Introduction

Speech in noise (SIN) perception is essential for day-to-day communication since everyday conversations seldom transpire in ideal silent environments. While it is speculated that musical training may aid in SIN discrimination via multiple neural pathways, the empirical evidence of musician-advantage on SIN has only been mixed [1-7]. Musical training has been associated with experience dependent brain plasticity leading to significant physiological changes in the auditory system [8], auditory cortico-fugal pathways [9], auditory-motor coupling [10,11], and multimodal sensorimotor integration system [12] thereby enhancing the mechanisms involved in auditory processing.

The alterations in the auditory system following musical training are attributed to stronger inter/intra-hemispheric connections of the auditory cortices, increased grey matter in the inferior temporal gyrus [13], pre-central gyri [14,15], intraparietal sulcus, and the inferior lateral temporal lobe [16]. Given the shared sensory activation of the brain regions for speech,
language and music processing, higher levels of musical training are associated with effective processing of acoustic information in musicians [10,15,16]. Such overlap and increased activation of brain regions in musicians maybe achieved via fine-tuned sensitivity and attention to variations in acoustic information, resulting in a musicianship-advantage to decipher speech in noise [17].

While the benefit of musical training on SIN has drawn much attention in the literature in recent years, whether musicianship advantage exists for “listening effort” (LE) required to understand speech in deteriorated situations remains less explored [7,18-21]. LE is defined as “the mental exertion required to attend to and understand an auditory message” [22-24]. The underlying premise of LE is that, as the listening situation becomes challenging, differentiating speech from noise entails greater LE as more cognitive resources are required to retain adequate comprehension to achieve successful SIN perception [19,22-24].

LE has been often measured using behavioral reactions to listening tasks and consequently, influenced by the task complexity and cognitive load of information in a task [21]. Dual-task paradigm (DTP) is one such well-established behavioral method that is employed to measure LE. DTP entails participants to execute two tasks at once, in which, the secondary task performance measures the amount of effort dedicated to the auditory task [19,25]. The current study also employs DTP technique to understand whether sustained training on South Indian classical (Carnatic) music for years translates to improved LE in SIN environment, not been explored in the literature as yet, to the best of our knowledge.

Prior research shows that the form of musical training influences perceptual skills and brain processing of acoustic features, offering unique sets of benefits [26]. The choice of Carnatic music in the present study is inspired by recent research that indicates that active and passive participation in music, notably Carnatic music [27] stimulates brain networks that execute spatial thinking, memory, and thus improve SIN perception. Studies document a 1 dB improvement in signal-to-noise ratio (SNR) or 6%–8% increase in SIN scores with duration of musical training [28] while specifically that of Carnatic music was 2–3 dB SNR or 16%–18% increase in scores of SIN perception [27].

Indian classical music particularly differs from other forms of music due to the presence of microtones, Gamakas, and the style of representation, suggesting dependence on memory, improvisation, and aural learning [8,27]. In this background, the current study investigates whether musicians trained in Carnatic music can decipher SIN better than non-musicians, and whether such musicians expend less effort compared to their non-musician counterparts in the process.

**Subjects and Methods**

**Participants**

A total of 60 participants, including 25 musicians and 35 non-musicians (34 females and 26 males) in the age range of 18 to 35 years (mean age 23.2 years), were recruited for the study. The participants were classified as musicians in line with earlier research [4,5] where the number of years of musical training in Carnatic music (vocal/instrumental) was a minimum of 5 years. In addition, all the musician participants also underwent a minimum of 3 years of continuous training in the last 6 years and 3.5 hours of practice/week in any 2 years for the last 3 years. Those who did not meet the criteria above were pooled as non-musicians. Table 1 shows the demographic information of the participants considered in the two groups of the study. Only participants who reported no history of any physical or neurological deficits were included. All individuals were free of otological disorders and underwent hearing screening and had pure-tone air conduction thresholds ≤25 dB HL across all the octave frequencies from 250 Hz to 8 kHz (ANSI S3.6, 1996).

A participant recruitment questionnaire was administered to ensure all the participants satisfy the inclusion criteria. Information including demographics, handedness, languages exposed to, education, musical training, socioeconomic status, occupation, noise exposure, living conditions, and general musical orientation was obtained. Musical orientation, which is commonly defined as an individual’s general attitude towards music [29] was elicited in our survey based on whether they belong to a family of musicians with prior music exposure.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Musicians (n=25)</th>
<th>Non-musicians (n=35)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (yr)</td>
<td>23.56 ± 4.22</td>
<td>23.0 ± 3.27</td>
</tr>
<tr>
<td>Male, n (%)</td>
<td>14 (56)</td>
<td>15 (44)</td>
</tr>
<tr>
<td>Duration of musical training (yr)</td>
<td>17.28 ± 8.90</td>
<td>NA</td>
</tr>
<tr>
<td>Age at onset of musical training (yr)</td>
<td>6.0 ± 2.58</td>
<td>NA</td>
</tr>
<tr>
<td>Primary instrument, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Vocal</td>
<td>8 (32)</td>
<td>NA</td>
</tr>
<tr>
<td>Instrumental</td>
<td>17 (68)</td>
<td>NA</td>
</tr>
<tr>
<td>Languages known, n (%)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bilingual</td>
<td>15 (60)</td>
<td>16 (46)</td>
</tr>
<tr>
<td>Multilingual</td>
<td>10 (40)</td>
<td>19 (54)</td>
</tr>
</tbody>
</table>

Data are presented as mean ± standard deviation unless otherwise indicated. SD, standard deviation; NA, not applicable.
sure (any genre), the type of music they listen to, and the number of hours they listen to music every week. Participation willingness and written consent were obtained as required by the institutional ethics committee prior to initiation of the data collection. The current study was approved by the Institutional Ethics Committee (IEC-NI/20/SEP/75/82).

Procedure

The study was carried out in three phases where the participants’ general musical ability, SIN, and LE were measured in each phase respectively. The entire test procedure was administered using circum-aural headphones (HDA 280 PRO, Sennheiser Electronic GmbH & Co. KG Wedemark, Germany) with a calibrated output routed through the laptop (Lenovo X1 Carbon with Intel i7 processor; Sichuan, China) in a quiet room. The average noise level in the testing room was around 30 dBA and never exceeded 40 dBA. This was ensured by continuous monitoring using the NIOSH Sound Level Meter app developed by National Institute of Occupational Safety and Health [30] throughout the testing duration.

Phase I: Assessment of musical ability and working memory

All the participants were administered the online Mini-PROMS (Mini-Profile of Music Perception Skills) to assess their musical ability. This test comprises of four subsets: melody, tuning, accent, and tempo [31]. All subsets consist of blocks of 3 musical tones (2 references and 1 comparison). The participants were instructed to identify if the comparison stimuli were same/different from the reference stimuli and choose the right option from the available 5 choices displayed. A correct response is given a score of “2” while correct but uncertain response is given a score of “1” and incorrect response is given a score of “0.” After completion of the test, the total score (maximum possible score 36) and the subset total score: melody (10), tuning (8), accent (10), and speed (8) were recorded.

The participants also underwent Backward Digit Span Test, to measure their working memory, which is demonstrated to have influence on SIN and LE [21]. This test required participants to repeat the increasingly longer sequences of digits in a backward order. Digit span was calculated as the longest sequence of digits repeated correctly in two consecutive presentations.

Phase II: Assessment of binaural SIN ability

SIN perception abilities was tested using phonemically balanced (PB) wordlist in Tamil [32] and Tamil Matrix Sentences Test (TMST) [33]. The words and sentences were mixed with Tamil multi-talker speech babble [34] using MATLAB version 2014b software (MathWorks, Natick, MA, USA) at SNRs of +5 dB, 0 dB, and -5 dB. Three lists of PB words (one list per SNR) with 25 words each were presented binaurally to the participants at an intensity level of 70 dB SPL. The participants were instructed to repeat the words heard in the presence of speech babble. Each correctly identified word was scored as 1 and 0 for incorrect/missed identification with a maximum possible score of 25 for each list.

The five-word sentences of TMST had a fixed semantic structure with nouns, numbers, adjectives, objects, and verbs. The order of these words in the sentence followed the sentence structure of Tamil language. Each word in the sentence list had 10 alternatives with a total of 50 words per sentence list. A total of 3 sentence lists were administered with one list at each of the SNRs selected. These were presented at an intensity level of 70 dB SPL binaurally. Each correctly identified word in a sentence was given a score of 1 with a total possible score of 50 per list.

Phase III: Assessment of LE

LE was measured using the DTP following a similar procedure used by Harvey, et al. [19] consisting of a baseline and an experiment phase where the latter involves a primary and a secondary task. In the baseline condition (only secondary task), participants were visually presented one random alphabet at a time on a computer monitor in a large black font against white background, developed in C# Windows Presentation Foundation (Bengaluru, India) using a visual studio tool. The display duration of the target alphabet was 200 ms with an inter-stimulus gap of 1 s. The participants were instructed to attend to the monitor and indicate via a button press whenever the target alphabet appeared from the randomized series of alphabets. The time taken to respond to the targets from the time of its appearance on screen, was measured as “baseline reaction time.” Correct identification of the target visual stimuli was considered “valid” response while incorrect/miss was considered as an “invalid” response. The total number of valid and invalid responses were recorded.

In the experimental condition, participants were asked to repeat the words and the sentences presented in the presence of noise (primary task) while simultaneously responding to the visual task (secondary task) outlined above. The reaction time to the secondary task was recorded and compared with the baseline measures. The difference in time duration between the baseline (only secondary task) and the task evoked (primary+secondary tasks), measured the effort exerted by the participant for SIN perception (primary task) and was considered as measure of LE. Only the correct responses to
the visual task were subjected to analysis. The premise behind
the DTP is that when speech comprehension becomes in-
creasingly difficult, performance in the secondary task gets
worse from the baseline.

Statistical analysis
A total of 60 individuals participated in the study. While
most of the studies in this strand of literature used group means
(ANOVA) to examine variations in outcomes between musi-
cians and non-musicians [4,21], the current study aimed to
advance the field by applying causal analysis that can control
for other variables that may influence the outcome of our in-
terest variables—namely, SIN and LE. The data obtained were
subjected to two causal-regression models: fractional logit
and linear regression, to account for the continuous and dis-
crete natures of the dependent outcome variables SIN and LE.
As per standard statistical convention in all our statistical anal-
ysis, we have treated p value <5% (p<0.05) as statistically sig-
nificant values. Specifically, linear regression was used for LE,
which was measured in milliseconds (hence continuous vari-
able), and fractional logit was used for SIN and Mini-PROMS,
which were measured as the percentage of correct responses
(and hence bounded between 0 and 100).

The generic functional form of fractional logit model is
\[
E \left( \log \frac{y}{1-y} \mid X \right) = X\beta,
\]
where y is the outcome variable—musical abilities (scores
from Mini-PROMS) and SIN perception measured in frac-
tion-percentage; Xs are the independent variables such as de-
mographics, economic status, living conditions, and musical
orientation; E is the exponential constant; and \( \beta \) is coefficient
estimates for each of the independent variable X. For ease of
interpretation of results, we converted the coefficient values to
their respective odds ratios, which will help understand the
relative advantage of musicians over non-musicians, if any.

The functional form for the linear regression model is
\[
Y_h = \beta_0 + \beta_1 \text{Musician} + \beta_2 \sum X_h + \epsilon,
\]
where \( Y_h \) is the LE measured as reaction time (ms), \( \beta_0 \) is the
coefficient estimate for musicianship, and \( \beta_2 \) is vector of coef-
fitient estimates from other variables such as demographics,
economic status, living conditions, and musical orientation.
In the model, “musician” is a binary variable that takes a value
of 1 if they met the conditions mentioned in the methods
section, and zero otherwise.

Results
Results of performance on working memory and
music ability test
Table 2 shows the mean and SD comparison of musicians
and non-musicians in overall Mini-PROMS score (and its
subsets) and the Backward Digit Span Test. The results from t-
test denote that while the scores of the musicians were higher
than that of non-musicians in all domains for Mini-PROMS,
there were no significant differences across the two groups on
working memory.

Table 3 presents the results that check for the effect of mu-
sicianship on the overall Mini-PROMS scores and each of its
components, while simultaneously controlling all other in-
dependent variables that can potentially affect Mini-PROMS
scores. Only for melody, tune, and the overall scores, the odds
were higher for musicians—for instance, the odds of getting a
higher overall score were 1.6 times greater for musicians than
non-musicians (similar interpretations follow for melody and
tune)—exhibiting a distinctive musicianship advantage. Statis-
tically no differences were observed for speed and accent across
the two groups.

Performance on SIN and LE
As mentioned earlier, the study aimed to understand if the
advantage that musicians have in music related auditory skills
translate to better SIN perception and reduced LE. To that ef-
effect the study employed six different models—one each for 3
different SNRs for words and sentences. Fig. 1 shows the av-

<table>
<thead>
<tr>
<th>Test</th>
<th>Musician</th>
<th>Non-musician</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mini-PROMS subsets (maximum score)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Melody (10)</td>
<td>6.6 ± 0.18</td>
<td>4.3 ± 0.15</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Tune (8)</td>
<td>6.6 ± 0.18</td>
<td>3.3 ± 0.16</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Accent (10)</td>
<td>5.5 ± 0.13</td>
<td>4.8 ± 0.10</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Speed (8)</td>
<td>6.6 ± 0.19</td>
<td>4.6 ± 0.21</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Overall (36)</td>
<td>6.3 ± 0.13</td>
<td>4.3 ± 0.10</td>
<td>&lt;0.01</td>
</tr>
<tr>
<td>Backward Digit Span Test</td>
<td>6.5 ± 1.08</td>
<td>5.9 ± 0.82</td>
<td>0.12</td>
</tr>
</tbody>
</table>

Data are presented as mean ± standard deviation. Mini-PROMS, Mini-Profile of Music Perception Skills

<table>
<thead>
<tr>
<th>Test</th>
<th>Musician</th>
<th>Non-musician</th>
<th>Odds ratio</th>
<th>p value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>1.59</td>
<td>1.63</td>
<td>1.36</td>
<td>1.46</td>
</tr>
<tr>
<td>Melody</td>
<td>1.90</td>
<td>1.63</td>
<td>1.94</td>
<td>1.68</td>
</tr>
<tr>
<td>Tune</td>
<td>2.09</td>
<td>1.46</td>
<td>2.09</td>
<td>1.36</td>
</tr>
<tr>
<td>Speed</td>
<td>2.30</td>
<td>2.00</td>
<td>2.09</td>
<td>1.68</td>
</tr>
<tr>
<td>Accent</td>
<td>2.51</td>
<td>1.75</td>
<td>2.51</td>
<td>1.75</td>
</tr>
</tbody>
</table>

Mini-PROMS, Mini-Profile of Music Perception Skills

Table 2. Comparison of Mini-PROMS and Backward Digit Span Test scores in musician and non-musician groups

Table 3. Results of fractional logit model analysis for Mini-PROMS scores of musician

The functional form for the linear regression model is
\[
Y = \beta_0 + \beta_1 \text{Musician} + \beta_2 \sum X + \epsilon,
\]
average score comparison across musicians and non-musicians for SIN and LE.

As like Mini-PROMS, Table 4 presents the results for SIN scores using fractional logit model, Table 5 for LE using linear regression. Results are presented for the base model, where none of the controls were included, and the controlled model, where all other variables were included (The study reports only the odds ratios and coefficient estimates on the interest variable—musical training—while Supplementary Tables 1 and 2 [in the online-only Data Supplement] have the entire set of results).

Tables 4 and 5 present limited evidence for musicianship advantage on LE and SIN that too in very specific conditions. For instance, Table 4 shows that the odds of obtaining the correct responses for musicians were twice that of non-musicians for sentences only at +5 dB SNR, and 6 times higher only for words at 0 dB SNR. And, Table 5 shows that only for words in noise in 0 dB, there was musicianship advantage; the reaction time was 8 seconds lesser for musicians compared to non-musicians.

Otherwise, as exhibited in Tables 4 and 5, the odds ratios or coefficient estimates on musicianship were not statistically significant to make inferences on the positive association of musical training on SIN or LE. While there was a clear musicianship advantage on Mini-PROMS, the gains seem to transfer only in very specific situations to SIN and LE, and hence not generalizable across SNRs.

**Discussion**

The current study aimed to determine if musical training enhances SIN perception and lowers LE in noisy circumstances. The current study aimed to determine if musical training enhances SIN perception and lowers LE in noisy circumstances.
and varies from recent research in this field in crucial ways. Regression techniques were employed to investigate the influence of musical training on SIN and LE to control for predisposition factors that can potentially affect the outcome, which is not possible in group mean differences employed by past studies. The study also attempted to add evidence to the current body of knowledge, notably with respect to the stimuli used, musical orientation, besides our modeling approach, and analyses. While much of past work predominantly used only sentences in an adaptive SNR procedure, here both words and sentences at fixed SNRs were used, exhibiting distinctive advantages in efficiency gains while having nearly equal correct and miss rates [35]. This advantage further reinforces the aim of unearthing group differences across musicians and non-musicians while controlling for individual and group characteristics through a regression model.

To overcome the influence of cognitive overload that prior studies have been affected from, the current study incorporated noise levels at 3 different SNRs: +5 dB (easy), 0 dB (moderate), and -5 dB (difficult) measuring the outcome in different levels of task complexity [21,25]. However, as the task complexity significantly influences LE, the DTP used in the current study with 3 levels of SIN degradation and the secondary task could have been less demanding to capture subtle group differences in sentences, but not for words [19]. Due to linguistic redundancy, a person less sensitive to fine temporal variations may nonetheless understand sentences [36]. Listeners rely more on knowledge-guided elements for speech perception as the burden on central cognitive resources grows. This reliance stems from a re-analysis of perceptual acoustic cues to word and phoneme identification weighing less than linguistic cues [37].

With these significant additions to the literature in this area, results from our study did not observe much difference between musicians’ and non-musicians’ SIN perceptions and LE, except for words at 0 dB SNR and sentences at +5 dB SNR. This limited musician-advantage at specific SNRs can be attributed to the demands for “high-precision” pitch-encoding during musical training. Speech perception requires sensitivity to pitch, spectral shape, and temporal variations, and pitch fluctuation conveys a range of information, including prosodic emphasis, phrase boundaries, etc. [38]. Musical training necessitates high-precision pitch encoding because it requires ongoing monitoring of intonation and note selection. This demand of precise pitch encoding paves way for musical training to refine pitch processing in the auditory system, enabling subcortical pitch-encoding networks to operate with greater accuracy for typical speech processing [35], thereby lowering LE in adverse conditions for musicians.

Overall findings that musical training has limited effect on understanding SIN suggests a lack of any systematic advantage for musicians’ over non-musicians perceiving speech in the background of noise. It is to be noted that, unlike most studies in the literature that predominantly included native English speakers trained in Western music, the current study included native Tamil speakers and musician participants’ who were trained in Carnatic music. Yet, the findings align with those of other past studies [5,18].

Our results exhibit that musical training did not bear a significant influence on LE. However, the comparable effort exerted by musicians and non-musicians indicates that listening to SIN not only impacts the primary task but also the listener’s

### Table 4. ORs of SIN scores for musicians at different SNRs using fractional logit model

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>PB words in noise (base model)</th>
<th>TMST in noise (base model)</th>
<th>PB words in noise (controlled model)</th>
<th>TMST in noise (controlled model)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OR (p value)</td>
<td>OR (p value)</td>
<td>OR (p value)</td>
<td>OR (p value)</td>
</tr>
<tr>
<td>+5</td>
<td>1.07 (0.06)</td>
<td>1.01 (0.33)</td>
<td>2.09 (0.39)</td>
<td>1.99 (&lt;0.01)</td>
</tr>
<tr>
<td>0</td>
<td>1.01 (0.63)</td>
<td>1.05 (0.29)</td>
<td>6.17 (0.04)</td>
<td>0.67 (0.70)</td>
</tr>
<tr>
<td>-5</td>
<td>1.03 (0.47)</td>
<td>1.01 (0.80)</td>
<td>0.54 (0.40)</td>
<td>0.89 (0.91)</td>
</tr>
</tbody>
</table>

OR, odds ratio; SIN, speech in noise; SNR, signal to noise ratio; PB, phonemically balanced; TMST, Tamil Matrix Sentences Test

### Table 5. Coefficient estimates of reaction time (ms) of musicians for SIN using linear regression model

<table>
<thead>
<tr>
<th>SNR (dB)</th>
<th>PB words in noise (base model)</th>
<th>PB words in noise (controlled model)</th>
<th>TMST in noise (base model)</th>
<th>TMST in noise (controlled model)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient (p value)</td>
<td>Coefficient (p value)</td>
<td>Coefficient (p value)</td>
<td>Coefficient (p value)</td>
</tr>
<tr>
<td>+5</td>
<td>39.21 (0.24)</td>
<td>249.01 (0.07)</td>
<td>37.74 (0.25)</td>
<td>4.98 (0.97)</td>
</tr>
<tr>
<td>0</td>
<td>14.43 (0.75)</td>
<td>8.058.12 (&lt;0.01)</td>
<td>30.08 (0.52)</td>
<td>212.81 (0.24)</td>
</tr>
<tr>
<td>-5</td>
<td>67.89 (0.16)</td>
<td>119.06 (0.49)</td>
<td>47.16 (0.43)</td>
<td>178.25 (0.44)</td>
</tr>
</tbody>
</table>

SIN, speech in noise; SNR, signal to noise ratio; PB, phonemically balanced; TMST, Tamil Matrix Sentences Test
The current study offers only limited evidence to support the claim that musical training improves speech perception in noisy environments or reduces listening effort. While the current study adds to literature by employing regression models to directly unearth this association by controlling for other variables, the results should not be interpreted to mean that music training can offer no benefit when listening in adverse conditions. Future work focusing on combining behavioral measures with self-report scales and physiological techniques in difficult listening conditions may assist in identifying more subtle distinctions between perception of SIN across musicians and non-musicians.

**Supplementary Materials**
The online-only Data Supplement is available with this article at https://doi.org/10.7874/jao.2023.00038.

**Acknowledgments**
The authors would like to acknowledge Mr. Udhyaa Kumar R, for helping in developing the speech in noise stimulus, Ms. Anagha for developing the stimulus software for Dual-task Paradigm and all the participants for participating in the study. Dr. Heramba Ganapathy is also acknowledged for the initial coordination and support for the experimental set-up. The Department of Audiology, Faculty of Audiology and Speech Language Pathology, SRHHER (DU) is acknowledged for providing the set-up to conduct the study.

**Conflicts of Interest**
The authors have no financial conflicts of interest.

**Author Contributions**
Conceptualization: Vallampati Lavanya, Ramaprasad Rajaram. Data curation: Ramaprasad Rajaram, Vallampati Lavanya. Formal analysis: Ramaprasad Rajaram, Vallampati Lavanya. Investigation: all authors. Methodology: Ramaprasad Rajaram, Vallampati Lavanya, Ramya Vaidyanath. Project administration: all authors. Software: Ajith Kumar Uppunda, Ramya Vaidyanath, Vallampati Lavanya. Supervision: Ramaprasad Rajaram. Validation: all authors. Writing—original draft: Vallampati Lavanya, Ramaprasad Rajaram. Writing—review & editing: all authors. Approval of final manuscript: all authors.

**ORCID iDs**
Vallampati Lavanya https://orcid.org/0000-0002-5332-6472
Ramaprasad Rajaram https://orcid.org/0000-0002-4763-9624
Ramya Vaidyanath https://orcid.org/0000-0003-4093-9424
Ajith Kumar Uppunda https://orcid.org/0000-0002-1368-9834

**REFERENCES**
18) Hornsby BW. The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands. Ear Hear 2013;34:523-34.
23) Hornsby BW. The effects of hearing aid use on listening effort and mental fatigue associated with sustained speech processing demands.
Supplementary Table 1. Coefficient estimates for controlled variables of each of SIN scores for PB words and TMST from fractional logit model

<table>
<thead>
<tr>
<th>Variables</th>
<th>PB words in noise</th>
<th>TMST in noise</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+5 dB SNR</td>
<td>0 dB SNR</td>
</tr>
<tr>
<td></td>
<td>Coefficient</td>
<td>(p value)</td>
</tr>
<tr>
<td>Age</td>
<td>0.00 (0.95)</td>
<td>0.01 (0.07)</td>
</tr>
<tr>
<td>Gender</td>
<td>0.04 (0.42)</td>
<td>0.73 (0.11)</td>
</tr>
<tr>
<td>Years of training</td>
<td>0.06 (0.42)</td>
<td>0.03 (0.60)</td>
</tr>
<tr>
<td>Hours of practice</td>
<td>0.17 (0.65)</td>
<td>0.04 (0.18)</td>
</tr>
<tr>
<td>Age of onset</td>
<td>0.00 (0.98)</td>
<td>0.11 (0.11)</td>
</tr>
<tr>
<td>Musicians in family</td>
<td>0.02 (0.70)</td>
<td>0.89 (0.15)</td>
</tr>
<tr>
<td>Noise exposure at work</td>
<td>0.02 (0.27)</td>
<td>0.01 (0.35)</td>
</tr>
<tr>
<td>Language expertise</td>
<td>0.03 (0.58)</td>
<td>0.11 (0.05)</td>
</tr>
<tr>
<td>Training in other disciplines</td>
<td>0.07 (0.15)</td>
<td>0.00 (0.89)</td>
</tr>
</tbody>
</table>

SIN, speech in noise; PB, phonemically balanced; TMST, Tamil Matrix Sentences Test; SNR, signal to noise ratio
Supplementary Table 2. Coefficient estimates for controlled variables of each of reaction time for SIN (PB words and TMST) from linear regression model

<table>
<thead>
<tr>
<th>Variables</th>
<th>PB words in noise</th>
<th></th>
<th>TMST in noise</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+5 dB SNR</td>
<td>0 dB SNR</td>
<td>-5 dB SNR</td>
<td>+5 dB SNR</td>
<td>0 dB SNR</td>
<td>-5 dB SNR</td>
</tr>
<tr>
<td></td>
<td>Coefficient (p value)</td>
<td>Coefficient (p value)</td>
<td>Coefficient (p value)</td>
<td>Coefficient (p value)</td>
<td>Coefficient (p value)</td>
<td>Coefficient (p value)</td>
</tr>
<tr>
<td>Age</td>
<td>6.40 (0.35)</td>
<td>5.22 (0.53)</td>
<td>12.10 (0.24)</td>
<td>6.99 (0.29)</td>
<td>1.14 (0.90)</td>
<td>0.29 (0.98)</td>
</tr>
<tr>
<td>Gender</td>
<td>2.10 (0.96)</td>
<td>30.98 (0.55)</td>
<td>12.52 (0.84)</td>
<td>37.48 (0.35)</td>
<td>91.65 (0.12)</td>
<td>80.91 (0.28)</td>
</tr>
<tr>
<td>Years of training</td>
<td>71.01 (0.95)</td>
<td>28.84 (0.73)</td>
<td>33.16 (0.74)</td>
<td>29.40 (0.65)</td>
<td>27.21 (0.77)</td>
<td>42.22 (0.72)</td>
</tr>
<tr>
<td>Hours of practice</td>
<td>0.37 (0.99)</td>
<td>63.08 (0.13)</td>
<td>7.91 (0.87)</td>
<td>56.16 (0.08)</td>
<td>25.18 (0.58)</td>
<td>47.89 (0.41)</td>
</tr>
<tr>
<td>Age of onset</td>
<td>34.16 (0.58)</td>
<td>61.60 (0.43)</td>
<td>16.51 (0.86)</td>
<td>24.77 (0.68)</td>
<td>53.91 (0.53)</td>
<td>53.93 (0.63)</td>
</tr>
<tr>
<td>Musicians in family</td>
<td>26.07 (0.64)</td>
<td>52.02 (0.46)</td>
<td>59.81 (0.48)</td>
<td>19.78 (0.71)</td>
<td>117.64 (0.14)</td>
<td>29.69 (0.77)</td>
</tr>
<tr>
<td>Noise exposure at work</td>
<td>1.30 (0.94)</td>
<td>11.27 (0.63)</td>
<td>16.03 (0.57)</td>
<td>29.71 (0.11)</td>
<td>2.17 (0.93)</td>
<td>38.20 (0.26)</td>
</tr>
<tr>
<td>Language expertise</td>
<td>62.06 (0.24)</td>
<td>22.05 (0.73)</td>
<td>29.56 (0.71)</td>
<td>75.08 (0.14)</td>
<td>34.09 (0.63)</td>
<td>52.14 (0.58)</td>
</tr>
<tr>
<td>Training in other disciplines</td>
<td>12.35 (0.76)</td>
<td>65.40 (0.20)</td>
<td>54.17 (0.38)</td>
<td>18.91 (0.44)</td>
<td>46.03 (0.41)</td>
<td>53.51 (0.46)</td>
</tr>
</tbody>
</table>

SIN, speech in noise; PB, phonemically balanced; TMST, Tamil Matrix Sentences Test; SNR, signal to noise ratio