Song Perception by Professional Singers and Actors: An MEG Study

Ken Rosslau[1]*, Sibylle C. Herholz[2,3], Arne Knief[1], Magdalene Ortmann[2,4], Dirk Deuster[1], Claus-Michael Schmidt[1], Antoinetteam Zehnhoff-Dinnesen[1], Christo Pantev[2], Christian Dobel[2,5]

1 Department of Phoniatrics and Pedaudiology, University Hospital Muenster, Muenster, Germany, 2 Institute for Biomagnetism and Biosignalanalysis, University of Muenster, Muenster, Germany, 3 German Center for Neurodegenerative Diseases (DZNE), Bonn, Germany, 4 Jean-Uhrmacher-Institute for Clinical ENT-Research, University Hospital Cologne, Cologne, Germany, 5 Department of Otorhinolaryngology, Friedrich-Schiller University Jena, Jena, Germany

* ken.rosslau@uni-muenster.de

Abstract

The cortical correlates of speech and music perception are essentially overlapping, and the specific effects of different types of training on these networks remain unknown. We compared two groups of vocally trained professionals for music and speech, singers and actors, using recited and sung rhyme sequences from German art songs with semantic and/ or prosodic/melodic violations (i.e. violations of pitch) of the last word, in order to measure the evoked activation in a magnetoencephalographic (MEG) experiment. MEG data confirmed the existence of intertwined networks for the sung and spoken modality in an early time window after word violation. In essence for this early response, higher activity was measured after melodic/prosodic than semantic violations in predominantly right temporal areas. For singers as well as for actors, modality-specific effects were evident in predominantly left-temporal lateralized activity after semantic expectancy violations in the spoken modality, and right-dominant temporal activity in response to melodic violations in the sung modality. As an indication of a special group-dependent audiation process, higher neuronal activity for singers appeared in a late time window in right temporal and left parietal areas, both after the recited and the sung sequences.

Introduction

Recent research has increased our knowledge about the organization of neuronal networks for speech and music perception and suggests the presence of training-induced and interdependent modulation of musical and speech abilities [1]. This research is based on many studies on brain morphology, training effects and receptive/expressive functions of music and speech processing comparing instrumental musicians to novices [2–5]. In contrast, there is little knowledge about training or profession-specific cortical processing of speech and music as, for example, in professional voice experts such as actors in comparison to opera singers. Both of these groups have comparable levels of voice training and practice on stage, but with different

emphasis on the specific type of voice training. Therefore, in order to investigate the training-related effects on brain function, it is much more informative to compare behavioral and neurophysiological data with respect to these fields of expertise, than to relate these groups to untrained novices. We regard singers and actors as a unique group of artists for comparative purposes, because both need to work on artistic expression through their voices as well as bodies and with high demands on self-perception. Furthermore, the comparison of these groups, using both a spoken and sung stimulus modality, addresses with very high specificity the questions of modality dependence, group dependence or any interaction in processing linguistic and musical content. Based on similar semantic and syntactic rule systems in language and in music, a complex and intertwined cerebral network for language and music processing is assumed [3,5–7]. Nevertheless, there is to date no study comparing two groups of experts who developed their expertise with very similar amounts and types of training with the same stimulus material, once in a sung and once in a recited modality.

An intertwined network for processing music and language

Previous research transferred experimental approaches, which were established to investigate different levels of language processing, into the field of music processing. Most notable were designs employing semantic and syntactic expectancy violations. Semantic expectancy violations in language result in a N400 component generated mainly in the left superior temporal lobe, as evidenced by electro- and magnetoencephalographic (EEG and MEG) studies [8–14]. Syntactic violations in spoken sentences are reflected in an early negative electrophysiological component (ELAN; early left anterior negativity) and/or a late positive centro-parietal component (P600) over left anterior temporal and left inferior frontal regions [15]. Early left anterior magnetic fields after syntactical violation were also detected by MEG [16], while there is to date no clear correlate for the electrophysiological P600 component in MEG. Similarly to these findings, semantic and syntactic expectancy violations in musical material elicit negative electrophysiological components in right anterior frontal and superior temporal regions that are homologous to the above-mentioned left-lateralized speech-related correlates [17,18], i.e. the regions are highly similar, but with different hemispheric dominance. In a magnetoencephalographic study, high neuronal activity after musical syntactical violation was found in temporal regions on both hemispheres [19]. Typical "language" regions seemed to be less language-specific than previously thought [20]. Still, assuming a relative dominance of hemispheres in musical versus linguistic contexts, right temporal areas are reported to be mainly involved in processing and analyzing musical sequences [21–23]. In this vein, several studies investigated pitch violations in music and speech. An increment or decrement of fundamental frequency (final pitch) at the end of a spoken or sung line may represent a prosodic or melodic violation, respectively. Both can be interpreted as a violation of the syntactical rule system, and thus several studies focused on such prosodic/ melodic differentiations. They found evidence for positive centro-parietal and temporal components peaking between 300–600 ms after stimulus onset, as described for syntactical violations [24,25]. The amplitude of these components depended on the strength of violation (weak or strong) and on the degree of musical education of the participant [26,27].

In order to test whether simultaneously presented linguistic parameters (represented by semantic violations) and musical parameters (represented by melodic pitch violations at the end of a sung melody line) are processed dependently or independently, a medium is required that combines these two aspects. Comparing different musicians and laymen, Bonnel and coauthors [28] prepared excerpts from French operatic songs by manipulating the final word in such a way that it was either semantically congruous (S+) or incongruous (S-) and/ or by

manipulating the final pitch of the melody line either in (P+) or out of key (P-). The simultaneous appearance of both an N400 and a P600 component in response to the combined prosodic and semantic violated condition (S-P-) suggested that semantic and syntactic aspects of language and music were processed by independent systems and did not compete for the same pool of mental resources in musicians and nonmusicians [28,29]. However, successive studies failed to find this division of labour, and rather presented evidence for more intertwined neuronal networks in bilateral middle and superior temporal gyri as well as inferior and middle frontal gyri, during combined musical and linguistic tasks [18,30,31]. Most of the above-mentioned studies compared the neurophysiological influence of linguistic content in divided sets of stimuli for language and music conditions, respectively. The advantage of using song lines performed by the human voice is that both linguistic and musical information are merged into one ecologically valid acoustic signal. The separation into a recited and a sung version allows a comparison of more linguistically-based and a more musically-based context with the same experimental material, which is a prerequisite for a study of highly professional artists. If there are interactions of semantic and syntactic processing in either a recited or sung modality, professional opera singers, as highly trained musical voice users, and professional actors as highly trained linguistic voice users, represent ideal subjects to search for neurophysiological correlates.

Cognitive and neuronal characteristics for singers and actors

During singing, professional singers display increased activation of bilateral primary somatosensory cortex (where cortical representations of the larynx are situated), inferior parietal lobe and dorsolateral prefrontal cortex, and at a subcortical level, increased activation in the basal ganglia, thalamus and cerebellum compared to nonmusicians. This is is generally interpreted as evidence for training-induced cortical plasticity [32,33]. To the best of our knowledge, there is only one study investigating training-induced plastic effects as a result of acting training, identifying high activation during speech perception in bilateral premotor regions that are commonly activated by mouth movements [34].

Regarding a specialization of higher-order cognitive skills, several findings over the last few years point towards an enhanced quality of auditory imagery in musicians [35,36]. Musical imagery preserves many structural and temporal properties of auditory stimuli and can facilitate auditory discrimination by, for instance, the integration of semantically interpreted information and expectancies [37]. A special form of imagery, so called "audiation", is described as an internal analog of aural music perception [38] and interpreted as a mental representation of music by internally "hearing" a music sequence that has just been auditorily or visually presented. It represents an integration of auditory, visual and/ or motor imagery in the brain and results in a cross-modal encoding of a unisensory input [39]. In line with such a description, audiation should be especially developed in musicians. However, the neural correlates of audiation haven't been investigated so far.

Aim and approach of the current study

The aim of our study was to investigate music and speech perception by voice experts, professional singers and actors, in order to disentangle the training-induced cortical networks for processing music and speech. To measure brain activity we used magnetoencephalography (MEG) due to its high sensitivity to time and its moderate to high accuracy in determining the underlying sources of brain activity [40]. This is the first study comparing these groups by using complex, but ecologically valid stimulus material in recited and sung modalities. Although all native-speaking participants are per definition linguistically highly educated in

speaking their mother tongue, we consider it important to compare singers with actors in order to control for long-term voice training. This would be not the case in participants without such experience. To stimulate at a high artistic level, we used rhyme sequences of German art songs by Franz Schubert. Importantly, the lyrical basis for these songs is similar in structure to material that actors recite in a dramatic performance. One characteristic of art songs is a close integration of music and lyrics, typically without singing several notes in one syllable, a frequent feature in operatic arias. Since the songs are based on poetry, it is feasible to present the material both in a spoken and in a sung condition, thus comparing modality-specific processing of semantic and syntactic aspects.

Based on the nature of the semantic and melodic/ prosodic violations and because of using a sung and spoken modality, we expected increased activity upon violations in temporal areas in both hemispheres. Additionally, we predicted higher sensitivity for melodic/prosodic violations in singers and vice versa for semantic violations in actors. If singers indeed display more long lasting representations of auditory stimuli after their offset (i.e. what was above called audiation), we expect long-lasting activity in temporal regions, possibly with a right-hemispheric dominance, due to musical training.

Material and Methods

Participants

Fifteen professional singers (mean age = 29.2 years; 8 female) and 15 professional actors (mean age = 32.4 years; 9 female) took part in the experiment. The singers and actors had passed a university final qualifying examination after at least 4 years of training. At the time of the study, they currently practiced singing or acting on stage or in rehearsal for a minimum of 4 hours a day. The actors had not received any additional musical education besides compulsory music classes in high school and the singers had received articulation training for one year at the beginning of their university studies.

As an inclusion criterion all participants were familiar with the German art song cycles "Beautiful Miller Girl" and "Winter Journey" by the composer Franz Schubert, but had not practiced or performed them in auditions or on stage. All participants were right handed, free of neurological or psychiatric disorders, native speakers of German and had normal hearing thresholds as assessed by clinical audiometry. All gave written consent to participate in the study. The study protocol was approved by the local ethics committee of the Medical Faculty. The study was conducted according to the Declaration of Helsinki.

Stimulus material

As indicated above, we used 30 short excerpts of songs from the romantic epoch (music by the German composer Franz Schubert, lyrics by Wilhelm Mueller) from the cycles "Beautiful Miller Girl" and "Winter Journey" for stimulation in the experiment [41,42]. The excerpts consisted of a rhyming couplet with a monosyllabic ending and the original melody line composed by Franz Schubert. For all excerpts, one version sung *a capella* (without accompaniment) and one recited, spoken version were recorded using a high-quality recording system and microphone (lingwaves software/ Wevosys 2010; microphone: 322 Datalogger, MK:216/ Voltcraft). For the recording, the same professionally educated singer sang and recited all excerpts. The duration of sung phrases ranged from 4.5 to 6.8 seconds (mean average 5.7 sec.), and the duration of recited phrases ranged from 4.2 to 5.4 seconds (mean average 4.8 sec.). In the same way, the mean length of the recited last words (452 ± 26 ms) differed from the mean of length of sung last words (710 ± 44 ms). For each modality (sung and spoken), the 30 excerpts were presented in four different conditions, resulting in 120 stimuli per modality. In the first condition,

the original line was presented in the correct sung/ recited version (S+P+, for correct **s**emantic and **p**itch information). In the second condition, the pitch of the last word was decreased or increased in the sung modality by a semitone out of key in compliance with the original melodic contour (melodic violation), and in the spoken modality, by an increase of fundamental frequency of 35% (prosodic violation), which represents a violation of the expected decrease of prosody for a clause of statement (S+P-). A different relation between the deviation of a fundamental frequency in music and speech was first claimed by Besson et al. [27]. The authors described the deviation of 1/5 tone in music and 35% increase of the fundamental frequency (quart interval) in speech to be appropriate for a "weak" incongruity, because it is much harder to recognize such a difference in speech compared to the harmonic context of music. This is probably due to the strong harmonic rule system for melody compared to the only sensational rule system for speech prosody. After piloting our stimulus material with a group of healthy musical students, we decided an interval of ½ tone compared to 35% increase of fundamental frequency in speech to be more appropriate for our study. In the third condition, the original last word of the excerpt was exchanged by a semantically incongruent word (S-P+). These semantically incongruent monosyllabic words fulfilled the original rhyme scheme. In the fourth condition, we presented a double incongruency at the end of the excerpt with an incorrect pitch ending (syntactic/ prosodic violation) and a semantically incongruent last word (S-P-). All pitch manipulations in the sung and spoken modality were performed on the original, digitally stored sound files using the software PRAAT (Version 5.3.34) to ensure the correct pitch violation (Fig 1).

Procedure

Subjects were comfortably seated in a magnetically shielded room and their head position was stabilized in the MEG scanner using soft pads. All stimuli were presented binaurally 60 dB above the individual hearing threshold of each ear, which was determined at the beginning of the experiment with an accuracy of at least 5 dB by reduction of the sound of one stimulus sentence to the minimal individual sound level both for the sung and spoken modality. Instructions, visual prompts and feedback were presented via back-projection on a screen in front of the subject that was adjusted in height to be comfortably visible for the subject.

Subjects worked through the experimental instructions and eight practice trials at their own pace. Stimuli used for the practice trials were not used again in the subsequent experiment. The 240 stimuli were presented in 4 experimental runs per 60 stimuli using the software Presentation (Neurobehavioral Systems Inc., Albany, CA, USA). Within each run, stimuli were presented in a pseudo-randomized order. Occurrence of the different versions of each excerpt were distributed equally across four runs, with the constraint that two versions of the same excerpt did not occur subsequently and not more than 3 stimuli from the same condition were played consecutively.

After the presentation of each stimulus, subjects had to judge the accuracy of the semantic congruence of the last word and the accuracy of the pitch of the last word, both for sung and for spoken stimuli. Subjects responded by means of successive button presses and were visually prompted to give their responses, with the prompt for the first judgment appearing 1500 ms after stimulus offset. They were instructed to respond within 2000 ms. The next prompt or trial was presented automatically after the subject's response or after a time lapse of 5 seconds. The order of the prompts (semantic and pitch judgments) and the assignment of buttons to responses (correct and incorrect) were balanced across participants and remained the same for each subject throughout the experiment. Each run took around 15 minutes and the entire
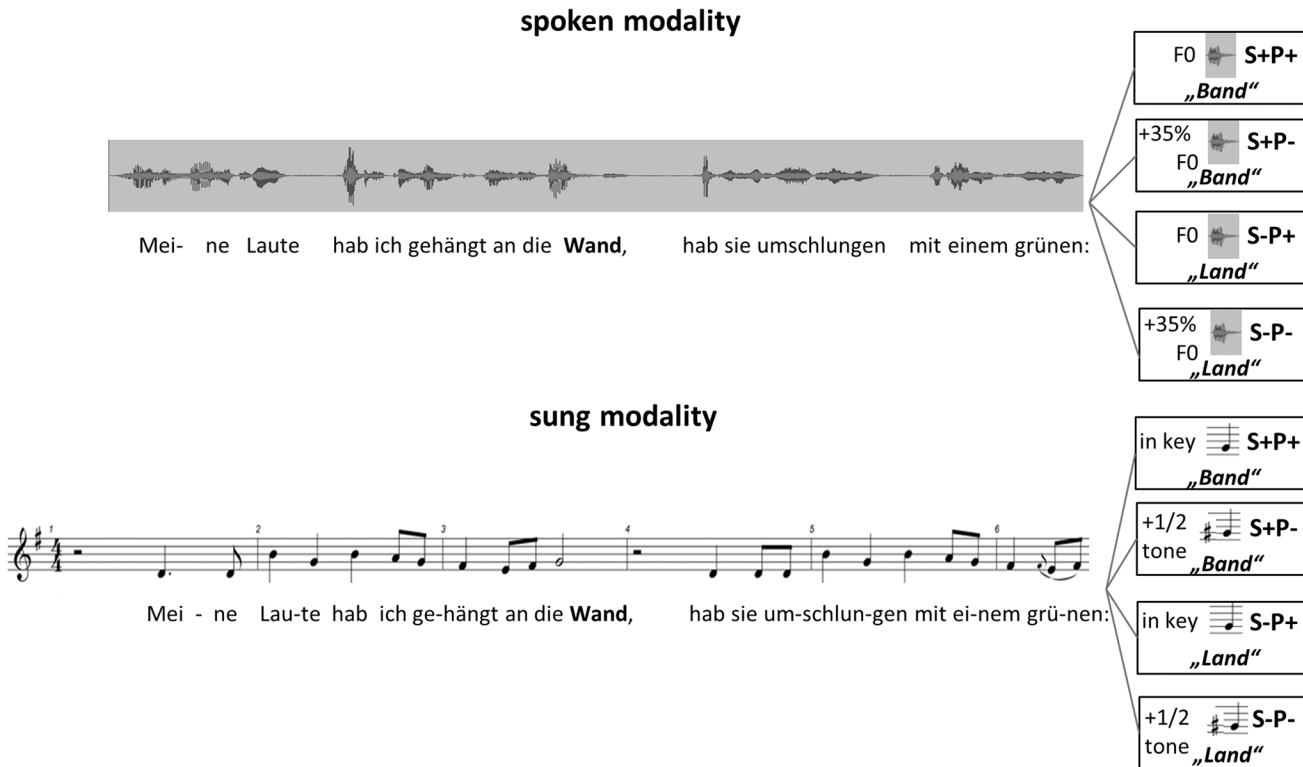
Fig 1. Example of a varied rhyme-couplet. Example of a varied rhyme-couplet from the song cycle "Beautiful Miller Girl" by Franz Schubert, poems by Wilhelm Müller: *Meine Laute hab ich gehängt an die **Wand**, hab sie umschlungen mit einem grünen **Band*** (semantic correct) / ***Land** (incorrect)*, (english translation by Emily Ezust: My lute I´ve hung upon the wall, I´ve tied it there with a green **band/ land**). Semantic variation of the last word and/ or prosodic/ melodic variation of the final pitch resulted in 4 different conditions (S+: correct semantic sense, S-: incorrect semantic sense, P+: correct fundamental frequency/ final pitch, P-: incorrect fundamental frequency/ final pitch) for both spoken and sung modalities.

measurement process including instructions, practice trials and pauses between runs, lasted about 90 minutes.

After the measurements, subjects took part in a semi-structured interview to summarize how attention-demanding they had found the tasks to be. For evaluation purposes, the answers were classified into three different categories from "low", "moderate" and "high" level of attention.

MEG recordings and data analysis

MEG signals were recorded continuously, using a whole-head device with 275 first-order axial SQUID gradiometers (Omega 275, CTF, VSM MedTech, Coquitlam, Canada), filtered online (150 Hz low-pass for aliasing, 50 Hz notch for European power grid) and sampled at 600 Hz. The continuous data were then band-pass filtered offline in a 0.1–48 Hz range, using a zero-phase second-order Butterworth filter. The triggers for data analysis were set at the beginning of the last word for each stimulus. For each trial, epochs ranging from 200 ms before acoustic trigger at the word onset to 2000 ms after onset were extracted from the continuous data. Artifact rejection and pre-processing, with baseline correction of the first 100 ms and rejection of sensor activity higher than 3000 fT, was performed with EMEGS 2.3 [43] running under MATLAB 7 SP3 (The MathWorks, Natick, MA, USA). Epochs for each condition were averaged. Individual averages were standardized on the mean MEG sensor configuration across all participants and runs, and thus corrected for differing head positions of the participants within

the MEG scanner. The amplitude and distribution of event-related magnetic fields depended on the individual head position within the sensor coordinate system, as well as individual head geometry, especially head size. An estimation of the underlying neuronal generators, such as the L2- Minimum-Norm Estimate, (L2-MNE; [40] however, is independent of such individual factors and enables statistical tests across participant groups and conditions. The L2-MNE served as an inverse-distributed source modeling method for examining the cortical generator of the magnetic field activity without a priori assumptions about the location and/or number of current sources. The present analyses were based on an isotropic spherical head model with 197 dipolar sources evenly distributed on an inner spherical shell. The sphere position and radius were estimated in order to optimally fit the digitized head shape of each participant. Across all participants and conditions, a Tikhonov regularization parameter of $k = 0.2$ was applied.

Dipole strength at a given dipole site was obtained as the square root of the sum of squared L2 values for each of the two tangential orientations, for each time point for each data set. The L2-MNE amplitudes were analyzed with a point-wise repeated measures ANOVA with the within-subject factor CONDITION and the between-subject factor GROUP separately for the spoken and sung modalities. To avoid false positives, a significance criterion of $p < 0.01$ was used, and significant effects were considered only when they were observed for at least 10 consecutive sampling points (e.g. around 15 ms) and at least 10 neighboring dipoles. The statistical parametric F values were mapped on a standard cortical surface in time slots of 50 ms in order to display the origin of effects in more detail. Such foci of high activity were further analyzed by averaging the mean activity as different clusters of both hemispheres. This type of analysis for multichannel recordings (EEG and MEG) has become an established procedure for sensor and source space (recent studies include [12,44].

For a comparison of local clusters with high activity, the relevant time windows for defining clusters were based on significant activation differences of the point-wise ANOVA. In line with the literature, we detected clusters of activity in an interval between 200 and 500 ms after onset of the last word in both temporal lobes, and we described foci of activity in this time window as "early" components. Because we also detected activation peaks in a second interval between 600 and 1700 ms with a significant dependence on the between-factor GROUP, we described these activations as "late" components. Even though we defined the clusters in a data-driven manner, the dipoles in these clusters overlapped substantially comparing the hemispheres separately for each modality (spoken modality: left hemisphere 22 dipoles and right hemisphere 22 dipoles, with 18 corresponding dipoles; sung modality: left hemisphere 16 dipoles and right hemisphere 19 dipoles, with 11 corresponding dipoles). For "early" cluster comparison, we calculated a repeated measures ANOVA including within-subject factors SEMANTIC VIOLA TION, MELODIC/ PROSODIC VIOLATION, HEMISPHERE and between-subjects factor GROUP. Because no corresponding dipoles were found in the "late" clusters, we calculated separately for each hemisphere a repeated measures ANOVA, including within-subject factors SEMANTIC VIOLATION, MELODIC/ PROSODIC VIOLATION and between-subject factor GROUP.

All analyses were conducted separately for the sung and spoken modality to minimize bias caused by differences in length of the last word for the sung and spoken versions and the different time window of resulting magnetic fields. Pairwise post-hoc comparisons between significant and relevant condition pairs were computed and thresholded by Bonferroni correction.

Analysis of behavioral data

In order to evaluate the behavioural data, we computed the mean values of hits and correct rejections concerning melodic/ prosodic correctness of the last word as the accuracy of pitches

and concerning the semantic congruence of the last word, as the accuracy of words, in order to obtain more detailed information of the kind of mistakes associated with the different conditions. In the same way as for the MEG data, performance scores were analyzed using a repeated measures ANOVA with the factors CONDITION and GROUP separately for the sung and spoken modalities and additionally, as mentioned before, each for accuracy of pitches and accuracy of words. The ANOVA results are reported when significant at p < = 0.05. All p values for results were adjusted, when necessary, with the Greenhouse-Geisser epsilon correction for nonsphericity. Pairwise post-hoc comparisons between significant and relevant condition pairs were computed and thresholded as before by Bonferroni correction.

Results

Behavioural data

Spoken modality. The results concerning the accuracy of the ending pitch of the line (melodic/ prosodic violation) revealed a significant main effect of CONDITION ($F_{(1.75, 49.11)}$ = 6.33, p = .005), and no main effect of and interaction with GROUP. Comparing the significant mean values for post-hoc analysis ([Fig 2]) showed that the performance was lowest for the condition with double incongruency S-P- (24.7 ± 5.2 correct responses). This condition differed from all other three (S+P-: 26.5 ± 3.6 c.r., S-P+: 27.1 ± 2.9 c.r. and S+P+: 28.0 ± 2.8 c.r.; post-hocs: S+P+ vs. S-P-, p = .003; S+P- vs. S-P-, p = .011; S-P+ vs. S-P-, p = .020), but they were not significantly different from each other.

Identifying the semantic accuracy of the last words (accuracy of words) in the spoken modality yielded a significant main effect CONDITION ($F_{(1.69, 47.36)}$ = 4.24, p = .026), and again, no main effect of and interaction with GROUP. Comparing the significant mean values
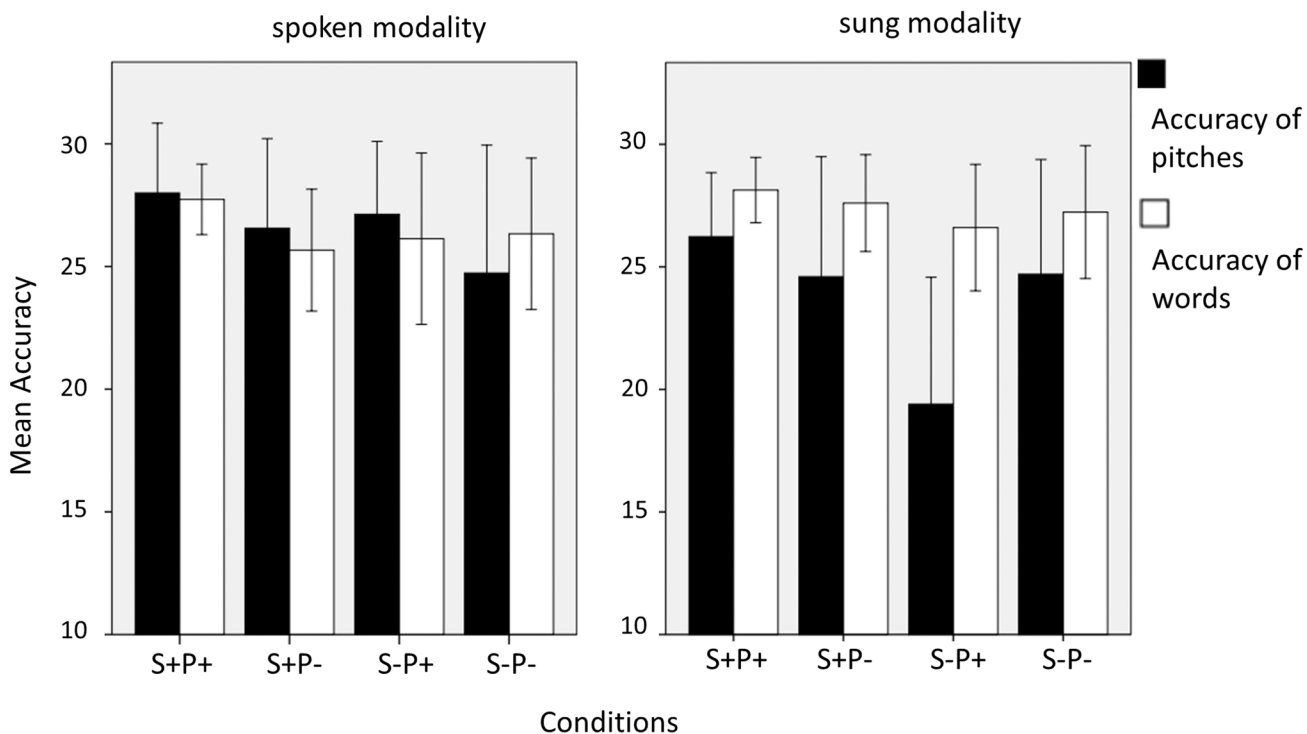

Fig 2. Mean values of accuracy of pitches and accuracy of words. Mean values of accuracy of pitches and accuracy of words (max. 30) for all conditions (S+: correct semantic sense, S-: incorrect semantic sense, P+: correct fundamental frequency/ final pitch, P-: incorrect fundamental frequency/ final pitch) for both spoken and sung modalities. Error bars indicate one standard deviation.

doi:10.1371/journal.pone.0147986.g002

for post-hoc analysis showed that the performance was nearly the same for all conditions that contained an expectancy violation (S+P-: 25.6 ± 2.4 c.r., S-P+: 26.1 ± 3.4 c.r., S-P-: 26.3 ± 3.0 c.r.), but they all yielded lower accuracy compared to the correct recited line (S+P+: 27.7 ± 1.4 c.r.; post-hocs: S+P+ vs. S+P-, p < .001; S+P+ vs. S-P+, p = .024; S+P+ vs. S-P-, p = .029).

Sung modality. In the sung modality, the singers reached a significantly higher accuracy regarding the judgment of pitches (25.0 ± 1.3 c.r) compared to actors (22.4 ± 1.6 c.r.; main effect GROUP: $F_{(1, 28)}$ = 11.05, p = .002) without an interaction with CONDITION ([Fig 3](#)). Also, we found a significant main effect of CONDITION ($F_{(1.57, 44.14)}$ = 14.33, p < .001). The post-hoc analysis ([Fig 2](#)) revealed a significant difference in recognizing the correct pitch combined with a semantic violation (S-P+: 19.4 ± 5.1 c.r.) compared to the other conditions (S+P-: 24.6 ± 4.8 c.r., S-P-: 24.7 ± 4.6 c.r., S+P+: 26.2 ± 2.6 c.r.; all post-hoc comparisons with S-P+ ≤ .001).

The accuracy of semantic sense of the last word in the sung modality again showed a significant main effect CONDITION ($F_{(1.96, 54.88)}$ = 3.19; p = .048) with no interaction with or main effect of GROUP. The post-hoc analysis of the main effect CONDITION ([Fig 2](#)) confirmed the S-P+ condition as most difficult to recognize in the sung modality (26.6 ± 2.5 c.r.), differing significantly from the correct line S+P+ (28.1 ± 1.3 c.r.; S-P+ vs. S+P+, p < .007) and differing in terms of non-significant trends and from S-P- (27.3 ± 2.7 c.r.) and from S+P- (27.6 ± 1.9 c.r.; S-P+ vs. S+P-, p = .051; S-P+ vs. S-P-, p = .054).

Group differences with regard to attention required and fatigue

We observed differences between singers and actors at the level of attention reported necessary to perform successfully. Summarized in semi-structured interviews after the measurements, subjects were categorized according to three different levels of effort. In the group of actors, nine subjects reported the need to pay substantial attention during the experiment and felt exhausted at the end. Four subjects described the required attention as "moderate" and two as "low". In contrast, ten singers reported a kind of "easy-flow" and a "inner rehearsal in the mind" described also in terms of internal repetition of the spoken and sung lines with focus on melodies but also on texts and no need for additional attention. These ten singers reported low attention demands, while four other singers reported a moderate level and one a high level.
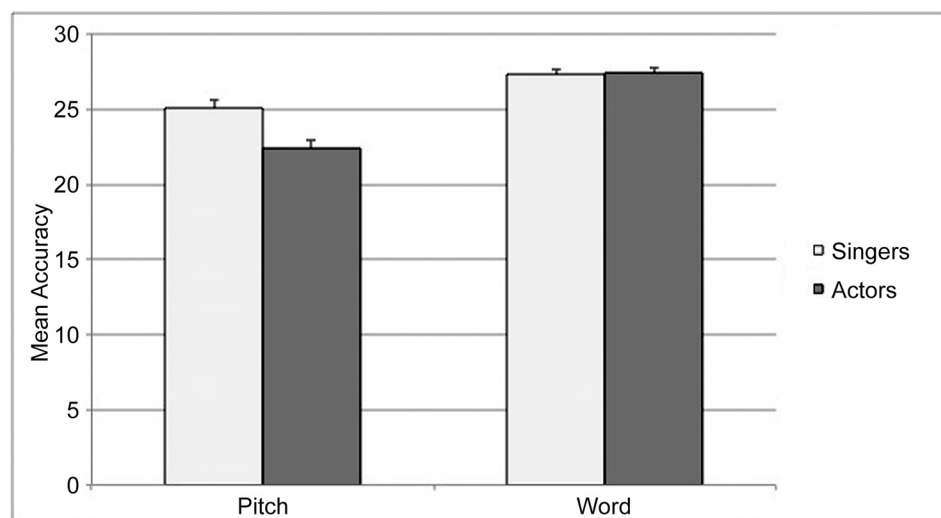

Fig 3. Mean values of accuracy of pitches and words in the sung modality. Mean values of accuracy of pitches and accuracy of words (max. 30) in the sung modality in comparison of singers and actors. Error bars indicate the standard error.

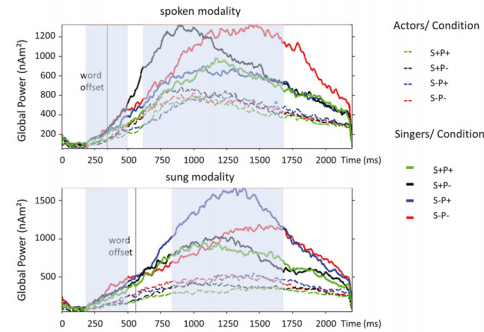doi:10.1371/journal.pone.0147986.g003

Fig 4. Global power of minimum norm estimates of all dipoles. Global power of minimum norm estimates of all dipoles, separated for the different conditions, for singers and actors, and for both spoken and sung modalities. Blue shadowed are the time windows of the analysis of the early and late activity.

doi:10.1371/journal.pone.0147986.g004

Magnetoencephalographic data

Global power.   Inspection of the Global Power of L2-MNE solutions (Fig 4) demonstrated a long interval of high activation, starting around 400 ms after last word onset/ peaking around 1000 ms for spoken stimuli and starting around 600 ms/ peaking around 1250 ms for sung stimuli, with much higher cortical activity for singers than for actors.

The results of the point-wise repeated measures ANOVA revealed two time windows for both modalities, in which the experimental manipulations resulted in different brain responses. Anticipating briefly the statistical results below, during the early time window (200–500 ms), brain responses differed in predicted ways between the violations, but independent of group. In the late time window (600–1700 ms), no difference between conditions was found, but singers displayed a unexpected long-lasting and substantial level of activity. Because the topography of evoked activity was rather stable and varied only minimally within these time windows, we will present only averaged responses for these intervals.

Statistical analysis of the early activity

The pointwise repeated measures ANOVA revealed that the factors SEMANTIC VIOLATION and PROSODIC (spoken modality)/ MELODIC (sung modality) VIOLATION differed significant in the left and right temporal regions, averaged for the time interval of 200–500 ms (Fig 5). Because of corresponding dipole groups in the temporal areas of both hemispheres, we included the factor HEMISPHERE in the analysis of early activity. For both modalities, no significant statistical effects were detectable for the between-subject factor GROUP or for any interaction with GROUP according to the thresholds described in the methods section.

Spoken modality.   Repeated measures ANOVAs for both temporal clusters (left temporal: 200-500ms, 22 dipoles; right temporal: 200–500 ms, 22 dipoles; with 18 corresponding dipoles) revealed significant main effects of the factor HEMISPHERE ($F_{(1,28)} = 14.73$, p = .001), representing higher activation on the right hemisphere (21.97 ± 5.45 nAm^2) than on the left hemisphere (17.39 ± 4.97 nAm^2). Furthermore, there was a main effect of the mean activity of PROSODIC VIOLATION ($F_{(1,28)} = 32.97$, p < .001) comparing prosodic violations (S+P- and S-P-: 21.99 ± 6.44 nAm^2) to conditions without violation (S+P+ and S-P+: 17.38 ± 5.77 nAm^2), but no effect of SEMANTIC VIOLATION. The only significant interaction SEMANTIC VIOLATION X HEMISPHERE ($F_{(1,28)} = 65.76$, p = .007) reflected higher activity for conditions with semantic violations (mean of: S-P+ and S-P-: 18.01 ± 5.46 nAm^2), compared to conditions without semantic violations (mean of: S+P+ and S+P-: 16.77 ± 5.48 nAm^2) on the left
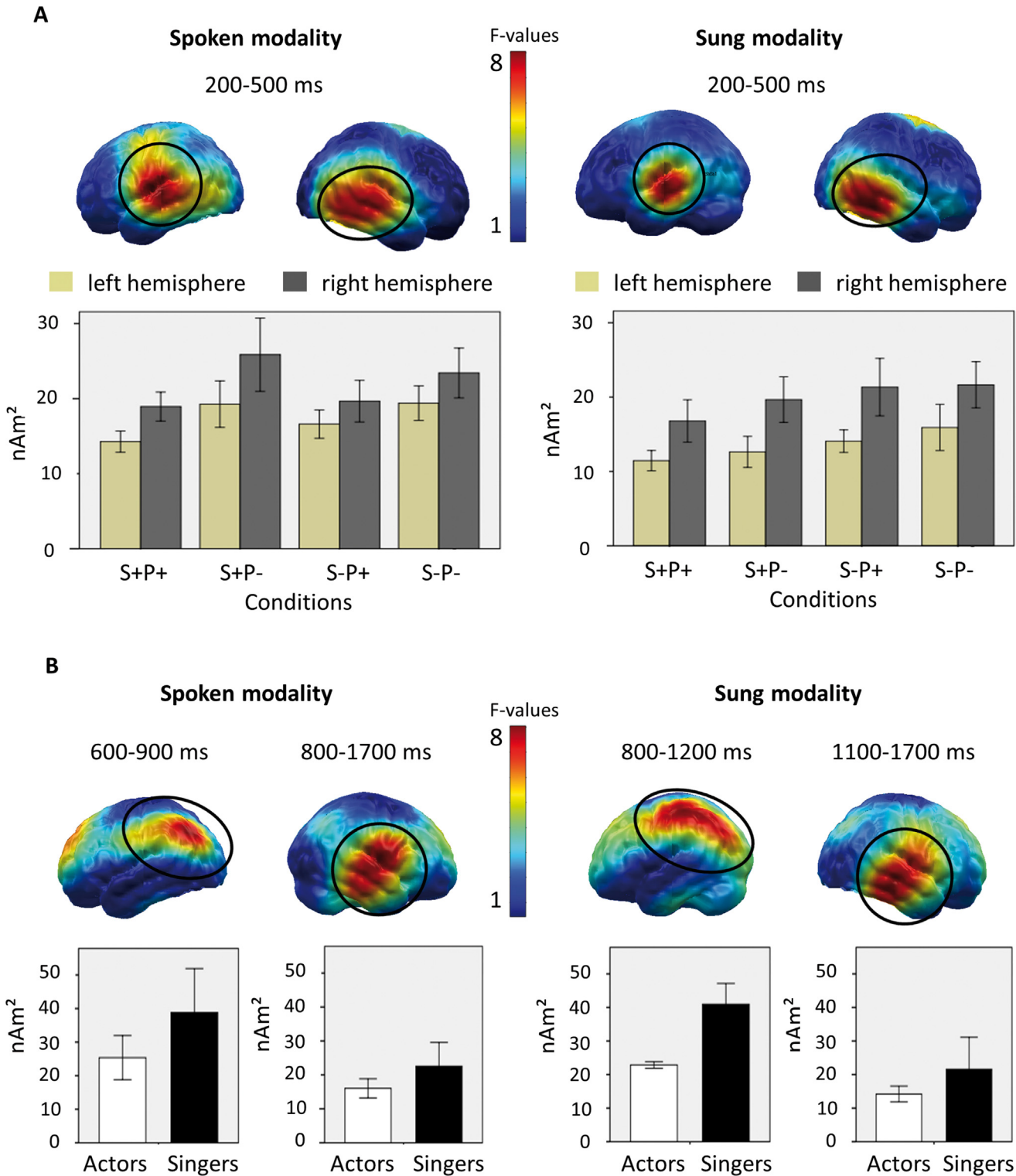
Fig 5. Maps and mean neural activity for respective clusters of dipoles. A. Top: Mapping of the F-values for the interaction CONDITION and HEMISPHERE on a cortical surface for average time intervals, for both spoken and sung modalities. Bottom: Mean neural activity for the respective clusters of dipoles for all conditions and for both spoken and sung modalities. Error bars denote one standard deviation. B. Top: Mapping of the F-values for the main effect group on a cortical surface for average time intervals, for both spoken and sung modalities. Bottom: Mean neural activity for the respective clusters of dipoles for both groups and for both spoken and sung modalities. Error bars denote one standard deviation.

doi:10.1371/journal.pone.0147986.g005

hemisphere ($t_{(29)}$ = 3.35, p = .002) and the opposite pattern (mean of: S-P+ and S-P-: 21.54 ± 7.93 nAm^2; mean of: S+P+ and S+P-: 25.86 ± 13.08 nAm^2) on the right hemisphere ($t_{(29)}$ = -3.35, p = .003).

Thus, for the spoken modality, there was generally higher activity on the right hemisphere. Both hemispheres were equally involved in the processing of prosodic violations, while there was a hemispheric specialization for processing semantic violations. For the latter, the left hemisphere displayed stronger activity as a response to violated sentences than correct sentences, i.e. a typical N400 effect.

Sung modality.   Repeated measures ANOVAs for both temporal Clusters (left temporal: 200-500ms, 16 dipoles; right temporal: 200–500 ms, 19 dipoles; both hemispheres: 11 corresponding dipoles) revealed significant main effects of the factor HEMISPHERE ($F_{(1,28)}$ = 34.23, p < .001) representing higher activation on the right hemisphere (19.86 ± 4.43 nAm^2) than on the left hemisphere (13.52 ± 3.77 nAm^2). Moreover, a significantly higher level of activity was detected for expectancy-violated conditions than those without MELODIC VIOLATION ($F_{(1,28)}$ = 10.17, p = .004; S+P- and S-P-: 17.46 ± 3.24 nAm^2 vs. S+P+ and S-P+: 15.91 ± 4.13 nAm^2) and SEMANTIC VIOLATION ($F_{(1,28)}$ = 29.17, p < .001; S-P+ and S-P-: 18.25 ± 3.42 nAm^2 vs. S+P+ and S+P-: 15.13 ± 2.81 nAm^2). These main effects were modulated by the significant interaction SEMANTIC VIOLATION X MELODIC VIOLATION X HEMISPHERE ($F_{(1,28)}$ = 39.44, p = .014), which themselves resulted from a significant interaction of SEMANTIC VIOLATION X MEDOLDIC VIOLATION only on the right hemisphere ($F_{(1,29)}$ = 49.71, p = .010), not present on the left hemisphere ($F_{(1,29)}$ = 1.29; n.s.). Post-hoc analysis for this interaction revealed a similar level of activity for the conditions with a semantic violation (S-P+: 21.34 ± 3.19 nAm^2, S-P-: 21.65 ± 4.32 nAm^2), while there significant difference emerged in comparing the original line (S+P+: 16.78 ± 2.07 nAm^2) to the condition that only contained a melodic violation (S+P-: 19.66 ± 3.11 nAm^2; $t_{(29)}$ = -3.98, p < .001).

Thus, for the sung modality, there was generally more activity on the right hemisphere. Both hemispheres responded to semantic and melodic violations, while there was an additional interaction of melodic and semantic violations only in the right hemisphere. Here, stronger activity was only found in response to a melodic violation, if there was no semantic violation.

Statistical analysis of the late activity

The pointwise repeated measures ANOVA revealed, for both modalities, a significant main effect of the between-subject factor GROUP, representing higher neuronal activity for the singers than the actors localized in right temporal and left parietal regions in a late time window (Fig 5). There was neither a significant effect for the factors SEMANTIC VIOLATION and PROSODIC/ MELODIC VIOLATION, nor for any interaction with GROUP, according to the thresholds mentioned in the above methods section. The absence of corresponding dipole groups on the homologous hemisphere did not warrant the additional factor HEMISPHERE in the analyses.

Spoken modality.   Repeated measures ANOVAs on time-averaged activity levels in the left parietal cluster (600-900ms, 10 dipoles) revealed a significant GROUP effect ($F_{(1,28)}$ = 4.68, p = .039), with higher activity for the singers (41.11 ± 13.2 nAm^2) than the actors (29.64 ± 5.82 nAm^2), independent of the stimulus condition. Additionally, in the right temporal cluster (800-1700ms, 19 dipoles), singers displayed significantly higher activity (22.56 ± 5.21 nAm^2) than actors (16.04 ± 2.34 nAm^2; $F_{(1,28)}$ = 11.20, p = .002).

Sung modality.   Repeated measures ANOVAs on time-averaged activity levels revealed a similar pattern of results to the spoken modality. In the left parietal cluster (800-1200ms, 21 dipoles), we found a significant GROUP effect ($F_{(1,28)}$ = 8.27, p = .008) with higher activity for

the singers (40.94 ± 4.42 nAm^2) than the actors (22.87 ± 0.91 nAm^2). Additionally, for the right temporal cluster (1100-1700ms, 18 dipoles), we found significantly higher activity for the singers (21.58 ± 8.41 nAm^2) compared to the actors (14.21 ± 2.14 nAm^2; $F_{(1,28)} = 8.44$, $p = .007$).

Discussion

The aim of this study was to compare linguistic and musical processing in two groups of highly trained voice users, i.e. professional singers and actors. We employed rhyme sequences from German art songs, and presented analogous semantic and/ or melodic/prosodic violations in sung and recited versions of the material. MEG measurements were implemented to identify functional brain activity with regard to the type of expertise. Behavioral data revealed greater accuracy of pitch detection in the sung modality for singers than for actors, while there were no detectable group-specific advantages for actors, neither for the sung nor for the recited material. Although previous studies referred to dependence in the neuronal processing of linguistic and musical dimensions, both in a spoken and sung modality [5,30,45], this is the first study presenting combinations of semantic and melodic/ prosodic expectancy violations for speaking and singing in a complex, but ecologically valid context. Confirming an intertwined neuronal network for music and speech, MEG data analysis disclosed condition- and modality-specific differences of "early" temporal activity (200–500 ms) on both hemispheres in homologous clusters, independent of the kind of expertise. Significant group differences appeared as "late" neuronal activity (600–1700 ms) for both stimulus modalities in right temporal and left parietal areas. We will discuss the results of the behavioural data and the two time windows in turn.

Behavioural data

In the behavioral data, we did not find an effect of modality-specific expertise, apart from for higher accuracy for pitch discrimination in singers for the sung modality, which can be explained with a higher sensation for musical patterns in their familiarized domain. Interestingly, while performance in terms of word accuracy in the sung and spoken modality was nearly at the same level for the different conditions, both groups performed lower for the discrimination of correct pitches in the case of a semantic violation (S-P+) in the sung modality. This might result from the need for high attention to recognize the semantic sense of the words, at the expense of the discriminatory power of pitches. Due to the different pronunciation of speech in singing compared to speaking, vowels gain stronger emphasis than consonants. This might have created particularly challenging conditions for the discrimination of closely related phonemes, such as in the rhyme words that were used to replace the original words.

Early neuronal activity related to linguistic and musical context

Summarizing the similarities of the early effects for the sung and spoken modalities, high neuronal activity was measured especially after melodic/prosodic violations in predominantly right temporal areas. Consequently, it seems that neuronal networks involved in processing both modalities exhibited higher neuronal activity for the expectancy violation of the final pitch deviation of the lines compared to semantic violations. Therefore, in the present design, the rule system for syntax—melodic/ prosodic aspect—represents a global characteristic for both sung and recited phrases and indicates a global syntactical system represented bilaterally with dominance of the right hemisphere if factors involfing pitch (melody, prosody) are violated (see e.g. [46]). In line with our findings, previous studies investigating linguistic aspects of speech revealed a dominance of left temporal areas [13,15,16], especially if linguistic stimuli

were presented in a complex syntactical structure. In comparison, a dominance of right temporal areas was found after violation of a musical order system such as a chord sequence or a melody line [17,19,23,47]. This was the case if linguistic and musical tasks were performed by human voice, mixed-animal or vocally similar sounds demanding high attention for frequency analysis [48,49].

The above-mentioned similarities during the processing of recited and sung phrases are contrasted by different interactions of effects for the sung and spoken modalities, presenting a more complex dependence of information processing for semantic and prosodic/ melodic content on both hemispheres. While there was a predominantly left-temporal lateralized activity after semantic expectancy violations for the recited sequences, which is in line with findings for the classical N400 effect [50,51], this was opposed by right-dominant temporal activity in response to melodic violations, but semantic correctness, in the sung version. The combination of the semantic content with the musical syntactical form of the melody line seems to be strongly connected in the modality of singing and represents intertwined networks reacting to different degrees after expectancy violation. These findings confirm recent research revealing a more bilateral temporal network system dependent on modality-specific aspects for sung and spoken units [5,30,45]. But compared to these findings, the present study combined syntactical and semantic violations in a complete rhyme sequence both in a spoken and a sung modality. To the best of our knowledge, the only design using also original excerpts to compare different musicians and laymen (i.e. different romantic opera composers), but only presented in a sung modality, was presented by Besson, Schon and Bonnel [28,29]. In their study, the simultaneous violation of the semantic and the syntactic sense at the end of a sung line resulted in a N400 and a P600 component, suggesting that semantic and syntactic aspects of language and music were processed by independent systems, which was, however, not confirmed in subsequent investigations. In contrast, vocally generated stimuli which create simultaneously high demands on linguistic and musical aspects seem to involve middle and superior temporal areas, acting as an intertwined network. This network is adapted in a modality dependent way to different conditions [5,52–54].

Group differences in late neuronal activity

During the analysis, we detected differences in brain activation in an unexpected late and long-lasting time window (up to 1700 ms after stimulus onset), with higher activation for singers than actors. The activity was localized in right temporal areas, similar to the early activation clusters generated by the semantic and syntactic incongruencies, as well as in parietal areas on the left hemisphere, which are known to be involved in higher-order music cognition [36].

The specific role of temporal areas on both hemispheres concerning speech and music processing was discussed before in the interpretation of "early" activation clusters. We interpret the renewed appearance of neuronal activity in the right temporal area as a special form of working memory function. In contrast to actors, in the semi-structured interviews, singers reported repeating a heard sequence in their minds (inner rehearsal). Thus, the late right temporal activity might stem from cognitive processes representing an illusionary perception of previously perceived auditory stimuli. Studies in the field of music psychology described findings of a mental representation of music in musicians, by internally hearing sounds after the onset of the physical stimulus [37,39], without any neurophysiological evidence of this phenomenon being reported so far. Right temporal lobe activity, as a correlate of a very vivid mental representation of music, is also supported by a recent fMRI study of musical imagery of familiar tunes. This study reported a relationship of activity in right secondary auditory areas and the subjective vividness of mental imagery [55]. Because we found the higher activity in

musicians for both the sung and the spoken modality, we assume that this is the result of a transfer effect after extensive training. There is some evidence for the existence of two separate working memory functions in musicians for music (nonspeech) and speech material: a phonological loop and a tonal loop. Thus, training effects of the tonal loop seem to carry over to the phonological loop, involving highly similar neural correlates [56].

The functional role of the inferior parietal cortex involves auditory-verbal working memory and short-term memory for musical pitch, especially on the right hemisphere of musically trained subjects [57]]. More generally, the inferior parietal cortex has been associated with the integration of sensory and motor signals for the somatosensory guidance of movements [58]. Previous neuroimaging studies have documented its response to both speech and music perception [49]. A model of speech motor control [59] posits a role of the parietal cortex (PC) in a feed forward control mechanism of articulatory motor commands. In this model, the PC acts as a control system for somatosensory feedback from the vocal tract by comparing the actual kinaesthetic feedback with the expectation of the pronounced sound. Accordingly, we assume that the increment in PC activity could reflect enhanced processing of a mismatch between intention, action and consequences and thus allows for more rapid sensorimotor adaptions/ corrections in singers than actors. Another previous study presented evidence that increased activity of receptive systems subserve the precise transformation of highly automatic speech motor sequences into appropriately adjusted motor patterns for singing [32]. Kleber and coauthors demonstrated that imagined and overt singing involves partly different brain systems in singers, with imagined singing activating a large frontal and parietal network, indicating increased involvement of higher-order cognitive processes during mental imagery [60]. These topical findings during receptive and expressive functions supplement our findings and suggest an important role of the parietal cortex for music processing in singers. Additionally, it has been shown that complex mental transformations of musical material, like the mental reversal of imagined melodies, is related to activity in the posterior parietal cortex [36]. Performing music in the mind is a technique used by professional musicians to rehearse various aspects of a musical piece, for example to mentally revise difficult parts of a previously executed musical passage. As such our unexpected findings integrate well into the existing literature, but it is yet an open question whether the phenomenon of audition and its neurophysiological counterpart comes into existence through training or whether this is prerequisite to become a professional musician.

Conclusions

In conclusion, our results of early and late neuronal activation are in line with studies emphasizing a bilateral neuronal network during linguistic and musical auditory processing, which can be tuned according to the level of mental demand [5,11,52,54,61]. Regarding the effect of experience, we did not find any early differences of neuronal activities evoked by a semantic and/or melodic/ prosodic violation. In contrast, rather late and long lasting time windows were characterized by strong activity in left parietal and right temporal areas. We propose that these are related to stronger mental imagery and higher-order music cognition in singers. This might constitute the effect of musical training or a prerequisite.

Acknowledgments

Author Contributions

Conceived and designed the experiments: KR SCH CD CMS CP AAZ DD. Performed the experiments: KR SCH. Analyzed the data: KR AK MO CD. Contributed reagents/materials/analysis tools: CP CD. Wrote the paper: KR AAZ CD CMS DD SCH.

References

1. Moreno S, Marques C, Santos A, Santos M, Castro SL, Besson M. Musical training influences linguistic abilities in 8-year-old children: more evidence for brain plasticity. Cereb Cortex 2009 Mar; 19(3):712–723. doi: 10.1093/cercor/bhn120 PMID: 18832336

2. Koelsch S, Kasper E, Sammler D, Schulze K, Gunter T, Friederici AD. Music, language and meaning: brain signatures of semantic processing. Nat Neurosci 2004 Mar; 7(3):302–307. PMID: 14983184

3. Peretz I, Gosselin N, Belin P, Zatorre RJ, Plailly J, Tillmann B. Music lexical networks: the cortical organization of music recognition. Ann N Y Acad Sci 2009 Jul; 1169:256–265. doi: 10.1111/j.1749-6632.2009.04557.x PMID: 19673789

4. Schneider S, Schonle PW, Altenmuller E, Munte TF. Using musical instruments to improve motor skill recovery following a stroke. J Neurol 2007 Oct; 254(10):1339–1346. PMID: 17260171

5. Schon D, Gordon R, Campagne A, Magne C, Astesano C, Anton JL, et al. Similar cerebral networks in language, music and song perception. Neuroimage 2010 May 15; 51(1):450–461. doi: 10.1016/j.neuroimage.2010.02.023 PMID: 20156575

6. Koelsch S. Neural substrates of processing syntax and semantics in music. Curr Opin Neurobiol 2005 Apr; 15(2):207–212. PMID: 15831404

7. Zatorre RJ, Peretz I, Penhune V. Neuroscience and Music ("Neuromusic") III: disorders and plasticity. Preface. Ann N Y Acad Sci 2009 Jul; 1169:1–2.

8. Kutas M, Hillyard SA. Reading senseless sentences: brain potentials reflect semantic incongruity. Science 1980 Jan 11; 207(4427):203–205. PMID: 7350657

9. Lau EF, Phillips C, Poeppel D. A cortical network for semantics: (de)constructing the N400. Nat Rev Neurosci 2008 Dec; 9(12):920–933. doi: 10.1038/nrn2532 PMID: 19020511

10. Van Petten C, Luka BJ. Neural localization of semantic context effects in electromagnetic and hemodynamic studies. Brain Lang 2006 Jun; 97(3):279–293. PMID: 16343606

11. Maess B, Herrmann CS, Hahne A, Nakamura A, Friederici AD. Localizing the distributed language network responsible for the N400 measured by MEG during auditory sentence processing. Brain Res 2006 Jun 22; 1096(1):163–172. PMID: 16769041

12. Dobel C, Junghofer M, Breitenstein C, Klauke B, Knecht S, Pantev C, et al. New names for known things: on the association of novel word forms with existing semantic information. J Cogn Neurosci 2010 Jun; 22(6):1251–1261. doi: 10.1162/jocn.2009.21297 PMID: 19583468

13. Hirschfeld G, Zwitserlood P, Dobel C. Effects of language comprehension on visual processing—MEG dissociates early perceptual and late N400 effects. Brain Lang 2011 Feb; 116(2):91–96. doi: 10.1016/j.bandl.2010.07.002 PMID: 20708788

14. Geukes S, Huster RJ, Wollbrink A, Junghofer M, Zwitserlood P, Dobel C. A large N400 but no BOLD effect—comparing source activations of semantic priming in simultaneous EEG-fMRI. PLOS One 2013 Dec 31; 8(12):e84029. doi: 10.1371/journal.pone.0084029 PMID: 24391871

15. Friederici AD, Kotz SA. The brain basis of syntactic processes: functional imaging and lesion studies. Neuroimage 2003 Nov; 20 Suppl 1:S8–17. PMID: 14597292

16. Wolters CH, Anwander A, Maess B, Macleod RS, Friederici AD. The influence of volume conduction effects on the EEG/MEG reconstruction of the sources of the Early Left Anterior Negativity. Conf Proc IEEE Eng Med Biol Soc 2004; 5:3569–3572. PMID: 17271062

17. Koelsch S, Gunter T, Friederici AD, Schroger E. Brain indices of music processing: "nonmusicians" are musical. J Cogn Neurosci 2000 May; 12(3):520–541. PMID: 10931776

18. Steinbeis N, Koelsch S. Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. Cereb Cortex 2008 May; 18(5):1169–1178. PMID: 17720685

19. Maess B, Koelsch S, Gunter TC, Friederici AD. Musical syntax is processed in Broca's area: an MEG study. Nat Neurosci 2001 May; 4(5):540–545. PMID: 11319564

20. Koelsch S. Toward a neural basis of music perception—a review and updated model. Front Psychol 2011; 2:110. doi: 10.3389/fpsyg.2011.00110 PMID: 21713060

21. Koelsch S, Maess B, Gunter TC, Friederici AD. Neapolitan chords activate the area of Broca. A magne-toencephalographic study. Ann N Y Acad Sci 2001 Jun; 930:420–421. PMID: 11458855

22. Hyde KL, Peretz I, Zatorre RJ. Evidence for the role of the right auditory cortex in fine pitch resolution. Neuropsychologia 2008 Jan 31; 46(2):632–639. PMID: 17959204

23. Warrier CM, Zatorre RJ. Right temporal cortex is critical for utilization of melodic contextual cues in a pitch constancy task. Brain 2004 Jul; 127(Pt 7):1616–1625. PMID: 15128620

24. Patel AD, Gibson E, Ratner J, Besson M, Holcomb PJ. Processing syntactic relations in language and music: an event-related potential study. J Cogn Neurosci 1998 Nov; 10(6):717–733. PMID: 9831740

25. Besson M, Macar F. An event-related potential analysis of incongruity in music and other non-linguistic contexts. Psychophysiology 1987 Jan; 24(1):14–25. PMID: 3575590

26. Besson M, Faita F. An Event-Related Potential (ERP) Study of Musical Expectancy: Comparison of Musicians with Nonmusicians. Journal of Experimental Psychology: Human Perception and Performance 1995; 21(6):1278–1296.

27. Schon D, Magne C, Besson M. The music of speech: music training facilitates pitch processing in both music and language. Psychophysiology 2004 May; 41(3):341–349. PMID: 15102118

28. Bonnel AM, Faita F, Peretz I, Besson M. Divided attention between lyrics and tunes of operatic songs: evidence for independent processing. Percept Psychophys 2001 Oct; 63(7):1201–1213. PMID: 11766944

29. Besson M, Schon D. Comparison between language and music. Ann N Y Acad Sci 2001 Jun; 930:232–258. PMID: 11458832

30. Kolinsky R, Lidji P, Peretz I, Besson M, Morais J. Processing interactions between phonology and melody: vowels sing but consonants speak. Cognition 2009 Jul; 112(1):1–20. doi: 10.1016/j.cognition.2009.02.014 PMID: 19409537

31. Schon D, Boyer M, Moreno S, Besson M, Peretz I, Kolinsky R. Songs as an aid for language acquisition. Cognition 2008 Feb; 106(2):975–983. PMID: 17475231

32. Kleber B, Veit R, Birbaumer N, Gruzelier J, Lotze M. The brain of opera singers: experience-dependent changes in functional activation. Cereb Cortex 2010 May; 20(5):1144–1152. doi: 10.1093/cercor/bhp177 PMID: 19692631

33. Zarate JM, Zatorre RJ. Experience-dependent neural substrates involved in vocal pitch regulation during singing. Neuroimage 2008 May 1; 40(4):1871–1887. doi: 10.1016/j.neuroimage.2008.01.026 PMID: 18343163

34. Dick F, Lee HL, Nusbaum H, Price CJ. Auditory-Motor Expertise Alters "Speech Selectivity" in Professional Musicians and Actors. Cereb Cortex 2010 Sep 9.

35. Herholz SC, Lappe C, Knief A, Pantev C. Neural basis of music imagery and the effect of musical expertise. Eur J Neurosci 2008 Dec; 28(11):2352–2360. doi: 10.1111/j.1460-9568.2008.06515.x PMID: 19046375

36. Zatorre RJ, Halpern AR, Bouffard M. Mental reversal of imagined melodies: a role for the posterior parietal cortex. J Cogn Neurosci 2010 Apr; 22(4):775–789. doi: 10.1162/jocn.2009.21239 PMID: 19366283

37. Hubbard TL. Auditory imagery: empirical findings. Psychol Bull 2010 Mar; 136(2):302–329. doi: 10.1037/a0018436 PMID: 20192565

38. Gordon EE. Learing sequences in music: Skill, content and patterns. Chicago: GIA Publications 1993.

39. Brodsky W, Rubinstein B. The Mental Representation of Music Notation: Notational Audiation. Journal of Experimental Psychology 2008; 34(2):427–445. doi: 10.1037/0096-1523.34.2.427 PMID: 18377180

40. Hämäläinen MS, Ilmoniemi RJ. Interpreting magnetic fields of the brain: Minimum-norm estimates. Medical & Biological Engineering & Computing 1994(32:):35–42.

41. Schubert F, Mueller W. Die schöne Müllerin (the beautiful miller-girl), D-major,795, op.25. C.F.Peters 1823;1; 6824:4–52.

42. Schubert F, Mueller W. Winterreise (winter journey), D-major, 911, op.89. C.F.Peters 1827;1; 6824:54–120.

43. Peyk P, De Cesarei A, Junghofer M. ElectroMagnetoEncephalography Software: Overview and Integration with Other EEG/MEG Toolboxes. Comput Intell Neurosci 2011; 2011:861705. doi: 10.1155/2011/861705 PMID: 21577273

44. Brockelmann AK, Steinberg C, Elling L, Zwanzger P, Pantev C, Junghofer M. Emotion-associated tones attract enhanced attention at early auditory processing: magnetoencephalographic correlates. J Neurosci 2011 May 25; 31(21):7801–7810. doi: 10.1523/JNEUROSCI.6236-10.2011 PMID: 21613493

45. Schon D, Gordon RL, Besson M. Musical and linguistic processing in song perception. Ann N Y Acad Sci 2005 Dec; 1060:71–81. PMID: 16597752

46. Kreitewolf J, Friederici AD, von Kriegstein K. Hemispheric lateralization of linguistic prosody recognition in comparison to speech and speaker recognition. Neuroimage 2014 Nov 15; 102 Pt 2:332–344. doi: 10.1016/j.neuroimage.2014.07.038 PMID: 25087482

47. Samson S, Zatorre RJ. Contribution of the right temporal lobe to musical timbre discrimination. Neuropsychologia 1994 Feb; 32(2):231–240. PMID: 8190246

48. Zatorre RJ, Belin P. Spectral and temporal processing in human auditory cortex. Cereb Cortex 2001 Oct; 11(10):946–953. PMID: 11549617

49. Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. Trends Cogn Sci 2002 Jan 1; 6(1):37–46. PMID: 11849614

50. Hagoort P, Brown CM. ERP effects of listening to speech: semantic ERP effects. Neuropsychologia 2000; 38(11):1518–1530. PMID: 10906377

51. Keuper K, Zwanzger P, Nordt M, Eden A, Laeger I, Zwitserlood P, et al. How 'love' and 'hate' differ from 'sleep': Using combined electro/magnetoencephalographic data to reveal the sources of early cortical responses to emotional words. Hum Brain Mapp 2012 Dec 26.

52. Sammler D, Koelsch S, Ball T, Brandt A, Elger CE, Friederici AD, et al. Overlap of musical and linguistic syntax processing: intracranial ERP evidence. Ann N Y Acad Sci 2009 Jul; 1169:494–498. doi: 10.1111/j.1749-6632.2009.04792.x PMID: 19673829

53. Warren JD, Scott SK, Price CJ, Griffiths TD. Human brain mechanisms for the early analysis of voices. Neuroimage 2006 Jul 1; 31(3):1389–1397. PMID: 16540351

54. Sammler D, Koelsch S, Ball T, Brandt A, Grigutsch M, Huppertz HJ, et al. Co-localizing linguistic and musical syntax with intracranial EEG. Neuroimage 2013 Jan 1; 64:134–146. doi: 10.1016/j.neuroimage.2012.09.035 PMID: 23000255

55. Herholz SC, Halpern AR, Zatorre RJ. Neuronal correlates of perception, imagery, and memory for familiar tunes. J Cogn Neurosci 2012 Jun; 24(6):1382–1397. doi: 10.1162/jocn_a_00216 PMID: 22360595

56. Schulze K, Zysset S, Mueller K, Friederici AD, Koelsch S. Neuroarchitecture of verbal and tonal working memory in nonmusicians and musicians. Hum Brain Mapp 2010 Jun 9.

57. Gaab N, Schlaug G. Musicians differ from nonmusicians in brain activation despite performance matching. Ann N Y Acad Sci 2003 Nov; 999:385–388. PMID: 14681161

58. Jancke L, Kleinschmidt A, Mirzazade S, Shah NJ, Freund HJ. The role of the inferior parietal cortex in linking the tactile perception and manual construction of object shapes. Cereb Cortex 2001 Feb; 11 (2):114–121. PMID: 11208666

59. Guenther FH. Cortical interactions underlying the production of speech sounds. J Commun Disord 2006 Sep-Oct; 39(5):350–365. PMID: 16887139

60. Kleber B, Birbaumer N, Veit R, Trevorrow T, Lotze M. Overt and imagined singing of an Italian aria. Neuroimage 2007 Jul 1; 36(3):889–900. PMID: 17478107

61. Gordon RL, Schon D, Magne C, Astesano C, Besson M. Words and melody are intertwined in perception of sung words: EEG and behavioral evidence. PLOS One 2010 Mar 31; 5(3):e9889. doi: 10.1371/journal.pone.0009889 PMID: 20360991