



Musicians use speech-specific areas when processing tones: The key to their superior linguistic competence?

Mariacristina Musso^{a,b,*}, Hannah Fürniss^a, Volkmar Glauche^a, Horst Urbach^{a,c},
Cornelius Weiller^{a,b}, Michel Rijntjes^a

^a Department of Neurology, Medical Center, University of Freiburg, Faculty of Medicine, University of Freiburg, Breisacher Strasse 64, 79106, Freiburg, Germany

^b BrainLinks-BrainTools Excellence Cluster, University of Freiburg, Germany

^c Department of Neuroradiology, University Medical Center Freiburg, Breisacherstrasse 64, 79106, Freiburg, Germany

ARTICLE INFO

Keywords:

Auditory processing
Domain-general
Speech processing
Music processing
Musicians

ABSTRACT

It is known that musicians compared to non-musicians have some superior speech and language competence, yet the mechanisms how musical training leads to this advantage are not well specified.

This event-related fMRI study confirmed that musicians outperformed non-musicians in processing not only of musical tones but also syllables and identified a network differentiating musicians from non-musicians during processing of linguistic sounds.

Within this network, the activation of bilateral superior temporal gyrus was shared with all subjects during processing of the acoustically well-matched musical and linguistic sounds, and with the activation distinguishing tones with a complex harmonic spectrum (bowed tone) from a simpler one (plucked tone). These results confirm that better speech processing in musicians relies on improved cross-domain spectral analysis.

Activation of left posterior superior temporal sulcus (pSTS), premotor cortex, inferior frontal and fusiform gyrus (FG) also distinguishing musicians from non-musicians during syllable processing overlapped with the activation segregating linguistic from musical sounds in all subjects. Since these brain-regions were not involved during tone processing in non-musicians, they could code for functions which are specialized for speech. Musicians recruited pSTS and FG during tone processing, thus these speech-specialized brain-areas processed musical sounds in the presence of musical training.

This study shows that the linguistic advantage of musicians is linked not only to improved cross-domain spectral analysis, but also to the functional adaptation of brain resources that are specialized for speech, but accessible to the domain of music in the presence of musical training.

1. Introduction

Numerous studies have found that musicians in comparison to non-musicians have not only superior competence in music but also in several linguistic areas. Among others, a clear causality has been shown between musical training and enhanced pre-attentive discrimination of speech [1], phonological awareness, and pitch discrimination in speech [2] as well as more efficient word discrimination [3], lexical knowledge [4,5], naming [6], writing and reading [2–4,6–8] and second language pronunciation accuracy [9]. Moreover, musical training improves (or positively correlates) with better speech perception [10–12], prosody [13], and language comprehension in noise [14]. However, musical

training does not improve semantic processing [15]. It does facilitate the acquisition of long-distance syntactic dependencies, but not of local syntactic dependencies [16].

The superiority of musicians in the language domain is, therefore, not in general language skills, but, as far as is known, restricted to specific linguistic sub-domains. It has been proposed that linguistic advantage in musicians is related to increased auditory competence, e.g. the ability to process the rapid spectro-temporal changes of speech sounds, and the ability to construct stable phonological representations [16–19]. Likewise, there is also some evidence that linguistic skills improve musical skills. Speaking Mandarin (a tonal language) can help an individual acquire absolute pitch and sensitivity of interval distances

* Corresponding author at: Breisacher Str. 64, 79106, Freiburg, Germany.

E-mail addresses: mariacristina.musso@uniklinik-freiburg.de (M. Musso), hannah.fuerniss@universitaets-herzzentrum.de (H. Fürniss), volkmar.glauche@uniklinik-freiburg.de (V. Glauche), horst.urbach@uniklinik-freiburg.de (H. Urbach), cornelius.weiller@uniklinik-freiburg.de (C. Weiller), michel.rijntjes@uniklinik-freiburg.de (M. Rijntjes).

<https://doi.org/10.1016/j.bbr.2020.112662>

Received 7 October 2019; Received in revised form 21 April 2020; Accepted 22 April 2020

Available online 19 May 2020

0166-4328/ © 2020 Elsevier B.V. All rights reserved.

[20,21], and language expertise modulates brain activity in fronto-parietal regions involved in the top-down regulation of auditory functions associated with more proficient tone processing [22]. The link between auditory competence and higher-level linguistic skills, like reading, is well-acknowledged [5,23,24]. The neural representation of essential sound elements are impaired in subjects with difficulties in language processing and reading [17,18,25–27] whereas they are enhanced in musicians.

Musicians spend many years practicing and acoustically training themselves, and therefore it is not surprising to see an increase in their musical skills. But how do musicians transfer their superior musical competence to another domain?

The most common explanation is that musicians do not really transfer music training effects to the language domain, rather musical training positively shapes processing of both domains at the same time [28–31]. Under this hypothesis, the parallel improvement of both musical and linguistic skills is mediated by shared cortical and subcortical resources, functionally related to cross-domain computations, as reflected by the large overlap shown in several neuropsychological and functional brain imaging studies in non-musicians during processing of both domains. This shared anatomical framework between music and language was shown for the fine-grained auditory skills pitch, timing and timbre [11,32–36] that are the basic components into which any sound that reaches the human ear (including music or speech) can be broken down, and that music and speech use to convey information [27]. Large overlap between both domains was evident also for the memory and attention processes related to language skills [14,36,37] and for the ability to integrate discrete acoustic events into a coherent perceptual stream according to specific syntactic rules [38–40]. Based on these results it could be argued that, through years of musical practice, musicians in comparison to non-musicians have developed an increased sensitivity to the acoustic features and the cognitive competence common to both domains of music and language, leading to more efficient processing of music, speech and language.

Until now, studies comparing behavioral and fMRI data in musicians versus non-musicians are scarce and predominantly relate to music processing only. In these studies, musicians show stronger activation of widespread cortical brain regions [41–45], some of them specifically linked to music processing, as the right anterior superior temporal gyrus [44], others also involved in language processing [45]. Magnetoencephalography and electroencephalography (EEG) methods provided evidence that the advantage in musicians during speech perception may be related to the more efficient processing of cross-domain auditory character traits such as pitch [30,32], and most importantly to spectral and temporal features of musical and linguistic sounds [10,12,27,33,45]. In line with this, Patel [31] proposes the OPERA hypothesis, according to which music-driven adaptive plasticity in the language network occurs when five essential conditions are met: anatomical overlap (O) in the brain networks that process both domains, precision of processing (P), emotion (E), repetition (R) and attention (A).

Therefore, more recently, a second explanation postulated that superior language-related functions in musicians could be also directly related to the functional adaptation of the speech-specialized network. Price et al. [47] suggested that speech-specific processing emerges at the level of functional connectivity among distributed brain regions, each of which participates in processes that are engaged by both speech and non-speech tasks. Thus, it might be possible that brain regions segregating speech from music processing, when not directly involved in processing a specific musical task, are accessible to the domain of music and can be modulated by musical training. In other words, musicians could benefit not only from cross-domain competence but also from enhanced speech-specialized skills due to a functional plasticity of the language-specialized network. In this case, it would indeed be appropriate to speak of transfer of training effects from the music domain to the language domain [27]. However, up to now, evidence in favor of

this account is very few and is limited to the clinical observation that experience in music facilitates not only domain-general but also domain-specific processing in language. The fact that musicians are better in segmental processing of a non-native language [48] is an example of transfer as defined in this framework. One recent study investigated speech perception in musicians and proposed that musical training improves speech perception due to more stable phoneme representations [19]. However, as the authors did not compare the speech domain with the music domain, their data cannot explain how music processing interacts with speech processing.

The identification of the relative contribution of these two proposed mechanisms explaining superior linguistic competence in musicians was the main goal of this event-related fMRI study. It focused on speech and music processing in subjects with and subjects without musical expertise to investigate whether the linguistic advantage in musicians may be related to functional plasticity of cross-domain brain regions (hypothesis1), and/or to the strengthening of functions that are characteristic for speech and language, due to the functional plasticity of brain areas that are part of the neuronal network specialized for language (hypothesis 2).

It is well known that standard hemodynamic functional imaging techniques cannot distinguish between same neuronal populations, or different, functionally independent, neuronal populations. In fMRI, the best design to assess context insensitive activation in a series of cognitive subtractions is the conjunction analysis, as it can identify areas activated equally by diverse tasks in different groups [49–51].

For evidence of hypothesis 1, areas activated more strongly in musicians than in non-musicians during processing of stimuli belonging to the speech domain should overlap with the same brain areas activated by processing of musical stimuli, independently from musical skills and training. Evidence in support of hypothesis 2 would be the finding that the areas activated more strongly in musicians than in non-musicians during processing of stimuli belonging to the speech domain overlap with speech specialized brain-regions.

This is the first study to directly test these hypotheses by comparing subjects with and without musical expertise using well-matched musical and linguistic stimuli

This event-related fMRI study compared musicians and non-musicians during the same linguistic and musical semantically neutral task, which can be performed easily also by people without musical training. More specifically, during scanning subjects performed an oddball paradigm where they had to attentively listen to the syllables /ba/, /ga/, /da/ and to a bowed or a plucked violin tone. In order to measure behavioral performance and to ensure auditive attention (Fig. 1), subjects had to detect the rare predefined targets: a syllable /pa/ and the piano tone, respectively. As in another study [28], the choice of stimuli in this experiment was dictated by the fact that they engage a similar frequency range and are clearly recognized as "music" for tones and "speech" for syllables. These syllables and tones were acoustically well-matched (Fig. 1) and are perceived as a single musical and linguistic sound [52,53], affording us a unique opportunity for investigating the case for the cross-domain hypothesis (H1).

Despite the established common features of our musical and speech stimuli, however, differences between them do exist. Syllables are considered a fundamental phonological building block of morpheme and words of every language (despite the great differences in the phonetic inventories between languages) [54,55]. At a functional level, speech processing is dependent on the precise temporal arrangement of spectro-temporal features, such as formant transitions or VOT (Voice onset), more than on other acoustic signals [56,57]. Moreover, a range of studies suggests that speech perception not only activates auditory cortex, but can engage motor cortex [58–62]. Active listening to speech in discrimination tasks was shown to recruit motor speech brain regions [60,64]; and articulation of syllables produces activity in posterior auditory areas, even when sound input is masked [65]. More recently, Elmer et al. 2017 showed that the recruitment of the left prefrontal

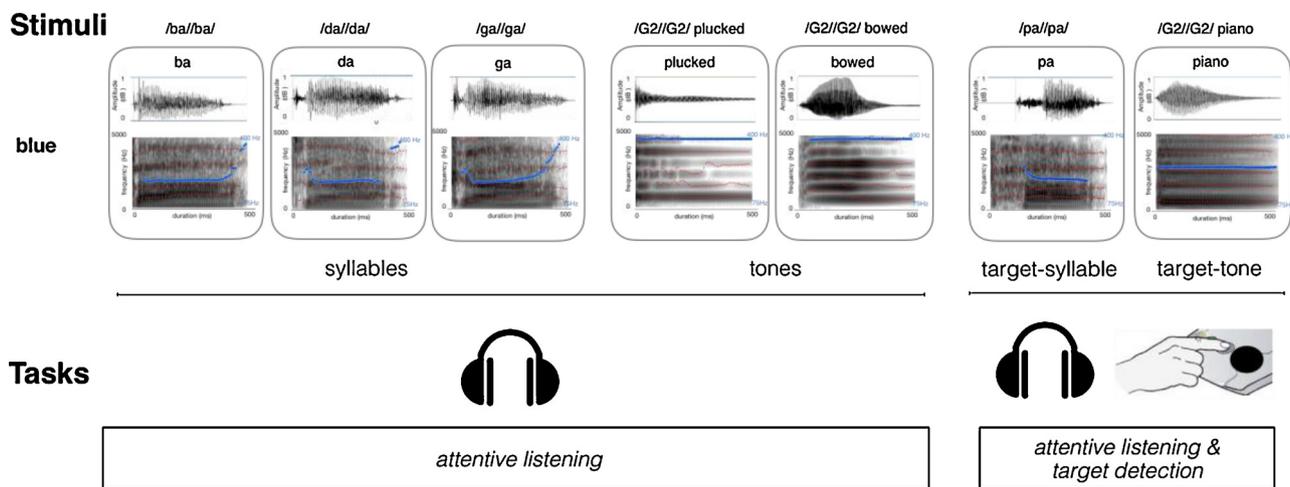


Fig. 1. Study design.

The stimuli consisted of linguistic and musical tokens relating to different phonetic and timbral categories: “plucked” and “bowed” violin tones were the non-target musical stimuli, and a piano tone was the target stimulus; the syllables /ba/, /ga/ and /da/ were the non-target speech stimuli and the syllable /pa/ was the target stimulus. For each stimulus, the respective oscillogram is displayed (above), presenting the waveform and amplitude of the sound (y-axis) over time (x-axis), as well as the spectrogram (below), consisting of a plot of the sound frequency (y-axis) over time (x-axis). The blue line in the spectrogram represents the course of the fundamental frequency (pitch contour), the red line represents the formants determined by PRAAT. The range of the frequencies of the harmonics H2–H5 are the same for all syllables and tones, while the relative amplitude and duration of these components, the pitch contour and the voice onset time differ. More specifically, for plucked tones, the high harmonics faded away quickly, leaving only the fundamental and some weak lower harmonics. The bowed tone maintained the rich harmonic spectrum over time. Subjects –musicians and non-musicians - were instructed to attentively listen to all stimuli and to detect the target stimuli pressing a button with their left hand, while refraining from doing so during the non-target stimuli.

regions is dependent on the temporal integrity of the speech signal decoded in the temporal cortex, and that it correlates with speech superiority through a more efficiently engaging of sound-to-motor mapping mechanism [66]. Thus, the integration of sensory-related and top-down information facilitates speech perception [58,66]. Auditory–motor interactions are also discussed to be relevant also for music perception in general [67], but until now there is no evidence of the relevance for musical tone perception. Therefore, the choice of these elementary units of speech and music allowed us to identify the overlap and the difference of the basic elements of universal linguistic and musical processing. Since there is a growing body of literature showing that the optimization of the auditory system in musicians leads to a finer-grained acoustic resolution in response to a variety of musical items, but that it also promotes other aspects of auditory processing, we needed to verify a superiority in musicians compared to non-musicians for detecting not only musical but also linguistic targets, as was already shown using very similar stimuli to ours by Musacchia et al. (2006) [28]. Comparing fMRI data from musicians to non-musicians during processing of elementary linguistic and musical units would allow us to identify the anatomical framework associated with the linguistic advantage in musicians. According to the first hypothesis, the use of the conjunction analysis would then identify to what extent brain activation in musicians in comparison to non-musicians during speech stimuli shared cross-domain resources, for instance related to spectral analysis. Considering that the bowed and plucked tones differ in the amplitude of their relative formants (Fig. 1), the conjunction analysis of musicians and non-musicians for the comparison between both tones would detect the brain regions functionally related to spectral analysis, independently of musical training. Most importantly, the conjunction analysis of first, the speech-related brain activation in musicians compared to non-musicians, and second, the speech-specific brain regions could find evidence for the second hypothesis, i.e. that better and faster syllable processing in musicians involves speech-specialized brain regions

2. Materials and methods

2.1. Subjects

In this fMRI study 15 subjects with professional musical training (“musicians”) and 15 subjects without musical training (“non-musicians”) were tested. Musicians (4 string players and 6 keyboard players, and 5 who played both strings and keyboard, 7 females, mean age: 23.66, range 19–33 years, education mean 16.3, range 13–22 years) had started their instrumental training at about 5 years of age (mean 5.13 years, range 3–6 years) and had continued with high practice intensity up to the present day with an average of 18.4 years of training. At the time of the experiment they were engaged in musical activities for an average of 14.06 h per week (range 6–30 hours/week). Non-musicians (8 females, mean age: 25.4 years, range: 20–31 years, education mean: 16.6, range: 13–19 years) had never participated in any extra-curricular music lessons or performances. History of music training was assessed by an in-house questionnaire that was specifically designed to evaluate the age of onset of music training, the instruments played, the number of years of music training, and the estimated number of training hours per day/week.

The two groups did not differ with respect to age ($t(30) = -1.867$, $p > 0.10$) [58], gender ($\chi^2(1) = 0.133$, $p > 0.50$) [59] or years of formal education ($t(30) = -0.506$, $p > 0.50$) [60].

All subjects were German native speakers, monolingual and right handed as assessed by the Edinburgh Handedness Inventory [68]. They were undergraduate students at, or recent graduates of the faculty of music or medicine Albert Ludwig University of Freiburg. All had normal hearing (5 dB pure-tone thresholds from 500 to 4000 Hz), and had no history of neurological or learning disorders. Furthermore, all subjects completed a form assessing habits and skills in language and music, and reported their feedback on the fMRI experiment (see supplementary materials).

The local ethics committee approved the study and subjects gave written informed consent.

2.2. Task and stimuli

During scanning, participants were instructed to listen attentively to all musical and linguistic stimuli and to press a button placed beneath their left forefinger as soon as they recognized a target stimulus.

The linguistic tokens consisted of two repetitions of a consonant-vowel (CV) syllable, resulting in the structure /CV-CV/ and representing different phonetic categories: /ba-/ba/, /da-/da/ and /ga-/ga/ were the non-target stimuli, while /pa-/pa/ was the target stimulus. The musical tokens consisted of two repetitions of the same tone, resulting in an identical stimulus structure to the speech stimuli. The two non-target music stimuli were the G2 tone (European term) of a violin being bowed (/G2b-/G2b/) and the same tone being plucked (/G2p-/G2p/), the target stimulus was the G2 key of a piano being played (/G2pi-/G2pi/). Both linguistic and musical tokens relate to different phonetic and timbral categories. Speech and music sounds were 500 ms in duration and similar to each other in envelope and spectral characteristics. As displayed in the spectrograms of each stimulus in Fig. 1, the range of the frequencies of harmonics H2–H5 were the same for speech and music stimuli (75–5000 Hz), although the relative amplitude of these components differed. More specifically, in plucked tones, the high harmonics faded away quickly, leaving only the fundamental and some weak lower harmonics. The bowed tone maintained its rich harmonic spectrum over time (for more details, see supplementary material).

All verbal stimuli (total of 191) were recorded beforehand and spoken by the same woman. The musical stimuli (total of 191) were generated on regular wooden instruments. The auditory stimuli were recorded and digitized at a professional recording studio in Freiburg i. Br., Germany. Target stimuli occurred with 19 % as in most of the oddball paradigms [28] to keep executive demands as low as possible, whilst ensuring that participants remained alert and attentive.

Stimulus presentation was pseudo-randomized according to stimulus category (target speech, target music, non-target speech, non-target music) in an event-related design. The inter-stimulus interval (ISI) varied in a pseudo-randomized manner from 3000 to 6000 ms, with shorter ISI between stimuli of the same category and longer ISI between stimuli of different categories.

2.3. Procedure

Before scanning, participants completed a short off-line trial of the experiment to ensure correct identification of the stimuli and understanding of the experimental tasks. The subjects were then placed in the scanner and equipped with auditory fMRI-compatible soundproof headphones (NordicNeuroLab AS, Bergen, Norway). The fMRI session lasted 14 min, during which participants lay in supine position and looked at a black screen via a mirror attached to the head coil. Two other sessions not relevant to this publication succeeded this auditory experiment before the measurement of anatomical MRI images concluded the experiment.

2.4. Image acquisition

The MRI measurements were conducted with a Siemens 3 Tesla TIM TRIO scanner using a Tx/Rx CP Head Coil (Siemens AG, Erlangen, Germany) in the University Hospital Freiburg. Stimuli were delivered by Presentation® (Neurobehavioral Systems, Inc.). Functional images were acquired using a BOLD-sensitive T2*-weighted multi-slice echoplanar imaging sequence (40 axial slices, thickness 2.4 mm, TR = 2640 ms; TE = 30 ms; flip angle = 75°; matrix size 64 × 64 pixels; voxel size 2.4 × 2.4 × 2.4 mm, interleaved acquisition of odd and even slices). We did not achieve full brain coverage, as the field of view was positioned to include the brainstem. Reaction time data were registered by Presentation® software and saved as log-files. Anatomical images were measured in a high-resolution T1-weighted sequence (176 sagittal

slices, thickness 1 mm, TR = 2200 ms; TE = 2.15 ms; flip angle = 12°; voxel size 1 × 1 × 1 mm). All images were corrected online for movement and distortion [69,70].

2.5. Data analysis

2.5.1. Behavioral data

All behavioral data were analyzed with IBM® SPSS Statistics 18. Reaction time was analyzed with dependent t-tests for within-group comparison of speech and music reaction times (RT), and a mixed design repeated measures ANOVA for between-group RT. One musician had to be excluded from the analysis of the RT data, as he was the first who was tested and received imprecise information on the task. Accuracy data were split into percent of correct reactions to target (button press, “correct positive responses”) and non-target (no button press, “correct negative responses”) conditions. False alarms were also detected.

2.5.2. MRI data

MRI data were analyzed using “Statistical Parametric Mapping” (SPM8; Wellcome Trust Centre for Neuroimaging, University College of London, Great Britain). The first five scans of each fMRI session were dummy scans required for MR equilibration and therefore excluded from the analysis.

To correct for different acquisition time within each TR, slice-time correction was performed on the middle slice. Co-registration of functional (T2*-weighted) and anatomical (T1-weighted) scans was then performed to allow mapping of the images onto a T1-template, followed by normalization into stereotaxic space (template provided by the Montréal Neurological Institute (MNI)). Lastly, smoothing with an 8-mm full width at half maximum Gaussian kernel was applied to achieve higher signal-to-noise ratios.

The aim of the fMRI analysis was to identify whether linguistic advantage in musicians is associated with an increased activation of cross-domain brain regions, for instance of brain region functionally related to broad spectral analysis (hypothesis 1) and/or speech-specialized brain regions (hypothesis 2). Cross-domain brain regions are the regions shared by processing of all musical and linguistic tokens, while the brain regions functionally related to spectral analysis were revealed contrasting the bowed tone with a rich harmonic spectrum with the plucked tone with weak lower harmonics. Speech-specialized brain regions were identified contrasting linguistic with musical tokens.

In the first level analysis, we estimated the baseline contrasts for speech and music non-target conditions, i.e. the contrasts non-target linguistic tokens and non-target musical tokens against implicit baseline. We also calculated the differential contrasts: bowed tone > plucked tones and linguistic > musical tokens.

In the second level analysis, we used two-sample t-tests to contrast speech or music processing in musicians versus non-musicians to identify the anatomical framework/ associated with linguistic advantage in musicians.

Overlap of brain activation was identified using the null-conjunction analysis [49,51] within the two sample t-tests when two factors were paired and with the full-factorial analysis when more than two factors were included [49,50].

According to hypotheses1, we performed a full-factorial analysis including all non-target conditions (i.e. linguistic and musical tokens in musicians and non-musicians) and calculated the conjunction of first, the contrast linguistic tokens in musicians > non-musicians - which identified, as mentioned, the specific activation in musicians during speech processing - and second, the baseline contrast of all the linguistic and musical non-target tokens in musicians and non-musicians, as their conjunction identified the cross-domain network. The brain regions functionally related to spectral analysis in both groups were identified using the conjunction of the contrasts bowed > plucked in both groups. Within this contrast we calculated also the difference between

musicians and non-musicians. The overlap of the contrast bowed > plucked in non-musicians with the contrast linguistic tokens in musicians > non-musicians was calculated within a conjunction of both contrasts within a two-sample *t*-test. This conjunction identified the brain regions within the distinctive network for syllables in musicians that were shared with the network for spectral analysis.

According to hypothesis 2, we calculated the conjunction of the contrast linguistic tokens in musicians > non-musicians and the contrast linguistic > musical tokens in both groups. This conjunction identified the speech-specific brain regions in both groups. Note that performing this last contrast should not necessarily identify brain regions that are not involved in music, i.e., these brain regions were not necessarily speech-unique. The within-subjects analysis shows if these brain regions are uniquely activated in speech or if they are involved also in music but with lesser significance than in speech. It would be possible that brain regions segregating linguistic from musical processing, even if not generally involved in processing bowed and plucked tones, are accessible to the domain of music and can be modulated by musical training. In this case, these brain regions should show significant activation during tone processing in musicians. Therefore, we performed a conjunction analysis of the contrast linguistic tokens in musicians > non-musicians with the baseline contrast tone processing in musicians.

For all conjunction analyses, statistical threshold at peak-level was set at $p < 0.05$, corrected for multiple comparisons (FWE corrected). In the supplementary-figure2 we also report the contrast tone processing in musicians > non-musicians at uncorrected level for multiple comparisons ($p < 0.001$), in order to show that some activation that were highly significantly related to speech-processing were actually not speech-unique.

3. Results

3.1. Behavioral results

The 2×2 ANOVA including all factors (subjects: musicians and non-musicians, domains: music and speech) showed that reaction time (RT) for tones was significantly shorter than for syllables $F(1, 30) = 6.812$, $p = 0.011$ (Fig. 2). Between groups, ANOVA showed that musicians had significantly shorter overall RT than non-musicians, $F(1, 30) = 7.445$, $p = 0.011$ (Fig. 2). Musicians were faster than non-musicians in the speech domain ($t(30) = 2.44$, $p = 0.022$) and in the music domain ($t(30) = 2.799$, $p = 0.009$).

Regarding accuracy, musicians had significantly more correct positive responses across domains than non-musicians, $H(1) = 4.444$, $p = 0.035$, while the groups did not differ in correct negative responses,

$H(1) = 0.659$, $p = 0.417$.

The mean of false alarm rate was very low: in musicians 0.005 (standard deviation (sd) 0.009, range 0.0–0.0100) for the target tone, 0.0007 (sd 0.002, range 0.0–0.0109) for the target syllable; in non-musicians 0.005 (sd 0.01, range 0.0–0.04) for the target tone and 0.006 (sd 0.01, range 0.0–0.03) for the target syllable.

3.2. MRI results

3.2.1. Linguistic and musical processing in musicians and non-musicians

Fig. 3 and Tables 1A and 1B show the results of the one-sample *t*-tests in musicians and non-musicians for processing linguistic (A) and musical (B) sounds. Musicians activated the same regions as non-musicians (Tables 1A and 1B) including bilateral Heschl's gyrus (HG) and the superior temporal gyrus (STG) for both domains, the left posterior superior temporal sulcus (STSp) and the fusiform gyrus (FG) as well as the bilateral cerebellum for syllables and left rolandic operculum for tones. Moreover, musicians additionally activated the left premotor cortex (PMC), the inferior frontal gyrus (IFG) and the right anterior superior temporal gyrus (STGa) for syllable processing and the left STGa, the posterior superior temporal sulcus (STSp), FG and cerebellum for tone processing.

3.2.2. The brain signature of syllable and tone perception in musicians

Musicians showed stronger activation in the syllable condition than non-musicians in bilateral superior temporal lobe, including HG, STG and left pSTS, the primary and secondary visual cortex, FG and superior cerebellum (lobule VI) as well as the left PMC and IFG (pars opercularis) (Fig. 3, Tables 1A and 1B).

Comparing musicians with non-musicians during processing of musical tokens, they activated more strongly left Heschl's gyrus, the STG, the bilateral primary visual cortex and the right rolandic operculum (Tables 1A and 1B). Supplementary-figure 2 shows the activation identified by the same comparison (tones in musicians > non-musicians) with a lower significance (at $p < 0.001$ uncorrected): during tone processing musicians activated more than non-musicians the same brain regions that a) they involved more than non-musicians during syllable processing and b) that both groups activated more during linguistic tokens than during musical tokens processing.

Hypothesis 1: Overlap between the distinctive brain network for syllable processing in musicians and cross-domain brain resources

The conjunction of the contrast linguistic tokens in musicians > non-musicians and the contrasts linguistic and musical tokens in musicians and non-musicians identified in left STG 1–2, STGp and right STG the cross-domain brain regions, as they were activated on one hand more strongly in musicians than in non-musicians during syllable

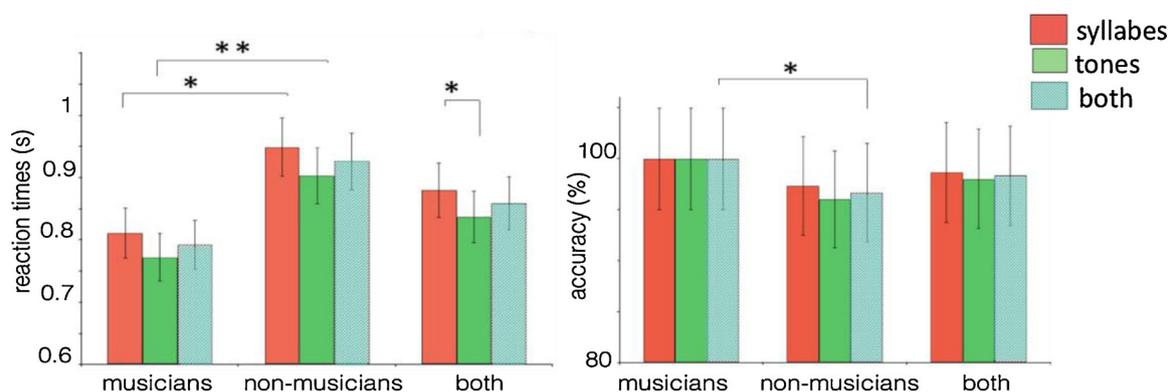


Fig. 2. Superior linguistic and musical competence in musicians.

Musicians were faster in detecting linguistic (red) and music (green) targets in comparison to non-musicians ($p = 0.022$ and $p = 0.009$, respectively). Detecting tone targets was faster than detecting syllable targets ($p = 0.01$) in musicians and in both groups ($p = 0.011$). Accuracy: both groups performed the tasks almost perfectly. Nevertheless, musicians showed better performance when comparing the whole sample (blue) ($p = 0.035$). * = $p < 0.05$, ** = $p < 0.01$. Error bars denote ± 1 standard deviation.

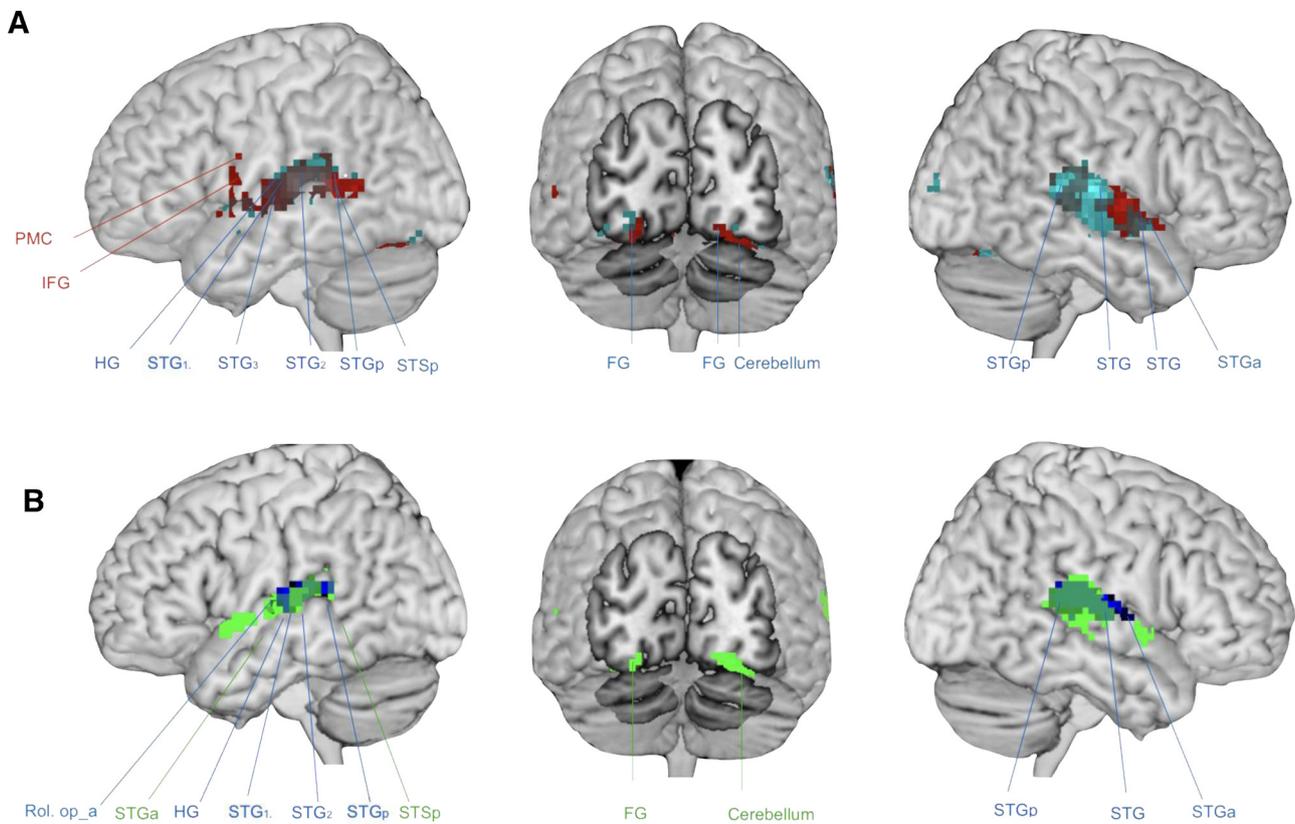


Fig. 3. Processing of linguistic (A) and musical (B) sounds in musicians and non-musicians. A. Processing of linguistic sounds in musicians (red) and non-musicians (light blue) involves bilateral STG in both groups. Musicians additionally activate left pre-rolandic regions. B. Processing of musical sounds in musicians (green) and non-musicians (dark blue) activates bilateral STG and the left pre-rolandic gyrus in both groups. Activation in left STGa, STSp, FG, and cerebellum are found in musicians only. Coordinates of activation: see table. Abbreviations: FG = fusiform gyrus, HG = Heschl's gyrus, IFGop = inferior frontal gyrus pars opercularis, PMC = premotor cortex, Rol. op = rolandic operculum, STG = superior temporal gyrus, STGp = posterior STG, STGa = anterior STG, STSp = posterior superior temporal sulcus.

Table 1A
Main effect of musicians for processing linguistic and musical tokens. Activation is displayed in Fig. 3.

Main effect musicians	a) linguistic tokens								b) musical tokens							
	left hemisphere				right hemisphere				left hemisphere				right hemisphere			
Brain regions	x	y	z	T	x	y	z	T	x	y	z	T	x	y	z	T
HG	-51	-13	7	4.5	48	-22	10	9.2	-51	-13	7	9.2	48	-22	10	8.5
STG	-33	-25	7	7.8	51	-4	1	7.8	-33	-25	7	7.3	51	-4	1	5.9
	-57	-13	7	8.6					-57	-13	7	7.8				
	-63	-22	7	8.6					-63	-22	7	9.3				
STGp	-60	-13	1	7.8	66	-34	7	9.1	-60	-40	7	7.1	66	-34	7	7.3
	-60	-40	7	12.1					-60	-40	7	7.1				
	57	-28	16	10.3					57	-28	16	10				
STGa					51	-4	1	8.4	-57	-1	1	9.3	51	-4	1	8.8
STSp	-54	-40	7	9.8					-54	-40	7	5.7				
IFGop (44)	-57	2	10	7												
PMC inferior	-60	2	22	7.1												
Rol op_p	-39	-31	13	8.8	60	5	13	6.8	-39	-25	19	7.5				
Rol op_a									-54	2	4	7				
TPJ					54	-28	19	10.3					54	-28	16	9.1
Lingual G	-21	-82	-14	7.3					-21	-85	-11	7.0				
FG	-27	-76	-14	7.4	24	-85	-11	6.2	-27	-76	-14	6.0				
Cerebellum	0	-73	-14	7.2					0	-73	-14	6				

processing and on the other hand were significantly involved also during tone processing in subjects with and without musical training (Fig. 3 blue, Table 3A).

The conjunction of the contrast linguistic tokens in musicians > non-musicians and the contrast “bowed” > “plucked” in non-musicians showed an overlap in diverse parts of bilateral STG (Fig. 3 yellow, Table 3B). Note that we did not find a difference between the two

groups during comparing “bowed” > “plucked” tones, rather a large overlap (supplementary-figure2, yellow).

Hypothesis 2: Overlap between the distinctive brain network during syllable processing in musicians and speech-specific brain resources

The conjunction of the contrasts linguistic tokens in musicians versus non-musicians and the contrast linguistic tokens versus musical tokens in musicians and non-musicians showed a relevant overlap

Table 1B

Main effect of non-musicians for processing linguistic and musical tokens. Activation is displayed in Fig. 3. Note that activation of left STSp, FG, STG3 during tone processing was not found also reducing the level of significance at $p < 0.001$.

Main effect non-musicians	a) linguistic tokens								b) musical tokens									
	left hemisphere				right hemisphere				left hemisphere				right hemisphere					
	x	y	z	T	x	y	z	T	x	y	z	T	x	y	z	T		
Brain regions																		
HG	-51	-13	7	10	48	-22	10	12.5	-51	-13	7	7.5	48	-22	10	10.5		
	-33	-25	7	6.7					-33	-25	7	5.5						
STG	1	-57	-13	7	7.7	51	-4	1	8.6	-57	-13	7	6.0	51	-4	1	8.1	
	2	-63	-22	7	11.3					-63	-22	7	7.5					
	3	-60	-13	1	7.0													
STGp		-60	-40	7	8.7	66	-22	16	10.5	-60	-40	7	7.0	66	-22	16	10.3	
						57	-28	16	12.9					57	-28	16	10.9	
STGa						51	-4	1	8.6					51	-4	1	8.2	
STSp		-54	-40	7	6.5													
Rol op_p		-39	-31	13	9.0						-39	-25	19	6.7				
Rol op_a											-54	2	4	6.7				
TPJ						54	-28	19	10.5					54	-28	16	9.7	
Lingual G.											-21	-85	-11	6.6				
FG		-27	-76	-14	6.9	24	-85	-11	7.2									
Cerebellum		0	-73	-14	5.7													

between the areas that were more strongly activated in musicians than in non-musicians during syllable processing and the areas specialized for syllable processing in both groups (independent of the musical training), in bilateral STG and left STSp (Fig. 3, red).

The conjunction of the contrasts linguistic tokens in musicians > non-musicians and musical tokens processing in musicians showed an overlap between the distinctive syllable network in musicians and the network for tone processing in subjects with musical training in bilateral STG, and the left STSp and FG (Fig. 3, green). The summary of the fMRI-results is displayed in Fig. 4 and Table 4.

4. Discussion

The present study confirms that long-term musical training is

beneficial for the auditory processing of elementary units of language and music, as musicians were faster and more accurate in the recognition of both musical and linguistic tokens compared to non-musicians (Fig. 2) [3,10–12,19,28–33]. The novelty of the present study comes from the fMRI results, showing that this linguistic advantage in musicians is associated with functional plasticity not only of cross-domain brain regions, but also of speech-specialized brain regions. This may be the key for superior linguistic competence in musicians.

Activation patterns for processing of basic linguistic and musical units differentiating subjects with musical training from subjects without musical training, but well balanced on factors known to influence cognitive functions, like age, gender and years of formal education [19], were located in a widespread bilateral network, including bilateral STG, left STSp, IFG, PMC (Fig. 4A). Within this network, a

Table 2

Brain activation distinguishing musicians from non-musicians during processing of linguistic (A) and musical (B) tokens. Abbreviations: see Tables 1A and 1B.

A		linguistic tokens musicians > non-musicians															
		left hemisphere								right hemisphere							
		x	y	z	T	x	y	z	T	x	y	z	T	x	y	z	T
Brain regions																	
HG		-51	-13	7	9	48	-22	10	6.1								
STG	1	-57	-13	7	6.7	51	-4	1	6								
	2	-63	-22	7	6.0	-60	-13	1	6.2								
STGp		-60	-40	7	9.8	66	-22	13	6								
STGa						57	-14	-8	5.8								
STSp		-54	-40	7	8.1												
IFGop		-54	-40	7	8.1												
PMC inferior		-60	2	22	5.8												
Rol op_p		-39	-31	13	8.8	60	5	13	5.9								
TPJ						54	-28	19	5.2								
Calcarine G		0	-94	10	5.7	3	-94	10	6.5								
Lingual G		-21	-82	-14	5.9												
FG		-27	-76	-14	5.7	24	-85	-11	5.6								
Cerebellum		-9	-73	-14	7.5	6	-49	-2	6								

B		linguistic tokens musicians > non-musicians															
		left hemisphere								right hemisphere							
		x	y	z	T	x	y	z	T	x	y	z	T	x	y	z	T
Brain regions																	
HG		-51	-13	7	6.4	48	-22	10	6.1								
STG	1	-57	-13	7	6.7	51	-4	1	6								
STGa		-57	-1	1	6.4												
Rol op_p		-54	2	4	6.3	60	5	13	5.9								
Lingual G		-21	-67	-11	5.6	15	-82	-11	7.5								
FG						24	-85	-11	6.0								
Cerebellum		-9	-70	-11	5.79	6	-49	-2	6.7								

Table 3
Overlap between musicians' distinctive brain activation during syllable processing and cross-domain brain resource.

Conjunction (\wedge) linguistic tokens musicians > non-musicians																	
a. syllables musicians, non-musicians \wedge tones musicians \wedge tones non-musicians								b. bowed > plucked tones non-musicians									
left hemisphere				right hemisphere				left hemisphere				right hemisphere					
Brain R.	x	y	z	T	x	y	z	T	x	y	z	T	x	y	z	T	
Rol op_p	-39	-31	13	8.8					6.1	-48	-13	1	6.0				
HG	-51	-13	7	7.3	48	-22	10	6.1	-48	-13	1	6.0					
STG 1	-57	-13	7	5.6	54	-19	1	7.8	-57	-13	7	5.6	60	-1	1	6.0	
2	-63	-22	7	6	63	-13	4	5.6	-63	-22	10	5.2					
STGp	-60	-40	-60	5.9													
STGa					51	-4	1	7.7									
c. syllables > tones musicians \wedge syllables > tones non-musicians								d. tones musicians									
left hemisphere				right hemisphere				left hemisphere				right hemisphere					
Brain R.	x	y	z	T	x	y	z	T	x	y	z	T	x	y	z	T	
Rol op_p									-39	-31	13	7.8	39	-25	19	5.7	
PMC	-57	2	10	6.0													
HG					54	-25	10	7.0	-51	-13	7	7.3					
STG 1	-57	-13	7	5.4	54	-19	1	5.3	-57	-13	7	5.3	48	-28	10	8.1	
2	-63	-22	7	6.0	63	-13	4	6.0	-63	-22	10	6.0	60	-1	4	5.4	
3	-60	-13	4	7.0					-60	-13	1	7.3					
STGp	-60	40	16	5.0	63	-28	4	6	-60	40	16	5.0	63	-28	4	5.6	
STSp	-54	-40	7	5.6					-54	-40	7	5.4					
STGa					51	-4	1	5.2									
TPJ													54	-28	19	8.3	
FG									-27	-76	-14	5.3					

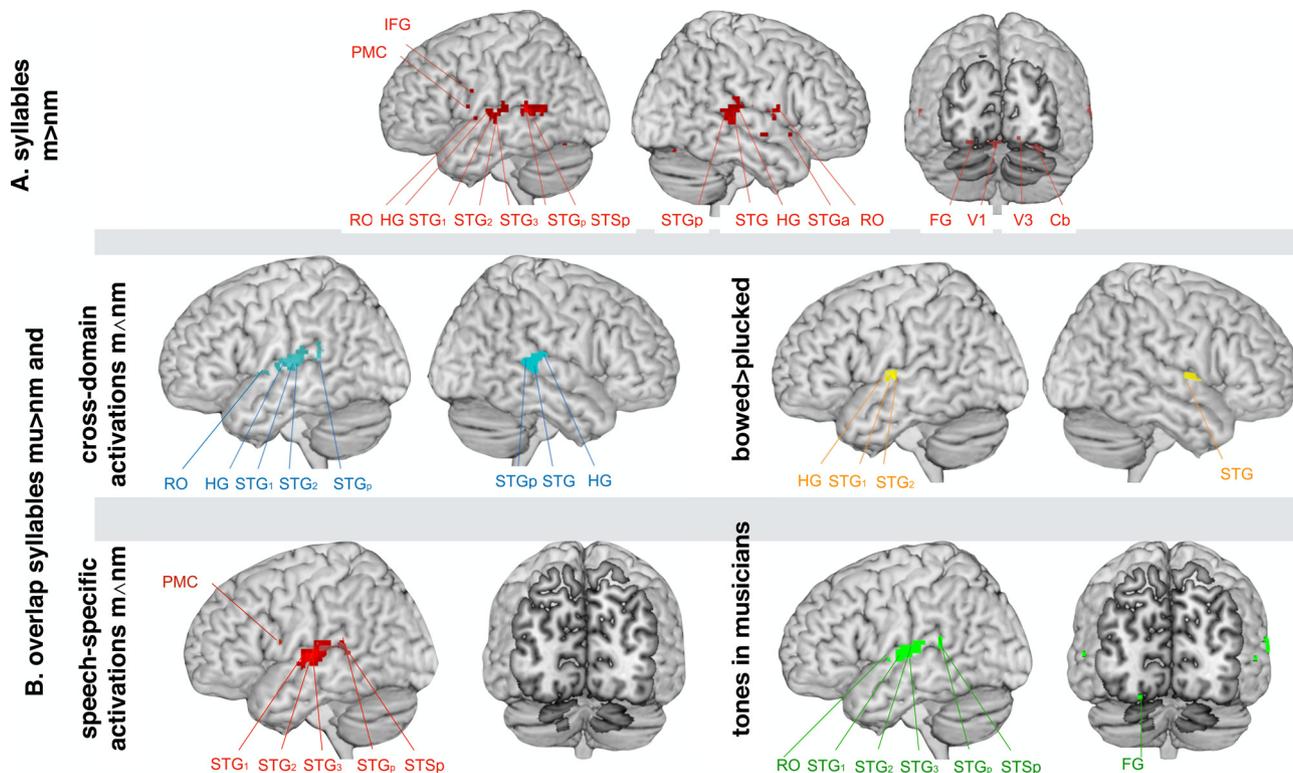


Fig. 4. The brain signature of processing linguistic sounds in musicians.
 A. Result of the contrast syllables musicians (m) > non-musicians (nm) (Table 2) identified the brain signature of processing linguistic sounds in musicians.
 B. Results of the conjunction analysis of the contrast linguistic tokens musicians > non-musicians from A and: 1) in blue all conditions and subjects. This conjunction identified within the distinctive brain regions during syllable processing in musicians (A) the brain regions that are cross-domain and training independent; 2) in yellow: the contrast bowed > plucked violin tones, showing that left HG, STG1,2 differentiates spectrally different sounds; 3) in red: syllables > tones in musicians and syllables > tones in non-musicians and musicians. The regions distinguishing musicians from non-musicians during syllable processing (A) overlapped with the activation that segregated processing of linguistic from musical sounds in all subjects. Note that the cross-domain regions (STG 1–2 and the STGp) are activated stronger during language than during tone processing, while left PMC, STG2 and STSp are specialized for syllables; 4) in green: the contrast musical tokens in musicians shows that the left STSp, STG3 and FG, regions that are activated during syllables processing in both groups and stronger in musicians, are significantly involved during tone processing in musicians only (Fig2B). Thus, the brain signature of processing linguistic sounds in musicians included both cross-domain regions and regions specialized for speech, but accessible to the domain of music in the presence of musical training.

Table 4
Schematic list of activation.

syllables_m>nm (LH)			syllables_m>nm, conjunction (\wedge) with:				Main effects for				
Brain regions	coordinates			syllables \wedge tones	bowed>plucked	syllables>tones	tones in musicians	syllables		tones	
	x	y	z	cross-domain	spectral analysis	speech-specialized	accessible to tones	musicians	non-musicians	musicians	non-musicians
rol op	-39	-25	19	✓			✓	✓	✓	✓	✓
LingualG	-21	-82	-14	✓			✓	✓	✓	✓	✓
HG	-51	-13	7	✓	✓		✓	✓	✓	✓	✓
STG1	-57	-13	7	✓	✓	✓	✓	✓	✓	✓	✓
STG2	-63	-22	7	✓	✓	✓	✓	✓	✓	✓	✓
STGp	-60	-40	7	✓		✓	✓	✓	✓	✓	✓
STG3	-60	-13	1			✓	✓	✓	✓	✓	✓
STSp	-54	-40	7			✓	✓	✓	✓	✓	✓
FG	-27	-76	-14			✓	✓	✓	✓	✓	✓
PMC	-60	2	22			✓	✓	✓	✓		
IFG	-57	5	10			✓	✓	✓	✓		

In the left column, the list of activation distinguishing musicians from non-musicians ($m > nm$) in the left hemisphere (LH) during syllable processing (Fig. 4A). The same regions are shared in different conditions as identified by a serial of conjunction analysis (see Fig. 4B and C) and involved in different contrasts (see Tables 1A and 1B, supplementary-figure 2). * indicates that IFG and FG are significant comparing linguistic versus musical sounds in musicians only (see S-table 1). Note that during tone processing STG3, STSp and FG are recruited in subjects with musical training, while PMC and IFG are inactive also in musicians.

series of conjunction analyses identified several brain regions that, according to hypothesis 1, were equally involved in both domains, others that were cross-domain but more strongly used for linguistic than musical processing, and brain areas that, supporting hypothesis 2, were speech-specific but accessible to the domain of music as they process musical sounds in the presence of musical training (Fig. 4, Tables 3 and 4).

The stronger activation of bilateral primary auditory cortex (HG) in musicians in comparison to non-musicians indicates a greater sensitivity in musicians for basic acoustic cues common to both domains. This result is in line with previous studies finding that the activity evoked in primary auditory cortex and the gray matter volume of HG were larger in musicians than in non-musicians [71,72]. Our data confirm that musicians involved the same part of bilateral HG activated in all subjects during processing of both domains more strongly than non-musicians (Fig. 3B). Since speech and music overlapped, according to the OPERA hypothesis, the finding of HG specialization may explain benefits in both music and speech [31]. This finding is also in line with the fact that this brain region is active during syllable and tone processing in humans and monkeys, despite their vastly different experience with human speech [73]. Our study-design, however, could not explain which cross-domain acoustic feature was better during processing in musicians. Since we did not find a difference of activation, particularly of the left HG during speech and music processing (Table 4), this region should code for a factor that is equally involved in both domains. Several studies identified HG as a specialized brain region in the auditory cortex that is involved in the representation of pitch (for review [74]), which was indeed roughly comparable in our stimuli (Fig. 1).

Improved processing of acoustic features in musicians common to both domains, however, should not be the only explanation of their superior linguistic competence, because the brain signature of syllable processing in musicians involved brain regions whose activation was specialized for linguistic sounds (Fig. 4, red, supplementary-figure 3A, tables 2 and 5).

The activation of the bilateral STG and left pSTG distinguishing musicians from non-musicians during syllable processing overlapped on the one hand with the activation segregating linguistic from musical sounds (Fig. 4B, red), on the other hand with the activation during tone processing (Fig. 4B, blue). Thus, musicians showed enhanced activation of brain-regions that code for cross-domain functions, but are more strongly involved during speech processing. Both tone and syllable perception relies on the analysis of similar (but not the same) spectral cues [27]. Fig. 1 shows that our linguistic and musical stimuli, although the range of the frequency of the harmonics H2–H5 is the same (100 Hz–5000 Hz), are characterized by diverse relative amplitude,

amplitude envelope and different voice onset time [61]. Since activation of the same part of STG differentiated tones with a complex harmonic spectrum from tones with a simpler spectrum over time (supplementary Fig. 3), we functionally linked these regions to spectral analysis. In line, higher levels of activity in the bilateral dorsal STG are observed for sounds with multiple spectral components [75]. Our data suggest that, although syllables may require stronger brain responses than violin sounds, the activation of these areas were not exclusive to speech. Considering previous EEG and multimodal MRI studies that found a relationship between musical expertise, processing of spectral features and speech processing performance [35,37,46,75], we propose that spending countless hours attending to spectrally rich musical sounds throughout years of musical training stimulated STG brain regions specialized in processing these acoustic features.

Within the brain regions differentiating syllable processing in musicians from non-musicians, the left STG3, STSp and PMC, whose activation overlapped with the network segregating processing basic linguistic units from musical tokens independently of the presence of musical training (Fig. 4B, red), were not involved in processing musical tones in subjects without musical training (Fig. 3B). Recent studies using auditory stimuli closely equated with speech with respect to temporal and spectral acoustic content [76,77] (for further discussion, see Price et al. [47] have identified the same temporal sub-regions (STG3 and STSp) as specific for speech sounds. These brain regions, therefore, should be functionally linked to competencies characteristic for the language domain, while absent during the processing of musical sounds in persons without musical training. More specifically, activation of the left pSTS was found to be crucial for perception of changes in VOT [11], voice-selective processing [78,79], for phonological processing [77] and for lexical and non-lexical lip reading [80,81]. Boatman et al. (1997) [82] reported that the stimulation of one temporal electrode in this region elicited a deficit in syllable discrimination in five subjects. These data were interpreted as meaning that this area, therefore, is exclusive for speech. However, some other studies using stimuli closely matched for rapid temporal acoustic cues, found a significant overlap in STG/STSp activation for speech and non-speech sounds [86; for further discussion, see 45]. Our speech stimuli are characterized by specific F1, F2 and F3 formant transitions indicating bilabial, alveolar and velar place for the /ba/, /da/ and /ga/ syllables respectively (see supplementary material). Such formant transitions were less differentiated in the bowed and plucked tones, explaining a possible reason of the lack of activation of the STG/STSp regions during music perception in non-musicians. In contrast, musicians recruited the speech-specialized left STSp and STG3 for processing of musical sounds (Fig. 3B). Their function, therefore, should be characteristic for processing linguistic sounds, but accessible to the domain of music, and

may be positively influenced by musical training. According to current studies, the left STSp could be sensitive to rapid changes in acoustic frequency [45,86,87]. This aspect is characteristic for speech, but not unique to it, as music processing also involves processing of rapid variation of sensory parameters over time. However, since people without musical training did not recruit this brain region, this aspect can account for processing tones only in presence of musical training. Musicians have trained temporal ordering of acoustic events more extensively than non-musicians, and therefore can better recruit this function for speech processing, and also use it for processing musical categories, facilitating their identification and recognition (Fig. 2).

This finding may have important repercussions. It has been discussed that the ability to track brief, rapidly successive acoustic changes within the complex acoustic waveform of speech should be crucial in phonological (and more general than for language) development [1,6,8,12]. Each language has its own set of phonemes (composed of complex acoustic spectra) that must be learned from experience and represented as neural firing patterns in auditory cortex [59,83,84]. Complex neural firing patterns of cell assemblies will become generalized to represent the individual phonemes and syllables of a language, regardless of the specific context or speaker [59]. Some studies suggest that the inability to process dynamic spectral and/or temporal change does co-occur in reading deficits observed in children with dyslexia and language impairments (LLI) [16–18]. Language training incorporating acoustically modified speech in which rapid spectrotemporal segments were amplitude-enhanced and extended in duration - as the Fast Forward training - can lead to substantial improvements in speech discrimination and language comprehension in children with LLI compared to a well matched control group [85]. Musical training induces the same positive effect on phonological awareness as a phonological skills program [1]. We speculate that the finding of an increased activation of STSp/STG in musicians may explain the association between musical training and phonological awareness.

Activation of FG in our data was very intriguing, first because it was significant during processing linguistic sounds in both groups (Figs. 3, 4A). This region is typically inactive during auditory stimulation [86]. However, FG is essential for lip reading [81,86,87] and it can be activated in a top-down manner during spoken language tasks as an orthographic code that can be mobilized when needed to facilitate speech processing [88]. In agreement with this interpretation, musicians and non-musicians might use FG as an orthographic code also for facilitating syllables processing. Second, musicians activated the same part of FG during processing of linguistic sounds more strongly than non-musicians, indicating that the superior ability in musicians to recognize speech categories can be explained by the stronger recruitment of the orthographic codes of speech. Considering the close interaction of orthographic processing with sub-lexical and phonological processing up to reading, functional plasticity of this brain region in musicians seems to be another key aspect for understanding their superior linguistic competence [16,17,86]. Third, similarly to the left STSp, the same part of FG processed musical sounds in musicians, but not in non-musicians (Fig. 3, green, Tables 1A and 1B and supplementary-figure 1.). It has been shown that increase of activation of the left FG during reading acquisition occurs only when there is systematic correspondences between print and speech sounds [88,89] and that the left FG and STSp work closely together, modulated by experience to recognize a sound [90]. Thus, this region should code for processes that are accessible to the domain of music and trainable in subjects with musical expertise. Dehaene and Cohen proposed that this region is “recycled from a prior cortical competence for invariant object recognition” [88]. FG is indeed activated during visually discriminating between exemplars of a particular category, for instance, faces, letters, and musical notation [86,91,92]. Thus, the activation in the FG during processing musical sounds observed in musicians might suggest that musical auditory processing is mediated by the recoding of visual musical form. Musicians train musical notation over long periods, extensively recruiting

FG, which may explain why musicians use this brain region more than those without musical training for processing speech as well as music categories, facilitating their identification and recognition.

Within the brain activation distinguishing musicians from non-musicians during syllables processing, we found also two left pre-rolandic brain regions.

The left IFG (p. opercularis), whose activation was completely absent in both non-musicians and musicians during tone processing (Fig. 3 table1), was recruited in musicians, but not in non-musicians, for syllable processing (Fig. 3 table1) and segregated in this group linguistic from music tokens (Table 4, supplementary-table 2). The role of this region in speech perception was demonstrated in studies with patients with Broca’s aphasia [60] and with healthy subjects, particularly when using specific tasks, as the speech-discrimination task or rhyme task, but not during passively listening to speech sounds [61–64]. It is discussed, therefore, that its role within speech perception may be related to sub-lexical phonological processing for facilitating speech [58,61,64,81], since a syllable is a phonological structure in speech [55]. Musicians could thus benefit from a more stable and proficient sub-lexical phonological processing within a very easy task as ours. According to our results, the identification of musical tones within a recognition task does not involve tonal rehearsal. Since a virtually an identical network centered on left IFG and PMC during verbal rehearsal was also found to be activated during the sub-vocally rehearse of musical pitches [36]. According to the assumption that speech-specific processing emerges at the level of functional connectivity, it is possible that musicians trained tonal rehearsal and thereby improve also verbal rehearsal [51].

Alternatively, activation of left IFG was linked to verbal working memory [93]. It is documented that musicians have better verbal working memory ability than non-musicians [30,37]. However, the casual link between musical expertise, working memory and speech processing has not been directly provided. Du and Zatorre [19], showed that working memory does not play a crucial role in speech perception, even in the case of ambiguous tasks (as in syllable in noise identification task). Finally, some authors, for instance Hansen et al. [94], observed that musical training improves temporal scanning of auditory stimuli and sub-lexical phonological processing, and this, in turn, might result in the improved verbal working memory storage capacity in musicians.

At last, the left PMC was significantly activated in both groups during processing of linguistic units only, i. e it was not involved in processing musical tokens, even in musicians (Tables 1A and 1B). Clearly, musical training also improves skills independently from those based on musical tone processing. Playing music is a very complex sensory–motor activity, requiring a precise control over pitch interval production [67]. Plasticity in the motor network is commonly found in musicians [95,96]. The syllable is, as mentioned in the introduction, a co-articulatory structure in speech [55] reflected by the activation of the left PMC [60–64]. Du and Zatorre [19] documented in musicians a greater activity in the left PMC and discussed that stronger phoneme representations in speech premotor cortex may better explain speech perception in the presence of noise. In our study the activation of the left PMC was speech-specific and associated with speech perception per se. The more robust motor phoneme representations in musicians may generate more accurate articulatory predictions that enhance the top-down modulation of speech perception, and so facilitating linguistic processing.

In summary, the present study shows that the faster and more accurate processing of elementary musical and speech units in musicians in comparison to well-matched non-musicians is related not only to more refined perception of acoustic stimuli, but also to a more proficient top-down system for speech, articulatory, and orthographic predictions.

A major question, of course, is whether this more refined perception is related to a more efficient attentional control system [14]. The state

of attention, indeed, is known to significantly modulate BOLD response amplitudes in auditory sensory as well as language-related brain areas [97]. Since a complete control over the applied attentional resources is not possible with fMRI, musicians could outperform non-musicians because through musical training they have learned to pay more attention to the presented sound signals [14,22,97]. In this case, our data could be interpreted such that learning to pay attention through intense musical training also improves speech-specific processing. There are, however, some arguments against the claim that the difference of activation between subjects with and without musical training is merely related to a different state of attention. First, focal auditory attention related to target detection task is associated with activation of posterior parietal lobe, areas of the lateral and medial frontal cortex and cingulate regions [98]. Accordingly, we did find significant activation of these brain regions during the detection of targets (supplementary-figure1) only, however without significant difference between musicians and non-musicians (supplementary results). In line with our results, Marie et al. showed that during speech tasks, independently of the direction of attention, the P2 component, which reflects perceptual processing, was larger in musicians than in non-musicians, while the N1 component that is known to be particularly sensitive to selective attention did not differ between musicians and non-musicians [48]. Moreover, it has been discussed that musical training-induced long-term modulation of the state of attention may shape auditory cortex (which indeed showed functional and anatomical adaptation) explaining the more fine perception in musicians, without the need for conscious attention [27–30]. According to these data, the musical training-induced long-term modulation of attention allocated to the auditory input may be underlay the strengthening of the sensory-related brain regions within STG that we observed in musicians in comparison of non-musicians, but it cannot actually explain how and why regions as STG4, STSp or FG, whose activation was completely lacking in non-musicians, are significantly involved in tone processing in musicians. Therefore, differences in state of auditory attention cannot explain a large portion of the results reported in the present study.

5. Conclusion

The present study confirms that in line with hypothesis 1, musicians shape temporal brain regions processing auditory traits common in both domains, for instance pitch and spectral analysis, and this, in turn, may allow for a better perception, not only of music but also of speech. Most importantly, our study showed that musicians in comparison to non-musicians have increased activation of brain regions segregating linguistic from musical sounds, giving first evidence for the validity of hypothesis 2. The superior linguistic competence in musicians are thus associated with regions that show a functional adaptation of a speech-specialized network, including auditory brain regions as the left STSp, which may be sensitive to the analysis of rapid changes in acoustic frequency, as well as brain regions functionally linked to phonological, articulatory or orthographic processing, such as left IFG, PMC and FG respectively. The finding that persons with musical training use regions for processing musical units that are not primarily recruited for such purposes (as left STSp and FG), but rather for speech, encourage the design of novel musical training based on musical hearing and practice for patients with deficits in processing auditory information, for children from low-socioeconomic backgrounds [8] or with developmental language disorders and dyslexia, but also for stroke patients with aphasia [16–18].

Acknowledgments

We would like to thank Hans-Jorg Mast, medical technical assistant of the Department of Diagnostic Imaging at University Hospital in Freiburg, Germany, and the study volunteers. Special thanks to Rebecca Sauter und Frederich Rombach for the assistance in preparing music

stimuli and in scanning volunteers. This work was (partly) supported by Brain Links-BrainTools Cluster of Excellence funded by the German Research Foundation (DFG, grant number EXC 1086).

References

- [1] F. Degé, G. Schwarzer, The effect of a music program on phonological awareness in preschoolers, *Front. Psychol.* 2 (2011) 124, <https://doi.org/10.3389/fpsyg.2011.00124>.
- [2] S. Moreno, C. Marques, A. Santos, M. Santos, S.L. Castro, M. Besson, Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity, *Cereb. Cortex* 19 (3) (2009) 712–723, <https://doi.org/10.1093/cercor/bhn120>.
- [3] Y. Nan, L. Liu, E. Geiser, H. Shu, C.C. Gong, Q. Dong, J.D.E. Gabrieli, R. Desimone, Piano training enhances the neural processing of pitch and improves speech perception in Mandarin-speaking children, *Proc. Natl. Acad. Sci. U. S. A.* 115 (28) (2018) E6630–E6639, <https://doi.org/10.1073/pnas.1808412115>.
- [4] M. Forgeard, E. Winner, A. Norton, G. Schlaug, Practicing a musical instrument in childhood is associated with enhanced verbal ability and nonverbal reasoning, *PLoS One* 3 (10) (2008) e3566, <https://doi.org/10.1371/journal.pone.0003566>.
- [5] S. Moreno, D. Friesen, E. Bialystok, Effect of music training on promoting pre-literacy skills: preliminary causal evidence, *Music Percept.: Interdiscip. J.* 29 (2011) 165–172.
- [6] L. Herrera, O. Lorenzo, S. Defior, G. Fernandez-Smith, G. Costa-Giomi, Effects of phonological and musical training on the reading readiness of native and foreign-Spanish-speaking children, *Psychol. Music* 39 (1) (2011) 68.
- [7] D. Register, The effects of an early intervention music curriculum on prereading/writing, *J. Music Ther.* Fall 38 (3) (2001) 239–248.
- [8] J. Slater, D.L. Strait, E. Skoe, S. O'Connell, E. Thompson, N. Kraus, Longitudinal effects of group music instruction on literacy skills in low-income children, *PLoS One* (2014).
- [9] L.R. Slevc, A. Miyake, Individual differences in second-language proficiency: does musical ability matter? *Psychol. Sci.* (8) (2006) 675–678.
- [10] J. Zuk, O. Ozernov-Palchik, H. Kim, K. Lakshminarayanan, J.D. Gabrieli, P. Tallal, N. Gaab, Enhanced syllable discrimination thresholds in musicians, *PLoS One* 8 (12) (2013) e80546, <https://doi.org/10.1371/journal.pone.0080546>.
- [11] J. Chobert, C. Marie, C. François, D. Schön, M. Besson, Enhanced passive and active processing of syllables in musician children, *J. Cogn. Neurosci.* 23 (12) (2011) 3874–3887, https://doi.org/10.1162/jocn_a_00088.
- [12] S. Elmer, C. Klein, J. Kühnis, F. Liem, M. Meyer, L. Jäncke, Music and language expertise influence the categorization of speech and musical sounds: behavioral and electrophysiological measurements, *J. Cogn. Neurosci.* 26 (2014) 2356–2369.
- [13] M. Besson, D. Schön, S. Moreno, A. Santos, C. Magne, Influence of musical expertise and musical training on pitch processing in music and language, *Restor. Neurol. Neurosci.* 25 (3) (2007) 399–410.
- [14] D.L. Strait, N. Kraus, A. Parbery-Clark, R. Ashley, Musical experience shapes top-down auditory mechanisms: evidence from masking and auditory attention performance, *Hear. Res.* 261 (1–2) (2010) 22–29, <https://doi.org/10.1016/j.heares.2009.12.021>.
- [15] G. Brod, B. Opitz, Does it really matter? Separating the effects of musical training on syntax acquisition, *Front. Psychol.* 3 (2012) 543, <https://doi.org/10.3389/fpsyg.2012.00543>.
- [16] P. Tallal, N. Gaab, Dynamic auditory processing, musical experience and language development, *Trends Neurosci.* 29 (7) (2006) 382–390, <https://doi.org/10.1016/j.tins.2006.06.003>.
- [17] K. Overy, Dyslexia and music. 2003 from timing deficits to musical intervention, *Ann. N. Y. Acad. Sci.* 999 (2003) 497–505.
- [18] N. Gaaba, J.D.E. Gabrieli, G.K. Deutsch, P. Tallal, E. Temple, Neural correlates of rapid auditory processing are disrupted in children with developmental dyslexia and ameliorated with training: an fMRI study, *Restor. Neurol. Neurosci.* 25 (2007) 295–310.
- [19] J. Du, R. Zatorre, Musical training sharpens and bonds ears and tongue to hear speech better, *Proc. Natl. Acad. Sci. U. S. A.* 114 (51) (2017) 13579–13584, <https://doi.org/10.1073/pnas.1712223114>.
- [20] R.J. Giuliano, P.Q. Pfordresher, E.M. Stanley, S. Narayana, N.Y. Wicha, Native experience with a tone language enhances pitch discrimination and the timing of neural responses to pitch change, *Front. Psychol.* 3 (2) (2011) 146, <https://doi.org/10.3389/fpsyg.2011.00146>.
- [21] D. Deutsch, T. Henthorn, E. Marvin, H. Xu, Absolute pitch among American and Chinese conservatory students: prevalence differences, and evidence for a speech-related critical period, *J. Acoust. Soc. Am.* 119 (2) (2006) 719–722.
- [22] S. Elmer, M. Meyer, L. Marrama, L. Jäncke, Intensive language training and attention modulate the involvement of fronto-parietal regions during a non-verbal auditory discrimination task, *Eur. J. Neurosci.* 34 (2011) 165–175.
- [23] B. Chandrasekaran, J.M. Hornickel, E. Skoe, T. Nicol, N. Kraus, Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: implications for developmental dyslexia, *Neuron* 64 (2009) 311–319.
- [24] J. Hornickel, E. Skoe, T. Nicol, S. Zecker, N. Kraus, Subcortical differentiation of stop consonants relates to reading and speech-in-noise perception, *Proc. Natl. Acad. Sci. U. S. A.* 106 (2009) 13022–13027.
- [25] S.H. Anvari, L.J. Trainor, J. Woodside, B.A. Levy, Relations among musical skills, phonological processing, and early reading ability in preschool children, *J. Exp. Child Psychol.* 83 (2) (2002) 111–130.
- [26] K. Banai, M. Ahissar, Musical experience, auditory perception and reading-related

- skills in children, *PLoS One* 8 (2013) e75876.
- [27] N. Kraus, B. Chandrasekaran, Music training for the development of auditory skills, *Nat. Rev. Neurosci.* 11 (8) (2010) 599–605, <https://doi.org/10.1038/nrn2882>.
- [28] G. Musacchia, M. Sams, E. Skoe, N. Kraus, Musicians have enhanced subcortical auditory and audiovisual processing of speech and music, *PNAS* 104 (40) (2007) 15894–15898, <https://doi.org/10.1038/nrn843>.
- [29] S.S. Asaridou, J.M. McQueen, Speech and music shape the listening brain: evidence for shared domain-general mechanisms, *Front. Psychol.* 4 (4) (2013) 321 doi: 10.3389.
- [30] M. Besson, J. Chobert, C. Marie, Transfer of training between music and speech: common processing, attention, and memory, *Front. Psychol.* 2 (2011), <https://doi.org/10.3389/fpsyg.2011.00094>.
- [31] A.D. Patel, Why would musical training benefit the neural encoding of speech? The OPERA hypothesis, *Front. Psychol.* 2 (2011) 142, <https://doi.org/10.3389/fpsyg.2011.00142>.
- [32] P.C.M. Wong, E. Skoe, N.M. Russo, T. Dees, N. Kraus, Musical experience shapes human brainstem encoding of linguistic pitch patterns, *Nat. Neurosci.* 10 (4) (2007) 420, <https://doi.org/10.1038/nrn1872>.
- [33] S. Elmer, J. Kühnis, P. Rauch, S. Abolfazl Valizadeh, L. Jäncke, The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: an EEG study, *Neuropsychologia* 51 (2013) 1608–1618.
- [34] C. Pantev, L.E. Roberts, M. Schulz, A. Engelien, B. Ross, Timbre-specific enhancement of auditory cortical representations in musicians, *NeuroReport* 12 (1) (2001) 169–174, <https://doi.org/10.1038/33918>.
- [35] R.J. Zatorre, P. Belin, V.B. Penhune, Structure and function of auditor cortex: music and speech, *Trends Cogn. Sci.* 6 (1) (2002) 37–46, [https://doi.org/10.1016/S1364-6613\(00\)01816-01817](https://doi.org/10.1016/S1364-6613(00)01816-01817).
- [36] K. Schulze, S. Koelsch, Working memory for speech and music, *Ann. N.Y. Acad. Sci.* 1252 (1) (2012) 229–236, <https://doi.org/10.1111/j.1749-6632.2012.06447>.
- [37] S. Elmer, M. Greber, A. Pushparaj, J. Kühnis, L. Jäncke, Faster native vowel discrimination learning in musicians is mediated by an optimization of mnemonic functions, *Neuropsychologia* 104 (2017) 64–75.
- [38] M. Musso, C. Weiller, A. Horn, V. Glauche, R. Umarova, J. Hennig, A. Schneider, M. Rijntjes, A single dual-stream framework for syntactic computations in music and language, *NeuroImage* 117 (2015) 267–283, <https://doi.org/10.1016/j.neuroimage.2015.05.020>.
- [39] A.D. Patel, Language, music, syntax and the brain, *Nat. Neurosci.* 6 (7) (2003) 674–681, <https://doi.org/10.1038/nrn1082>.
- [40] S. Koelsch, T. Fritz, K. Schulze, D. Alsop, G. Schlaug, Adults and children processing music: an fMRI study, *NeuroImage* 25 (4) (2005) 1068–1076, <https://doi.org/10.1093/cercor/bhu003>.
- [41] T. Ohnishi, H. Matsuda, T. Asada, M. Aruga, M. Hirakata, M. Nishikawa, A. Katoh, E. Imabayashi, Functional anatomy of musical perception in musicians, *Cereb. Cortex* 11 (8) (2001) 754–760 doi: 0.1093/cercor/11.8.754.
- [42] V.J. Schmithorst, S.K. Holland, The effect of musical training on music processing: a functional magnetic resonance imaging study in humans, *Neurosci. Lett.* 348 (2) (2003) 65–68, [https://doi.org/10.1016/s0304-3940\(03\)00714-6](https://doi.org/10.1016/s0304-3940(03)00714-6).
- [43] E. Fedorenko, J.H. McDermott, S. Norman-Haignere, N. Kanwisher, Sensitivity to musical structure in the human brain, *J. Neurophysiol.* 108 (12) (2012) 3289–3300, <https://doi.org/10.1152/jn.00209.2012>.
- [44] D.A. Abrams, A. Bhatarra, S. Ryal, E. Balaban, D.J. Levitin, V. Menon, Decoding temporal structure in music and speech relies on shared brainresources but elicits different fine-scale spatial patterns, *Cereb. Cortex* 21 (7) (2011) 1507–1518, <https://doi.org/10.1093/cercor/bhq198>.
- [45] J. Pa, G. Hickok, A parietal-temporal sensory-motor integration area for the human vocal tract: evidence from an fMRI study of skilled musicians, *Neuropsychologia* 46 (1) (2008) 362–368 Epub 2007 Aug 20.
- [46] J. Kühnis, S. Elmer, M. Meyer, L. Jäncke, The encoding of vowels and temporal speech cues in the auditory cortex of professional musicians: an EEG study, *Neuropsychologia* 51 (2013) 1608–1618.
- [47] C. Price, G. Thierry, T. Griffiths, Speech-specific auditory processing: where is it? *Trends Cogn. Sci.* 9 (6) (2005) 271–276.
- [48] C. Marie, F. Delogu, G. Lampis, M.O. Belardinelli, M. Besson, Influence of musical expertise on segmental and tonal processing in Mandarin Chinese, *J. Cogn. Neurosci.* 23 (10) (2011) 2701–2715, <https://doi.org/10.1162/jocn.2010.21585> Epub 2010 Oct 14.
- [49] K.J. Friston, W.D. Penny, D.E. Glaser, Conjunction revisited *neuroimage*, *Neuroimage* 25 (3) (2005) 661–667 15.
- [50] K.J. Friston, A.P. Holmes, C.J. Price, C. Büchel, K.J. Worsley, Multisubject fMRI studies and conjunction analyses, *Neuroimage* 10 (4) (1999) 385–396.
- [51] C.J. Price, K.J. Friston, Cognitive conjunction: a new approach to brain activation experiments, *Neuroimage* 5 (4 Pt 1) (1997) 261–270.
- [52] S. Greenberg, The ears have it: the auditory basis of speech perception. International congress of phonetic, Stockholm 3 (1995) 34–41.
- [53] Kenneth De Jong, Temporal constraints and characterizing syllable structuring, in: John Local, Richard Ogden, Rosalind Temple (Eds.), *Phonetic Interpretation: Papers in Laboratory Phonology VI*, Cambridge University Press, 2003, pp. 253–253268, <https://doi.org/10.1017/CBO9780511486425.015> ISBN 978-0-521-82402-6. Page 254.
- [54] Adam I. Cooper, Draga Zec, Syllables, in: Mark Aronoff (Ed.), *Oxford Bibliographies Online: Linguistics*, Oxford University Press, New York, 2013.
- [55] C.J.A. Alvarez, M. Carreiras, M. Perea, Are syllables phonological units in visual word recognition? *Lang. Cogn. Process.* 19 (3) (2004) 427–452.
- [56] A.D. Patel, *Music, Language, and the Brain*, Oxford University Press, New York, NY, 2008.
- [57] M. Steinschneider, K.V. Nourski, Y.I. Fishman, Representation of speech in human auditory cortex: Is it special? *Hear. Res.* 305 (2013) doi:10.1016/j.heares.2013.05.013.
- [58] A.M. Liberman, I.G. Mattingly, The motor theory of speech perception revised, *Cognition* 21 (1) (1985) 1–36.
- [59] R. Näätänen, A. Lehtokoski, M. Lennes, M. Cheour, M. Huotilainen, A. Iivonen, M. Vainio, P. Alku, R.J. Ilmoniemi, A. Luuk, J. Allik, J. Sinkkonen, K. Alho, Language-specific phoneme representations revealed by electric and magnetic brain responses, *Nature* 385 (6615) (1997) 432–434, <https://doi.org/10.1038/385432a0>.
- [60] S.E. Blumstein, W.E. Cooper, E.G. Zurif, A. Caramazza, The perception and production of voice-onset time in aphasia, *Neuropsychologia* (3) (1977) 371–383.
- [61] S.M. Wilson, A.P. Saygin, M.I. Sereno, M. Iacoboni, Listening to speech activates motor areas involved in speech production, *Nat. Neurosci.* 7 (7) (2004) 701.
- [62] S.E. Blumstein, E.B. Myers, J. Rissman, The perception of voice onset time: an fMRI investigation of phonetic category structure, *J. Cogn. Neurosci.* 17 (9) (2005) 1353–1366, <https://doi.org/10.1162/0899929054985473>.
- [63] L. Fadiga, L. Craighero, G. Buccino, G. Rizzolatti, Speech listening specifically modulates the excitability of tongue muscles: a TMS study, *Eur. J. Neurosci.* (2) (2002) 399–402.
- [64] F. Pulvermüller, M. Huss, F. Kherif, F. Moscoso del Prado Martin, O. Hauk, Y. Shtyrov, Motor cortex maps articulatory features of speech sounds, *Proc. Natl. Acad. Sci. U. S. A.* 103 (20) (2006) 7865–7870 Epub 2006 May 8.
- [65] T. Paus, D.W. Perry, R.J. Zatorre, K.J. Worsley, A.C. Evans, Modulation of cerebral blood flow in the human auditory cortex during speech: role of motor to-sensory discharges, *Eur. J. Neurosci.* 8 (1996) 2236–2246.
- [66] S. Elmer, J. Kühnis, P. Rauch, S. Abolfazl Valizadeh, L. Jäncke, Functional connectivity in the dorsal stream and between bilateral auditory-related cortical areas differentially contribute to speech decoding depending on spectro-temporal signal integrity and performance, *Neuropsychologia* 106 (2017) 398–406.
- [67] 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 (7) (2007) 547–558.
- [68] R.C. Oldfield, The assessment and analysis of handedness: the Edinburgh inventory, *Neuropsychologia* 9 (1) (1971) 97–113.
- [69] M. Zaitsev, C. Dold, J. Hennig, O. Speck, Prospective real-time slice-by-slice 3D motion correction for EPI using an external optical motion tracking system, *Proc. Intl. Soc. Mag. Reson. Med.* 11 (2004).
- [70] M. Zaitsev, J. Hennig, O. Speck, Point spread function mapping with parallel imaging techniques and high acceleration factors: fast, robust, and flexible method for echo-planar imaging distortion correction, *Magn. Reson. Med.* 52 (5) (2004) 1156–1166, <https://doi.org/10.1002/mrm.20261>.
- [71] C.G.M. Ott, N. Langer, M.S. Oechslin, M. Meyer, L. Jäncke, Processing of voiced and unvoiced acoustic stimuli in musicians, *Front. Psychol.* 6 (2) (2011) 195, <https://doi.org/10.3389/fpsyg.2011.00195>.
- [72] P. Schneider, M. Scherg, H.G. Dosh, H.J. Specht, A. Gutschalk, A. Rupp, Morphology of Heschl's gyrus reflects enhanced activation in the auditory cortex of musicians, *Nat. Neurosci.* 5 (7) (2002) 688–694, <https://doi.org/10.1038/nrn871>.
- [73] M. Steinschneider, V. Nourski, Y.I. Fishman, Representation of speech in human auditory cortex: is it special? *Hear. Res.* 305 (2013) doi:10.1016/j.heares.2013.05.013.
- [74] D.I. Bendor, X. Wang, Cortical representations of pitch in monkeys and humans, *Curr. Opin. Neurobiol.* 16 (4) (2006) 391–399 Epub 2006 Jul 13.
- [75] N. Mesgarani, C. Cheung, K. Johnson, E.F. Chang, Phonetic feature encoding in human superior temporal gyrus, *Science* 343 (6174) (2014) 1006–1010, <https://doi.org/10.1126/science.1245994>.
- [76] R.R. Benson, M. Richardson, D. Whalen, S. Lai, Phonetic processing areas revealed by sinewave speech and acoustically similar non-speech, *NeuroImage* 31 (1) (2006) 342–353, <https://doi.org/10.1016/j.neuroimage.2005.11.029>.
- [77] L. Jäncke, T. Wüstenberg, H. Scheich, H.J. Heinze, Phonetic perception and the temporal cortex, *NeuroImage* 15 (4) (2002) 733–746, <https://doi.org/10.1006/nimg.2001.1027>.
- [78] E. Liebenthal, J.R. Binder, S.M. Spitzer, E.T. Possing, D.A. Medler, Neural substrates of phonemic perception, *Cereb. Cortex* 15 (10) (2005) 1621–1631, <https://doi.org/10.1093/cercor/bhi040>.
- [79] P. Belin, R.J. Zatorre, P. Lafaille, P. Ahad, B. Pike, Voice-selective areas in human auditory cortex, *Nature* 403 (6767) (2000) 309–312, <https://doi.org/10.1038/35002078>.
- [80] G.A. Calvert, E.T. Bullmore, M.J. Brammer, R. Campbell, S.C. Williams, P.K. McGuire, P.W. Woodruff, S.D. Iversen, A.S. David, Activation of auditory cortex during silent lipreading, *Science* 276 (5312) (1997) 593–596, <https://doi.org/10.1126/science.276.5312.593>.
- [81] E. Paulsen, D. Perani, V. Blasi, G. Silani, N.A. Borghese, U. De Giovanni, S. Sensolo, F. Fazio, A functional-anatomical model for lip-reading, *J. Neurophysiol.* 90 (3) (2003) 2005–2013, <https://doi.org/10.1152/jn.00926.2002>.
- [82] D. Boatman, C. Hall, M.H. Goldstein, R. Lesser, B. Gordon, Neuropsychological differences in consonant and vowel discrimination: as revealed by direct cortical electrical interference, *Cortex* 33 (1) (1997) 83–98.
- [83] P.K. Kuhl, K.A. Williams, F. Lacerda, K.N. Stevens, B. Lindblom, Linguistic experience alters phonetic perception in infants by 6 months of age, *Science* 255 (5044) (1992) 606–608, <https://doi.org/10.1126/science.1736364>.
- [84] G.H. Recanzone, C.E. Schreiner, M.M. Merzenich, Plasticity in the frequency representation of primary auditory cortex following discrimination training in adult owl monkeys, *J. Neurosci.* 13 (1) (1993) 87–103, <https://doi.org/10.1523/JNEUROSCI.13-01-00087.1993>.
- [85] R.B. Gillam, D.F. Loeb, L.M. Hoffman, T. Bohman, C. Champlin, L. A. Thibodeau, J. Widen, J. Brandel, S. Friel-Patti, The efficacy of fast for word language intervention in school-age children with language impairment: a randomized controlled

- trial, *J. Speech Lang. Hear. Res.* (2008), [https://doi.org/10.1044/1092-4388\(2008/007\)](https://doi.org/10.1044/1092-4388(2008/007)).
- [86] S. Dehaene, F. Pegado, L.W. Braga, P. Ventura, G. Nunes Filho, A. Jobert, G. Dehaene-Lambertz, R. Kolinsky, J. Morais, L. Cohen, How learning to read changes the cortical networks for vision and language, *Science* 330 (6009) (2010) 1359–1364, <https://doi.org/10.1126/science.1194140>.
- [87] G.A. Calvert, R. Campbell, Reading speech from still and moving faces: the neural substrates of visible speech, *J. Cogn. Neurosci.* 15 (1) (2003) 57–70, <https://doi.org/10.1162/089892903321107828>.
- [88] S. Dehaene, L. Cohen, The unique role of the visual word form area in reading, *Trends Cogn. Sci.* 15 (6) (2011) 254–262, <https://doi.org/10.1016/j.tics.2011.04.003>.
- [89] R. Hashimoto, K.L. Sakai, Learning letters in adulthood: direct visualization of cortical plasticity for forming a new link between orthography and phonology, *Neuron* 42 (2004) 311–322.
- [90] K. Von Kriegstein, A. Kleinschmidt, P. Sterzer, A.L. Giraud, Interaction of face and voice areas during speaker recognition, *J. Cogn. Neurosci.* 17 (3) (2005) 367–376.
- [91] J.E. Goold, M. Meng, Categorical learning revealed in activity pattern of left fusiform cortex, *Hum. Brain Mapp.* 38 (7) (2017) 3648–3658, <https://doi.org/10.1002/hbm.23620>.
- [92] Y.K. Wong, I. Gauthier, A multimodal neural network recruited by expertise with musical notation, *J. Cogn. Neurosci.* 22 (4) (2010) 695–713, <https://doi.org/10.1162/jocn.2009.21229>.
- [93] E. Feredoes, G. Tononi, B.R. Postle, Direct evidence for a prefrontal contribution to the control of proactive interference in verbal working memory, *Proc. Natl. Acad. Sci. U. S. A.* 103 (51) (2006) 19530–19534.
- [94] M. Hansen, M. Wallentin, P. Vuust, Working memory and musical competence of musicians and non-musicians, *Psychol. Music* 41 (6) (2012) 779–793.
- [95] P. Bermudez, J.P. Lerch, A.C. Evans, R.J. Zatorre, Neuroanatomical correlates of musicianship as revealed by cortical thickness and voxel-based morphometry, *Cereb. Cortex* 19 (7) (2008) 1583–1596, <https://doi.org/10.1093/cercor/bhn196>.
- [96] C. Gaser, G. Schlaug, Brain structures differ between musicians and non-musicians, *J. Neurosci.* 23 (27) (2003) 9240–9245, <https://doi.org/10.1523/JNEUROSCI.23-27-09240.2003>.
- [97] D.L. Strait, N. Kraus, Can you hear me now? Musical training shapes functional brain networks for selective auditory attention and hearing speech in noise, *Front. Psychol.* 13 (2) (2011) 113, <https://doi.org/10.3389/fpsyg.2011.00113> eCollection 2011.
- [98] K.R. Pugh, B.A. Shaywitz, S.E. Shaywitz, R.K. Fulbright, D. Byrd, P. Skudlarski, D.P. Shankweiler, L. Katz, R.T. Constable, J. Fletcher, Auditory selective attention: an fMRI investigation, *Neuroimage* 4 (1996) 159–173.