Audiology and Neurotology

Research Article

Audiol Neurotol 2026;31:1–11
DOI: 10.1159/000547205

Received: October 9, 2024
Accepted: June 25, 2025
Published online: August 25, 2025

Learning Effects of the Dutch/Flemish Matrix Test for Bimodal Cochlear Implant Users

Nienke Cornelia Langerak[a] Hendrik Christiaan Stronks[a, b]
Jeroen Johannes Briaire[a] Johan Hubertus Maria Frijns[a, b, c]

[a]Department of Otorhinolaryngology and Head & Neck Surgery, Leiden University Medical Center, Leiden, The Netherlands; [b]Leiden Institute for Brain and Cognition, Leiden University, Leiden, The Netherlands; [c]Department of Bioelectronics, Delft University of Technology, Delft, The Netherlands

Keywords
Cochlear implants · Sensorineural hearing loss · Learning effects · Speech in noise · Speech intelligibility

Abstract
Introduction: Cochlear implantation (CI) is the standard treatment for severe-to-profound sensorineural hearing loss, but CI users often struggle with speech understanding in noisy environments. The Dutch/Flemish Matrix test is frequently used to evaluate speech-in-noise performance due to its assumed immunity to learning effects. However, studies challenge this assumption, revealing significant learning effects that can confound research outcomes. In this study, we modeled the learning curves of the Dutch/Flemish Matrix test to assess the influence of both between-session and between-test effects. We hypothesized that a exponential model would describe the learning effects more accurately than a linear model. ***Methods:*** The perceptual learning effects associated with the Dutch/Flemish Matrix test were assessed in 17 bimodal CI users. All participants performed the Matrix speech-in-noise tests across four sessions, with 13 randomized tests per session. The tests were conducted in a soundproof booth with an eight-speaker babble noise. The outcome parameter was the speech recognition threshold and was analyzed with a linear mixed model to account for confounders. ***Results:*** The results showed a statistically significant learning effect between sessions that added up to a speech intelligibility increase of 1.3 dB signal-to-noise ratio (SNR) (equivalent to ~10% word score) between the first and second sessions, 0.86 dB SNR (~7%) between the second and third sessions and 0.67 dB SNR (~5%) between the third and fourth sessions. In addition, a statistically significant within-session learning effect (i.e., between tests) was observed with a linear slope of −0.11 dB SNR/test (~0.9% word score/test), which accumulates to a total of 1.7 dB SNR (13%) between session start and end. The between-session learning curve was described more accurately with an exponential fit than with a linear fit. The between-test learning curve can be described equally well with a linear and an exponential fit. ***Conclusion:*** A robust between-test learning effect was observed, which could be accurately modeled using either a linear or exponential learning curve. Additionally, a between-session learning effect was evident and was best described by an exponential learning curve. This study provides an important handle for correcting these learning effects in future studies.

karger@karger.com
www.karger.com/aud

Karger
OPEN ACCESS

Correspondence to:
Hendrik Christiaan Stronks, h.c.stronks@lumc.nl

Plain Language Summary

This study investigated how people with cochlear implants improve when taking a common test that measures their ability to understand speech in noisy environments, called the Dutch/Flemish Matrix test. The test is often assumed to not be affected by learning, but other studies show that people do get better at it with practice, which can affect the results of studies. We wanted to investigate the between-session and between-test learning in more detail. A total of 17 participants with cochlear implants were included and performed a speech-in-noise test over four sessions, with 13 tests in each session. To reduce learning effects, two recommended practice tests were performed at the start of each session. The researchers found that participants showed significant improvement between sessions and between tests. The between-session improvement was better explained using an exponential model (which shows rapid initial learning that slows down over time) rather than a linear model (which would suggest steady, constant improvement). The between-test improvement could be described equally well with a linear or exponential fit. The findings suggest that learning effects need to be taken into account when using this test in studies. This could be useful when randomization of test conditions is not possible.

Introduction

The standard treatment nowadays for people with severe to profound sensorineural hearing loss is cochlear implantation (CI). A CI bypasses the damaged hair cells and directly stimulates the auditory nerve cells electrically. In general, CI users understand speech in quiet situations well, but communication in noise remains challenging [1]. Therefore, much ongoing research is directed at improving listening in noise for CI users. When new front-end processing algorithms or novel speech coding strategies are tested in this regard, a test is needed to validate these adaptations. One of the regularly used speech-in-noise tests in the Netherlands is the Matrix test.

The Matrix test is developed in several languages and is widely used in audiology and hearing research [2]. One of the main reasons for its extensive use is the assumption that the Matrix test is free of learning effects because the speech material contains a closed set of a limited number of 50 words. It is assumed that after the two recommended practice lists, there are minimal learning effects [2, 3]. Open-set sentence tests are prone to learning effects, and it is common practice to wait 1 year before reusing the same lists [4]. For Dutch CI users, not many sentence-in-noise tests are available, and hence, the Matrix test is an attractive alternative because it can be used repetitively and indefinitely.

Learning effects can be divided into procedural and perceptual learning effects, where procedural learning is defined as the process of getting familiar with the procedure, and perceptual learning is defined as the phenomenon of getting better at the task because of experience [5]. In this study, the task is speech understanding in noise. The procedural learning effect can be minimized by performing practice tests as described in [2, 3, 6]. However, perceptual learning effects cannot be easily reduced and are, therefore, important to take into account [7]. Learning effects can be depicted in a so-called learning curve that takes the shape of an exponential function [8] or of a standard psychometric function such as the Weibull equation [9, 10]. Regardless of their exact shape, all learning curves are characterized by an initial steep increase in performance in the beginning that tapers off to asymptotic values later.

The assumption that the Matrix test is free of learning effects has been challenged by multiple studies, which consistently showed significant learning effects across consecutive test sessions (here referred to as "between-session learning") as well as learning effects across consecutive lists of the Matrix test ("between-test learning") [3, 11–15]. Learning effects are an important confounder in many psychophysical studies because the magnitude of this learning effect can easily exceed the effect under study [16, 17]. As a consequence, there is a need for a reliable description of the learning effects associated with the Matrix test, both between sessions and between tests. This can greatly aid in interpreting study data, where randomization of conditions is not possible, such as study protocols where cohorts of study participants are followed that are consecutively exposed to a series of treatments in a fixed sequence [6]. Unfortunately, many studies investigating perceptual learning curves associated with the Matrix test have done so by only taking selected sessions or tests into account without building a model (e.g., Schlueter et al. [13]), modeling the data with linear fits without an examination of the underlying data [18] or by disregarding the consensus [2, 3] that at least two practice tests need to be included when performing Matrix testing in each session to familiarize the participant with the test [3, 14]. Characterizing the learning effect over a selected number of tests and sessions restricts the usefulness of the results for use in other studies with different numbers of tests or sessions. Assuming a linear trend simplifies the learning effect, which is by definition nonlinear [9] and underestimates initial learning effects while overestimating them later on. The omission of training tests can result in substantially increased effects of learning early in the session because of procedural learning effects adding to the perceptual learning process [19, 20]. In this study, we characterized learning effects associated with the Dutch/Flemish Matrix test in a group of bimodal listeners. Bimodal listeners are unilateral CI users with a hearing aid (HA) fitted in the other ear. Because implantation guidelines for CI are progressively becoming less strict [21, 22], the population of bimodal listeners is growing, and especially so in countries where bilateral implantation is not the standard of care. We used a data set collected in a study where over 40 Matrix tests were collected per participant, divided over four sessions. As a result, this data set provided an excellent opportunity to investigate the learning effect of the Matrix test both between sessions and between tests. The goal of this study was to provide a detailed description of the underlying function of the learning curves between sessions and between tests. We hypothesized that an exponential fit would describe the data more accurately than a linear fit. Based on the existing evidence, we expected a prominent and significant learning effect between sessions, and especially so between the first and the second session [3, 13, 18]. Our expectation of the learning effect between tests was that they were more subtle than the learning effect between visits [18].

Materials and Methods

Participants

Seventeen bimodal CI users, implanted with a CI on one side and wearing a HA on the other, were included in this crossover observational study. All participants were older than 18 years of age, had experience with their CI and their HA for longer than 6 months, and were native Dutch speakers. Some of the participants had participated in previous studies with the Dutch/Flemish matrix test, but this was at least 6 months before the start of this study. Demographics of the study participants can be found in Table 1. All study participants provided written informed consent. This study was approved by the Medical Ethics Committee Leiden, The Hague, Delft (METC number P20.018), and it was included in the clinical trial register of the Dutch Central Committee on Research Involving Human Subjects (https://www.onderzoekmetmensen.nl/nl/trial/49290).

Study Design

Learning effects were investigated across four sessions, each comprising 13 tests. There were six different microphone settings and three origins of speech (explained under "listening conditions" below), resulting in 18 conditions. The test and retest of the conditions were randomized over the four sessions for all participants. A blocked randomization design was used, and test and retest of a condition were never in the same session. The time between sessions differed per participant, but was at least 1 week (median = 14 days, IQR = 20 days). At the beginning of each session, two practice tests of 20 sentences with the Dutch/Flemish matrix test were performed, as recommended [2, 3] and, subsequently, the 13 tests were completed. The two practice tests were excluded from the analysis of learning effects. One break of at least 15 min was held in the middle of the session (typically after test number 6). More breaks were introduced when needed.

Fitting of CI and HA

During a session, the participant was fitted with a Harmony™ speech processor (Advanced Bionics LLC, Valencia, CA, USA) with their personal clinical program. Contralaterally, a hearing aid (Phonak Audéo M90-312, Stäfa, Switzerland) was fitted with the adaptive Phonak Digital Bimodal fitting formula (Advanced Bionics, LLC, Valencia, CA, USA) [23], based on the in situ audiogram and feedback test in Phonak Target (version 6.2.5; Sonova, Stäfa, Switzerland). The type of HA and its fitting were potentially different from their own device. Any other front-end processing algorithms were turned off (e.g., SoundRelax™, NoiseBlock™, and WindBlock™).

Speech Test

The Matrix speech corpus comprises 13 lists, each containing 20 short sentences of 5 words. Each sentence has a fixed syntax, namely, a name followed by a verb, number, color, and an object drawn from a closed set of 50 words [20]. The corpus is voiced by a female Flemish speaker. Some lists had to be repeated within a session (15 lists were needed, including the two practice tests), but care was taken that lists were not used more than two times, and were not directly used after each other. Sentences were presented using an MSP5A monitor speaker (Yamaha Corp., Japan) that was situated approximately 1 m away from the participant at a height of 1.2 m. The average signal-to-noise ratio (SNR) of the last eight sentences was used as

Table 1. Demographics of the study participants

Subject	Sex	Age	CI use, years	Duration of deafness, years	$PTA_{125\text{-}500}$, dB	Etiology	Clinical CVC score, %
S02	F	67	7	[a]	65	DFNA9	96
S03	F	69	5	24	35	Unknown progressive	85
S04	F	65	6	[b]	22	Unknown progressive	84
S05	M	72	7	10	37	Unknown progressive	86
S06	M	54	6	14	43	Unknown progressive	93
S07	F	66	11	[a]	63	Usher syndrome	93
S08	F	71	7	[a]	53	DFNA22	86
S09	M	58	9	[b]	42	Usher syndrome	98
S10	M	77	1	5	43	Unknown progressive	89
S11	M	70	1	[a]	53	DFNA9	91
S12	F	62	1	7	30	Unknown	86
S13	M	80	1	34	48	Unknown	96
S14	M	76	2	13	52	Unknown progressive	87
S15	M	63	1	[b]	45	DFNA9	95
S16	F	65	2	7	53	Waardenburg syndrome	98
S17	M	67	1	7	65	DFNA9	90
S18	M	72	0.5	18	55	Unknown progressive	92
		Mean	Mean		Median		Median
		68	4		48		91

PTA, pure tone average of 125 Hz, 250 Hz, and 500 Hz of non-implanted ear; CVC, phoneme score on the consonant-vowel-consonant test. [a]Unknown. [b]Still able to make phone calls at the time the CI was implanted.

an estimate for the speech recognition threshold (SRT). The adaptivity of the test was based on the algorithm used in the Oldenburg Sentence Test [24] using a MATLAB programming environment (R2017b; MathWorks, Inc., Natick, MA, USA). Participants were allowed to guess, and no feedback was provided.

The background noise was generated in a homogeneous noise field of 60 dBA using eight loudspeakers that were equally distributed in two planes, below and above the listener. Each of these was individually calibrated with a sound meter (Rion NA-28; Rion Co. Ltd., Tokyo, Japan). Noise was based on the male 2-talker babble noise of the International Collegium of Rehabilitative Audiology [25]. It was adapted to create eight uncorrelated sources of babble noise, one for each of the eight loudspeakers (Control 1; JBL Corp., Los Angeles, CA, USA). See previous studies for more information on this setup [18, 26, 27].

Listening Conditions

During the measurements, speech was presented from different directions, and the microphone settings were changed. Speech originated from the front, left or right (HA side or CI side). These angles are expected only to affect the sound level of the speech and thus the SNR, which makes a strong argument that the different angles of speech did not influence learning effects. The tested microphone settings were beamforming algorithms, i.e., front-end, multiple-microphone processing algorithms that also mainly impact the SNR. Angle and microphone settings were not the topic of this work and will be published elsewhere. In this study, these factors were treated as possible confounders that had to be accounted for with a linear mixed model (LMM) to extract the learning effect, as described below in Statistical Analysis. The microphone settings included RealEarSound [28], a binaural beamformer (StereoZoom™) [18], synchronized automatic gain control [29], and experimental beamformers pointing to the sides, similar to the technique described by Dieudonné and Francart [30]. Only the researcher was aware of the microphone setting. A few additional tests were performed that were not included in the analysis of this article, namely, fittings with only their CI or HA.

Statistical Analysis

SPSS version 29 for Windows (IBM Corp., Armonk, NY) was used to build two LMMs to analyze the learning effects on the SRT between tests and between sessions. In both models, SRT was the outcome parameter, subject number was a random effect, and microphone setting and angle were included as fixed effects. The distinction between the two models was the expression of session number and test number in the models. In the first model, session and test number were included as fixed effects (Eq. 1). With these fixed effects, the marginal means of the SRT for every session and test could be estimated, while accounting for the confounders (microphone setting and angle). With these estimates, the learning effects could be plotted, and the underlying relation of the learning effect between sessions and tests characterized. We did not investigate interaction factors between the fixed effects microphone setting and angle because we did not expect an interaction since microphone settings only affected SNR. We did include the interaction between the factors of interest, namely, session * test. The covariance matrix of the first model was optimized by comparing the Bayesian and Akaike's information criteria of 7 structures suitable for longitudinal data, i.e., "scaled identity," "compound symmetry," "correlation compound symmetry," "heterogeneous compound symmetry," "unstructured," and "unstructured correlation." For optimization, the restricted maximum likelihood (REML) procedure was used. The optimized model included the "scaled identity" covariance matrix. After optimization, the interaction factor session * test did not prove significant and was subsequently excluded. The final LMM with fixed effects is presented in equation (1).

$$\begin{aligned}\mathrm{SRT} = {} & \mathrm{Intercept} + \mathrm{MicSetting} + \mathrm{Angle} + \mathrm{Session} + \mathrm{Test} \\ & + 1 \,\big|\, \mathrm{Subject}\end{aligned} \tag{1}$$

In a second LMM (Eq. 2), session and test number were included as covariates. This model was built for statistical significance testing and was used to determine the marginal mean estimates of the slopes of the linear learning curves. An optimization of the covariance matrix for the second model was performed in the same way as for the first model, by comparing the Bayesian and Akaike's information criteria of 7 structures for longitudinal data. The optimized second model also included the "scaled identity" covariance matrix.

$$\begin{aligned}\mathrm{SRT} = {} & \mathrm{Intercept} + \mathrm{MicSetting} + \mathrm{Angle} + (a \times \mathrm{Session}) \\ & + (b \times \mathrm{Tests}) + 1 \,\big|\, \mathrm{Subject}\end{aligned} \tag{2}$$

To fit the learning curves, an exponential decay function based on the one described by Leibowitz et al. [8] (Eq. 3 was used by deploying a nonlinear least-squares approach implemented with the *lsqcurvefit* function in MATLAB.

$$\mathrm{SRT} = a \cdot e^{-b \cdot t} + c \tag{3}$$

Where a represents the initial SRT, b is the decay rate, t is the session or test number and c is the minimum SRT value when test or session number goes to infinity.

Results

Between-Session Learning Effect

To extract the learning effect and visualize the data across sessions, we used the LMM, where session and test number were included as fixed factors (Eq. 1). By including them as fixed effects, the SRT at each session could be estimated. These estimated marginal mean SRTs are presented in Table 2. Figure 1 shows the estimated marginal mean SRTs with 95% confidence intervals as a function of session number. The confidence interval was the same for all four sessions, ±1.06 dB SNR. The total learning effect, i.e., the mean difference between sessions one and four, was 2.9 dB SNR. Table 2 shows that the learning effect decreases over time, namely, from 1.32 dB SNR to 0.86 dB SNR and 0.67 dB SNR between sessions one and two, two and three, and three and four, respectively.

To model the data with a linear fit, a second LMM (Eq. 2) was used where session was included as a covariate to obtain the slope of the curve, reflecting the change in SRT over time. The resulting linear fit is presented in Figure 1. Session had a statistically significant effect on SRT ($F = 136.0$, $p < 0.001$). The coefficient of the covariate *session* in this model yielded the linear regression coefficient of the regression line plotted in Figure 1. For every session, a decrease in SRT of 0.94 dB SNR per session was found for the linear scale (standard error [SE] = 0.08, $t = 11.7$, $p < 0.001$, degrees of freedom [df] = 683). The estimated marginal mean SRT for sessions one and four are situated above the linear regression line, whereas those for sessions two and three are below it.

Table 2. Estimated marginal (EM) mean SRT with the between-session learning effect

Session number	EM mean SRT, dB SNR	Between-session learning effect, dB SNR|%
Session 1	2.21	
Session 2	0.89	1.32|10.2[a]
Session 3	0.03	0.86|6.6[a]
Session 4	−0.64	0.67|5.2[a]

[a]This is an estimation based on the psychometric functions for CI users of the Matrix test described by Stronks et al. [18].

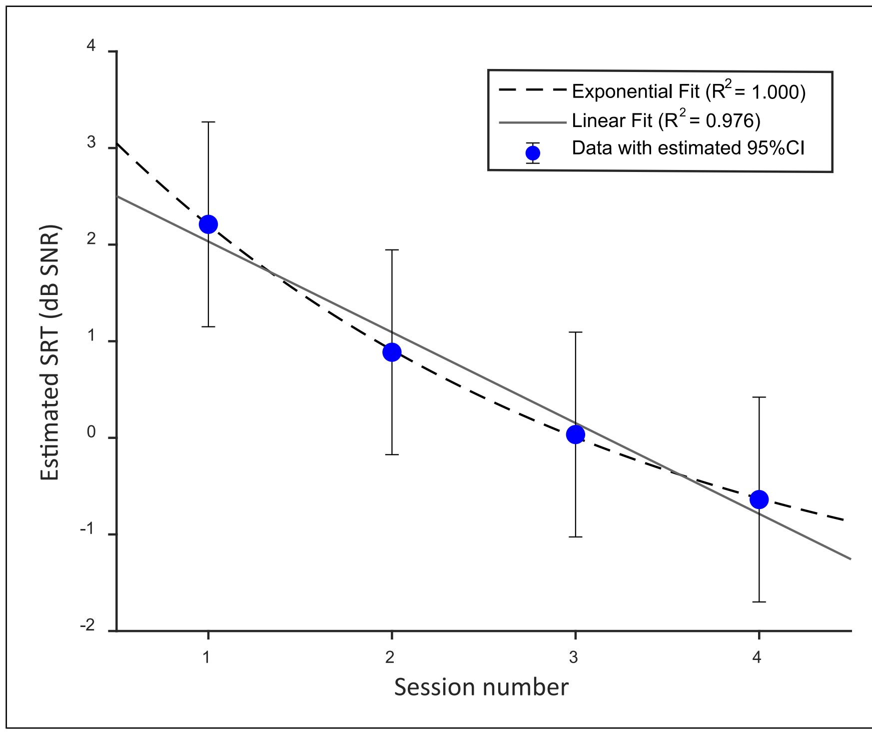

Fig. 1. Estimated marginal mean SRT for four different sessions with their 95% confidence intervals. Including a linear and exponential fit of the data.

This reflects the nonlinear trend observed in Table 2. Therefore, we additionally fitted the data with an exponential decay function (Fig. 1). This yielded a slightly better goodness of fit (linear: R^2 = 0.976; exponential: R^2 = 1,000), and, more importantly, a better distribution of the data around the curve. Fitting the data with an exponential decay function yielded the parameters a = 6.147, b = 0.359, c = −2.088. By using these values in equation (3), the learning effect can be approximated.

Between-Test Learning Effect

We also analyzed the learning effects between consecutive speech tests using equation 1. The corresponding estimated marginal mean SRT and the 95% confidence interval of 13 tests are presented in Figure 2. The mean difference between the first and last test was 1.66 dB SNR. The 95% confidence intervals for all tests were ±1.2 dB SNR, except for test 13, where the confidence interval was ±1.4 dB SNR. When observing the data graphically, it was not readily apparent whether a linear or an exponential model fitted the data better. After test 7, the results became more variable. In this regard, it is important to note that 73% of the breaks were held after test 6 (Fig. 3).

The second LMM showed a statistically significant effect for test number (F = 19.5, p < 0.001). Every consecutive test led to a decrease in SRT of 0.11 dB SNR with the linear approach (SE = 0.025, t = 4.41, p < 0.001,

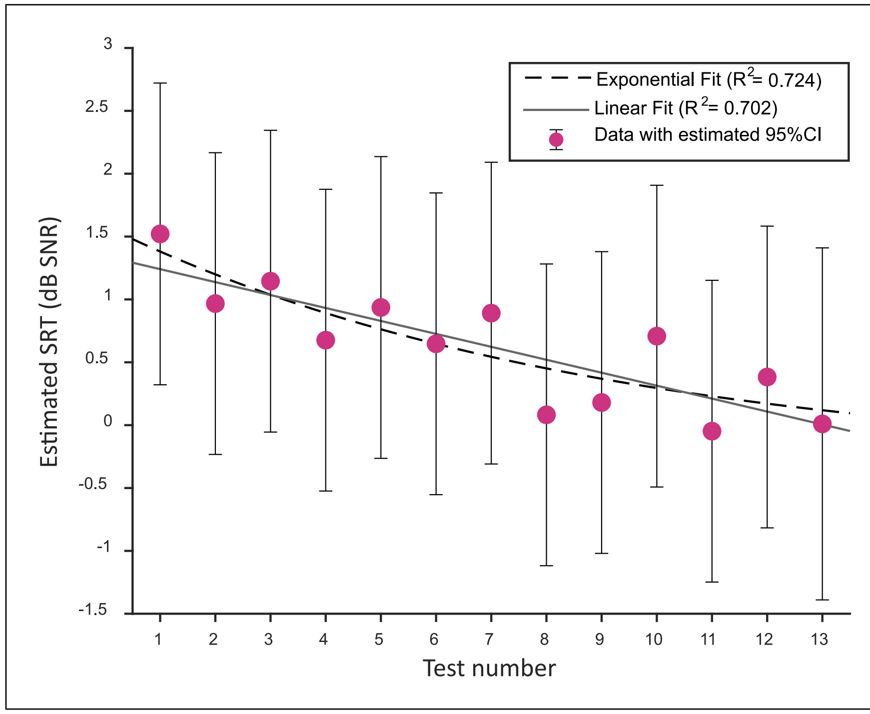

Fig. 2. Estimated marginal means of the SRTs as a function of test number with their 95% confidence intervals. Including a linear and exponential fit of the data.

df = 683) (Fig. 2). The exponential fit yielded parameters for equation 3 of a = 1.903, b = 0.113, c = −0.319. The goodness-of-fit was comparable for both the linear and exponential model with R^2 values of 0.702 and 0.724, respectively, and the distribution of the data around the fitted line was comparable for the two models.

Discussion

In this study, the learning effect of the Dutch/Flemish Matrix test was assessed between tests and between sessions by performing speech-in-noise measurements in 17 bimodal CI users. We reported results where test and session were included in the LMM as fixed effects as well as covariates. The model with fixed effects allowed us to separate the learning effects from the confounders and graphically plot the learning effects per session and per test. We show that an exponential model fits the learning effects between consecutive sessions substantially better than a linear fit. The between-test learning effects could be fitted equally well with a linear fit and an exponential fit. However, we expect an exponential decaying trend beyond the range of the number of tests investigated in this study. Therefore, we conclude that the linear model suffices to model between-test learning effects up to 13 tests.

Both between-session and between tests learning effects were significant and added up to a speech intelligibility increase of 2.9 dB SNR between the first and last sessions, and an increase of 1.7 dB SNR from the start of a session to the end of the session. These learning effects expressed as SRT values can be converted to changes in percent correct scores by using the slope of the psychometric function of the Dutch/Flemish Matrix test, which is reported to be 7.7% for CI users [31]. Hence, we estimate ~10% improvement in performance between session one and two, ~7% improvement between session two and three and ~5% improvement between session three and four (Table 2). Moreover, the 0.11 dB SNR/test corresponds to an estimate improvement of 0.9% per test resulting in a total improvement of 13% between the first and last tests within a session. It is important to note, however, that the psychometric functions used for this conversion were derived from data collected with monaural CI users, and that bimodal CI users may exhibit different psychometric slopes, potentially leading to greater observed percentual learning effects across sessions and tests.

There are three studies with which we can compare our results, as other studies either included procedural learning effects in their analysis or compared only specific test numbers to calculate the learning effect. The study of Stronks et al. [18] had a similar approach as our

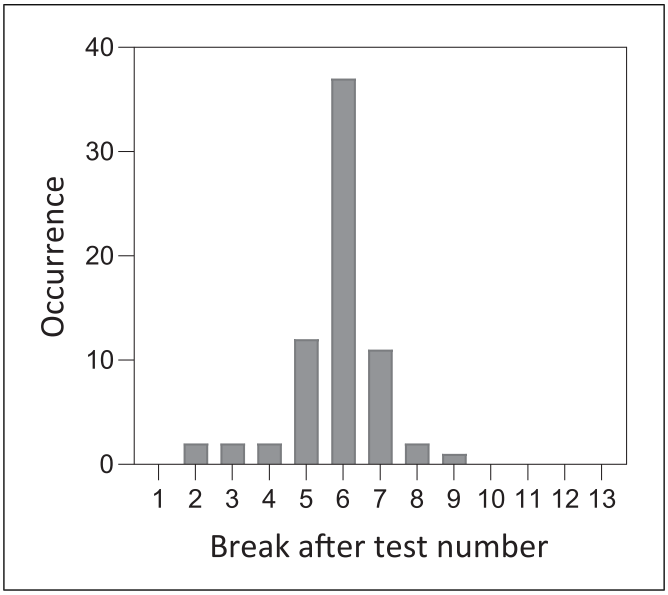

Fig. 3. Histogram showing the cumulative number of breaks introduced after each test, across sessions, and participants. Most of the participants (15 out of 18) had a single break after test number 6 or 7. Three of the participants needed two breaks per session.

study, in that they characterized the learning effects of the Dutch-Flemish Matrix test in a group of bimodal CI users with a similar number of sessions and tests. In that study, learning effects were regarded as confounders, however, and the quantification of the between-session effect was limited to a fixed-effects model, and the between-test effect to a simple linear covariate model. They reported a statistically significant learning effect of 1.5 dB SNR between the first and second sessions, and subsequent improvements of 0.5 dB SNR between the second and third sessions and 0.7 dB between the third and fourth sessions. Our findings corroborate theirs, in that we report an improvement of 1.3 dB SNR between the first and second sessions and 0.7 dB between the third and fourth sessions. However, our improvement between sessions two and three was higher (0.9 dB SNR), which adheres better to the expected trend of a nonlinear learning curve [9, 10]. In the study of Stronks et al. [18] the simple linear covariate model implied an SRT improvement of 0.08 dB SNR/test, which corresponds closely to our finding of 0.11 dB SNR/test in the linear model.

The study of Stronks et al. [18] did not include an analysis of the interaction between session and test. We report that this interaction is not statistically significant and hence, that learning effects between tests are similar across sessions. In other words, our model of the between-test learning is valid for any session, at least up to the four sessions tested here. These findings, however, are in contrast to those of Willberg et al. [14] who found a statistically significant interaction of the learning effects between tests and sessions for the Finnish Matrix test in typical hearing participants. They reported a between-test decrease (difference between the first and last tests) of 1.1 dB SNR in the first session to only a between-test decrease of 0.2 dB SNR in the last session. By contrast, we find an overall between-test effect of 1.7 dB SNR for bimodal CI users across one session. One explanation for the significant interaction in their study is that they only measured four tests instead of the 13 tests included per session here. Indeed, another study from Hernvig et al. [3], where linear mixed modeling was applied, yielded a significant interaction between session and test. In that study, only 4 trials were included per session. Our findings of the overall between-test learning effect correspond well to their effect in the first session (1.7 dB SNR), but their second session only showed a learning effect of 0.7 dB SNR. The number of tests appears to determine the interaction between session and trial effects.

Because of the procedural differences between different studies, the comparison of our results with those of others is less straightforward than it seems. None of the studies discussed properly fitted learning curves, restricting comparison between studies. In addition, the language of the Matrix test, study population, and test paradigms differed across studies. The findings described above of Hernvig et al. [3], obtained with aided hearing-impaired participants, correspond to our between-test results observed during the first session. Similarly, the results reported by Stronks et al. [18], mentioned above, who also investigated bimodal CI users, match our data well. It is important to understand whether we can generalize our findings for different hearing populations. Perceptual learning, which presumably underlies the learning effects described here, involves several mechanisms including attentional weighting (focusing on the most relevant cues), stimulus imprinting (forming long-term memory traces of specific stimuli), differentiation (discriminating between similar stimuli through increased exposure), and unitization (perceiving complex stimuli as integrated wholes, such as recognizing entire words rather than decoding individual phonemes) [7]. These mechanisms likely vary across hearing populations as the nature of auditory input differs, for example, in terms of frequency range and temporal resolution. However, bimodal and monaural CI users may share similar patterns of perceptual learning, as both primarily rely on their CI ear for auditory information. Nevertheless, further research is necessary to determine whether the predictive models presented here extend to populations beyond bimodal CI users and to languages other than Dutch.

We observed some variance in the later tests, which could be a consequence of the breaks introduced during the session (Fig. 3). Additionally, fatigue effects may vary between patients and contribute further variability. Only three participants mentioned fatigue. However, we did not systematically ask the participants about the occurrence of listening fatigue, and we expect that more participants may have dealt with it [32].

The Dutch/Flemish Matrix test can be executed as an open or closed speech in noise test. In this study, the test was executed as an open-set speech in noise test, meaning that the participants did not have a visual presentation of the words in front of them. With a closed-set paradigm, the learning effects could be less than with an open-set paradigm because every participant knows the words in the test to the same extent from the beginning of the session until the end [33]. A few studies investigated Matrix tests in different languages in a closed-set format for typical hearing participants (i.e., participants had access to the word matrix). These studies found statistically significant between-test learning effects during one session. These learning effects were smaller than the test with an open-set format, but the interaction between format type and measurement number was not statistically significantly different. Hence, the response format had no significant impact on the learning effect [34–37]. For some studies, the use of a visual presentation of the words is not possible, for example, during pupillometry when stable ambient light conditions are critical [38]. Therefore, we evaluated the learning effect of the test without visual presentation of the words.

Conclusion

An evident between-test and between-session learning effect was observed for the open-set Dutch/Flemish matrix test in bimodal CI users. The between-session learning effects can be accurately presented with an exponential model as the learning effect decreases substantially between later sessions. By contrast, the between-test learning effect can be accurately described by a simple linear model since the outcome of the exponential model resulted in a similar goodness of fit. The linear model is easier to interpret clinically than the exponential one. This study provides an important handle for correcting these learning effects in future studies with bimodal CI users, especially when randomization is not possible.

Acknowledgments

We thank the study participants for their time and dedication, and Advanced Bionics for technical support.

Statement of Ethics

This study protocol was reviewed and approved by the Medical Ethics Committee Leiden, The Hague, Delft (METC protocol number P20.018). All participants in this study signed written informed consent before participation, and the study was conducted ethically in accordance with the World Medical Association Declaration of Helsinki.

Conflict of Interest Statement

This study was financially and technically supported by Advanced Bionics (European Research Center, Hannover, Germany).

Funding Sources

This publication is part of the INTENSE project, a crossover research program partly financed by the Dutch Research Council (NWO, project number 17619).

Author Contributions

Nienke Cornelia Langerak contributed to the study design, data collection, analyses, and draft writing. Hendrik Christiaan Stronks obtained ethical approval for the study, designed the study, performed analyses, and critically revised the work. Jeroen Johannes Briaire supervised the project and revised the manuscript. Johan Hubertus Maria Frijns supervised the project and revised the manuscript. All authors read and approved the final manuscript.

Data Availability Statement

All data generated or analyzed during this study are included in this article. Further inquiries can be directed to the corresponding author.

References

1 Vroegop JL, Homans NC, Goedegebure A, Dingemanse JG, van Immerzeel T, van der Schroeff MP. The effect of binaural beamforming technology on speech intelligibility in bimodal cochlear implant recipients. Audiol Neurootol. 2018;23(1):32–8. https://doi.org/10.1159/000487749

2 Kollmeier B, Warzybok A, Hochmuth S, Zokoll MA, Uslar V, Brand T, et al. The multilingual matrix test: principles, applications, and comparison across languages: a review. Int J Audiol. 2015;54(Suppl 2):3–16. https://doi.org/10.3109/14992027.2015.1020971

3 Hernvig LH, Olsen SO. Learning effect when using the Danish Hagerman sentences (Dantale II) to determine speech reception threshold. Int J Audiol. 2005;44(9):509–12. https://doi.org/10.1080/14992020500189997

4 van Wieringen A, Wouters J. LIST and LINT: sentences and numbers for quantifying speech understanding in severely impaired listeners for Flanders and The Netherlands. Int J Audiol. 2008;47(6):348–55. https://doi.org/10.1080/14992020801895144

5 Zhen LQ, Pratt SR. Perceptual, procedural, and task learning for an auditory temporal discrimination task. J Acoust Soc Am. 2023;153(3):1823. https://doi.org/10.1121/10.0017548

6 de Jong MAM, Briaire JJ, Frijns JHM. Learning effects in psychophysical tests of spectral and temporal resolution. Ear Hear. 2018;39(3):475–81. https://doi.org/10.1097/AUD.0000000000000499

7 Goldstone RL. Perceptual learning. Annu Rev Psychol. 1998;49:585–612. https://doi.org/10.1146/annurev.psych.49.1.585

8 Leibowitz N, Baum B, Enden G, Karniel A. The exponential learning equation as a function of successful trials results in sigmoid performance. J Math Psychol. 2010;54(3):338–40. https://doi.org/10.1016/j.jmp.2010.01.006

9 Culler E, Girden E. The learning curve in relation to other psychometric functions. Am J Psychol. 1951;64(3):327–49. https://doi.org/10.2307/1418996

10 Gallistel CR, Fairhurst S, Balsam P. The learning curve: implications of a quantitative analysis. Proc Natl Acad Sci U S A. 2004;101(36):13124–31. https://doi.org/10.1073/pnas.0404965101

11 Hagerman B. Clinical measurements of speech reception threshold in noise. Scand Audiol. 1984;13(1):57–63. https://doi.org/10.3109/01050398409076258

12 Hagerman B, Kinnefors C. Efficient adaptive methods for measuring speech reception threshold in quiet and in noise. Scand Audiol. 1995;24(1):71–7. https://doi.org/10.3109/01050399509042213

13 Schlueter A, Lemke U, Kollmeier B, Holube I. Normal and time-compressed speech:how does learning affect speech recognition thresholds in noise? Trends Hearing. 2016;20:2331216516669889. https://doi.org/10.1177/2331216516669889

14 Willberg T, Sivonen V, Hurme S, Aarnisalo AA, Löppönen H, Dietz A. The long-term learning effect related to the repeated use of the Finnish matrix sentence test and the Finnish digit triplet test. Int J Audiol. 2020;59(10):753–62. https://doi.org/10.1080/14992027.2020.1753893

15 Stronks HC, Briaire JJ, Frijns JHM. Residual hearing does not influence the effectiveness of beamforming when using a cochlear implant in conjunction with contralateral routing of signals. Audiol Neurotol. 2023;28(4):262–71. https://doi.org/10.1159/000528767

16 Geißler G, Arweiler I, Hehrmann P, Lenarz T, Hamacher V, Büchner A. Speech reception threshold benefits in cochlear implant users with an adaptive beamformer in real life situations. Cochlear Implants Int. 2015;16(2):69–76. https://doi.org/10.1179/1754762814Y.0000000088

17 Devocht EM, Janssen AM, Chalupper J, Stokroos RJ, George EL. Monaural beamforming in bimodal cochlear implant users: effect of (A)symmetric directivity and noise type. PLoS One. 2016;11(8):e0160829. https://doi.org/10.1371/journal.pone.0160829

18 Stronks HC, Briaire J, Frijns J. Beamforming and single-microphone noise reduction: effects on signal-to-noise ratio and speech recognition of bimodal cochlear implant users. Trends Hear. 2022a;26:23312165221112762. https://doi.org/10.1177/23312165221112762

19 Hagerman B. Sentences for testing speech intelligibility in noise. Scand Audiol. 1982;11(2):79–87. https://doi.org/10.3109/01050398209076203

20 Luts H, Jansen S, Dreschler W, Wouters J. Development and normative data for the Flemish/Dutch Matrix test. 2014.

21 Sampaio ALL, Araújo MFS, Oliveira CACP. New criteria of indication and selection of patients to cochlear implant. Int J Otolaryngol. 2011;2011(1):573968. https://doi.org/10.1155/2011/573968

22 Zeitler DM, Prentiss SM, Sydlowski SA, Dunn CC. American cochlear implant alliance task force: recommendations for determining cochlear implant candidacy in adults. Laryngoscope. 2024;134(Suppl 3):S1–4. https://doi.org/10.1002/lary.30879

23 Cuda D, Murri A, Mainardi A, Chalupper J. Effectiveness and efficiency of a dedicated bimodal fitting formula. Audiol Res. 2019;9(1):219. https://doi.org/10.4081/audiores.2019.219

24 Dyballa KH, Hehrmann P, Hamacher V, Nogueira W, Lenarz T, Büchner A. Evaluation of a transient noise reduction algorithm in cochlear implant users. Audiol Res. 2015;5(2):116. https://doi.org/10.4081/audiores.2015.116

25 Dreschler WA, Verschuure H, Ludvigsen C, Westermann S. ICRA Noises: Artificial Noise Signals with Speech-like Spectral and Temporal Properties for Hearing Instrument Assessment: Ruidos ICRA: Señates de ruido artificial con espectro similar al habla y propiedades temporales para pruebas de instrumentos auditivos. Audiology. 2001;40(3):148–57. https://doi.org/10.3109/00206090109073110

26 Soede W, Bilsen FA, Berkhout AJ. Assessment of a directional microphone array for hearing-impaired listeners. J Acoust Soc Am. 1993;94(2 Pt 1):799–808. https://doi.org/10.1121/1.408181

27 Langerak NC, Stronks HC, Briaire JJ, Frijns JHM. The benefit of bimodal hearing and beamforming for cochlear implant users. Audiol Neurootol. 2024;29(4):297–305. https://doi.org/10.1159/000536431

28 Xu J, Han W. Improvement of adult BTE hearing aid wearers' front/back localization performance using digital pinna-cue preserving technologies: an evidence-based review. Korean J Audiol. 2014;18(3):97–104. https://doi.org/10.7874/kja.2014.18.3.97

29 Boyle PJ, Büchner A, Stone MA, Lenarz T, Moore BCJ. Comparison of dual-time-constant and fast-acting Automatic Gain Control (AGC) systems in cochlear implants. Int J Audiol. 2009;48(4):211–21. https://doi.org/10.1080/14992020802581982

30 Dieudonné B, Francart T. Head shadow enhancement with low-frequency beamforming improves sound localization and speech perception for simulated bimodal listeners. Hear Res. 2018;363:78–84. https://doi.org/10.1016/j.heares.2018.03.007

31 Stronks HC, van Deurzen R, Jansen PL, Briaire JJ, Frijns JHM. Effect of speech material and scoring method on psychometric curves for cochlear implant users and typical hearing listeners. Ear Hear. 2025. https://doi.org/10.1097/AUD.0000000000001672

32 Philips C, Jacquemin L, Lammers MJW, Mertens G, Gilles A, Vanderveken OM, et al. Listening effort and fatigue among cochlear implant users: a scoping review. Front Neurol. 2023;14:1278508. https://doi.org/10.3389/fneur.2023.1278508

33 Clopper CG, Pisoni DB, Tierney AT. Effects of open-set and closed-set task demands on spoken word recognition. J Am Acad Audiol. 2006;17(5):331–49. https://doi.org/10.3766/jaaa.17.5.4

34 Hochmuth S, Brand T, Zokoll MA, Castro FZ, Wardenga N, Kollmeier B. A Spanish matrix sentence test for assessing speech reception thresholds in noise. Int J Audiol. 2012;51(7):536–44. https://doi.org/10.3109/14992027.2012.670731

35 Puglisi GE, Warzybok A, Hochmuth S, Visentin C, Astolfi A, Prodi N, et al. An Italian matrix sentence test for the evaluation of speech intelligibility in noise. Int J Audiol. 2015;54(Suppl 2):44–50. https://doi.org/10.3109/14992027.2015.1061709

36 Warzybok A, Zokoll M, Wardenga N, Ozimek E, Boboshko M, Kollmeier B. Development of the Russian matrix sentence test. Int J Audiol. 2015;54(Suppl 2): 35–43. https://doi.org/10.3109/14992027.2015.1020969

37 Zokoll MA, Fidan D, Türkyılmaz D, Hochmuth S, Ergenç İ, Sennaroğlu G, et al. Development and evaluation of the Turkish matrix sentence test. Int J Audiol. 2015; 54(Suppl 2):51–61. https://doi.org/10.3109/14992027.2015.1074735

38 Winn MB, Wendt D, Koelewijn T, Kuchinsky SE. Best practices and advice for using pupillometry to measure listening effort: an introduction for those who want to get started. Trends Hear. 2018;22:2331216518800869. https://doi.org/10.1177/2331216518800869