reported similar correlation (r = 0.81) between the French DTT and the French matrix sentence test in hearing-impaired adults [30].

Since all of our participants had at least four months of CI experience, it was not surprising that CI experience did not correlate with the SIN perception. What was striking, however, was the lack of correlation between the patients' age and SIN perception (see Figure 2). The current policy in Finland is to provide two CI for patients under the age of 65 if the benefit from bimodal fitting is insufficient. With some exceptions (i.e., extremely poor vision), patients over the age of 65 typically only receive one CI. This is reflected in our data, as the proportion of bilateral CI users in participants under the age 65 (33%) was twice of that in older participants (16%). Despite this, the differences in the SRT estimates between the two age-groups were small and statistically insignificant. Multiple studies have reported excellent outcomes in elderly CI users [5,59], and our data corroborate these findings.

Ceiling effects are a well-known problem with speech perception tests in quiet and SIN tests that used fixed presentation levels [6], and we also observed a ceiling effect for the FMST presented at +10 dB SNR. For tests that use a fixed presentation level, a more difficult SNR can be used to prevent ceiling effects, but, on the other hand, this predisposes patients with poor speech perception to floor effects [32]. Adaptive test procedures can more easily accommodate a large performance range, as our data demonstrates. None of our participants reached the reference range for NH listeners (Figure 4), and only three participants had to be excluded for having inadequate speech perception to conduct the SIN tests with the adaptive test procedure. With the adaptive test procedure, the FMST was also more accurate in detecting individual differences between the participants; Figure 4 demonstrates how, despite reaching perfect or near perfect speech perception at +10 dB SNR, the speech perception significantly deteriorated at more adverse SNRs, leaving some patients performing very poorly (SRT of −1.2 dB SNR), while others reached nearly the performance range of NH listeners (SRT of −7.7 dB SNR).

5. Limitations

Studies on SIN tests often include a speech intelligibility test in quiet with short words (CNC words, for example) for comparison. We were not able to include any such speech intelligibility measures, since the only available word-level speech intelligibility test in Finnish is outdated, and the test lists differ in intelligibility [39]. Additionally, since the FMST is the only sentence-level SIN test in Finnish, we were not able to include a SIN test with everyday sentences for comparison. An interesting addition to the current data set would have been a subjective evaluation of hearing and hearing related quality of life to assess how well the test results correlate with CI users' subjective evaluation of their functional hearing.

Since the main aim of our study was not to compare different device modalities, but to assess how well the three SIN test results correlate in different test conditions, we allowed our participants to use their preferred device configuration without assessing objectively, whether it was their best aided condition. All of our bimodal users used their HAs regularly, and they all reported significant subjective benefit from their HAs, which were fitted to the

prescriptive target recommended by the device specific fitting software. The HA fitting was confirmed with real-ear measurements as required. However, no free field thresholds were measured for the HAs in order to quantify the bimodal benefit. Therefore, if assessed for the best aided condition, some of our bimodal users might have been categorized as unilateral users, which could affect the comparative analysis between listening modalities.

6. Conclusions

Our data confirm that it is possible to create a comprehensive SIN test battery that includes different types of SIN tests, and still produces comparable results in a diverse patient population. The data reported here provide references for the FMST, FINSIMAT and FDTT in CI users, and are the first data for simplified matrix tests in adult CI users. Similar CI user reference data for other SIN tests would provide valuable information for clinicians assessing the rehabilitation results or CI candidacy, and for the scientific community at large in comparing results from different centers and studies.

Author Contributions: Conceptualization, T.W., A.D. and V.S.; methodology, all authors; data curation, all authors; formal analysis, T.W.; resources, A.D. and V.S.; writing—original draft preparation, T.W.; writing—review and editing, all authors; visualization, T.W.; supervision, A.D.; project administration, A.D. and V.S.; funding acquisition, A.D., T.W. All authors have read and agreed to the published version of the manuscript.

Institutional Review Board Statement: The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board and Ethics Com-mittee of Northern Savo Hospital District (Kuopio University Hospital) (protocol code 5551822, 204/2015).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Boisvert, I.; Reis, M.; Au, A.; Cowan, R.; Dowell, R.C. Cochlear Implantation Outcomes in Adults: A Scoping Review. *PLoS ONE* **2020**, *15*, e0232421. [CrossRef] [PubMed]
2. Buchman, C.A.; Herzog, J.A.; McJunkin, J.L.; Wick, C.C.; Durakovic, N.; Firszt, J.B.; Kallogjeri, D. CI532 Study Group Assessment of Speech Understanding After Cochlear Implantation in Adult Hearing Aid Users: A Nonrandomized Controlled Trial. *JAMA Otolaryngol. Neck Surg.* **2020**, *146*, 916. [CrossRef] [PubMed]
3. Dunn, C.; Miller, S.E.; Schafer, E.C.; Silva, C.; Gifford, R.H.; Grisel, J.J. Benefits of a Hearing Registry: Cochlear Implant Candidacy in Quiet Versus Noise in 1,611 Patients. *Am. J. Audiol.* **2020**, *29*, 851–861. [CrossRef] [PubMed]
4. Layfield, E.; Hwa, T.P.; Naples, J.; Maina, I.; Brant, J.A.; Eliades, S.J.; Bigelow, D.C.; Ruckenstein, M.J. Failure and Revision Surgery after Cochlear Implantation in the Adult Population: A 10-Year Single-Institution Retrospective and Systematic Re-view of the Literature. *Otol. Neurotol.* **2021**, *42*, 408–413. [CrossRef]
5. Lally, J.W.; Adams, J.K.; Wilkerson, B.J. The Use of Cochlear Implantation in the Elderly. *Curr. Opin. Otolaryngol. Head Neck Surg.* **2019**, *27*, 387–391. [CrossRef]
6. Gifford, R.H.; Shallop, J.K.; Peterson, A.M. Speech Recognition Materials and Ceiling Effects: Considerations for Cochlear Implant Programs. *Audiol. Neurotol.* **2008**, *13*, 193–205. [CrossRef]
7. Hey, M.; Hocke, T.; Hedderich, J.; Müller-Deile, J. Investigation of a Matrix Sentence Test in Noise: Reproducibility and Discrimination Function in Cochlear Implant Patients. *Int. J. Audiol.* **2014**, *53*, 895–902. [CrossRef]
8. Zwolan, T.A.; Kallogjeri, D.; Firszt, J.B.; Buchman, C.A. Assessment of Cochlear Implants for Adult Medicare Beneficiaries Aged 65 Years or Older Who Meet Expanded Indications of Open-Set Sentence Recognition: A Multicenter Nonrandomized Clinical Trial. *JAMA Otolaryngol. Neck Surg.* **2020**, *146*, 933. [CrossRef]
9. Dorman, M.F.; Gifford, R.H. Speech Understanding in Complex Listening Environments by Listeners Fit With Cochlear Implants. *J. Speech Lang. Hear. Res.* **2017**, *60*, 3019–3026. [CrossRef]
10. Minimum Speech Test Battery. 2011. Available online: www.auditorypotential.com/MSTBfiles/MSTBManual2011-06-20%20.pdf (accessed on 30 June 2021).

11. British Society of Audiology. Practice Guidance. *Assessment of Speech Understanding in Noise in Adults with Hearing Difficulties.* 2019. Available online: www.thebsa.org.uk/wp-content/uploads/2019/03/BSA-Practice-Guidance-Speech-in-Noise-FINAL. Feb-2019.pdf (accessed on 30 June 2021).

12. Adunka, O.F.; Gantz, B.J.; Dunn, C.; Gurgel, R.K.; Buchman, C.A. Minimum Reporting Standards for Adult Cochlear Implantation. *Otolaryngol. Neck Surg.* **2018**, *159*, 215–219. [CrossRef]

13. Theunissen, M.; Swanepoel, D.W.; Hanekom, J. Sentence Recognition in Noise: Variables in Compilation and Interpretation of Tests. *Int. J. Audiol.* **2009**, *48*, 743–757. [CrossRef]

14. Plomp, R.; Mimpen, A.M. Improving the Reliability of Testing the Speech Reception Threshold for Sentences. *Audiology* **1979**, *18*, 43–52. [CrossRef] [PubMed]

15. Smits, C.; Kapteyn, T.S.; Houtgast, T. Development and Validation of an Automatic Speech-in-Noise Screening Test by Telephone. *Int. J. Audiol.* **2004**, *43*, 15–28. [CrossRef] [PubMed]

16. 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. Development and Validation of the AzBio Sentence Lists. *Ear Hear.* **2012**, *33*, 112–117. [CrossRef]

17. Kollmeier, B.; Warzybok, A.; Hochmuth, S.; Zokoll, M.A.; Uslar, V.; Brand, T.; Wagener, K.C. The Multilingual Matrix Test: Principles, Applications, and Comparison across Languages: A Review. *Int. J. Audiol.* **2015**, *54*, 3–16. [CrossRef]

18. Spahr, A.J.; Dorman, M.F.; Litvak, L.M.; Cook, S.J.; Loiselle, L.M.; DeJong, M.D.; Hedley-Williams, A.; Sunderhaus, L.S.; Hayes, C.A.; Gifford, R.H. Development and Validation of the Pediatric AzBio Sentence Lists. *Ear Hear.* **2014**, *35*, 418–422. [CrossRef] [PubMed]

19. Holder, J.T.; Sheffield, S.W.; Gifford, R.H. Speech Understanding in Children with Normal Hearing: Sound Field Normative Data for BabyBio, BKB-SIN, and QuickSIN. *Otol. Neurotol.* **2016**, *37*, e50–e55. [CrossRef] [PubMed]

20. Coco, L. Teleaudiology: Strategies, Considerations during a Crisis and Beyond. *Hear. J.* **2020**, *73*, 26, 28, 29. [CrossRef]

21. Bennett, R.; Eikelboom, R.; Swanepoel, D.W.; Manchaiah, V. Tele-audiology services in Australia: A shift in clinical practices. *Audiol. Now* **2020**, *81*, 11–13.

22. Potgieter, J.-M.; Swanepoel, D.W.; Myburgh, H.C.; Hopper, T.C.; Smits, C. Development and Validation of a Smartphone-Based Digits-in-Noise Hearing Test in South African English. *Int. J. Audiol.* **2016**, *55*, 405–411. [CrossRef]

23. Cullington, H.E.; Aidi, T. Is the Digit Triplet Test an Effective and Acceptable Way to Assess Speech Recognition in Adults Using Cochlear Implants in a Home Environment? *Cochlear Implant. Int.* **2017**, *18*, 97–105. [CrossRef]

24. de Graaff, F.; Huysmans, E.; Merkus, P.; Theo Goverts, S.; Smits, C. Assessment of Speech Recognition Abilities in Quiet and in Noise: A Comparison between Self-Administered Home Testing and Testing in the Clinic for Adult Cochlear Implant Users. *Int. J. Audiol.* **2018**, *57*, 872–880. [CrossRef]

25. Nilsson, M.; Soli, S.D.; Sullivan, J.A. Development of the Hearing in Noise Test for the Measurement of Speech Reception Thresholds in Quiet and in Noise. *J. Acoust. Soc. Am.* **1994**, *95*, 1085–1099. [CrossRef]

26. Dietz, A.; Buschermöhle, M.; Aarnisalo, A.A.; Vanhanen, A.; Hyyrynen, T.; Aaltonen, O.; Löppönen, H.; Zokoll, M.A.; Kollmeier, B. The Development and Evaluation of the Finnish Matrix Sentence Test for Speech Intelligibility Assessment. *Acta Otolaryngol.* **2014**, *134*, 728–737. [CrossRef]

27. Gilbert, J.L.; Tamati, T.N.; Pisoni, D.B. Development, Reliability, and Validity of PRESTO: A New High-Variability Sentence Recognition Test. *J. Am. Acad. Audiol.* **2013**, *24*, 026–036. [CrossRef]

28. Smits, C.; Theo Goverts, S.; Festen, J.M. The Digits-in-Noise Test: Assessing Auditory Speech Recognition Abilities in Noise. *J. Acoust. Soc. Am.* **2013**, *133*, 1693–1706. [CrossRef] [PubMed]

29. Wilson, R.H.; McArdle, R.A.; Smith, S.L. An Evaluation of the BKB-SIN, HINT, QuickSIN, and WIN Materials on Listeners with Normal Hearing and Listeners with Hearing Loss. *J. Speech Lang. Hear. Res.* **2007**, *50*, 844–856. [CrossRef]

30. Jansen, S.; Luts, H.; Wagener, K.C.; Kollmeier, B.; Del Rio, M.; Dauman, R.; James, C.; Fraysse, B.; Vormès, E.; Frachet, B.; et al. Comparison of Three Types of French Speech-in-Noise Tests: A Multi-Center Study. *Int. J. Audiol.* **2012**, *51*, 164–173. [CrossRef] [PubMed]

31. Dillon, M.T.; Buss, E.; King, E.R.; Deres, E.J.; Obarowski, S.N.; Anderson, M.L.; Adunka, M.C. Comparison of Two Cochlear Implant Coding Strategies on Speech Perception. *Cochlear Implant. Int.* **2016**, *17*, 263–270. [CrossRef]

32. Kelsall, D.; Lupo, J.; Biever, A. Longitudinal Outcomes of Cochlear Implantation and Bimodal Hearing in a Large Group of Adults: A Multicenter Clinical Study. *Am. J. Otolaryngol.* **2021**, *42*, 102773. [CrossRef]

33. Prentiss, S.; Snapp, H.; Zwolan, T. Audiology Practices in the Preoperative Evaluation and Management of Adult Cochlear Implant Candidates. *JAMA Otolaryngol. Neck Surg.* **2020**, *146*, 136. [CrossRef]

34. Holder, J.T.; Levin, L.M.; Gifford, R.H. Speech Recognition in Noise for Adults with Normal Hearing: Age-Normative Performance for AzBio, BKB-SIN, and QuickSIN. *Otol. Neurotol.* **2018**, *39*, e972–e978. [CrossRef]

35. Roland, J.T.; Gantz, B.J.; Waltzman, S.B.; Parkinson, A.J. Long-term Outcomes of Cochlear Implantation in Patients with High-frequency Hearing Loss. *Laryngoscope* **2018**, *128*, 1939–1945. [CrossRef]

36. Hey, M.; Neben, N.; Stöver, T.; Baumann, U.; Mewes, A.; Liebscher, T.; Schüssler, M.; Aschendorff, A.; Wesarg, T.; Büchner, A.; et al. Outcomes for a Clinically Representative Cohort of Hearing-Impaired Adults Using the Nucleus® CI532 Cochlear Implant. *Eur. Arch. Otorhinolaryngol.* **2020**, *277*, 1625–1635. [CrossRef] [PubMed]

37. Dornhoffer, J.R.; Reddy, P.; Meyer, T.A.; Schvartz-Leyzac, K.C.; Dubno, J.R.; McRackan, T.R. Individual Differences in Speech Recognition Changes After Cochlear Implantation. *JAMA Otolaryngol. Neck Surg.* **2021**, *147*, 280. [CrossRef]

38.   Biever, A.; Amurao, C.; Mears, M. Considerations for a Revised Adult Cochlear Implant Candidacy Evaluation Protocol. *Otol. Neurotol.* **2021**, *42*, 159–164. [CrossRef]

39.   Dietz, A.; Buschermöhle, M.; Sivonen, V.; Willberg, T.; Aarnisalo, A.A.; Lenarz, T.; Kollmeier, B. Characteristics and International Comparability of the Finnish Matrix Sentence Test in Cochlear Implant Recipients. *Int. J. Audiol.* **2015**, *54*, 80–87. [CrossRef] [PubMed]

40.   Willberg, T.; Kärtevä, K.; Zokoll, M.; Buschermöhle, M.; Sivonen, V.; Aarnisalo, A.; Löppönen, H.; Kollmeier, B.; Dietz, A. The Finnish Simplified Matrix Sentence Test for the Assessment of Speech Intelligibility in the Elderly. *Int. J. Audiol.* **2020**, *59*, 763–771. [CrossRef] [PubMed]

41.   Willberg, T.; Buschermöhle, M.; Sivonen, V.; Aarnisalo, A.A.; Löppönen, H.; Kollmeier, B.; Dietz, A. The Development and Evaluation of the Finnish Digit Triplet Test. *Acta Otolaryngol.* **2016**, *136*, 1035–1040. [CrossRef] [PubMed]

42.   Wagener, K.C.; Brand, T.; Kollmeier, B. Entwicklung und Evaluation eines Satztests für die deutsche Sprache Teil III: Evaluation des Oldenburger Satztests. (Development and evaluation of a German speech intelligibility test. Part III: Evaluation of the Oldenburg sentence test.). *Z. Audiol.* **1999**, *38*, 86–95.

43.   Wagener, K.C.; Brand, T.; Kollmeier, B. Entwicklung und Evaluation eines Satztests in deutscher Sprache Teil II: Optimierung des Oldenburger Satztests (Development and evaluation of a German sentence test Part II: Optimization of the Oldenburg sentence test.). *Z. Audiol.* **1999**, *38*, 44–56.

44.   Wagener, K.C.; Kühnel, V.; Kollmeier, B. Entwicklung und Evaluation eines Satztests in deutscher Sprache I: Design des Oldenburger Satztests. (Development and evaluation of a German sentence test I: Design of the Oldenburg sentence test.). *Z. Audiol.* **1999**, *38*, 4–15.

45.   Wagener, K.C.; Kollmeier, B. Evaluation des Oldenburger Satztests mit Kindern und Oldenburger Kinder-Satztest. (Evaluation of the Oldenburg sentence test with children and the Oldenburg children's sentence test). *Z. Audiol.* **2005**, *44*, 134–143.

46.   Brand, T.; Kollmeier, B. Efficient Adaptive Procedures for Threshold and Concurrent Slope Estimates for Psychophysics and Speech Intelligibility Tests. *J. Acoust. Soc. Am.* **2002**, *111*, 2801–2810. [CrossRef]

47.   Sivonen, V.; Willberg, T.; Aarnisalo, A.A.; Dietz, A. The Efficacy of Microphone Directionality in Improving Speech Recognition in Noise for Three Commercial Cochlear-Implant Systems. *Cochlear Implant. Int.* **2020**, *21*, 153–159. [CrossRef] [PubMed]

48.   Iso-Mustajärvi, M.; Sipari, S.; Löppönen, H.; Dietz, A. Preservation of Residual Hearing after Cochlear Implant Surgery with Slim Modiolar Electrode. *Eur. Arch. Otorhinolaryngol.* **2020**, *277*, 367–375. [CrossRef]

49.   Buss, E.; Hodge, S.E.; Calandruccio, L.; Leibold, L.J.; Grose, J.H. Masked Sentence Recognition in Children, Young Adults, and Older Adults: Age-Dependent Effects of Semantic Context and Masker Type. *Ear Hear.* **2019**, *40*, 1117–1126. [CrossRef] [PubMed]

50.   AuBuchon, A.M.; Pisoni, D.B.; Kronenberger, W.G. Evaluating Pediatric Cochlear Implant Users' Encoding, Storage, and Retrieval Strategies in Verbal Working Memory. *J. Speech Lang. Hear. Res.* **2019**, *62*, 1016–1032. [CrossRef] [PubMed]

51.   Chandramouli, S.H.; Kronenberger, W.G.; Pisoni, D.B. Verbal Learning and Memory in Early-Implanted, Prelingually Deaf Adolescent and Adult Cochlear Implant Users. *J. Speech Lang. Hear. Res.* **2019**, *62*, 1033–1050. [CrossRef] [PubMed]

52.   Pichora-Fuller, M.K.; Schneider, B.A.; Daneman, M. How Young and Old Adults Listen to and Remember Speech in Noise. *J. Acoust. Soc. Am.* **1995**, *97*, 593–608. [CrossRef] [PubMed]

53.   Dubno, J.R.; Dirks, D.D.; Morgan, D.E. Effects of Age and Mild Hearing Loss on Speech Recognition in Noise. *J. Acoust. Soc. Am.* **1984**, *76*, 87–96. [CrossRef] [PubMed]

54.   Goossens, T.; Vercammen, C.; Wouters, J.; van Wieringen, A. Masked Speech Perception across the Adult Lifespan: Impact of Age and Hearing Impairment. *Hear. Res.* **2017**, *344*, 109–124. [CrossRef]

55.   Rader, T.; Fastl, H.; Baumann, U. Speech Perception with Combined Electric-Acoustic Stimulation and Bilateral Cochlear Implants in a Multisource Noise Field. *Ear Hear.* **2013**, *34*, 324–332. [CrossRef]

56.   Stronks, H.C.; Apperloo, E.; Koning, R.; Briaire, J.J.; Frijns, J.H.M. SoftVoice Improves Speech Recognition and Reduces Listening Effort in Cochlear Implant Users. *Ear Hear.* **2021**, *42*, 381–392. [CrossRef] [PubMed]

57.   Grossmann, W.; Brill, S.; Moeltner, A.; Mlynski, R.; Hagen, R.; Radeloff, A. Cochlear Implantation Improves Spatial Release From Masking and Restores Localization Abilities in Single-Sided Deaf Patients. *Otol. Neurotol.* **2016**, *37*, 658–664. [CrossRef] [PubMed]

58.   Kaandorp, M.W.; Smits, C.; Merkus, P.; Goverts, S.T.; Festen, J.M. Assessing Speech Recognition Abilities with Digits in Noise in Cochlear Implant and Hearing Aid Users. *Int. J. Audiol.* **2015**, *54*, 48–57. [CrossRef]

59.   Rohloff, K.; Koopmann, M.; Weiß, D.; Rudack, C.; Savvas, E. Cochlear Implantation in the Elderly: Does Age Matter? *Otol. Neurotol.* **2017**, *38*, 54–59. [CrossRef] [PubMed]