COGNITIVE NEUROSCIENCE

Successful aging of musicians: Preservation of sensorimotor regions aids audiovisual speech-in-noise perception

Lei Zhang[1,2], Xiuyi Wang[1], Claude Alain[3,4], Yi Du[1,2,5,6]*

Musicianship can mitigate age-related declines in audiovisual speech-in-noise perception. We tested whether this benefit originates from functional preservation or functional compensation by comparing fMRI responses of older musicians, older nonmusicians, and young nonmusicians identifying noise-masked audiovisual syllables. Older musicians outperformed older nonmusicians and showed comparable performance to young nonmusicians. Notably, older musicians retained similar neural specificity of speech representations in sensorimotor areas to young nonmusicians, while older nonmusicians showed degraded neural representations. In the same region, older musicians showed higher neural alignment to young nonmusicians than older nonmusicians, which was associated with their training intensity. In older nonmusicians, the degree of neural alignment predicted better performance. In addition, older musicians showed greater activation in frontal-parietal, speech motor, and visual motion regions and greater deactivation in the angular gyrus than older nonmusicians, which predicted higher neural alignment in sensorimotor areas. Together, these findings suggest that musicianship-related benefit in audiovisual speech-in-noise processing is rooted in preserving youth-like representations in sensorimotor regions.

INTRODUCTION

Aging is associated with myriad cognitive deficits. According to the revised model of the scaffolding theory of aging and cognition (STAC-r) (*1*, *2*), both aging and life course experience influence brain structure and function (e.g., cortical thickness, white matter integrity, and neural specificity) as well as the development of compensatory scaffolding (*2*–*4*). Compensatory scaffolding, e.g., enhanced recruitment of frontal-parietal regions and additional circuits, could shore up declining structures whose functioning has become noisy, inefficient, or both (*1*, *2*). Positive lifestyle choices, e.g., long-term musical training, which contribute to neural resource enrichment, could mitigate the negative impacts of aging (e.g., structural and functional decline) and further enhance the compensatory scaffolding process (*1*, *2*). However, it remains to be determined whether musicianship mitigates cognitive aging more through functional preservation or functional compensation, which are the key components in the STAC-r model (*1*).

Among age-related cognitive deficits, difficulty in understanding speech in noisy environments is one of the most prevalent, even in older adults who show normal hearing for their age. Although musicians' advantage in speech perception in noise is not robust in young and middle-aged adults (*5*–*7*), accruing evidence suggests that age-related decline in speech-in-noise perception can be mitigated by long-term or short-term musical training (*8*–*12*). However, such benefits have not yet been demonstrated in audiovisual speech-in-noise scenarios in aging populations. According to the overlap, precision, emotion, repetition, and attention hypothesis, musicians' advantage in speech processing is driven by the anatomical overlap in the brain networks that process music and speech as well as the more precise, emotional, repetitive, and attentive engagement of the brain networks during musical training (*13*). In older musicians (OM), the shared brain networks for speech and music processing may provide a mean by which musicianship generates the cross-domain transfer benefit in processing audiovisual speech in noise.

The benefit of musical expertise on older adults' audiovisual speech-in-noise perception could be related to functional preservation. For instance, compared to older nonmusicians (ONM), OM have enhanced central auditory processing functions and preserved cognitive abilities such as auditory attention and working memory, which may contribute to their comparable performance to young adults (*8*, *9*, *11*, *12*, *14*, *15*). Musical training also improves the neural specificity of young adults' speech representations in both auditory and speech motor regions [i.e., inferior frontal gyrus (IFG) and ventral precentral gyrus (prCG)] (*16*). Notably, neural dedifferentiation, i.e., declined specificity of neural representations, is less observed in speech motor regions than in auditory areas in older adults during speech-in-noise perception (*17*). Therefore, OM appear to maintain relatively intact speech representations, especially in speech motor areas, as young adults do, whereas ONM may suffer from degraded neural representations of speech sounds.

OM' advantage in mitigating the aging effect on audiovisual speech-in-noise perception may also arise from functional compensation. That is, OM might recruit additional sensory and/or frontal regions to compensate. When processing audiovisual stimuli, older adults tend to maximize the use of multimodal information to mitigate the decline in unimodal perception (*18*, *19*). This strategy might be amplified by musical training, which emphasizes multisensory (auditory, visual, and motor) integration (*12*–*16*). For instance, young musicians show more robust auditory-motor integration during speech-in-noise perception (*16*); stronger

[1]CAS Key Laboratory of Behavioral Science, Institute of Psychology, Chinese Academy of Sciences, Beijing 100101, China. [2]Department of Psychology, University of Chinese Academy of Sciences, Beijing 100049, China. [3]Rotman Research Institute, Baycrest Centre for Geriatric Care, Toronto, ON M6A 2E1, Canada. [4]Department of Psychology, University of Toronto, ON M8V 2S4, Canada. [5]CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai 200031, China. [6]Chinese Institute for Brain Research, Beijing 102206, China.
*Corresponding author. Email: duyi@psych.ac.cn

audiovisual integration (i.e., narrower temporal binding window) when processing flash-beep, music, and speech stimuli (20–22); and more robust brainstem response to audiovisual speech stimuli than young nonmusicians (YNM) (23). Moreover, older adults recruit more frontal regions (e.g., IFG, anterior insula, and prCG) than young adults when perceiving speech in adverse conditions (17, 24–26), which is consistent with the posterior-anterior shift in aging and decline-compensation hypothesis (27, 28). The greater recruitment of frontal areas might represent a compensatory scaffolding for greater listening effort, cognitive control, and sensorimotor integration. Besides, older adults with higher frontal activation show better speech specificity in both frontal and auditory regions, suggesting that frontal speech motor up-regulation could also provide functional reserve for older adults (17).

Compensation may also be reflected in stronger deactivation of default mode network (DMN) regions. DMN is a set of widely distributed brain regions, showing reductions in activity during attention-demanding tasks but increased activity during tasks linked to memory (29–31). The enhanced activity in the older adults' DMN is thought to reflect disinhibition and has been shown to be detrimental to task performance (32, 33). Notably, musicians show stronger structural and resting-state functional connectivity in DMN regions than nonmusicians (34–36). Musical training appears to strengthen the brain's executive function (37), which, in turn, could be reflected in OM' DMN activity. That is, when performing the speech-in-noise task, OM may be able to minimize interference from task-unrelated long-term memory (e.g., intrusive thought and mind wandering) and focus their attention on external auditory information, which may be associated with greater suppression in DMN regions.

In this functional magnetic resonance imaging (fMRI) study, we tested three hypotheses about how life span musical experience mitigates the age-related speech-in-noise perception deficits: (i) OM show preserved speech representation patterns that are similar to young adults, (ii) OM recruit additional sensory and frontal speech motor regions as a compensation, and (iii) OM deactivate DMN regions to a greater extent as a compensation. Twenty-five OM, 25 ONM, and 24 YNM (see Table 1 for details) identified syllables in noise under three signal-to-noise ratios (SNRs) with congruent visual lip movements inside the scanner. As expected, ONM performed worse than YNM, but OM performed better than ONM and equally well as YNM. First, we used multivariate pattern analysis (MVPA) and recently developed neural alignment measurement defined by intersubject spatial pattern correlation (38) between each older individual and young brain average to examine neural dedifferentiation in ONM and the youth-like brain activity pattern in OM. Results revealed degraded speech

representations in bilateral sensorimotor areas in ONM compared to YNM but reserved neural specificity of speech representations and high neural alignment to young in the same areas in OM. Higher neural alignment in sensorimotor areas significantly predicted better speech-in-noise performance in ONM. Next, we conducted univariate analysis and found that OM had stronger activation in frontal-parietal regions and greater deactivation in left angular gyrus (AG; a core DMN region) than ONM. Moreover, greater deactivation of AG correlated with better behavioral performance in ONM, while higher activation in frontal-parietal areas and stronger AG deactivation supported the youth-like activity patterns in sensorimotor regions. Our findings thus provide a strong linkage between functional preservation and compensation underlying aging and musical plasticity and highlight the preservation of similar activity patterns in bilateral sensorimotor regions as in young adults as a central mechanism in supporting speech-in-noise perception in OM.

RESULTS

Musical expertise mitigates age-related decline in audiovisual speech-in-noise perception

Figure 1 shows the group mean accuracy in identifying the speech sound as a function of SNR. The mixed-design analysis of variance (ANOVA) revealed significant main effects of group ($F_{2,71} = 22.89$, $P < 0.001$) and SNR ($F_{2,71} = 133.94$, $P < 0.001$) and their interaction ($F_{4,71} = 7.10$, $P < 0.001$) in perceiving audiovisual speech-in-noise. Post hoc analysis showed that ONM performed significantly worse than YNM at all SNRs [−8 dB: $t(71) = 6.26$, $P < 0.001$; 0 dB: $t(71) = 5.91$, $P < 0.001$; and 8 dB: $t(71) = 5.86$, $P < 0.001$], and OM significantly outperformed ONM at all SNRs [−8 dB: $t(71) = 2.60$, $P = 0.034$; 0 dB: $t(71) = 4.71$, $P < 0.001$; and 8 dB: $t(71) = 5.32$, $P < 0.001$; all Bonferroni corrected]. Although OM still performed significantly worse than YNM at the lowest SNR [−8 dB: $t(71) = -3.69$, $P = 0.001$], OM achieved a similar level of performance as YNM at moderate to high SNR levels [0 dB: $t(71) = 1.25$, $P = 0.648$ and 8 dB: $t(71) = 0.59$, $P = 1$; all Bonferroni corrected]. This suggests that musical expertise can mitigate age-related decline in audiovisual speech-in-noise perception, especially when the noise is not overwhelming. In OM, the correlation between years or intensity (practice hours per week) of musical training and behavioral performance was not significant [correlation coefficient ($|r|$) < 0.36, $P > 0.09$].

Table 1. The group mean (SD) values and statistics of age, education, PTA at 250 to 4000 Hz, the MoCA score, age of training onset, and years of musical training in each group. Independent two-sample t tests were used for examining the group difference between older groups. PTA, pure tone average; MoCA, Montreal Cognitive Assessment; OM, older musicians; ONM, older nonmusicians; YNM; young nonmusicians; NA, not applicable.

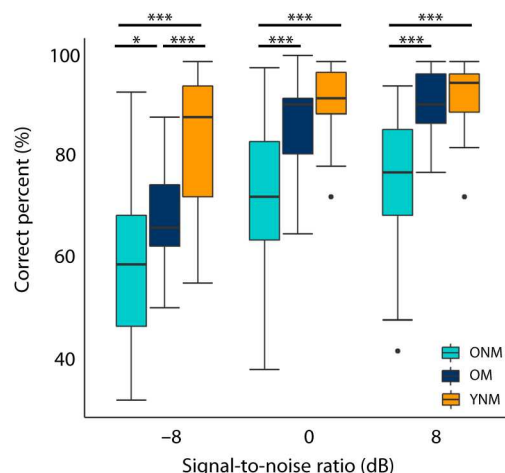| Group	Age	Education	PTA	MOCA	Age of onset	Years of training
OM	65.12 (4.06)	13.02 (2.96)	12.38 (5.39)	27.92 (1.19)	10.90 (4.56)	50.88 (8.75)
ONM	66.64 (3.40)	9.50 (3.02)	12.58 (3.58)	27.52 (1.39)	NA	NA
t (P)	−1.43 (0.158)	4.15 (<0.001)	−0.15 (0.881)	1.09 (0.280)		
YNM	23.13 (2.38)	16.50 (1.67)	0.71 (3.32)	NA	NA	NA

Fig. 1. Behavioral performance. Older musicians (OM) outperformed older nonmusicians (ONM) on audiovisual syllable-in-noise perception at all signal-to-noise ratios (SNRs = −8, 0, and 8 dB) and performed equally well as young nonmusicians (YNM) at moderate to high SNRs (0 and 8 dB). *$P < 0.05$ and ***$P < 0.001$ by post hoc analysis of mixed-design analysis of variance (ANOVA), Bonferroni corrected.

Neural representations of audiovisual syllables are degraded in ONM but preserved in OM in visual and sensorimotor areas

After showing that OM discriminated syllables equally well as YNM but better than ONM, we determined in which regions OM would show similar neural representations as YNM but better neural representations than ONM. We conducted whole-brain searchlight classification analysis and then compared the classification accuracy across groups to reveal the regions that showed similar or different neural specificity of phonemes. The interaction between group and SNR was not significant, and the primary interest of this study was the group effect on audiovisual speech-in-noise processing. Hence, we focused on the main effect of group on neural activity in the following analysis, and the main effect of SNR was not reported further, although it was significant ($P_{fwe} < 0.05$).

Compared to YNM, ONM showed lower neural specificity of audiovisual phoneme representations in left visual regions, bilateral sensorimotor areas including prCG, postcentral gyrus (poCG), and supplementary motor area ($P_{fwe} < 0.05$; Fig. 2, B and D, and table S1). However, no significant difference was found between OM and YNM. Meanwhile, OM showed better phoneme specificity than ONM in left visual areas and speech motor regions, including the ventral premotor area in prCG and the opercular part of IFG (IFGop) ($P_{fwe} < 0.05$; Fig. 2, A and C, and table S1). Region of interest (ROI) analysis confirmed lower classification accuracy in ONM than YNM and OM in all significant clusters (all $t > 2.78$, $P < 0.021$, Bonferroni corrected) and no significant difference was found between YNM and OM (all $t < 2.06$, $P > 0.13$, Bonferroni corrected). Note that the classification accuracy in each mentioned region was significantly above the chance level in each group (Fig. 2, C and D).

OM show higher neural alignment to young brains than ONM in bilateral sensorimotor areas

Because OM showed equivalent classification accuracy as YNM and both showed higher classification accuracy than ONM, we

investigated whether musicianship mitigated the aging effect by preventing age-related declines in sensorimotor regions (i.e., preservation account). We hypothesized that compared with ONM, OM would exhibit more similar activation patterns to YNM. To test this hypothesis, we calculated the intersubject spatial pattern correlation between the activation pattern of each older adult and the mean activation pattern of young adults for each audiovisual syllable in the brain regions (see Fig. 3A and Materials and Methods for details) that showed significantly higher classification accuracy in YNM or OM than ONM (Fig. 2). This index measures the spatial neural alignment of multivoxel activation patterns of the older brain to that observed in younger brains. As shown in Fig. 3B, we found significantly higher neural alignment to young in OM than ONM in bilateral sensorimotor areas [regions showing better neural representations of syllables in YNM than ONM, left sensorimotor area: $t(46) = 2.61$, $P_{fdr} = 0.039$ and right sensorimotor area: $t(46) = 2.51$, $P_{fdr} = 0.039$], but not in left visual regions or speech motor regions (all $|t| < 1.09$, $P_{fdr} > 0.470$).

Higher neural alignment to young in bilateral sensorimotor areas predicts better audiovisual speech-in-noise performance

In older adults, we tested whether higher neural alignment to young would predict better behavioral performance by calculating the correlation between neural alignment and audiovisual speech-in-noise performance (Fig. 3C). Because neural alignment in left and right sensorimotor areas were highly correlated across older adults ($r = 0.67$, $P < 0.001$), we performed principal components analysis (PCA) to extract the component that best represented neural alignment in bilateral sensorimotor areas (for details, see Materials and Methods). This PC entered the latter correlation analysis.

We found that the neural alignment to young in sensorimotor areas positively predicted audiovisual speech-in-noise performance in ONM ($r = 0.57$, $P = 0.004$; Fig. 3C) but not in OM ($r = 0.13$, $P = 0.53$; Fig. 3C). The lack of relationship between neural alignment and performance in OM could be accounted for by the small variance in the OM group. Furthermore, musical training intensity during the last 3 years was positively correlated with neural alignment to young in bilateral sensorimotor areas ($r = 0.48$, $P = 0.017$; Fig. 3D). These findings suggest that musicianship may help older adults maintain a higher level of youth-like neural activity patterns in bilateral sensorimotor regions, which supports speech-in-noise perception.

OM show stronger activation in frontal-parietal regions and stronger deactivation in AG compared to ONM

To determine whether musicianship might also mitigate the aging effects via compensatory strategy, we compared the task-related blood oxygenation level–dependent (BOLD) responses between groups. This analysis reveals regions showing differences in neural activation and deactivation during audiovisual speech perception. Compared to ONM, OM showed stronger activation in left frontal-parietal regions, including the auditory-motor interfaces in the posterior superior temporal gyrus and supramarginal gyrus, visual dorsal pathway areas including poCG, inferior and superior parietal lobule (IPL and SPL), and speech motor regions including the IFGop and ventral prCG ($P_{fwe} < 0.05$; Fig. 4A and table S2). Musical training intensity of the first 3 years significantly correlated with the activation of the left IFGop/prCG ($r = 0.51$,
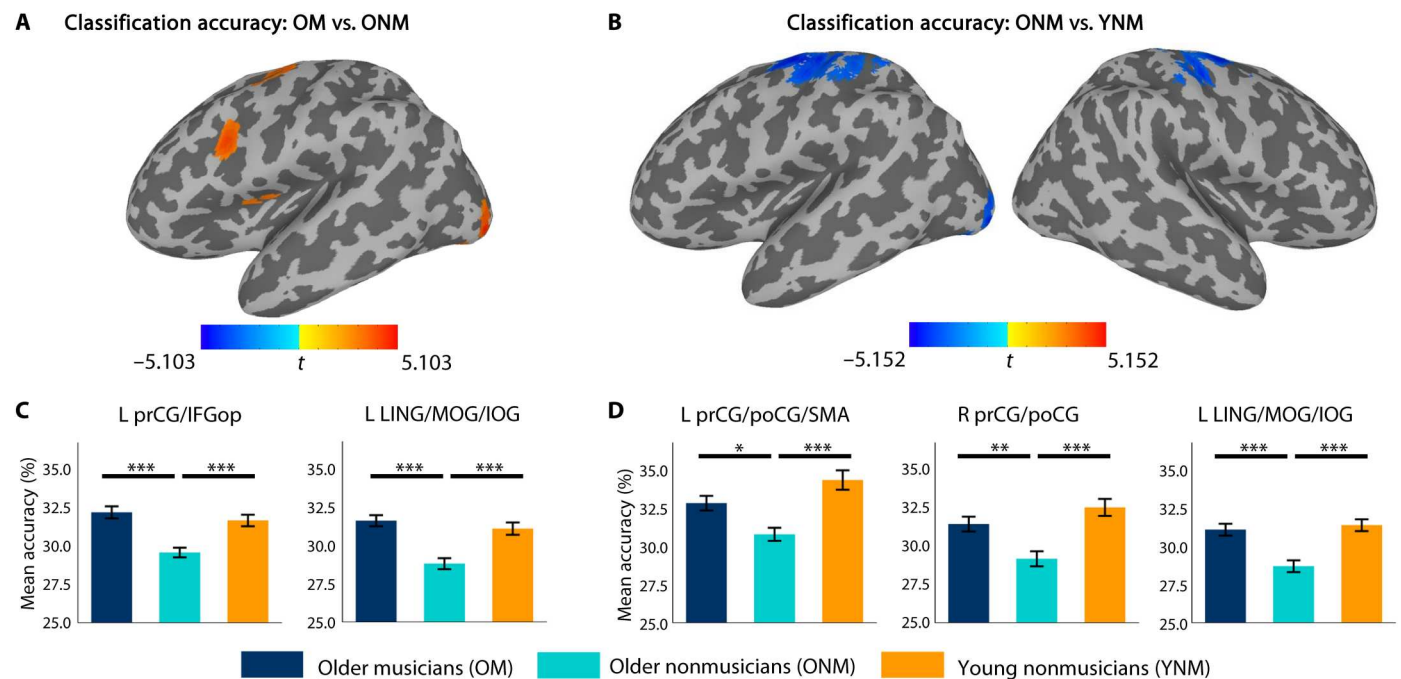
Fig. 2. Neural classification performance. ONM showed degraded neural representation of audiovisual syllables, while OM kept similar neural specificity to YNM. (**A** and **B**) Clusters that showed significant difference ($P_{fwe} < 0.05$) in classification accuracy of syllables between OM and ONM (A) and between ONM and YNM (B). No classification difference was observed between OM and YNM. (**C** and **D**) Bar displays the group mean classification accuracy in regions (i.e., clusters) showing significant group difference [(C): OM versus ONM and (D): ONM versus YNM]. The chance level was 0.25. Error bars indicate the SEM. *$P < 0.05$, **$P < 0.01$, and ***$P < 0.001$ by t tests after Bonferroni correction. L, left; R, right; prCG, precentral gyrus; IFGop, inferior frontal gyrus, opercular part; LING, lingual gyrus; MOG, middle occipital gyrus; IOG, inferior occipital gyrus; poCG, postcentral gyrus; SMA, supplementary motor area.

$P = 0.011$). In addition, OM showed stronger deactivation in the left AG, a typical DMN region, than ONM ($P_{fwe} < 0.05$; Fig. 4A and table S2). OM also showed greater deactivation in the left AG than YNM, but no difference was found between ONM and YNM (fig. S1; see the Supplementary Materials for more information about the aging effects on BOLD activity). ROI analysis confirmed that OM and YNM showed greater activation in frontal-parietal regions than ONM (all $t > 3.19$, $P < 0.006$, Bonferroni corrected). Moreover, OM showed greater deactivation in left AG than both ONM and YNM (both $t > 4.26$, $P < 0.001$, Bonferroni corrected) (Fig. 4A). Thus, long-term musical expertise was associated with the additional engagement of frontal-parietal regions and stronger deactivation of AG in processing audiovisual speech in noise.

Stronger AG deactivation correlates with better audiovisual speech-in-noise performance

We examined whether the changes of activation/deactivation in the above regions were valid mechanisms to mitigate the aging effect on performance by calculating the correlation between regional activation/deactivation and speech-in-noise performance. There was no significant correlation between activation in speech motor regions and speech-in-noise performance in the ONM group ($r = 0.20$, $P_{fdr} = 0.343$) or in the OM group ($r = -0.38$, $P_{fdr} = 0.124$). In both groups, the correlation between speech-in-noise performance and activation in auditory-motor interfaces was not significant (both $|r| < 0.16$, $P_{fdr} > 0.596$) nor was the correlation with the visual dorsal pathway areas (both $|r| < 0.24$, $P_{fdr} > 0.457$). However, deactivation in the left AG was negatively correlated with speech-in-noise

performance in the ONM group ($r = -0.62$, $P_{fdr} = 0.005$) but not in the OM group ($r = 0.46$, $P_{fdr} = 0.093$; Fig. 4B). The correlation between DMN deactivation and behavioral performance suggests that DMN deactivation indexes a compensatory mechanism for speech comprehension in older adults under adverse listening conditions.

Higher neural alignment to young in sensorimotor areas is correlated with stronger frontal-parietal activation and stronger AG deactivation

Last, we investigated whether the functional preservation (neural alignment to young) in sensorimotor areas and compensatory scaffolding of frontal speech motor and parietal regions were related or independent mechanisms. We performed correlation analyses to examine the relationships between the neural alignment index and BOLD activation in frontal speech motor and parietal regions that showed significant group effect. We found that stronger deactivation in AG was correlated with higher neural alignment to young in sensorimotor areas in the ONM group ($r = -0.66$, $P_{fdr} = 0.001$; Fig. 4C, right) but not in the OM group ($r = -0.28$, $P_{fdr} = 0.251$). A similar pattern was also observed between speech motor (IFGop/prCG) activation and neural alignment in ONM ($r = 0.50$, $P_{fdr} = 0.017$; Fig. 4C, left) but not in OM ($r = 0.12$, $P_{fdr} = 0.588$), while activation in left poCG/SPL and neural alignment were positively correlated in both groups (ONM: $r = 0.65$, $P_{fdr} = 0.001$ and OM: $r = 0.53$, $P_{fdr} = 0.031$; Fig. 4C, middle). Therefore, stronger AG deactivation and stronger frontal-parietal engagement supported older adults in maintaining better neural alignment to young in
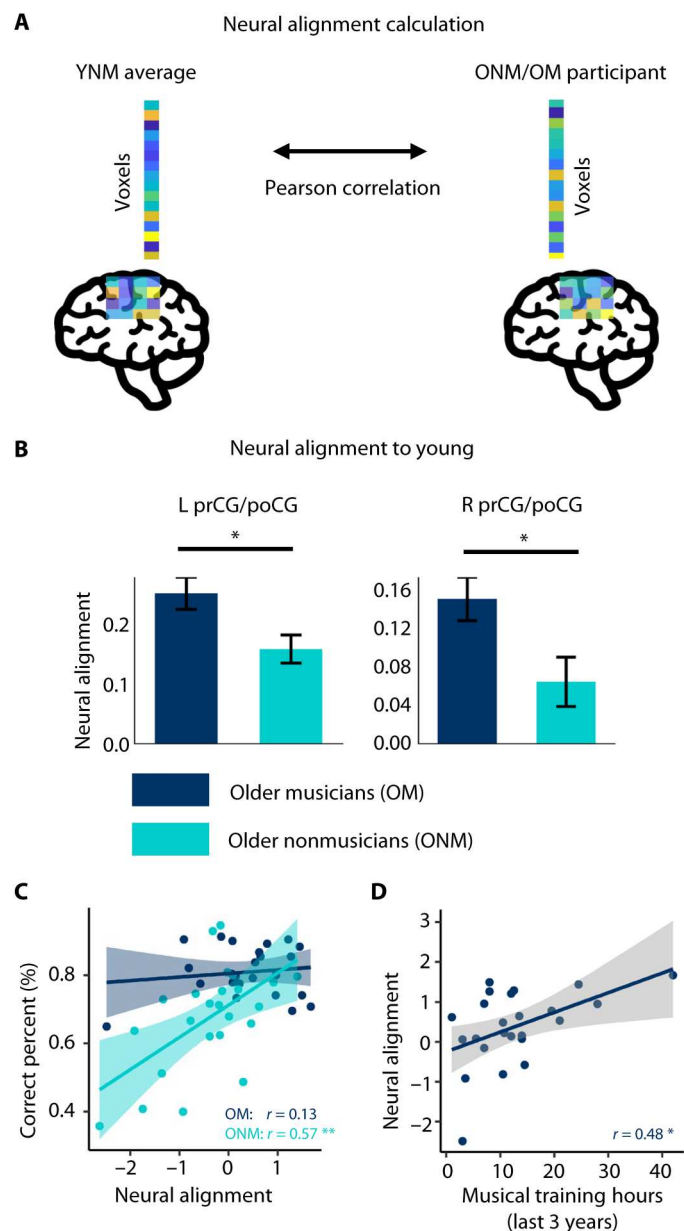
Fig. 3. Neural alignment measurement. OM had higher neural alignment to young than ONM in sensorimotor areas. (**A**) Calculation of neural alignment to young. For each region with significant multivariate pattern analysis (MVPA) group difference, correlation between the mean spatial activation pattern of young adults and the pattern of each older adult was calculated for each syllable at each SNR and averaged across conditions to derive the neural alignment-to-young index. (**B**) OM showed higher neural alignment to young than ONM in bilateral sensorimotor areas where significant group difference was found between ONM and YNM. Error bars indicate the SEM. *$P < 0.05$ by two-sample t tests after false discovery rate (FDR) correction. (**C**) Correlation between neural alignment and behavioral performance. (**D**) Correlation between neural alignment and average weekly musical training hours over the last 3 years. *$P < 0.05$ and **$P < 0.01$ by Pearson's correlation.

sensorimotor regions, which, in turn, predicted better audiovisual speech-in-noise performance.

DISCUSSION

This study demonstrates that long-term musicianship mitigates age-related deficits in audiovisual speech-in-noise perception through both functional preservation and compensation. In line with and extending previous studies using audio-only speech stimuli (*8–10, 12, 39*), we found that lifelong music playing lessens age-related declines in audiovisual speech-in-noise perception. OM showed equivalent phoneme representations in visual and sensorimotor regions to that of YNM, while the phoneme representations in these regions in the ONM were dedifferentiated. To our knowledge, we are the first to show that compared to ONM, OM showed greater similar representation patterns in bilateral sensorimotor areas to those of YNM. Higher neural alignment to young brains in these regions was strongly correlated with better behavioral performance in ONM. OM also showed stronger activation in frontal-parietal regions and stronger deactivation in left AG than ONM. The deactivation of AG was correlated with behavioral performance, while both greater activation and greater deactivation predicted higher neural alignment in sensorimotor regions. Our findings thus link compensatory scaffolding in frontal speech motor and DMN regions to the functional preservation of youthful representation patterns in bilateral sensorimotor areas, which ultimately serves as a core mechanism of how musical expertise dampens the effects of aging in speech perception.

Our results reveal that OM had better speech-in-noise perception performance through functional preservation, which is maintaining similar speech representation patterns as young adults. It has been shown that amateur musicians have younger "brain age," which is defined by a machine learning algorithm compared to nonmusicians (*40*). As expected, compared with YNM, ONM showed decreased neural specificity of phoneme representation (i.e., neural dedifferentiation) (*3, 4*) in bilateral sensorimotor regions. However, in the same regions, OM showed similar neural specificity of phonemes to YNM, and OM showed higher neural alignment to young than ONM. The motor and somatosensory cortexes are involved not only in speech production but also involved in speech perception (*41, 42*). Evidence from neuroimaging studies suggests that phonemic features that facilitate speech perception in adverse listening conditions are represented in the motor and somatosensory cortex (*41, 43, 44*). We found that musical training was associated with reduced age-related neural dedifferentiation and that OM maintained more similar activity patterns as young adults than ONM in sensorimotor areas. Greater neural alignment to young adults in sensorimotor regions was strongly correlated to better performance in ONM, supporting that older adults may rely more on the functional reserve of key brain regions in maintaining speech perception function. These results thus support the functional preservation hypothesis that higher neural alignment to young gives rise to better performance in older adults.

Our results also show that OM had better speech-in-noise perception performance by functional compensation in two manners. The first compensation mechanism is to recruit frontal-parietal regions. According to the STAC model, neural resources are depleted with aging, while compensatory scaffolding provides additional protection against cognitive decline (*1*). In this study, compared to
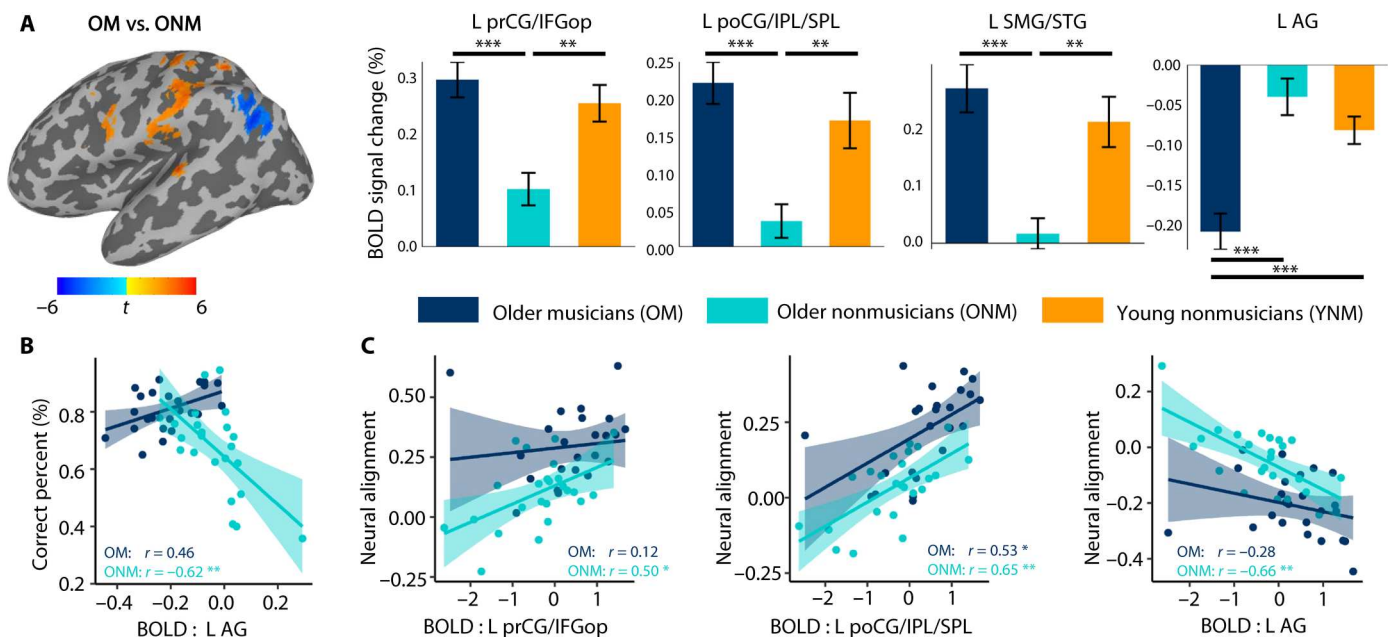
Fig. 4. BOLD activation pattern. OM showed greater activation in frontal-parietal areas and greater deactivation in angular gyrus (AG) compared to ONM. (**A**) Clusters that showed significant differences in blood oxygenation level–dependent (BOLD) activity between OM and ONM ($P_{fwe} < 0.05$). Bar plots show the group mean BOLD signal changes in specific regions. Error bars indicate the SEM. **P < 0.01 and ***P < 0.001 by t tests after Bonferroni correction. (**B**) Correlation between BOLD activity in left AG and behavioral performance. (**C**) Correlation between BOLD activity in left frontal-parietal and AG regions and neural alignment. *P < 0.05 and **P < 0.01 by Pearson's correlation after FDR correction. IPL, inferior parietal lobule; SMG, supramarginal gyrus; SPL, superior parietal lobule; STG, superior temporal gyrus.

ONM, OM exhibited higher activation in frontal-parietal regions belonging to the auditory and visual dorsal stream and better audio-visual consonant-vowel representations in speech motor areas. The visual dorsal stream is regarded as a "vision-for-action" stream that mediates motor programming and control based on current visual information (*45, 46*), while the dorsal auditory stream is considered a "sound-to-action" stream that maps auditory features to motor articulatory sequences (*47*). Speech perception under adverse conditions recruits speech motor areas in the dorsal stream, including the premotor cortex in prCG and IFG, to compensate for deteriorated auditory processing via sensorimotor integration (*44*). With visual lip movements, better speech encoding and tightened functional connectivity along the auditory and visual dorsal stream contribute to enhanced speech-in-noise perception (*48, 49*). Because musical training requires intensive interaction between sensory and motor modalities, musicians are found to exhibit stronger sensorimotor integration than nonmusicians during perception (*16, 50–55*). Therefore, our results suggest that OM better use internal motor plans and external visual lip movements to aid speech-in-noise perception through strengthened visual/sound-to-action mapping and sensorimotor integration. However, activity in speech motor areas was not correlated with behavioral performance in either of the older groups. This suggests that other brain areas may play a role in compensating task performance in older adults.

The second compensation mechanism may be indexed by the deactivation of a core DMN region, AG. To our knowledge, this is the first report of greater deactivation in left AG in OM than ONM during a speech-in-noise task. The DMN regions, which are involved in internal cognitive processes related to mind wandering, memory, and abstract thought, are found to be typically deactivated when participants perform external goal-directed tasks (*29*),

with an anticorrelation between DMN activity and task performance (*56*). DMN deactivation has been associated with active suppression of task-irrelevant sensory information rather than a relative increase in metabolic activity during rest (*57, 58*). DMN deactivation has been found to support attentional reallocation to external stimuli, which is critical for goal-directed activities (*30, 59*). Here, the higher deactivation level in AG was correlated with better performance for ONM. That is, DMN deactivation may index older adults' ability to allocate attentional resources to external speech stimuli and reduce internal interference, thereby achieving better performance. Overall, stronger DMN deactivation during audiovisual speech processing is consistent with a compensatory scaffolding strategy against aging, which is enhanced by lifelong musical experience.

Furthermore, rather than demonstrating multiple neural mechanisms associated with lifelong musicianship separately, this study clarified the relationships among frontoparietal activation, DMN deactivation, youth-like neural activity patterns in sensorimotor areas, and behavioral performance. We found that stronger activation in speech motor areas and stronger deactivation in AG were positively correlated with better alignment to young in ONM. Thus, the additional engagement of frontal speech motor areas and the DMN deactivation coincide to maintain neural alignment-to-young in sensorimotor areas, which predicts audiovisual speech-in-noise performance in older adults. We failed to find a significant correlation between the BOLD activity and the alignment-to-young or between the alignment and behavior in OM. This may be due to the small variance of the alignment index in OM. Our findings are in line with previous studies showing that: (i) different people, i.e., "successful" and "average" agers, may be associated with different patterns of neurocognitive aging; and (ii) different brain

regions within the same individual could exhibit different patterns of brain aging (28, 60). Consistent with the STAC-r model (1), long-term musical training supports both functional preservation in sensorimotor areas and compensatory scaffolding in frontoparietal and DMN regions. Moreover, we extend the model by showing the correlation, rather than independence, between functional preservation and compensatory scaffolding, with the former mechanism being more central.

There are several limitations in this study. First, we used a cross-sectional study design and, consequently, we cannot rule out the possibility that musicians may have better predisposition to musical training, which could partly account for the group difference in performance and brain activity (61, 62). In addition, the number of years of education was not perfectly balanced. Although in the current study the number of years of education did not significantly influence behavioral performance and brain responses to audiovisual speech, the level of education may protect the aging brain in other ways (1). Therefore, a longitudinal study needs to be carried out in the future to examine the causal effect of musical expertise. Second, because OM who were eligible and willing to participate in our fMRI study are rare, OM with different types of musical expertise were combined into one group (e.g., piano, singing, violin, etc.). However, comparing the effect of different types of lifelong musical training that older adults major in could further reveal how musical expertise promotes multisensory speech-in-noise perception. More specifically, different training types likely require different kinds of sensorimotor integration (e.g., singing engages articulation system, playing piano engages fingers, etc.). Hence, we expect that vocalists may have better speech-in-noise perception, since the brain regions engaged in singing and speaking are highly overlapped (63) and percussionists may also perform better in continuous audiovisual speech-in-noise perception because of their superior rhythm skills (64, 65). Moreover, further studies should investigate the interaction effect between musical expertise and tone language background, because evidence from a recent study suggests that the effect of musical expertise on speech processing is modulated by listeners' tone language experience (66). Last, in the current study, only local activation and speech representation were analyzed. To further unravel experience-related functional alterations of the neural networks during audiovisual speech-in-noise perception in older adults, the group difference in functional connectivity needs to be investigated in further studies. We will focus on the functional connectivity between sensorimotor areas that showed functional preservation and frontoparietal and DMN regions that showed functional compensation by musical expertise.

In summary, this study took a hypothesis-driven approach to assess not just the effects of long-term musical training on multisensory speech perception in noise in older adults but also the possible underlying mechanisms. Our findings provide an integrated view showing that lifelong musical experience mitigates age-related declines in audiovisual speech-in-noise perception by preserving youth-like speech representation patterns in bilateral sensorimotor regions and enhancing compensatory scaffolding through frontal-parietal up-regulation and DMN deactivation. The functional preservation of youth-like representation patterns in sensorimotor areas serves as a key mechanism, with activity patterns more similar to those of young adults associated with better performance. Moreover, the functional preservation is closely interlinked with compensatory scaffolding, with additional recruitment of frontal speech motor areas and DMN deactivation contributing to better maintenance of youthful activity patterns in sensorimotor areas. Our findings provide insights into adaptive brain reorganization in the aging populations and how lifelong musical training experience leads to "successful aging." The functional preservation of sensorimotor regions, along with compensatory DMN deactivation, also suggests avenues for more targeted training regimens to protect speech functions in the elderly.

MATERIALS AND METHODS
Participants
Twenty-five OM (65.12 ± 4.06 years old, 11 female), 25 ONM (66.64 ± 3.40 years old, 16 female), and 24 YNM (23.13 ± 2.38 years old, 12 female) completed this study. One OM with excessive head motion (more than 50% of data were censored in the 3dDeconvolve) and one ONM who was left-handed were excluded from the fMRI analysis. The rest of the participants were all healthy, right-handed, and native Mandarin speakers with no history of neurological disorder and normal hearing (average pure tone threshold <20-dB hearing level from 250 to 4000 Hz) in both ears. OM were recruited from the conservatory of music, chorus, and orchestras. OM had started training before 23 years old (mean = 10.90 ± 4.56 years old), had at least 32 years of training (mean = 50.88 ± 8.75 years), and practiced consistently in recent 3 years (1 to 42 hours per week, mean = 12.70 ± 8.99 hours per week). Nonmusicians reported less than 2 years of musical training experience. All older adults passed the Montreal Cognitive Assessment (MoCA) of the Beijing version (≥26 scores) (67). Age, pure tone average (PTA), and MoCA score were balanced between OM and ONM (see Table 1). However, OM showed more years of education than ONM ($t = 4.15$, $P < 0.001$). Note that years of education of older adults here included informal education like on-the-job training. We conducted a few supplementary analyses to exclude the effect of education on our results (for details, see Supplementary Text of the Supplementary Materials). All participants had signed the written consent before the measurement, and all procedures were approved by the ethical committee of the Institute of Psychology, Chinese Academy of Sciences.

Experimental design
The audiovisual stimuli comprised four naturally pronounced consonant-vowel syllables (ba, da, pa, and ta) uttered by a young Chinese female. The utterances were videoed in a soundproof room. Videos of each syllable lasted 1 s and were digitized at 29.97 frames per second in 1024-by-768 pixels. The pictures of the videos were cut, retaining the mouth and the neck part. The audios were low-pass filtered (4 kHz), matched for average root mean square sound pressure level, and aligned in time by lip movement onset (started around the 10th frame of the video). The audiovisual syllables were masked by a speech spectrum-shaped noise (4-kHz low-pass, 10-ms rise-decay envelope) that represents the spectrum of 113 different sentences by 50 Chinese young female speakers at three SNRs: −8, 0, and 8 dB (44). In the fMRI scanner, participants were instructed to listen to the speech signals, watch the mouth on the screen, and identify the syllables by pressing the corresponding button using their right-hand fingers (index to little fingers in response to ba, da, pa, and ta in half of the

participants as well as pa, ta, ba, and da in the other half of the participants sequentially). Each participant completed four blocks. Each block contained 60 stimuli (20 trials × 3 SNRs), which were pseudo-randomly presented with an average interstimuli interval of 5 s (4 to 6 s, 0.5-s step). Stimuli were presented via Psychtoolbox (68).

Behavioral analysis

To examine the group difference in behavioral performance, a three-by-three mixed-design ANOVA was performed (within-subject variable: SNR and between-subject variable: subject group). Post hoc analyses were performed to further investigate the differences between the two groups. The Bonferroni correction was applied to correct for multiple comparisons. Statistical analysis was conducted in R (69) with the package bruceR (70) and visualized using the package ggplot2 (71).

Functional imaging data acquisition and preprocessing

Imaging data were collected by a 3-T MRI system (Siemens Magnetom Trio). T_1-weighted images were acquired using the magnetization-prepared rapid gradient-echo (MPRAGE) sequence [repetition time (TR) = 2200 ms, echo time (TE) = 3.49 ms, field of view (FOV) = 256 mm, and voxel size = 1 by 1 by 1 mm]. T_2-weighted images were acquired using the multiband-accelerated EPI sequence (acceleration factor = 4, TR = 640 ms, TE = 30 ms, slices = 40, FOV = 192, and voxel sizes = 3 by 3 by 3 mm).

The fMRI data were preprocessed using Analysis of Functional NeuroImages (AFNI) software (72). The first eight volumes were removed. For univariate analysis, the following preprocessing steps included slice timing, motion correction, aligning the functional image with anatomy, spatial normalization (MNI152 space), spatial smoothing with a 6-mm full width at half maximum (FWHM) isotropic Gaussian kernel, and scaling each voxel time series to have a mean of 100. The fMRI data were not smoothed and scaled for MVPA and neural alignment analysis in the preprocessing steps.

Searchlight MVPA

An exploratory whole-brain searchlight MVPA with a 6-mm radius was conducted to decode the four syllables under each SNR using the Decoding Toolbox (73). The classifiers were trained using a support vector machine algorithm with a linear kernel. The cost parameter C was set to 1. The input feature was univariate trial-wise β coefficients that were estimated using AFNI program 3dLSS, which was recommended for performing MVPA in fast event-related designs (74). The leave-one-run-out cross-validation was used to evaluate classification performance, which was measured by mean classification accuracy. Then, accuracy maps were smoothed with an 8-mm FWHM Gaussian kernel. The smoothed accuracy maps were entered into the group analysis using the AFNI program 3dMVM. Group-level analysis was performed using the AFNI program 3dMVM. To investigate how aging affects the neural representation of audiovisual syllables among musicians and nonmusicians, two separate models were constructed with the same within-subject factor (SNR) and different between-subject factors (ONM versus YNM and OM versus YNM). To examine the musical training effect among older adults, one within-subject factor (SNR) and one between-subject factor (subject group: OM versus ONM) were put into the model. Multiple comparisons were corrected

using 3dClustSim ("fixed" version) with real smoothness of data estimated by 3dFWHMx (acf method) (75). A total of 10,000 Monte Carlo simulations were performed to get the cluster threshold (α = 0.05 family-wise error corrected). Results were visualized onto an inflated cortical surface using surface mapping with AFNI (SUMA).

Neural alignment-to-young analysis

The same single-subject multiple-regression modeling procedure as the univariate analysis was performed using functional data that were not smoothed and scaled. T-statistic maps of each syllable under three SNRs for each participant were used in the following neural alignment analysis.

Five significant clusters in the MVPA group analysis were used as ROIs in the neural alignment analysis. The neural responses (t value) of each voxel within the ROIs to each syllable under three SNRs were extracted for each participant. Then, the neural responses of each voxel within the ROIs were averaged across all YNM, which yielded the mean activation pattern of young adults to each syllable under three SNRs. Then, for each older participant and each activation pattern within the ROI, we calculated the alignment-to-young measure, which is the spatial activation similarity between each older participant and young average when they perceive each audiovisual syllable, using Pearson correlation (see Fig. 3A). Twelve correlation coefficients (four syllables * 3 SNRs) were averaged for each older participant. Such an intersubject pattern correlation framework has been used in previous studies related to memory and learning (76–78). The alignment-to-young measure of OM and ONM were compared using the two-sample t test. False discovery rate (FDR) correction was performed to correct multiple comparisons.

Univariate general linear model analysis

Single-subject multiple-regression modeling was performed using the AFNI program 3dDeconvolve. Four syllables under three SNRs and six regressors corresponding to motion parameters were entered into the analysis. TRs were censored if the motion derivatives exceeded 0.3. For each SNR, the four syllables were grouped and contrasted against the baseline. The group analysis was the same as the procedure in the multivariate analysis.

Conjunction analysis was performed to determine the regions that showed common and different aging effects between musicians and nonmusicians. Corrected group analysis results masks (ONM versus YNM and OM versus YNM) were entered into the conjunction analysis, and the conjunction map was created by 3dcalc.

Correlation analysis

Pearson correlation analyses were conducted among BOLD responses, alignment-to-young measurement, and speech-in-noise performance in two older groups. Behavioral performance and neural index in the ROIs that exhibited significant group differences between OM and ONM in BOLD activity or alignment-to-young entered correlation analysis. Because the neural alignment to young in bilateral sensorimotor areas across older adults were highly correlated ($r = 0.67$, $P < 0.001$), PCA with the varimax rotation was used to obtain the first PC that represented the neural alignment to young in bilateral sensorimotor areas, and the PC entered the correlation analysis. FDR correction was performed for the multiple comparisons.

Supplementary Materials

This PDF file includes:
Supplementary Text
Fig. S1
Tables S1 to S4

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. P. A. Reuter-Lorenz, D. C. Park, How does it STAC up? Revisiting the scaffolding theory of aging and cognition. *Neuropsychol. Rev.* **24**, 355–370 (2014).

2. D. C. Park, P. Reuter-Lorenz, The adaptive brain: Aging and neurocognitive scaffolding. *Annu. Rev. Psychol.* **60**, 173–196 (2009).

3. C. Grady, The cognitive neuroscience of ageing. *Nat. Rev. Neurosci.* **13**, 491–505 (2012).

4. J. D. Koen, M. D. Rugg, Neural dedifferentiation in the aging brain. *Trends Cogn. Sci.* **23**, 547–559 (2019).

5. I. Yeend, E. F. Beach, M. Sharma, H. Dillon, The effects of noise exposure and musical training on suprathreshold auditory processing and speech perception in noise. *Hear. Res.* **353**, 224–236 (2017).

6. D. Boebinger, S. Evans, S. Rosen, C. F. Lima, T. Manly, S. K. Scott, Musicians and non-musicians are equally adept at perceiving masked speech. *J. Acoust. Soc. Am.* **137**, 378–387 (2015).

7. A. Parbery-Clark, E. Skoe, C. Lam, N. Kraus, Musician enhancement for speech-in-noise. *Ear Hear.* **30**, 653–661 (2009).

8. C. Alain, B. R. Zendel, S. Hutka, G. M. Bidelman, Turning down the noise: The benefit of musical training on the aging auditory brain. *Hear. Res.* **308**, 162–173 (2014).

9. E. Dubinsky, E. A. Wood, G. Nespoli, F. A. Russo, Short-term choir singing supports speech-in-noise perception and neural pitch strength in older adults with age-related hearing loss. *Front. Neurosci.* **13**, 1153 (2019).

10. M. Perron, J. Vaillancourt, P. Tremblay, Amateur singing benefits speech perception in aging under certain conditions of practice: Behavioural and neurobiological mechanisms. *Brain Struct. Funct.* **227**, 943–962 (2022).

11. B. R. Zendel, G. L. West, S. Belleville, I. Peretz, Musical training improves the ability to understand speech-in-noise in older adults. *Neurobiol. Aging* **81**, 102–115 (2019).

12. L. Zhang, X. Fu, D. Luo, L. Xing, Y. Du, Musical experience offsets age-related decline in understanding speech-in-noise: Type of training does not matter, working memory is the key. *Ear Hear.* **42**, 258–270 (2021).

13. A. D. Patel, Why would musical training benefit the neural encoding of speech? The OPERA hypothesis. *Front. Psychol.* **2**, 142 (2011).

14. G. M. Bidelman, C. Alain, Musical training orchestrates coordinated neuroplasticity in auditory brainstem and cortex to counteract age-related declines in categorical vowel perception. *J. Neurosci.* **35**, 1240–1249 (2015).

15. A. Parbery-Clark, S. Anderson, E. Hittner, N. Kraus, Musical experience offsets age-related delays in neural timing. *Neurobiol. Aging* **33**, 1483. e1–1483. e4 (2012).

16. Y. Du, R. J. Zatorre, Musical training sharpens and bonds ears and tongue to hear speech better. *Proc. Natl. Acad. Sci. U.S.A.* **114**, 13579–13584 (2017).

17. Y. Du, B. R. Buchsbaum, C. L. Grady, C. Alain, Increased activity in frontal motor cortex compensates impaired speech perception in older adults. *Nat. Commun.* **7**, 12241 (2016).

18. A. L. de Dieuleveult, P. C. Siemonsma, J. B. F. van Erp, A.-M. Brouwer, Effects of aging in multisensory integration: A systematic review. *Front. Aging Neurosci.* **9**, 80 (2017).

19. S. A. Jones, U. Noppeney, Ageing and multisensory integration: A review of the evidence, and a computational perspective. *Cortex* 1–23 (2021).

20. H. Lee, U. Noppeney, Long-term music training tunes how the brain temporally binds signals from multiple senses. *Proc. Natl. Acad. Sci. U.S.A.* **108**, E1441–E1450 (2011).

21. H. Lee, U. Noppeney, Music expertise shapes audiovisual temporal integration windows for speech, sinewave speech, and music. *Front. Psychol.* **5**, 868 (2014).

22. G. M. Bidelman, Musicians have enhanced audiovisual multisensory binding: experience-dependent effects in the double-flash illusion. *Exp. Brain Res.* **234**, 3037–3047 (2016).

23. G. Musacchia, M. Sams, E. Skoe, N. Kraus, Musicians have enhanced subcortical auditory and audiovisual processing of speech and music. *Proc. Natl. Acad. Sci. U.S.A.* **104**, 15894–15898 (2007).

24. J. Erb, J. Obleser, Upregulation of cognitive control networks in older adults' speech comprehension. *Front. Syst. Neurosci.* **7**, 116 (2013).

25. P. C. M. Wong, J. X. Jin, G. M. Gunasekera, R. Abel, E. R. Lee, S. Dhar, Aging and cortical mechanisms of speech perception in noise. *Neuropsychologia* **47**, 693–703 (2009).

26. K. I. Vaden Jr., S. E. Kuchinsky, J. B. Ahlstrom, J. R. Dubno, M. A. Eckert, Cortical activity predicts which older adults recognize speech in noise and when. *J. Neurosci.* **35**, 3929–3937 (2015).

27. R. Cabeza, N. D. Anderson, J. K. Locantore, A. R. McIntosh, Aging gracefully: Compensatory brain activity in high-performing older adults. *Neuroimage* **17**, 1394–1402 (2002).

28. S. W. Davis, N. A. Dennis, S. M. Daselaar, M. S. Fleck, R. Cabeza, Que PASA? the posterior-anterior shift in aging. *Cereb. Cortex* **18**, 1201–1209 (2008).

29. M. E. Raichle, The brain's default mode network. *Annu. Rev. Neurosci.* **38**, 433–447 (2015).

30. J. Smallwood, B. C. Bernhardt, R. Leech, D. Bzdok, E. Jefferies, D. S. Margulies, The default mode network in cognition: A topographical perspective. *Nat. Rev. Neurosci.* **22**, 503–513 (2021).

31. Y. Yeshurun, M. Nguyen, U. Hasson, The default mode network: Where the idiosyncratic self meets the shared social world. *Nat. Rev. Neurosci.* **22**, 181–192 (2021).

32. J. Persson, C. Lustig, J. K. Nelson, P. A. Reuter-Lorenz, Age differences in deactivation: A link to cognitive control? *J. Cogn. Neurosci.* **19**, 1021–1032 (2007).

33. F. Sambataro, V. P. Murty, J. H. Callicott, H. Y. Tan, S. Das, D. R. Weinberger, V. S. Mattay, Age-related alterations in default mode network: Impact on working memory performance. *Neurobiol. Aging* **31**, 839–852 (2010).

34. D. M. Bashwiner, C. J. Wertz, R. A. Flores, R. E. Jung, Musical creativity "revealed" in brain structure: Interplay between motor, default mode, and limbic networks. *Sci. Rep.* **6**, 20482 (2016).

35. A. Belden, T. Zeng, E. Przysinda, S. A. Anteraper, S. Whitfield-Gabrieli, P. Loui, Improvising at rest: Differentiating jazz and classical music training with resting state functional connectivity. *Neuroimage* **207**, 116384 (2020).

36. L. Chaddock-Heyman, P. Loui, T. B. Weng, R. Weisshappel, E. McAuley, A. F. Kramer, Musical training and brain volume in older adults. *Brain Sci.* **11**, 50 (2021).

37. L. Moradzadeh, G. Blumenthal, M. Wiseheart, Musical training, bilingualism, and executive function: A closer look at task switching and dual-task performance. *Cognit. Sci.* **39**, 992–1020 (2015).

38. S. A. Nastase, V. Gazzola, U. Hasson, C. Keysers, Measuring shared responses across subjects using intersubject correlation. *Soc. Cogn. Affect. Neurosci.* **14**, 669–687 (2019).

39. B. R. Zendel, C. Alain, Musicians experience less age-related decline in central auditory processing. *Psychol. Aging* **27**, 410–417 (2012).

40. L. Rogenmoser, J. Kernbach, G. Schlaug, C. Gaser, Keeping brains young with making music. *Brain Struct. Funct.* **223**, 297–305 (2018).

41. M. R. Schomers, F. Pulvermüller, Is the sensorimotor cortex relevant for speech perception and understanding? An integrative review. *Front. Hum. Neurosci.* **10**, 435 (2016).

42. J. I. Skipper, J. T. Devlin, D. R. Lametti, The hearing ear is always found close to the speaking tongue: Review of the role of the motor system in speech perception. *Brain Lang.* **164**, 77–105 (2017).

43. S. Evans, M. H. Davis, Hierarchical organization of auditory and motor representations in speech perception: Evidence from searchlight similarity analysis. *Cereb. Cortex* **25**, 4772–4788 (2015).

44. Y. Du, B. R. Buchsbaum, C. L. Grady, C. Alain, Noise differentially impacts phoneme representations in the auditory and speech motor systems. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 7126–7131 (2014).

45. J. P. Rauschecker, Where, when, and how: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex* **98**, 262–268 (2018).

46. A. D. Milner, M. A. Goodale, Two visual systems re-viewed. *Neuropsychologia* **46**, 774–785 (2008).

47. G. Hickok, D. Poeppel, The cortical organization of speech processing. *Nat. Rev. Neurosci.* **8**, 393–402 (2007).

48. L. Zhang, Y. Du, Lip movements enhance speech representations and effective connectivity in auditory dorsal stream. *Neuroimage* **257**, 119311 (2022).

49. J. E. Peelle, B. Spehar, M. S. Jones, S. McConkey, J. Myerson, S. Hale, M. S. Sommers, N. Tye-Murray, Increased connectivity among sensory and motor regions during visual and audiovisual speech perception. *J. Neurosci.* **42**, 435–442 (2022).

50. E. M. Anaya, D. B. Pisoni, W. G. Kronenberger, Visual-spatial sequence learning and memory in trained musicians. *Psychol. Music.* **45**, 5–21 (2017).

51. E. B. J. Coffey, N. B. Mogilever, R. J. Zatorre, Speech-in-noise perception in musicians: A review. *Hear. Res.* **352**, 49–69 (2017).

52. S. C. Herholz, R. J. Zatorre, Musical training as a framework for brain plasticity: Behavior, function, and structure. *Neuron* **76**, 486–502 (2012).

53. X. Li, R. J. Zatorre, Y. Du, The microstructural plasticity of the arcuate fasciculus undergirds improved speech in noise perception in musicians. *Cereb. Cortex* **31**, 3975–3985 (2021).

54. J. V. Strong, B. T. Mast, The cognitive functioning of older adult instrumental musicians and non-musicians. *Neuropsychol. Dev. Cogn. B Aging Neuropsychol. Cogn.* **26**, 367–386 (2019).

55. R. J. Zatorre, J. L. Chen, V. B. Penhune, When the brain plays music: Auditory-motor interactions in music perception and production. *Nat. Rev. Neurosci.* **8**, 547–558 (2007).

56. M. D. Fox, A. Z. Snyder, J. L. Vincent, M. Corbetta, D. C. Van Essen, M. E. Raichle, The human brain is intrinsically organized into dynamic, anticorrelated functional networks. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 9673–9678 (2005).

57. L. J. Stiernman, F. Grill, A. Hahn, L. Rischka, R. Lanzenberger, V. P. Lundmark, K. Riklund, J. Axelsson, A. Rieckmann, Dissociations between glucose metabolism and blood oxygenation in the human default mode network revealed by simultaneous PET-fMRI. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2021913118 (2021).

58. X. Wang, Z. Gao, J. Smallwood, E. Jefferies, Both default and multiple-demand regions represent semantic goal information. *J. Neurosci.* **41**, 3679–3691 (2021).

59. E. Dohmatob, G. Dumas, D. Bzdok, Dark control: The default mode network as a reinforcement learning agent. *Hum. Brain Mapp.* **41**, 3318–3341 (2020).

60. X. Chen, M. M. Rundle, K. M. Kennedy, W. Moore, D. C. Park, Functional activation features of memory in successful agers across the adult lifespan. *Neuroimage* **257**, 119276 (2022).

61. K. Mankel, G. M. Bidelman, Inherent auditory skills rather than formal music training shape the neural encoding of speech. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 13129–13134 (2018).

62. R. J. Zatorre, Predispositions and plasticity in music and speech learning: Neural correlates and implications. *Science* **342**, 585–589 (2013).

63. D. E. Callan, V. Tsytsarev, T. Hanakawa, A. M. Callan, M. Katsuhara, H. Fukuyama, R. Turner, Song and speech: Brain regions involved with perception and covert production. *Neuroimage* **31**, 1327–1342 (2006).

64. J. Slater, A. Azem, T. Nicol, B. Swedenborg, N. Kraus, Variations on the theme of musical expertise: Cognitive and sensory processing in percussionists, vocalists and non-musicians. *Eur. J. Neurosci.* **45**, 952–963 (2017).

65. J. Slater, N. Kraus, The role of rhythm in perceiving speech in noise: A comparison of percussionists, vocalists and non-musicians. *Cogn. Process.* **17**, 79–87 (2016).

66. X. R. Toh, F. Lau, F. C. K. Wong, Individual differences in nonnative lexical tone perception: Effects of tone language repertoire and musical experience. *Front. Psychol.* **13**, 940363 (2022).

67. J. Yu, J. Li, X. Huang, The Beijing version of the montreal cognitive assessment as a brief screening tool for mild cognitive impairment: A community-based study. *BMC Psychiatry* **12**, 156 (2012).

68. D. H. Brainard, The psychophysics toolbox. *Spat. Vis.* **10**, 433–436 (1997).

69. R Core Team, R: A Language and Environment for Statistical Computing (2017); www.r-project.org.

70. H.-W.-S. Bao, bruceR: BRoadly Useful Collections and Extensions of R functions (2020); https://github.com/psychbruce/bruceR.

71. H. Wickham, in *ggplot2: Elegant Graphics for Data Analysis* (Springer-Verlag New York, 2009); http://ggplot2.tidyverse.org.

72. R. W. Cox, AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Comput. Biomed. Res.* **29**, 162–173 (1996).

73. M. N. Hebart, K. Görgen, J.-D. Haynes, The decoding toolbox (TDT): A versatile software package for multivariate analyses of functional imaging data. *Front. Neuroinform.* **8**, 88 (2015).

74. J. A. Mumford, B. O. Turner, F. G. Ashby, R. A. Poldrack, Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* **59**, 2636–2643 (2012).

75. R. W. Cox, G. Chen, D. R. Glen, R. C. Reynolds, P. A. Taylor, FMRI clustering in AFNI: False-positive rates redux. *Brain Connect.* **7**, 152–171 (2017).

76. J. Chen, Y. C. Leong, C. J. Honey, C. H. Yong, K. A. Norman, U. Hasson, Shared memories reveal shared structure in neural activity across individuals. *Nat. Neurosci.* **20**, 115–125 (2017).

77. J. F. Cantlon, R. Li, Neural activity during natural viewing of sesame street statistically predicts test scores in early childhood. *PLoS Biol.* **11**, e1001462 (2013).

78. M. Meshulam, L. Hasenfratz, H. Hillman, Y. F. Liu, M. Nguyen, K. A. Norman, U. Hasson, Neural alignment predicts learning outcomes in students taking an introduction to computer science course. *Nat. Commun.* **12**, 1922 (2021).