

Rhythm Structure Influences Auditory-Motor Interaction During Anticipatory Listening to Simple Singing

Jungblut Monika^{1*}, Huber Walter² and Schnitker Ralph³

¹Interdisciplinary Institute for Music- and Speech-Therapy, Duisburg, Germany

²Clinical Cognition Research, University Hospital Aachen, RWTH Aachen, Germany

³Department of Neurology, Section Neuropsychology, University Hospital Aachen, RWTH Aachen, Germany

*Corresponding author: Jungblut Monika, Interdisciplinary Institute for Music and Speech-Therapy Am Lipkamp 14 Duisburg, Germany, Tel: +49-203-711319; E-mail: msjungblut@t-online.de

Received date: September 28, 2015, Accepted date: April 14, 2016, Published date: April 26, 2016

Copyright: © 2016 Monika J, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Purpose: Our objective was to investigate if auditory-motor interactions during anticipatory listening to simple singing also vary depending on rhythm structure as we recently demonstrated for singing production.

Methods: In an event-related fMRI procedure 28 healthy subjects listened to vowel changes with regular groupings (1), regular groupings with rests (2), and irregular groupings (3) in anticipation of repeating the heard stimuli during the latter portion of the experiment.

Results: While subtraction (2) minus (1) yielded activation in the left precentral gyrus, both subtractions from condition (3), resulted in activation of bilateral putamen and caudate.

Only subtraction (3) minus (1) yielded activation of bilateral pre-SMA, and precentral gyrus more distinct in the left hemisphere. Middle, superior, and transverse temporal gyri as well as ventrolateral prefrontal cortex and the insula were activated most pronouncedly in the left hemisphere.

Conclusions: The higher the requirement of temporal grouping, the more necessary auditory-motor fine tuning might be resulting in a more distinct and left lateralized temporal, premotor, and prefrontal activation.

If it were possible to support programming and planning of articulatory gestures by anticipatory listening to specifically targeted vocal exercises, this could be beneficial for patients suffering from motor speech disorders as well as aphasia as our current research with patients substantiates.

Keywords: fMRI; Anticipatory listening; Singing; Rhythm; Therapy; Aphasia; Apraxia of speech

Introduction

An engagement of the auditory and the motor system or a communication between both systems is essential for example for musical performance but also for human speech to which we will refer later in the course of our introduction.

With regard to music, this applies to anticipating the rhythmic accents in a piece of music when a listener taps to the beat, as well as to listening to each note produced when an instrumentalist or singer adjusts precisely timed motor execution.

Such auditory-motor interactions were described in studies investigating musical discrimination [1,2], vocal imagery [3,4] as well as perception and production of musical rhythm [5-9].

The study of Chen et al. [6] demonstrates a close link between auditory and motor systems in the context of rhythm. The authors compared listening to 3 types of rhythms in two experiments. First, subjects listened in anticipation of tapping. Secondly, subjects naively

listened to rhythms without knowing that they had to tap along with them in the latter portion of the experiment. The same regions i.e. mid-premotor cortex, supplementary motor area, and cerebellum were engaged in both experiments. The ventral premotor cortex, however, was only activated when subjects listened in anticipation of tapping.

To the best of our knowledge, there are no imaging studies available as yet, which have investigated anticipatory listening to sung material.

Singing as a unique human ability requires complex mechanisms for sensory-motor control (for review, see [10]) on the one hand and combines language as well as musical capabilities on the other.

The observation that singing capabilities are often spared in patients suffering from aphasia [11-14] prompted many researchers and therapists to implement singing - although in very different ways - in the treatment of patients suffering from both motor speech disorders as well as aphasia [12,15-24].

One reason for the beneficial effect of singing in these patient groups is the greater bihemispheric organization for singing as opposed to speech, as we pointed out in detail in our previous studies [25-28].

Singing combines pitch and intonation processing but rhythmic and temporal processing as well. These last deserve specific attention in particular with regard to treatment. Temporal organization is an essential characteristic of language and speech processing and seems to be extremely susceptible to left-hemisphere brain damage. Lesion studies from the field of language as well as music demonstrate that patients with left-hemisphere lesions have problems with rhythm and time perception [29-32]. Many studies confirm deficits in aphasia patients with regard to temporal structuring of speech but also in apraxia of speech (AOS), a dysfunction of higher-order aspects of speech motor control characterized by deficits in programming or planning of articulatory gestures [33-37].

Our interest in examining this subject results on the one hand from theoretical considerations, on the other hand our research findings might be useful for practical application e.g. specifically targeted therapy interventions. As regards theoretical considerations, research carried out on processing of temporal organization or rhythm points to a dissociation of the two components: meter and grouping (phrasing). While the former comprises regularly repeating patterns of strong and weak beats, the latter comprises a segmentation of an ongoing sequence into temporal groups of events, phrases, and motifs [38-43]. Findings from rhythm discrimination research with brain-damaged patients demonstrated a left hemisphere specialization for temporal grouping [44,45] and a right hemisphere specialization for meter [46,47].

Given the temporal organization of speech, which is usually characterized by irregular rhythmic patterns and uneven timing, we are particularly interested in temporal grouping.

The impact of rhythm on language and speech recovery has been neglected for a long time.

An exception is the research of Stahl et al. [48]. The authors investigated the importance of melody and rhythm for speech production in an experimental study with 17 nonfluent aphasia patients. Based on their conclusions, rhythmic speech but not singing may support speech production. This applied particularly to patients with lesions including the basal ganglia.

Over the past 10 years, we have performed several studies with patients suffering from chronic aphasia, which have demonstrated that especially non-fluent aphasia patients and also patients with concomitant apraxia of speech could benefit remarkably from rhythmic-melodic voice training SIPARI developed for language rehabilitation [25,26,28,49,50]. The main part of this treatment is based on specific use of the voice. Focusing initially on melodic speech components, which are mainly processed in the right hemisphere, a step-by-step change to temporal-rhythmic speech components is carried out with the objective to stimulate phonological and segmental capabilities of the left hemisphere.

In an fMRI-study with 30 healthy subjects, we demonstrated that rhythmic structure is a decisive factor as regards lateralization and activation of specific areas during simple singing of sub-lexical material i.e. vowel changes [27]. According to our findings increasing demands on motor and cognitive capabilities resulted in additional activation of inferior frontal areas of the left hemisphere, particularly in those areas, which are described in connection with temporal processing and sequencing [51-53]. These activations do not only comprise brain regions, whose lesions are causally connected with language disorders but also regions of the left hemisphere (Broca's area,

insular cortex, inferior parietal cortex), whose lesions are reported to cause apraxia of speech (AOS) [54,55].

In a subsequent therapy study, three patients with severe nonfluent aphasia and concomitant AOS underwent the same fMRI-procedure before and after therapy as the healthy control subjects in our pre-study just mentioned.

A main finding was that post-minus pre-treatment imaging data yielded significant peri-lesional activations in all patients in the left superior temporal gyrus, whereas the reverse subtraction revealed either no significant activation or right hemisphere activation [28]. Likewise, pre and post-treatment assessments of patients' vocal rhythm production, language, and speech motor performance yielded significant improvements for all patients. The left superior temporal gyrus is assumed to interface with motor planning systems for sub-lexical aspects of speech [56,57]. This auditory-motor circuit provides the essential neural mechanisms for phonological short-term memory [58,59]. Functional reintegration of this region is mentioned in the literature in connection with language improvement [60-64].

It should be emphasized, though, that the literature just cited, refers to studies investigating the effects of language therapy. As far as we know, comparable studies from the field of music therapy are not available as yet.

The fact that these results were achieved by specific rhythmic-melodic voice training and included improvements in propositional speech may provide additional evidence by means of functional imaging for the hypothesis put forward by Stahl et al. [65].

According to their two-path model of speech recovery improvements of propositional speech may engage peri-lesional regions achieved through standard speech therapy. Formulaic speech may be trained rhythmically and engage right corticostriatal regions. Stahl and colleagues hypothesize that, at least theoretically, singing could act as an intermediary between these two paths.

Our present objective is to investigate auditory-motor interaction in healthy subjects during anticipatory listening to chanted vowel changes with different rhythmic structures.

Our considerations are as follows: if it were possible to influence auditory-motor interaction during listening to specifically targeted vocal exercises this should have positive effects on planning and programming of articulatory gestures. A transition between sound pattern (phonological form) e.g. of a syllable or word and the respective motor program might thus be supported, an interface which is assumed to represent one of the main problems particularly of patients with AOS [66,67].

In order to further substantiate our assumptions, we want to refer to the motor theory of speech perception [68], which claims that phonetic perception is perception of gesture, implying a close link between perception and production. This idea is compatible with the more recent discovery of an "echo-neuron system" i.e. a system that motorically reacts when subjects listen to verbal material [69,70]. Meanwhile, several studies confirmed that active listening to speech recruits speech-related motor regions of the brain e.g. posterior temporal and parietal regions, the superior part of the ventral premotor cortex, and Broca's area [71-73].

Over the past years, the dual-stream model of speech processing [56,74] with a ventral stream for speech recognition and a dorsal stream for speech perception has been suggested. The neural substrates

of the latter involve posterior frontal lobe structures, portions of the posterior temporal lobe and the parietal operculum. The dorsal stream, which is said to be strongly left-dominant, has an auditory-motor integration function by translating acoustic speech signals into articulatory representations of the frontal lobe [57]. Production deficits but also impairments in speech perception in patients suffering from left-hemisphere injury may be explained by disruption of the dorsal stream resulting from dorsal temporal and frontal lesions.

Based on our findings mentioned above, we assume that also during anticipatory listening to sub-lexical material rhythm structure is of particular importance with regard to demands on cognitive capacities like attention, short-term memory, and motor planning. Therefore, 28 healthy subjects will listen in anticipation (i.e. knowing that they have to repeat the heard stimuli immediately after presentation) to vowel changes with regular groupings, regular groupings with rests, and irregular groupings.

We hypothesize that the more auditory-motor fine tuning with regard to temporal grouping is required, the more distinct and left lateralized temporal, premotor, and prefrontal activation will occur.

This imaging study is intended to gain an insight into the relationship between rhythmic structure and auditory-motor interactions during anticipatory listening to simple singing and intended to form the basis for further studies with patients.

Methods

Participants

A total of 28 healthy German non-musicians (14 male, 14 female, mean age: 26.3, range 21-41) participated in the present study.

In order to avoid the reported sex differences at the level of phonological processing [75] and vocal and verbal production [76,77] only women who used contraceptive devices (birth control pill) were included. All subjects were right-handed as determined by means of the Edinburgh Handedness Scale. Inclusion criterion was a lateralization index >70%. None of the participants had a history of neurologic, psychiatric or medical diseases or any signs of hearing disorders.

Apart from general school education none of the participants had any special musical training. All participants gave written informed consent in line with the Declaration of Helsinki and the Ethics Committee of the RWTH Aachen. This study was undertaken in compliance with national legislation.

Stimuli and procedure

Tasks of our fMRI-paradigm comprised anticipatory listening to chanted vowel changes in rhythm sequences with differently demanding complexity levels. Chanting is a rudimentary or simple form of singing e.g. on one pitch only and facilitates evaluating the influence of rhythm structure because melodic components are reduced. Rhythmic chanting (e.g. the vowel change /a/i/) requires exact temporal coordination and sequencing of speech movements. By focusing on sub-lexical processing with a single vowel change, we minimized the influence of semantic and lexical components of speech processing. Our focus was on phonological processing as the basis for further research with patients. Furthermore, we decided for a vowel change with regard to our study with patients because even severely impaired patients are mostly able to cope with this task (Figure 1).

Stimuli consisted of quadruple measure groupings with duration of 4 sec. (8 vowel changes, alternately /a/i/) and differed as follows:

Condition (1): vowel changes with regular groupings: This condition comprises either no vowel change within one beat or the same tone durations and regular changes within one beat. From beat to beat tone durations change in even-numbered ratios.

These groupings allow smooth (legato) vocalization without break between notes.

Condition (2): vowel changes with regular groupings and rests: This condition contains the same tone durations and regular changes within the individual beat. Tone durations change from beat to beat in even-numbered ratios. The implementation of rests draws attention to forthcoming motor execution because legato (smooth) and staccato (short) vocalization have to be changed from beat to beat.

Condition (3): vowel changes with irregular groupings: This condition consists of varying durations (including odd numbers of syllables within one beat) and irregular changes. Tone durations change from beat to beat in odd-numbered ratios. Syncopations, dottings, and rests bring about accent shifts (off-beats), which increase demands on attention and short-term memory but also on subsequent sequencing. This assumption is corroborated by studies researching on movement sequence learning [78-81].

Figure 1 shows three musical notations in bass clef, 4/4 time. Example 1: A regular sequence of quarter notes: A, I, A, I, A, I, A, I. Example 2: A regular sequence of quarter notes with rests: A, I, A, I, A, I, A, I. Example 3: An irregular sequence of quarter notes: A, I, A, I, A, I, A, I.

Figure 1: Stimuli examples in musical notation Quadruple measure groupings with a duration of 4 sec. (1 beat per sec.) sung at a frequency of 220 Hz (A3) (1) Regular vowel changes (2) Regular vowel changes with rests (3) Irregular vowel changes.

Stimuli were sung by a female voice with the vowel change /a/i/ at a frequency of 220 Hz. The length of each stimulus was electronically set to 4 sec. with a max. deviation of 0.05 sec. For each condition (1 -3) four grouping variations with the same complexity level were available.

Subjects had to listen in anticipation i.e. they knew that they had to repeat the heard stimuli immediately after presentation.

We used an event-related design where a total of 25 trials per condition were presented and 25 null-events were randomly included.

The stimuli were presented in a pseudo-randomized order with a mean interstimulus interval of 9 sec. (jittered between 8 and 10 sec.). The presentation time for each stimulus took 4 sec.

The paradigm was implemented in Presentation (Neurobehavioral Systems) and synchronized to the scanner.

We want to emphasize that our tasks are also meant for following patient studies. Since they are essentially demanding with regard to cognitive abilities (attention, short-term memory), which are often impaired in patients with frontal lobe damage, a sparse temporal scanning design was not used in this study. We wanted to avoid attention loss and consequently lower functional response caused by relatively long inter-scan-intervals, which are required in sparse temporal schemes [82]. Moreover, stimuli were constantly sung at a frequency of 220 Hz, which is outside the main frequency peaks of the scanner spectrum.

Stimuli were presented binaurally through MR-compatible headphones with a sound absorption rate of 30 dBA (Resonance Technology). All conditions were performed with eyes closed.

Participants were instructed to listen to the stimuli and refrain from any motor or cognitive responses such as fingerlifting and silent counting.

Image acquisition

Functional images were obtained with a whole-body 3 T Siemens Trio MRI-system. Participants were fixated in the head coil using Velcro straps and foam paddings to stabilize head position and minimize motion artefacts. After orienting the axial slices in the anterior-posterior commissure (AC-PC) plane functional images were acquired using a T2*-weighted echo planar imaging (EPI) sequence with a repetition time (TR) of 2200 ms, an echo time (TE) of 30 ms and a flip angle (FA) of 90 degrees. 640 volumes consisting of 41 contiguous transversal slices with a thickness of 3.4 mm were measured. A 64 × 64 matrix with a field of view (FOV) of 220 mm was used, yielding an effective voxel size of 3.44 × 3.44 × 3.74 mm.

Image analysis

Functional images were pre-processed and analyzed using SPM8 (Wellcome Department of Cognitive Neurology London UK).

Image Pre-processing: Images were realigned and unwarped in order to correct for motion and movement-related changes in magnetic susceptibility. Translation and rotation correction did not exceed 1.7 mm and 1.6° respectively for any of the participants. Thereafter, images were spatially normalized into the anatomical space of the MNI brain template (Montreal Neurologic Institute) to allow pixel-by-pixel averaging across subjects with a voxel size of 4 × 4 × 4 mm in the x, y, and z dimensions. Finally, all images were smoothed using a Gaussian filter of 8 × 8 × 8 mm to accommodate intersubject variation in brain anatomy and to increase signal-to-noise ratio in the images.

Statistical analysis

First, a random effects analysis was performed to search for significantly activated voxels in the individual pre-processed data using the general linear model approach for time-series-data suggested by Friston and co-workers [83-86] and implemented in SPM8. Contrast-images were computed after applying the hemodynamic response function. A random effects analysis was performed using a one-sample t-test for each condition, where the reported contrasts are inclusively masked by the minuend with $p=0.05$ to eliminate deactivations of the subtrahend becoming significant because of the subtraction. Statistical

parametric maps (SPMs) were evaluated and voxels were considered significant if their corresponding linear contrast t-values were significant at a voxelwise threshold of $p=0.05$ (FWE-corrected). Only regions comprising at least 10 voxels are reported. Finally, coordinates of activations were transformed from MNI to Talairach space [87] using the matlab function `mni2tal.m` implemented by Matthew Brett (<http://www.mrc-cbu.cam.ac.uk/Imaging/mnispace.html>).

Results

Behavioral analysis

As already described, this study investigated anticipatory listening and subjects had to produce the heard stimuli in the latter portion of the experiment. The data from the second experiment (rhythmic singing) were commonly analyzed by 2 professional musicians (singer and percussionist) post hoc. They assessed each stimulus repetition with either correct (score 1) or incorrect (score 0). Only unanimous assessments that both matching the pitch and rhythm production had been performed without error were scored 1. Data were only included in the study if at least 75% of the stimuli were repeated correctly in each condition.

We put this high limit in order to make sure that differences regarding the results can be attributed to the specific rhythm structure and in order to eliminate activations, which resulted from incorrect reactions.

Using the Friedman-test for paired samples of non-normally distributed data no significant difference could be demonstrated concerning performance of the participants in the four grouping variations within each condition. However, complexity of the condition seems to influence the performance significantly. An additional Wilcoxon-test post-hoc analysis confirmed this result for all three paired comparisons (3-1; 2-1; 3-2) using a Bonferroni-corrected threshold of $p<0.05$.

Since anticipatory listening is a prerequisite for correct repetition of which the data were already published [27] we can be sure that subjects must have listened attentively.

FMRI data

Auditory stimulation was regarded as separately modeled condition in this design, reproduction data were already published [27]. Auditory presentation and reproduction were time-shifted; subjects did not sing along but after the presentation had stopped. Hence, only the expected activations in the auditory areas caused by the auditory stimulus presentation will be present in the reported results (Figure 2).

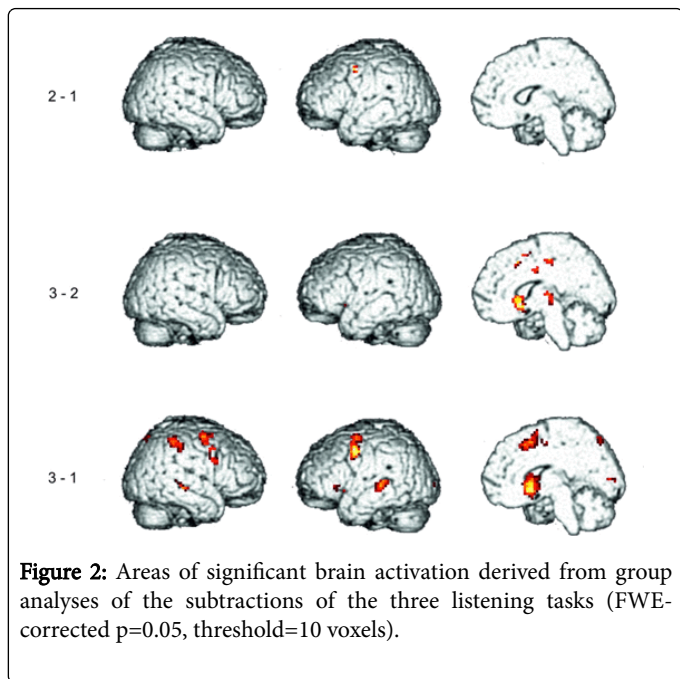


Figure 2: Areas of significant brain activation derived from group analyses of the subtractions of the three listening tasks (FWE-corrected $p=0.05$, threshold=10 voxels).

BA	left hemisphere					right hemisphere					
	x	y	z	cluster size k	z-scores	x	y	z	cluster size k	z-scores	
Subtraction (2) - (1)											
Frontal lobe											
Precentral gyrus	9 / 6	-45	4	42	13	5.81					
Subtraction (3) - (1)											
Frontal lobe											
Pre-SMA		-3	11	56	138	6.45	6	16	47	138	5.52
Precentral	6 / 9	-48	4	42	185	6.68	30	-5	52	131	6.29
Inferior frontal	47	-32	25	0	30	5.56					
Inferior frontal	45	-31	25	3	30	5.56					
Insula	13	-30	25	0	30	5.56					
Middle frontal	8						33	22	0	8	5.38
Cingulate gyrus	32 / 24						53	13	39	39	5.58
							21	8	47	131	5.73
Temporal lobe											
Middle temporal	22	-59	-30	2	59	6.32	59	-24	0	13	5.73
Superior temporal	42	-49	-31	10	59	6.32					
Transverse temporal	41	-47	-29	10	59	6.32					
Parietal lobe											
Inferior	40	-39	-40	39	10	5.15	45	-36	50	103	6.30
Superior	7	-15	-59	56	8	5.24	9	-65	54	13	5.18
Subcortical											
Putamen		-15	16	-5	162	7.76	15	16	-8	176	7.14
Caudate		-13	18	-5	162	7.76	13	16	-6	176	7.14
Subtraction (3) - (2)											
Frontal lobe											
Superior frontal	6						21	10	50	12	5.62
Cingulate gyrus	32 / 24	-2	19	41	11	5.20	3	-2	34	20	5.64
Precentral lobe	31	-6	-19	46	30	5.61	9	-13	46	30	5.21
Subcortical											
Putamen	40	-39	-40	39	10	5.15	45	-36	50	103	6.30
Caudate	7	-15	-59	56	8	5.24	9	-65	54	13	5.18
Thalamus							18	-21	7	25	5.79

Figure 3: Regions of peak activity. Coordinates are reported in Talairach space [87] BA=Brodmann Area; pre-SMA=pre-supplementary motor area.

Subtraction

To determine how neural activity differed between conditions subtractions between conditions were performed. Significant activations remained for the listening tasks concerning following subtractions:

Regions activated in the “regular with rests” condition compared to the “regular” condition

Subtraction of condition (2) minus (1) yielded significant activation of left precentral gyrus (BA 9, 6).

Regions activated in the “irregular” condition compared to the “regular” condition

Subtraction of condition (3) minus (1) yielded significant activation of bilateral putamen and caudate. Pre-Supplementary motor area (BA 6) was activated bilaterally. Precentral gyrus (BA 6,9) was activated more distinctly in the left hemisphere. Middle frontal gyrus (BAS 8) and cingulate gyrus (BA 32,24) were activated in the right hemisphere. While ventrolateral prefrontal cortex (BA 47,45) was activated in the left hemisphere, insula (BA 13) was activated bilaterally, however, most pronouncedly in the left hemisphere. The same holds true for middle temporal gyrus (BA 22), while superior and transverse temporal gyri (BA 42,41) were activated in the left hemisphere.

Inferior and superior parietal gyri (BA 40,7) were activated most pronouncedly in the right hemisphere.

Regions activated in the “irregular” condition compared to the “regular with rests” condition.

Subtraction of condition (3) minus (2) yielded significant activation of bilateral putamen and caudate. Thalamus was activated in the right hemisphere. Superior frontal gyrus (BA 6) was activated in the right hemisphere. Cingulate gyrus (BA 24) was activated bilaterally as well as paracentral lobe (BA 31).

Discussion

Our results demonstrate that rhythmic structure seems to influence auditory-motor interaction during anticipatory listening to monotonously sung vowel changes decisively. We investigated three grouping variations, which were sung as vowel changes with regular groupings, regular groupings with rests, and vowel changes with irregular groupings.

Rhythmic chanting (e.g. the vowel change /a/i/) requires exact temporal coordination and sequencing of speech movements. By focusing on sub-lexical processing, we minimized the influence of semantic and lexical components of speech processing. Our focus was on phonological processing as the basis for further research with patients. Furthermore, we decided for a vowel change with regard to our study with patients because even severely impaired patients are mostly able to cope with this task.

Since anticipatory listening is a prerequisite for correct repetition of which the data were already published [27], we can be sure that subjects must have listened attentively.

To the best of our knowledge, similarly demanding tasks using untrained material with varying demands on temporal grouping and short-term memory capacity have not been investigated as yet. Therefore, we will try to outline potential connections by discussing the main effects of complexity based on the involvement of specific brain regions.

Subcortical structures

A main effect of both subtractions (3) minus (2) and (3) minus (1) is the significant bilateral activation of the basal ganglia, however, subtraction (3) minus (1) yielded a z-score 50% higher than in subtraction (3) minus (2), indicating that irregularity seems to be a decisive factor.

The involvement of the basal ganglia in keeping internal representations of time is well documented from studies with healthy subjects [88,89] but also from research with patients suffering from Parkinson's disease (PD) [90-92]. In her review article, Schirmer [37] describes PD patients, who have problems detecting temporal cues in speech, in speech production, being too fast or too slow, and in using pauses during speaking.

According to the studies of Grahn and Brett [93] and Grahn [94], listening to rhythms in which internal beat generation is required results in greater basal ganglia activation than listening to rhythms in which beats are externally cued. Although it has to be mentioned, that these studies investigated neither singing nor rhythmic grouping.

All of our listening tasks demand generation of an internal beat as an aid to chunk the respective sequences and thus support memorizing. However, only subtractions from the most complex i.e. irregular rhythmic sequences resulted in basal ganglia activation, which implies that generating an internal beat while listening to irregular rhythmic groupings seems to be the decisive factor rather than internal beat generation per se. Possibly, this points to additional and specific demands on temporal grouping or sequencing capabilities concerning motor but also attentional control. According to Graybiel [95], the basal ganglia serve to organize neural activity in connection with action-oriented cognition and are involved in the chunking of action sequences [96]. Perception of predictable cues (regular beats, meter, and temporal chunks) is supposed to be closely linked to sequencing, an idea which is important for auditory language perception (for review see [97]).

Apart from the fact that grouping or chunking (i.e. bundling events together into larger units) serves to enhance maintenance of information in short-term and working memory [98-100], temporal-rhythmic chunking promotes speech motor processes by training intersyllabic programming, which is considered to play an important role in phonetic planning [101].

When applied in a musical context, it is important to mention that imaging studies indicate that regions activated in orienting attention in time for the most part overlap those underlying sequencing behaviors with basal ganglia activation in the service of attention (maintain representations of time intervals) as well as chunking of action representations [102].

Pre-Supplementary Motor Area (pre-SMA)

Further evidence for higher demands on sequence chunking due to increasing rhythmic complexity arises from additional activation of pre-SMA, which occurs only in subtraction (3) minus (1). This applies to the process of initiating new sets of sequences in contrast to automated sequences [78], and seems particularly clear when syncopation is required in contrast to synchronization with a metronome [81].

Since in our condition (3), we included syncopations but also dottings and rests the assumption of Mayville and colleagues seems to be convincing. The authors hypothesize that syncopations demand each motor response to be planned and executed separately. Synchronization as a continuous motor sequence, however, involves planning and execution of just a single plan of action adjusted by sensory feedback. Interestingly, the authors observed the same effect when only imagination of syncopation and synchronization with a metronome was required [103]. According to the authors, timing

strategy and planning seem to influence neural mechanisms decisively rather than motor execution.

Precentral Gyrus

Given additional activation of the precentral gyrus in subtraction (2) minus (1) and particularly in subtraction (3) minus (1), our results indicate that this area becomes activated already during anticipatory listening to chanted vowel changes. As an essential vocalization area, the precentral gyrus corresponds to the somatotopic orofacial region of the motor and premotor cortices including the large mouth and larynx representations. Brown et al. [104] report on activation of this specific region, called larynx area, which is commonly activated during passive perception, perceptual discrimination as well as vocal imagery, and vocal production and thus mediates audio-motor integration.

However, according to our results the degree of precentral gyrus activation seems to be task- dependent. Our subtraction (2) minus (1) underlines the difference between legato (smooth) and staccato (short) vocalization compared to regular, legato vocalization. Changing between legato and staccato vocalization requires precise initiation of articulatory movements especially in the initial phase of vocal preparation, which is reported to be dominated by the left hemisphere [105]. This is corroborated by our results for this subtraction with additional left hemisphere activation of precentral gyrus. Implementation of syncopations, dottings, and rests within one quadruple measure grouping like in our subtraction (3) minus (1) increases the demands regarding short-term memory as well as planning and programming for the upcoming motor execution. This might possibly explain the additional precentral gyrus activation most prominently in the left hemisphere

Ventrolateral prefrontal cortex

Our hypothesis of task-dependence is further supported by additional activations of left ventrolateral prefrontal cortex (BA 47,45) and the insula, most pronouncedly in the left hemisphere, which only occur in this subtraction.

Area 47 activation is reported from studies investigating working memory for pitch [73,106], auditory attention to space and frequency [107], divided attention for pitch and rhythm [108], meter discrimination [109] or rhythm production to visual inputs [110].

As regards our focus, it is important to note that studies searching for the parallels between music and language [53,111-113] have in common the description of BA 47 involvement in temporal perceptual organization. BA 47 seems to be activated when temporal expectancies and temporal coherence are involved [53,111].

The study of Chen et al. [6] is at least in parts comparable with our research because the authors also investigated listening to 3 types of rhythms, however, in two experiments. According to their results the ventral premotor cortex was only recruited when subjects listened in anticipation of tapping. In our study subjects also listened in anticipation, although in anticipation of rhythmic singing. The fact that only subtraction (3) minus (1) yielded activation of ventral premotor areas seems to indicate that listening to sub-lexical material in anticipation per se is not sufficient to cause these activations. Specific action-related demands caused by an irregular rhythmic structure possibly increased the demands on articulatory recoding capacity with regard to vocal planning. Additional activation in part of Broca's area may suggest that irregularity promotes sub-vocal

rehearsal, one of the main components of the articulatory or phonological loop, a subsystem of the working memory concept developed by Baddeley and colleagues [114-116]. Retention of verbal items in memory is supposed to be based on an interactive cooperation between a short-term “phonological store” and this inner rehearsal process, which refreshes the decaying memory traces in the phonological store [114].

Insular cortex

Furthermore, only subtraction (3) minus (1) yielded additional insula activation, most prominently in the left hemisphere. Left insula activation is reported from studies concerned with musical rhythm discrimination and attention to changes in rhythm respectively [108]. The authors report that rhythm judgment, in particular, whether the length of the intervals and notes in the given music sequence were regular or irregular resulted in activation of the left insula and left Broca's area. While activation of the left insula in the rhythm task is attributed to memory processing of sound sequences, activation of Broca's area (BA 6 and 44) is explained by the subjects' inner strategy to recall the sound by articulating.

Bamiou et al. ([51], see for review) emphasize the role of both insulae for many aspects of auditory temporal processing such as sequencing of sounds and detection of a moving sound. With regard to our research with patients, a further study of Bamiou et al. [52] provides evidence that insular lesions affect central auditory function, temporal processing and especially sequencing. As an integral component of the central auditory pathway the insula has connections to the auditory cortex and subcortex.

The important role of the insula in audio-vocal integration was explained in several studies concerned with singing [10,117,118], though, neither anticipatory listening to singing nor temporal aspects were investigated.

Riecker et al. [119] compared overt and covert continuous recitation of the months of a year and overt and covert reproduction of a well-known non-lyrical tune. The authors reach the conclusion that left insula supports coordination of speech articulation, while right insula mediates temporo-spatial control of vocal tract musculature during singing. Muscle coordination engaged in articulation and phonation was assumed to be the main reason for left insula activation rather than pre-articulatory processing since activation of the insula was restricted to overt tasks [120]. In accordance with the double filtering by frequency theory [121], the authors emphasize the role of the insular cortex operating across different time domains with a left hemisphere specialization for segmental information and a right hemisphere specialization for suprasegmental information i.e. intonation contours and musical melodies. Given, that only our subtraction (3) minus (1) yielded activation of ventrolateral prefrontal cortex areas in the left hemisphere and the insula, most pronouncedly in the left hemisphere, we assume that this may indicate pre-articulatory processing due to rhythm complexity. Our tasks consisted of sub-lexical material (vowel changes), which was sung on a single pitch. While pitch information and number of vowel changes were kept constant, temporal information changed resulting in increasing demands on auditory perception, short-term memory capacity and articulatory preparation with following motor execution in mind. Our assumption is supported by the study of Zaehle et al. [122], who investigated functional organization for sub-lexical auditory perception regarding auditory spectro-temporal processing in speech and non-speech sounds. Based on their results, the authors suggest

that segmental sub-lexical analysis of speech sounds but also segmental acoustic analysis of non-speech sounds with the same spectro-temporal characteristics involves the dorsal processing network.

Sylvian-parietal-temporal areas

Only subtraction (3) minus (1) yielded activation of middle, superior, and transverse temporal gyrus most distinctly in the left hemisphere. In our opinion, this might provide further evidence that specific motor and cognitive demands seem to be necessary to promote auditory-motor interaction during anticipatory listening. Our results also correspond with models of speech processing already mentioned in our introduction, describing the auditory dorsal pathway of speech perception, which maps sounds into articulatory-based representations [56-59,74,123]. Posterior auditory regions at the parieto-temporal boundary are suggested to be decisive nodes in the auditory-motor interface, serving linguistic but also non-linguistic processes like perception and production (covert humming) of tonal melodies [59]. Warren et al. [123] proposed a general model for auditory-motor transformations characterizing the dorsal stream as the auditory ‘dorsal pathway’. The planum temporale (PT), which occupies the superior temporal plane posterior to Heschl's gyrus, generally agreed to represent the auditory association cortex, is assumed to act as a computational hub [124], analyzing incoming complex sounds and transforming those that are of motor relevance into motor representations via the dorsal pathway in prefrontal, premotor and motor regions. Left posterior PT, known informally as Area Spt (Sylvian-parietal-temporal region), has been described as a cortical area, which is actively involved in integrating auditory inputs with vocal tract gestures [59,74,125]. Area Spt is involved in covert rehearsal in tests of phonological short-term memory [126,127] as well as in humming music and silent lip reading [125,128].

Since only subtraction (3) minus (1) also yielded these additional auditory-motor activations corresponding to the dorsal pathway, we assume that this further corroborates our hypothesis of task-dependence.

Conclusion

Our findings suggest that specific motor but also cognitive demands seem to be prerequisites for supportive paving of articulatory gestures during anticipatory listening to sung vowel changes. According to our results, sung vowel changes with an irregular rhythmic structure seem to meet these requirements. Irregularity is possibly a critical aspect because it resembles the temporal organization of spoken language most. In a further study we will elucidate potential influences in greater detail, for example, the relationship between rhythmic structure, grouping strategy, and phonological working memory. Nonetheless, our results demonstrate that auditory-motor interaction might be influenced by listening to specifically targeted vocal exercises, an evidence of particular importance for therapeutical implications, particularly in the treatment of patients suffering from AOS.

Provisional analyses of behavioral and imaging data of our ongoing study with patients indicate that such training seems to support short-term storage of sub-lexical material, possibly due to improved internal chunking of temporal events, resulting in improved temporal sequencing in the subsequent repetition.

References

1. Gaab N, Gaser C, Zaehle T, Jancke L, Schlaug G (2003) Functional anatomy of pitch memory--an fMRI study with sparse temporal sampling. *Neuroimage* 19: 1417-1426.
2. Brown S, Martinez MJ (2007) Activation of premotor vocal areas during musical discrimination. *Brain Cogn* 63: 59-69.
3. Halpern AR, Zatorre RJ (1999) When that tune runs through your head: a PET investigation of auditory imagery for familiar melodies. *Cereb Cortex* 9: 697-704.
4. Callan DE, Tsytarev V, Hanakawa T, Callan AM, Katsuhara M, et al. (2006) Song and speech: brain regions involved with perception and covert production. *Neuroimage* 31: 1327-1342.
5. Chen JL, Zatorre RJ, Penhune VB (2006) Interactions between auditory and dorsal premotor cortex during synchronization to musical rhythms. *Neuroimage* 32: 1771-1781.
6. Chen JL, Penhune VB, Zatorre RJ (2008) Listening to musical rhythms recruits motor regions of the brain. *Cereb Cortex* 18: 2844-2854.
7. Chen JL, Penhune VB, Zatorre RJ (2008) Moving on time: brain network for auditory-motor synchronization is modulated by rhythm complexity and musical training. *J Cogn Neurosci* 20: 226-239.
8. Chen JL, Penhune VB, Zatorre RJ (2009) The role of auditory and premotor cortex in sensorimotor transformations. *Ann N Y Acad Sci* 1169: 15-34.
9. Zatorre RJ, Chen JL, Penhune VB (2007) When the brain plays music: auditory-motor interactions in music perception and production. *Nat Rev Neurosci* 8: 547-558.
10. Zarate JM (2013) The neural control of singing. *Front Hum Neurosci* 7: 237.
11. Benton AL, Joynt RJ (1960) Early descriptions of aphasia. *Arch Neurol* 3: 205-222.
12. Keith RL, Aronson AE (1975) Singing as therapy for apraxia of speech and aphasia: report of a case. *Brain Lang* 2: 483-488.
13. Ustvedt HJ (1937) Über die Untersuchung der musikalischen Funktionen bei Patienten mit Gehirnleiden, besonders bei Patienten mit Aphasie. *Mercators Tryckeri, Acta Medica Scand, Suppl LXXXVI*, Helsingfors.
14. Yamadori A, Osumi Y, Masuhara S, Okubo M (1977) Preservation of singing in Broca's aphasia. *J Neurol Neurosurg Psychiatry* 40: 221-224.
15. Albert ML, Sparks RW, Helm NA (1973) Melodic intonation therapy for aphasia. *Arch Neurol* 29: 130-131.
16. Belin P, Van Eeckhout P, Zilbovicius M, Remy P, François C, et al. (1996) Recovery from nonfluent aphasia after melodic intonation therapy: a PET study. *Neurology* 47: 1504-1511.
17. Cohen, NS, Ford J (1995) The effect of musical cues on the nonpurposive speech of persons with aphasia. *J Music Ther*: 46-57.
18. Pilon MA, McIntosh KW, Thaut MH (1998) Auditory vs. visual speech timing cues as external rate control to enhance verbal intelligibility in mixed spastic-ataxic dysarthric speakers: a pilot study. *Brain Inj* 1: 793-803.
19. Schlaug G, Marchina S, Norton A (2008) From Singing to Speaking: Why Singing May Lead to Recovery of Expressive Language Function in Patients with Broca's Aphasia. *Music Percept* 25: 315-323.
20. Schlaug G, Marchina S, Norton A (2009) Evidence for plasticity in white-matter tracts of patients with chronic Broca's aphasia undergoing intense intonation-based speech therapy. *Ann N Y Acad Sci* 1169: 385-394.
21. Schlaug G, Norton A, Marchina S, Zipse L, Wan CY (2010) From singing to speaking: facilitating recovery from nonfluent aphasia. *Future Neurol* 5: 657-665.
22. Sparks R, Helm N, Albert M (1974) Aphasia rehabilitation resulting from melodic intonation therapy. *Cortex* 10: 303-316.
23. Sparks RW, Deck JW (1994) Melodic Intonation Therapy. In: Chapey, R. (Ed.), *Language intervention strategies in aphasic adults*. 3rd ed. Williams & Wilkins, Baltimore, USA. pp: 368-379.
24. Van Eeckhout P, Honrado C, Bhatt P, Deblais JC (1983) De la T.M.R. et de sa pratique. *Rééducation Orthophonique* 305-316.
25. Jungblut M (2005) Music therapy for people with chronic aphasia: a controlled study. In Aldridge, D. Ed.; *Music therapy and neurological rehabilitation*. Performing health. Jessica Kingsley Publishers London and Philadelphia: 189-211.
26. Jungblut M, Suchanek M, Gerhard H (2009) Long-term recovery from chronic global aphasia: a case report. *Music and Medicine* 1: 61-69.
27. Jungblut M, Huber W, Pustelniak M, Schnitker R (2012) The impact of rhythm complexity on brain activation during simple singing: an event-related fMRI study. *Restor Neurol Neurosci* 30: 39-53.
28. Jungblut M, Mais C, Huber W, Schnitker R (2014) Paving the way for speech: voice-training induced plasticity in chronic aphasia and apraxia of speech – three single cases. *Neural Plasticity*.
29. Efron R (1963) Temporal Perception Aphasia and D'e'ja vu. *Brain* 86: 403-424.
30. Liégeois-Chauvel C, de Graaf JB, Laguiton V, Chauvel P (1999) Specialization of left auditory cortex for speech perception in man depends on temporal coding. *Cereb Cortex* 9: 484-496.
31. Robin DA, Tranel D, Damasio H (1990) Auditory perception of temporal and spectral events in patients with focal left and right cerebral lesions. *Brain Lang* 39: 539-555.
32. Tallal P, Miller S, Fitch RH (1995) Neurobiological basis of speech: a case of preeminence of temporal processing. *Ir J Psychol* 16: 195-219.
33. Baum SR, Boyczuk JP (1999) Speech timing subsequent to brain damage: effects of utterance length and complexity. *Brain Lang* 67: 30-45.
34. Danly M, Shapiro B (1982) Speech prosody in Broca's aphasia. *Brain Lang* 16: 171-190.
35. Maas E, Robin DA, Wright DL, Ballard KJ (2008) Motor programming in apraxia of speech. *Brain Lang* 106: 107-118.
36. Wambaugh JL, Martinez AL (2000) Effects of rate and rhythm control treatment on consonant production accuracy in apraxia of speech. *Aphasiology* 14: 851-871.
37. Schirmer A (2004) Timing speech: a review of lesion and neuroimaging findings. *Brain Res Cogn Brain Res* 21: 269-287.
38. Altenmüller E (2001) How many music centers are in the brain? In: Zatorre, R.J. & Peretz, I. (Edn.), *The biological foundations of music*. The New York Academy of Sciences: New York, Vol. 930, pp. 273-281.
39. Clarke EF (1999) Rhythm and timing in music. In: Deutsch, D (Edn.) *The psychology of music*, 2nd ed. Academic Press: San Diego, pp: 473-500.
40. Gabrielsson A (1993). The complexities of rhythm. In: Thige, T.J., Dowling, W.J. (Eds.), *Psychology and music. The understanding of melody and rhythm*. Lawrence Erlbaum Associates, Publishers: Hillsdale, New Jersey, pp: 93-121.
41. Lehrdahl F, Jackendoff R (1983) *A Generative Theory of Tonal Music*. MIT Press: Cambridge, MA.
42. Povel DJ, Essens P (1985) Perception of temporal patterns. *Music Percept* 2: 411- 440.
43. Todd NP, McAngus (1994) The auditory "primal sketch": a multiscale model of rhythmic grouping. *J New Music Res* 2: 25-70.
44. Di Pietro M, Laganaro M, Leemann B, Schnider A (2004) Receptive amusia: temporal auditory processing deficit in a professional musician following a left temporo-parietal lesion. *Neuropsychologia* 4: 868-877.
45. Vignolo LA (2003) Music agnosia and auditory agnosia. Dissociations in stroke patients. *Ann N Y Acad Sci* 999: 50-57.
46. Fries W, Swihart AA (1990) Disturbance of rhythm sense following right hemisphere damage. *Neuropsychologia* 28: 1317-1323.
47. Wilson SJ, Pressing JL, Wales RJ (2002) Modelling rhythmic function in a musician post-stroke. *Neuropsychologia* 40: 1494-1505.
48. Stahl B, Kotz SA, Henseler I, Turner R, Geyer S (2011) Rhythm in disguise: why singing may not hold the key to recovery from aphasia. *Brain* 134: 3083-3093.
49. Bradt J, Magee WL, Dileo C, Wheeler BL, McGiloway E (2010) Music therapy for acquired brain injury. *Cochrane Database Syst Rev* : CD006787.

50. Jungblut M (2009) SIPARI® A music therapy intervention for patients suffering with chronic, nonfluent aphasia. *Music and Medicine* 1: 102-105.
51. Bamiou DE, Musiek FE, Luxon LM (2003) The insula (Island of Reil) and its role in auditory processing. Literature review. *Brain Res Brain Res Rev* 42: 143-154.
52. Bamiou DE, Musiek FE, Stow I, Stevens J, Cipolotti L, et al. (2006) Auditory temporal processing deficits in patients with insular stroke. *Neurology* 67: 614-619.
53. Levitin DJ (2009) The neural correlates of temporal structure in music. *Music and Medicine* 1: 9-13.
54. Dronkers NF (1996) A new brain region for coordinating speech articulation. *Nature* 384: 159-161.
55. Ogar J, Slama H, Dronkers N, Amici S, Gorno-Tempini ML (2005) Apraxia of speech: an overview. *Neurocase* 11: 427-432.
56. Hickok G, Poeppel D (2004) Dorsal and ventral streams: a framework for understanding aspects of the functional anatomy of language. *Cognition* 92: 67-99.
57. Hickok G, Poeppel D (2007) The cortical organization of speech processing. *Nat Rev Neurosci* 8: 393-402.
58. Buchsbaum BR, Olsen RK, Koch P, Berman KF (2005) Human dorsal and ventral auditory streams subservise rehearsal-based and echoic processes during verbal working memory. *Neuron* 48: 687-697.
59. Hickok G, Buchsbaum B, Humphries C, Muftuler T (2003) Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J Cogn Neurosci* 15: 673-682.
60. Breier JI, Juranek J, Maher LM, Schmadeke S, Men D, et al. (2009) Behavioral and neurophysiologic response to therapy for chronic aphasia. *Arch Phys Med Rehabil* 90: 2026-2033.
61. Cardebat D, Demonet JF, De Boissezon X, Marie N, Marie RM, et al. (2003) Behavioral and neurofunctional changes over time in healthy and aphasic subjects: a PET language activation study. *Stroke* 34: 2900-2906.
62. Crosson B, McGregor K, Gopinath KS, Conway TW, Benjamin M, et al. (2007) Functional MRI of language in aphasia: a review of the literature and the methodological challenges. *Neuropsychol Rev* 17: 157-177.
63. Karbe H, Thiel A, Weber-Luxemburger G, Herholz K, Kessler J, et al. (1998) Brain plasticity in poststroke aphasia: what is the contribution of the right hemisphere? *Brain Lang* 64: 215-230.
64. Sauer D, Lange R, Baumgaertner A, Schraknepper W (2006) Dynamics of language reorganization after stroke. *Brain* 129: 1371-1384.
65. Stahl B, Henseler I, Turner R, Geyer S, Kotz SA (2013) How to engage the right brain hemisphere in aphasics without even singing: evidence for two paths of speech recovery. *Frontiers in Human Neuroscience*.
66. Levelt WJM, Roelofs A, Meyer AS (1999) A theory of lexical access in speech production. *Behav Brain Sci*: 1-75.
67. Ziegler W (2002) Psycholinguistic and motor theories of apraxia of speech. *Semin Speech Lang* 23: 231-244.
68. Liberman AM, Mattingly IG (1985) The motor theory of speech perception revised. *Cognition* 21: 1-36.
69. Fadiga L, Craighero L, Buccino G, Rizzolatti G (2002) Speech listening specifically modulates the excitability of tongue muscles: a TMS study. *Eur J Neurosci* 15: 399-402.
70. Rizzolatti G, Craighero L (2004) The mirror-neuron system. *Annu Rev Neurosci* 27: 169-192.
71. Burton MW, Small SL, Blumstein SE (2000) The role of segmentation in phonological processing: an fMRI investigation. *J Cogn Neurosci* 12: 679-690.
72. Wilson SM, Saygin AP, Sereno MI, Iacoboni M (2004) Listening to speech activates motor areas involved in speech production. *Nat Neurosci* 7: 701-702.
73. Zatorre RJ, Evans AC, Meyer E, Gjedde A (1992) Lateralization of phonetic and pitch discrimination in speech processing. *Science* 256: 846-849.
74. Hickok G (2012) The cortical organization of speech processing: feedback control and predictive coding the context of a dual-stream model. *J Commun Disord* 45: 393-402.
75. Shaywitz BA, Shaywitz SE, Pugh KR, Constable RT, Skudlarski P, et al. (1995) Sex differences in the functional organization of the brain for language. *Nature* 37: 607-609.
76. Hough MS, Daniel HJ, Snow MA, O'Brien KE, Hume WG (1994) Gender differences in laterality patterns for speaking and singing. *Neuropsychologia* 32: 1067-1078.
77. Weis S, Hausmann M, Stoffers B, Vohn R, Kellermann T (2008) Estradiol modulates functional brain organization during the menstrual cycle: an analysis of interhemispheric inhibition. *J Neurosci* 28: 13401-13410.
78. Kennerley SW, Sakai K, Rushworth MF (2004) Organization of action sequences and the role of the pre-SMA. *J Neurophysiol* 91: 978-993.
79. Lewis PA, Miall RC (2003) Distinct systems for automatic and cognitively controlled time measurement: evidence from neuroimaging. *Curr Opin Neurobiol* 13: 250-255.
80. Macar F, Lejeune H, Bonnet M, Ferrara A, Pouthas V, et al. (2002) Activation of the supplementary motor area and of attentional networks during temporal processing. *Exp Brain Res* 142: 475-485.
81. Mayville JM, Jantzen KJ, Fuchs A, Steinberg FL, Kelso JAS (2002) Cortical and subcortical networks underlying syncopated and synchronized coordination revealed using fMRI. *Hum Brain Mapp* 17: 214-229.
82. Shah NJ, Steinhoff S, Mirzazade S, Zafiris O, Grosse-Ruyken ML (2000) The effect of sequence repeat time on auditory cortex stimulation during phonetic discrimination. *NeuroImage* 1:100-108.
83. Friston KJ, Frith CD, Frackowiak RS, Turner R (1995) Characterizing dynamic brain responses with fMRI: a multivariate approach. *Neuroimage* 2: 166-172.
84. Friston KJ, Frith CD, Turner R, Frackowiak RS (1995) Characterizing evoked hemodynamics with fMRI. *Neuroimage* 2: 157-165.
85. Friston KJ, Holmes AP, Poline JB, Grasby PJ, Williams SC, et al. (1995) Analysis of fMRI time-series revisited. *Neuroimage* 2: 45-53.
86. Friston KJ, Zarahn E, Josephs O, Henson RN, Dale AM (1999) Stochastic designs in event-related fMRI. *Neuroimage* 10: 607-619.
87. Talairach P, Tournoux J (1988) Co-planar stereotaxic atlas of the human brain. Thieme Verlag, Stuttgart, New York.
88. Belin P, McAdams S, Thivard L, Smith B, Savel S, et al. (2002) The neuroanatomical substrate of sound duration discrimination. *Neuropsychologia* 40: 1956-1964.
89. Rao SM, Harrington DL, Haaland KY, Bobholz JA, Cox RW, et al. (1997) Distributed neural systems underlying the timing of movements. *J Neurosci* 17: 5528-5535.
90. Artieda J, Pastor MA, Lacruz F, Obeso JA (1992) Temporal discrimination is abnormal in Parkinson's disease. *Brain* 115 Pt 1: 199-210.
91. Grahn JA, Brett M (2009) Impairment of beat-based rhythm discrimination in Parkinson's disease. *Cortex* 45: 54-61.
92. Harrington DL, Haaland KY, Knight RT (1998) Cortical networks underlying mechanisms of time perception. *J Neurosci* 18: 1085-1095.
93. Grahn JA, Brett M (2007) Rhythm and beat perception in motor areas of the brain. *J Cogn Neurosci* 19: 893-906.
94. Grahn JA (2009) The role of basal ganglia in beat perception. *The Neurosciences and Music III – Disorders and Plasticity: Ann N Y Acad Sci* 1169: 35-45.
95. Graybiel AM (1997) The basal ganglia and cognitive pattern generators. *Schizophr Bull* 23: 459-469.
96. Graybiel AM (1998) The basal ganglia and chunking of action repertoires. *Neurobiol Learn Mem* 70: 119-136.
97. Kotz SA, Schwartze M, Schmidt-Kassow M (2009) Non-motor basal ganglia functions: a review and proposal for a model of sensory predictability in auditory language perception. *Cortex* 45: 982-990.
98. Miller GA (1956) The magical number seven, plus or minus two: some limits on our capacity for processing information. *The Psychological Review* 63: 81-97.

99. Glenberg AM, Jona M (1991) Temporal coding in rhythm tasks revealed by modality effects. *Mem Cognit* 19: 514-522.
100. Payne MC, Holzman TG (1986) Rhythm as a factor in memory. In Evans, J.R.; Clynes, M., Edn.; *Rhythm in psychological, linguistic and musical processes*. Charles C Thomas Publisher: Springfield, IL, USA: 41-54.
101. Deger K, Ziegler W (2002) Speech motor programming in apraxia of speech. *J Phon* 200: 321-335.
102. Janata P, Grafton ST (2003) Swinging in the brain: shared neural substrates for behaviors related to sequencing and music. *Nat Neurosci* 6: 682-687.
103. Mayville JM, Cheyne D, Deecke L, Ding M, Fuchs A (2000) Desynchronization of MEG (15-30 Hz) associated with overt and imagined sensorimotor coordination reflects task difficulty. *Soc Neurosci Abstr* 26: 2214.
104. Brown S, Ngan E, Liotti M (2008) A larynx area in the human motor cortex. *Cereb Cortex* 18: 837-845.
105. Jürgens U (2002) Neural pathways underlying vocal control. *Neurosci Biobehav Rev* 26: 235-258.
106. Zatorre RJ, Evans AC, Meyer E (1994) Neural mechanisms underlying melodic perception and memory for pitch. *J Neurosci* 14: 1908-1919.
107. Zatorre RJ, Mondor TA, Evans AC (1999) Auditory attention to space and frequency activates similar cerebral systems. *Neuroimage* 10: 544-554.
108. Platel H, Price C, Baron JC, Wise R, Lambert J, et al. (1997) The structural components of music perception. A functional anatomical study. *Brain* 120: 229-243.
109. Parsons LM (2001) Exploring the functional neuroanatomy of music performance, perception, and comprehension. *Ann N Y Acad Sci* 930: 211-231.
110. Penhune VB, Zatorre RJ, Evans AC (1998) Cerebellar contributions to motor timing: a PET study of auditory and visual rhythm reproduction. *J Cogn Neurosci* 10: 752-765.
111. Brown S, Martinez MJ, Parsons LM (2006) Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *Eur J Neurosci* 23: 2791-2803.
112. Levitin DJ, Menon V (2003) Musical structure is processed in "language" areas of the brain: a possible role for Brodmann Area 47 in temporal coherence. *Neuroimage* 20: 2142-2152.
113. Steinbeis N, Koelsch S (2008) Shared neural resources between music and language indicate semantic processing of musical tension-resolution patterns. *Cereb Cortex* 18: 1169-1178.
114. Baddeley AD (1986). *Working memory* (Oxford, Oxfordshire: Clarendon Press).
115. Baddeley A (1998) Recent developments in working memory. *Curr Opin Neurobiol* 8: 234-238.
116. Baddeley AD, Hitch GJ (1974) Working memory. In: *The Psychology of Learning and Motivation*, G. Bower, edn. (New York, NY: Academic press): 47-90.
117. Zarate JM, Zatorre RJ (2008) Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage* 40: 1871-1887.
118. Zarate JM, Wood S, Zatorre RJ (2010) Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia* 48: 607-618.
119. Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W (2000). Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *NeuroReport* 11: 1997-2000.
120. Ackermann H, Riecker A (2004) The contribution of the insula to motor aspects of speech production: a review and a hypothesis. *Brain Lang* 89: 320-328.
121. Ivry RB, Robertson LC (1998) *The Two Sides of Perception*. MIT Press, Cambridge, MA.
122. Zaehle T, Geiser E, Alter K, Jäncke L, Mayer M (2008) Segmental processing in the human auditory dorsal stream. *Brain Research* 1220: 170-190.
123. Warren JE, Wise RJS, Warren JD (2005) Sounds do-able: auditory-motor transformations and the posterior temporal plane. *Trends Neurosci* 28: 636-643.
124. Griffiths TD, Warren JD (2002) The planum temporal as a computational hub. *Trends Neurosci* 25: 348-353.
125. Hickok G, Okada K, Serences JT (2009) Area Spt in the human planum temporale supports sensory-motor integration for speech processing. *J Neurophysiol* 101: 2725-2732.
126. Jacquemot C, Scott SK (2006) What is the relationship between phonological short-term memory and speech processing? *Trends Cogn Sci* 10: 480-486.
127. Koelsch S, Schulze K, Sammler D, Fritz T, Müller K, et al. (2009) Functional architecture of verbal and tonal working memory: an fMRI study. *Hum Brain Mapp* 30: 859-873.
128. Pa J, Hickok G (2008) A parietal-temporal sensory-motor integration area for the human vocal tract: evidence from an fMRI study of skilled musicians. *Neuropsychologia* 46: 362-368.